

Universiteit
Leiden
The Netherlands

# Virtual Neanderthals : a study in agent-based modelling Late Pleistocene hominins in western Europe

Scherjon, F.

**Citation**

Scherjon, F. (2019, May 28). *Virtual Neanderthals : a study in agent-based modelling Late Pleistocene hominins in western Europe*. Global Academic Press, Vianen. Retrieved from https://hdl.handle.net/1887/73639

| | |
|---|---|
| Version: | Not Applicable (or Unknown) |
| License: | Leiden University Non-exclusive license |
| Downloaded from: | https://hdl.handle.net/1887/73639 |

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden




The handle http://hdl.handle.net/1887/73639 holds various files of this Leiden University dissertation.

**Author**: Scherjon, F.
**Title**: Virtual Neanderthals : a study in agent-based modelling Late Pleistocene hominins in western Europe
**Issue Date**: 2019-05-28

# Virtual Neanderthals

A study in agent-based modelling
Late Pleistocene hominins in western Europe

**Fulco Scherjon**



# Virtual Neanderthals

## A study in agent-based modelling Late Pleistocene hominins in western Europe

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 28 mei 2019
klokke 15.00 uur

door

**Fulco Scherjon**

Geboren te Amsterdam, Nederland
in 1969

**Promotor**:

Prof. Dr. Wil Roebroeks (Universiteit Leiden)

**Co-promotor:**

Dr. Katherine MacDonald (Universiteit Leiden)

**Promotiecommissie:**

Prof. Dr. Corrie C. Bakels (Universiteit Leiden)

Prof. Dr. Paul Bons (Eberhard Karls University Tübingen, Germany)

Dr. Hans Kamermans (Universiteit Leiden)

Prof. Dr. Jan Kolen (Universiteit Leiden)

Dr. Iza Romanowska (Centro Nacional de Supercomputación, Barcelona, Spain

Prof. Dr. Marie Soressi (Universiteit Leiden)

Dr. Alexander Verpoorte (Universiteit Leiden)

## PREFACE

Wil Roebroeks has been kind enough to allocate part of his SPINOZA grant from the Netherlands Organisation for Scientific Research (NWO) to make this research possible. The research described in this thesis was carried out under the supervision of Prof. Dr. Wil Roebroeks, Dr. Kathy MacDonald, Dr. Alexander Verpoorte, and Dr. Hans Kamermans, from the Faculty of Archaeology of Leiden University, The Netherlands. I hereby certify that the research presented in this thesis is my own original work except where specific acknowledgments or references are made. All data and information including full source code necessary for confirming the conclusions presented in the thesis are represented fully within this thesis or in the Supplementary Materials on the accompanying data disk. The title of this thesis is shortened to 'Virtual Neanderthals'. Virtual as in "being on or simulated on a computer or computer network", as defined by the Merriam-Webster dictionary on-line: https://www.merriam-webster.com/dictionary/virtual, accessed 25 February 2018.

Fulco Scherjon
Leiden

**Cover design**

The front page is the illustration 'Cyborg with human skull in his hand', by Tatiana Shepeleva, used with permission and modified by me. The skull studied by the cyborg is a late *Homo neanderthalensis* skull and mandible from western Europe (La Ferrassie I), skull photo credited to Puwadol Jaturawutthichai. Photo composition designed using a blue sky picture as background (from ZaZa Studio, 'Blue sky with clouds and sun'). Reverse depicts an early hunter facing a mammoth: 'Mammoth and primitive man against a starry sky', illustration by Iurii. All images used for the cover retrieved from https://www.shutterstock.com, accessed 10 April 2019.

**Copyright**

The copyright of the content of this thesis and the associated data carrier remains with the author. All information derived from this thesis and/or the supporting information including software and data files must be acknowledged appropriately. The HominінSpace modelling and simulation software is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, any version 3.

**ISBN/EAN:** 978-94-6380-344-1

**Publisher:** Global Academic Press

**Printed by:** ProefschriftMaken || www.proefschriftmaken.nl

**Cover realization:** ProefschriftOntwerp.nl

This work is dedicated to my son, Thor Scherjon, whose innocence in modelling the world and creativity in deconstructing the true meanings of life keep surprising me.

# Contents

**Extended contents** ..... **ix**

**List of Tables** ..... **xiii**

**List of Figures** ..... **xv**

**List of Equations** ..... **xviii**

**1. Introduction** ..... **1**

**Part one: The Lay of the Land** ..... **9**

**2. Agent Based Modelling** ..... **11**

**3. Case Study: The Late Pleistocene in western Europe** ..... **25**

**Part two: Creating the Actors** ..... **35**

**4. The Modelled Environment** ..... **37**

**5. The Modelled Hominins** ..... **65**

**6. The Archaeological Data in the Model** ..... **85**

**Part Three: Setting the Stage** ..... **97**

**7. Overview, Design concepts and Details (ODD) of HominínSpace** ..... **99**

**8. Building the Simulation System** ..... **117**

**9. Activating Parts of the Model via Settings** ..... **143**

**Part Four: And Action!** ..... **161**

**10. Simulation Results** ..... **163**

**11. Analysis** ..... **187**

**12. Discussion** ..... **225**

**13. Conclusions** ..... **251**

**References** ..... **267**

**Appendix 1: Abbreviations** ..... **293**

**Appendix 2: Supplementary Materials** ..... **297**

**Summary** ..... **307**

**Samenvatting** ..... **308**

**Acknowledgements** ..... **309**

**Curriculum vitae** ..... **311**

# Extended contents

**List of Tables** ..... **xiii**

**List of Figures** ..... **xv**

**List of Equations** ..... **xviii**

**1. Introduction** ..... **1**

1.1 Background ..... 1
1.2 Study objectives ..... 2
1.3 Methods and approaches ..... 3
1.4 Thesis structure ..... 6

**Part one: The Lay of the Land** ..... **9**

**2. Agent Based Modelling** ..... **11**

2.1 Introduction ..... 11
2.2 Why model? ..... 12
*2.2.1 In or out: what to include in a model* ..... *14*
*2.2.2 Model credibility* ..... *14*
2.3 Techniques for modelling dispersal in landscapes ..... 16
*2.3.1 Agent Based Modelling* ..... *17*
2.4 HomininSpace – a model of the past world of hunter-gatherers ..... 19

**3. Case Study: The Late Pleistocene in western Europe** ..... **25**

3.1 Introduction ..... 25
3.2 The environment: climate and topography ..... 26
3.3 The hominins: Neanderthals ..... 31

**Part two: Creating the Actors** ..... **35**

**4. The Modelled Environment** ..... **37**

4.1 Introduction ..... 37
4.2 Reconstructing topography with fluctuating global sea levels ..... 39
4.3 Distribution of reconstructed temperature and precipitation ..... 44
*4.3.1 Reconstruction via interpolation between climate extremes* ..... *47*
4.4 An energy landscape in two forms ..... 52
*4.4.1 Calculating primary productivity* ..... *54*
*4.4.2 Secondary biomass: the continuous energy model* ..... *56*
*4.4.3 Secondary biomass: the reconstructed habitat model* ..... *57*
*4.4.4 The number of edible ungulates: carrying capacity* ..... *61*

**5. The Modelled Hominins** ..... **65**

5.1 Introduction ..... 65
5.2 Demography ..... 67
5.3 What does it take to be a Neanderthal - energy requirements ..... 72
*5.3.1 Temperature tolerance* ..... *75*
5.4 Are Hominins social animals? - hominin groups ..... 76
*5.4.1 Group demographics - new groups, group extinction, merging of groups* ..... *76*
*5.4.2 Social interaction - home range and foraging range* ..... *78*
*5.4.3 Mobility* ..... *80*
5.5 Overview: building blocks for the underlying model ..... 81

**6. The Archaeological Data in the Model ..... 85**

6.1 Introduction ..... 85
6.2 The archaeology of presence and absence ..... 86
6.3 What dating information to use? ..... 87
*6.3.1 Database structure ..... 88*
6.4 Checkpoints in Space and Time (CSTs) ..... 91

**Part Three: Setting the Stage ..... 97**

**7. Overview, Design concepts and Details (ODD) of HominSpace ..... 99**

7.1 Introduction ..... 99
7.2 Process overview and scheduling ..... 99
7.3 Simulation time and stochasticity ..... 102
7.4 Model parameters and the ViabilityIndex ..... 103
7.5 How to match the archaeology ..... 105
*7.5.1 Translating simulation results into numbers: MatchedIntervalCoverage ..... 106*
7.6 Genetic Algorithms - searching through N-dimensions ..... 108
7.7 Traversing the parameter space ..... 111
*7.7.1 Generating the Standard parameter value set ..... 111*
*7.7.2 Searching the highest scoring parameter value set with Genetic Algorithms ..... 112*

**8. Building the Simulation System ..... 117**

8.1 Introduction ..... 117
8.2 Agent-based modelling tools ..... 117
*8.2.1 Repast Simphony version 2.2 ..... 118*
8.3 The development environment ..... 120
*8.3.1 The source code ..... 120*
*8.3.2 Software development tools ..... 122*
8.4 The runtime environment ..... 123
*8.4.1 Batch execution of simulations ..... 124*
8.5 Data files defining a case study ..... 126
*8.5.1 Simulation grid, topography and climate data ..... 126*
*8.5.2 Reconstructed global mean temperature and sea level ..... 127*
*8.5.3 Checkpoints ..... 129*
8.6 Input files defining a simulation ..... 133
*8.6.1 Model parameter values ..... 134*
*8.6.2 Simulation settings ..... 134*
8.7 Simulation output ..... 135
*8.7.1 Logging and debug information ..... 136*
*8.7.2 Maps, bar charts and time series ..... 137*
*8.7.3 Simulation statistics ..... 141*

**9. Activating Parts of the Model via Settings ..... 143**

9.1 Introduction ..... 143
9.2 Model elements of the reconstructed environment ..... 144
*9.2.1 Energy versus habitat reconstruction ..... 144*
*9.2.2 Coastal resources ..... 146*
9.3 Model elements of hominin behaviour ..... 147
*9.3.1 Mobility type: dynamic versus static hominins ..... 147*
*9.3.2 Imposing a maximum foraging range ..... 149*
*9.3.3 Allowing the crossing of open water systems ..... 151*
9.4 Model elements of population distribution ..... 153
*9.4.1 Using hominin absence data ..... 153*
*9.4.2 Using population core areas (hominin factories) ..... 156*
9.5 Neutral models - levelling the playing field ..... 157

*9.5.1 Implementing topographical differences in the landscape ... 158*
*9.5.2 Distributing energy equally throughout the grid ... 159*
*9.5.3 Randomly walking hominins ... 159*

**Part Four: And Action! ... 161**

**10. Simulation Results ... 163**

10.1 Overview ... 163
*10.1.1 Design of the Experiments (DOE) ... 165*
10.2 Scenario Habitat-A (HabitatDynamic) ... 166
10.3 Scenario Habitat-B (HabitatDynamicAbsence) ... 167
10.4 Scenario Habitat-C (HabitatDynamicMaxrange) ... 168
10.5 Scenario Habitat-D (HabitatDynamicMaxrangeCoreAbsence) ... 169
10.6 Scenario Habitat-E (HabitatDynamicCoastalMaxrangeCore) ... 170
10.7 Scenario Habitat-F (HabitatDynamicCoastalMaxrangeCoreAbsence) ... 171
10.8 Scenario Habitat-G (HabitatStatic) ... 172
10.9 Scenario Habitat-H (HabitatStaticCrosswaterCore) ... 173
10.10 Scenario Habitat-I (HabitatStaticMaxrange) ... 174
10.11 Scenario Habitat-J (HabitatStaticCoastalMaxrange) ... 175
10.12 Scenario Energy-A (EnergyDynamic) ... 176
10.13 Scenario Energy-B (EnergyDynamicCoastalMaxrange) ... 177
10.14 Scenario Energy-C (EnergyDynamicCoastalMaxrangeAbsence) ... 178
10.15 Scenario Energy-D (EnergyDynamicCoastalMaxrangeCrosswater) ... 179
10.16 Scenario Energy-E (EnergyStaticCoastalMaxrange) ... 180
10.17 Additional simulations ... 181
*10.17.1 Three simulation scenarios with Random walk activated ... 181*
*10.17.2 Duplication of scenario Energy-A for verification purposes ... 184*

**11. Analysis ... 187**

11.1 Introduction ... 187
11.2 Correlating input to input and output to output ... 188
11.3 Sensitivity analysis in scenario exploration ... 190
*11.3.1 Variability in scenarios ... 191*
*11.3.2 Correlations between parameter values and simulation results ... 194*
*11.3.3 Relative Variable Importance (RVI) - what makes a Neanderthal tick? ... 198*
*11.3.4 Monotonic trends in parameter values ... 200*
*11.3.5 Creating impossible Neanderthals and other cheats ... 201*
11.4 Trends in the simulation output ... 203
*11.4.1 The simulation score ... 203*
*11.4.2 Characterizing the best models from the Standard set ... 206*
*11.4.3 Causes of life and death ... 207*
*11.4.4 Influence of data quality on the simulation score ... 212*
11.5 Analysing the modelled environment ... 213
*11.5.1 Topography ... 214*
*11.5.2 Energy via habitat reconstruction ... 218*
*11.5.3 Energy via direct extrapolation ... 221*

**12. Discussion ... 225**

12.1 Introduction ... 225
12.2 System development ... 225
*12.2.1 Characterizing Neanderthal presence using radiometrically dated archaeology 227*
*12.2.2 Genetic Algorithms – do we need them? ... 229*
*Realistic ... 232*
*12.2.3 environment reconstruction ... 232*
*12.2.4 Probability in HomininSpace ... 233*
12.3 Answering questions with HomininSpace ... 235

*12.3.1 Parameter value evolution in all scenarios* ... *236*
*12.3.2 Q1 - How does the Energy reconstruction of the environment compare to Habitat reconstruction?* ... *238*
*12.3.3 Q2 - What is the effect of adding coastal resources to the model?* ... *239*
*12.3.4 Q3 - How does the addition of ebb and flow dispersal with the available resources influences the match with the archaeology?* ... *241*
*12.3.5 Q4 - What is an optimal maximum foraging range?* ... *242*
*12.3.6 Q5 - What happens if we assume that Neanderthals were able to cross large open water systems?* ... *244*
*12.3.7 Q6 - What influence has adding absence data to the model?* ... *245*
*12.3.8 Q7 – What is the effect of using population core areas?* ... *246*
*12.3.9 Q10 - What are the results if the Neanderthals are implemented with random movement?* ... *248*

**13. Conclusions** ... **251**

13.1 Developing HomininSpace – a tool for modelling hominin mobility ... 253
*13.1.1 Usability of the tool* ... *254*
*13.1.2 Who were those Neanderthals?* ... *256*
13.2 Implications of this study ... 259
13.3 Avenues for further research ... 261
*13.3.1 Recommendations when reconstructing past environments* ... *262*
*13.3.2 Suggestions for simulating hominins* ... *263*
13.4 Last words ... 265

**References** ... **267**

**Appendix 1: Abbreviations** ... **293**

***Abbreviations*** ... *293*
***Bioclimatic parameters*** ... *294*
***Biome (Mega) classifications*** ... *294*

**Appendix 2: Supplementary Materials** ... **297**

***Contents of the HomininSpace data disk*** ... *297*
***Poster: Neandertals on the move. Or not. – ESHE 2015*** ... *301*
***Poster: How (not) to model Neandertal extinction. –ESHE 2016*** ... *302*
***Comparing past climate models: CCSM versus MIROC*** ... *303*

**Summary** ... **307**

**Samenvatting** ... **308**

**Acknowledgements** ... **309**

**Curriculum vitae** ... **311**

# LIST OF TABLES

Table 1: Common characteristics of agents in Agent Based Modelling. ..... 18
Table 2: Local and global minimum and maximum temperatures from Bintanja and van de Wal (2008), supplementary information. Temperatures are offset in degrees Celsius relative to modern day yearly mean values. ..... 46
Table 3: Color coding of grid cells for temperature and precipitation. ..... 49
Table 4: Example values for primary productivity for given temperatures. ..... 54
Table 5: Different types of climate and their defining boundaries, derived from Binford (2001), Table 4.04, page 70. ..... 59
Table 6: Calculating PB and SB from PP. Values derived from Table 3, page 284 (Kelly 1983, 284, Table 3). Type climate from Table 5 matched against biome type. ..... 60
Table 7: Calculated secondary biomass per grid cell in kilograms per year for different mean yearly temperatures (theoretical values) in different biomes. ..... 61
Table 8: Different mortality rates (in %) and the compounded effects on a starting population of 100 with no replenishment. Bold rows are results after 15 and 20 years. ..... 68
Table 9: Variables on hominin behaviour included in the HomininSpace model. ..... 82
Table 10: Data fields in the archaeological data database. ..... 89
Table 11: Regular Checkpoints in Space and Time with geographical position. ..... 92
Table 12: Checkpoints with x and y grid cell locations and their interval ranges. ..... 93
Table 13: The 16 parameters that can be changed in the HomininSpace model. ..... 104
Table 14: Minimum, maximum, step size, and default values for the Standard parameter set. ..... 112
Table 15: Structure of the climate data file after preprocessing by SAM. ..... 127
Table 16: Structure of input file “Bintanja2008.txt”. Timestamp (1), reconstructed temperatures (5) and sea levels (9) are used by HomininSpace. ..... 128
Table 17: Layout of the checkpoint input file. Fields are separated by tabs. ..... 130
Table 18: Default starting locations for groups of the initial population. ..... 131
Table 19: Monitoring checkpoints located at example positions. ..... 132
Table 20: The climate monitoring checkpoints with their locations. ..... 133
Table 21: The header and the first ten parameter value sets (comma separated) from the Standard parameter file. ..... 134
Table 22: Color coding of reconstructed habitats in simulation maps. ..... 138
Table 23: A list of the settings in the system. Each setting is a Boolean value, and must be activated or deactivated by assigning a value of True or False. ..... 143
Table 24: Relative scoring schema when matching real world absence and presence versus modelling results. A ‘+’ indicating a positive contribution of the match to an overall modelling score. ..... 154
Table 25: Summary of the simulated scenarios. ..... 163
Table 26: Overview of the setting values in all scenarios. ..... 165
Table 27: Best three simulation results and parameter values for scenario Habitat-A. ..... 166
Table 28: Best three simulation results and parameter values for scenario Habitat-B. ..... 167
Table 29: Best three simulation results and parameter values for scenario Habitat-C. ..... 168
Table 30: Best three simulation results and parameter values for scenario Habitat-D. ..... 169
Table 31: Best three simulation results and parameter values for scenario Habitat-E. ..... 170
Table 32: Best three simulation results and parameter values for scenario Habitat-F. ..... 171
Table 33: Best three simulation results and parameter values for scenario Habitat-G. ..... 172
Table 34: Best three simulation results and parameter values for scenario Habitat-H. ..... 173
Table 35: Best three simulation results and parameter values for scenario Habitat-I. ..... 174

Table 36: Best three simulation results and parameter values for scenario Habitat-J. ..... 175
Table 37: Best three simulation results and parameter values for scenario Energy-A. ..... 176
Table 38: Best three simulation results and parameter values for scenario Energy-B. ..... 177
Table 39: Best three simulation results and parameter values for scenario Energy-C. ..... 178
Table 40: Best three simulation results and parameter values for scenario Energy-D. ..... 179
Table 41: Best three simulation results and parameter values for scenario Energy-E ..... 180
Table 42: Best three simulation results and parameter values for scenario Energy-BR, (with randomwalk) for Standard and Evolved parameter value sets. ..... 182
Table 43: Best three simulation results and parameter values for scenario Energy-CR, with randomwalk. ..... 183
Table 44: Best three simulation results and parameter values for scenario Energy-ER, with randomwalk. ..... 184
Table 45: Best three simulation results for scenario Energy-A, batch 2. ..... 185
Table 46: For all Standard sets the significant Spearman correlation coefficients for parameters versus matchedIntervalCoverage. Green are positive correlations, red negative. Darker colors indicate stronger correlations. ..... 195
Table 47: The Spearman correlation coefficients of the Evolved sets for parameters versus matchedIntervalCoverage. Green are positive, red are negative correlations. ..... 198
Table 48: Relative variable importance assuming a linear model for all Standard sets. ..... 199
Table 49: Relative Variable Importance for the Evolved sets from all scenarios. ..... 200
Table 50: Correlations between model input parameters and scenario number, for the Evolved datasets. Positive correlations are green colored and darker colors indicate stronger correlations. ..... 201
Table 51: Number of model parameter value sets per scenario with parameter values manipulated beyond the initial limits. Darker colors identify larger table values. ..... 203
Table 52: Presenting the best performing models from the Standard set. The first column gives the number of times the simulation number appears in the top-10 scores over all scenarios. 207
Table 53: Maxima for the numbers of created hominins and groups, and maxima for the number of different possible deaths in any simulation. Color coding is per column and darker colors indicate higher scores within the column. ..... 210
Table 54: Relevant points in time (climate extremes) for screen capturing. ..... 214
Table 55: The locations of the climate recording checkpoints and their modelled LGM and current day precipitation (P) and temperature (T) values. ..... 214
Table 56: Contents of the HominінSpace disk. ..... 297

# LIST OF FIGURES

Figure 1: Simulation scores in HomininSpace 1.0. Blue circles indicate some local maxima where the dynamic or 'ebb and flow' hominin implementation (in red) score higher than the static or 'source and sinks' hominins (in green). ..... 21
Figure 2: Components and data flow in the HomininSpace simulation system. ..... 26
Figure 3: Maximum ice sheet coverages (adapted from van Gijssel 2006, Figure 3.4). The solid red line follows the maximum limits of the Pleistocene glaciations. ..... 29
Figure 4: Topography for western Europe, sea level lowered with 120 meters. Pink areas are now submerged. Adapted from http://topex.ucsd.edu/marine_topo/mar_topo.html, accessed 8 April 2013. ..... 30
Figure 5: The simulation area divided into 14.948 grid cells in the SAM tool. ..... 40
Figure 6: Bathymetry data for the simulation area (the red square), taken from the ETOPO1 dataset. ..... 42
Figure 7: Reconstructed global sea level in meters (Y-axis), from today to 140 ka (from Bintanja and van de Wal (2008)). ..... 43
Figure 8: Red Sea level until 140 ka adapted from the supplementary materials from Grant *et al.* (2012) converted to equidistant horizontal axis. ..... 43
Figure 9: Overview of reconstructed global mean temperatures (Y-axis) for the last 140ky, drawn according to Bintanja and van de Wal (2008). ..... 45
Figure 10: Interpolation of temperature values between two given data points: LGM and modern day values, with horizontally the reconstructed global mean temperature used as index for the linear interpolation. ..... 48
Figure 11: Mean annual temperature distribution for modern day conditions. ..... 49
Figure 12: Mean annual temperature distribution during the LGM (source CCSM). ..... 50
Figure 13: Annual precipitation distribution for the modern day conditions. ..... 50
Figure 14: Annual precipitation distribution during the LGM (source CCSM). ..... 51
Figure 15: Overview of modern day (left) and LGM temperatures of the warmest month (top) and coldest month (bottom). ..... 51
Figure 16: Relative frequency distribution for the three LGM temperature climate parameters: coldest month (top), yearly mean, and hottest month (bottom). ..... 52
Figure 17: Schema illustrating both pathways to compute carrying capacity. ..... 53
Figure 18: Datasets and derived equations for productivity based on temperature (left) and precipitation values. Figures reproduced from Lieth (1973), Figures 5 and 6. ..... 55
Figure 19: Relationship between herbivore biomass (B) and net primary productivity (NAP) for different types of ecosystem (from McNaughton *et al.* (1989, Figure 1)). ..... 56
Figure 20: Population calculations for a group with initial size 25: Pre=8, Fertile=11 and Post = 6. Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.33. ..... 71
Figure 21: Population composition after 50 years (left) and after 2000 years (right) with unlimited carrying capacity. Graphs visualize typical exponential growth and are created with "HomininSpace Population Calculator.xlsx", see Supplementary Materials. ..... 72
Figure 22: Foraging ranges and resource acquisition. Larger circles are home ranges with the value of the foraging range of the groups in the centre. A value of '2' means two grid cells in all directions (horizontal, vertical and diagonal) are included when foraging for resources. Black are shared grid cells. ..... 79
Figure 23: Example illustrating the interval construction. Port Racine, two layers with dates, taken from Cliquet *et al.* (2003) ..... 90
Figure 24: Overview of all checkpoints. Red are regular checkpoints, white monitoring checkpoints, green starting locations, pink climate monitoring points and purple plus signs are core areas. ..... 91
Figure 25: Flowchart presenting the Initialization process in HomininSpace ..... 100
Figure 26: Flowchart of the Simulation process. ..... 102
Figure 27: Calculating the ViabilityIndex for a group with 25 individuals. Code fragment is taken from the "CreateHSParameters.R" source file. ..... 105

Figure 28: Counting matching visits. A visit at simulation time *t1* matches a dated interval for a single archaeological layer (A, C and D). Only in B there is no match. ..... 107
Figure 29: Counting matching visits. A visit at time *t1* can match two dated intervals that represent two archaeological layers (blue dots for layer X, red for layer Y). ..... 107
Figure 30: Functional representation of the implemented GA. ..... 114
Figure 31: The Repast Simphony development environment. 1 = main menu; 2 = project directory structure; 3 = source editor; 4 = file structure; 5 = log information window. ..... 119
Figure 32: Source code excerpt from the "HomininSpaceContextBuilder.java" class. ..... 121
Figure 33: The Java source files in the HomininSpace development environment. ..... 122
Figure 34: User interface of HomininSpace after initialization. On the right the Topography layer for the simulation area, on the left the scenario tree. ..... 124
Figure 35: Values for the batch parameters provided in the Batch Parameters tab. ..... 125
Figure 36: Reconstructed global mean temperature relative to modern day temperatures. Tick 0 is the start of the simulations (131 ka). ..... 128
Figure 37: Global sea level relative to the modern day sea levels Values in meters *below* modern day level. Tick 0 is the starting point of the simulations (131 ka) with a global sea level of almost 60 meters below the level of today (0). ..... 129
Figure 38: Google Earth view of the starting locations for the initial population. ..... 131
Figure 39: Climate recording checkpoint locations (pink crosses numbered 1-6). ..... 133
Figure 40: Flowchart - main simulation routine. ..... 136
Figure 41: Different map types at the start of a simulation: topography (a), energy levels (b), habitat reconstruction (c) and habitat reconstruction with beaches (d). ..... 139
Figure 42: Death density map at the end of a simulation. Darker colors indicate more perished hominins. ..... 139
Figure 43: Example output at the end of a simulation: hominin count (totals in dark blue, different cohort sizes underneath) through time. ..... 140
Figure 44: Example output at the end of a simulation: group count through time. ..... 141
Figure 45: Example output from a Climate Monitor object. Taken from "HS Climate Monitor 3 Mediterranean Habitat [x=94, y=7, 0 intervals] 21 2018-03-23 1303.txt". ..... 142
Figure 46: Habitat reconstruction (left) versus energy level reconstruction. ..... 145
Figure 47: Illustrating the effect of coastal resources on population distribution. ..... 147
Figure 48: Screenshot of simulation 1384 illustrating the patch wise foraging behavior with a varying foraging range (different sizes for the foraging squares). ..... 151
Figure 49: Simulation results for scenario Habitat-A. ..... 166
Figure 50: Simulation results for scenario Habitat-B. ..... 167
Figure 51: Simulation results for scenario Habitat-C. ..... 168
Figure 52: Simulation results for scenario Habitat-D. ..... 169
Figure 53: Simulation results for scenario Habitat-E. ..... 170
Figure 54: Simulation results for scenario Habitat-F. ..... 171
Figure 55: Simulation results for scenario Habitat-G. ..... 172
Figure 56: Simulation results for scenario Habitat-H. ..... 173
Figure 57: Simulation results for scenario Habitat-I. ..... 174
Figure 58: Simulation results for scenario Habitat-J. ..... 175
Figure 59: Simulation results for scenario Energy-A. ..... 176
Figure 60: Simulation results for scenario Energy-B. ..... 177
Figure 61: Simulation results for scenario Energy-C. ..... 178
Figure 62: Simulation results for scenario Energy-D. ..... 179
Figure 63: Simulation results for scenario Energy-E. ..... 180
Figure 64: Simulation results for scenario Energy-BR, with random walk. ..... 181
Figure 65: Simulation results for scenario Energy-CR, with random walk. ..... 182
Figure 66: Simulation score for scenario Energy-ER (with random walk). ..... 183
Figure 67: Simulation results for scenario Energy-A, batch 2. ..... 184
Figure 68: Correlations between the different visit counting output variables for scenario Habitat-A. Top right part of the matrix contains the correlations. Three red stars indicate a significant correlation. Bottom left part are plotted values. ..... 190
Figure 69: Explaining three parameter values used to construct Chernoff faces. ..... 191

Figure 70: Chernoff faces representing model parameter values for all 18 scenarios, illustrating variability within the top-10 best performing models per scenario (line), with the best scoring model numbered 1 (leftmost). ..... 193
Figure 71: Overview of sorted MatchedIntervalCoverage values for different scenarios. Simulations are sorted on score. Scores from the Standard set are colored cyan. ..... 205
Figure 72: Overview of created_hominins and created_groups versus SimulationNumber from scenario Habitat-A. The first 1500 simulations are the Standard set. ..... 208
Figure 73: Relation between the number of created hominins per simulation and the simulation score (scenario Habitat-A). ..... 209
Figure 74: Depicted the numbers of groups that perished due to cold temperatures versus temperature tolerance. In orange evolved hominins, in cyan results from the Standard set (scenario Habitat-E). ..... 211
Figure 75: Average group age for scenario Energy_B. ..... 211
Figure 76: Output of the simulation variables matchingVisitsWithConfidence (left) and matchingVisits (both from Scenario Energy-B). ..... 212
Figure 77: Reconstruction of the topography during the simulation period for selected moments in time. Lighter blue indicate higher elevation. ..... 216
Figure 78: Location of Climate Monitor 6 Channel (left) and the topographical output for the simulation period. Topography type 0 in the graph is water, 2 is beach and 3 is plain. ... 217
Figure 79: Habitat reconstruction for the simulation period. ..... 219
Figure 80: Geographical positions and Habitat types for the checkpoints Climate Monitor 1 Atlantic (bottom left) and Climate Monitor 3 Mediterranean (right). Habitat type 2 is boreal forest, 3 is cool forest, 4 is cool grass, and 6 is warm grass. ..... 220
Figure 81: Reconstructing habitat types for the Scladina area. Results from HomininSpace (left, with HabitatType 2 = booreal forest, and 3 = cool forest) and López-García *et al.* (2017). ..... 221
Figure 82: Reconstruction of the energy levels in the landscape for the complete simulation period. Darker colors indicate less energy. ..... 222
Figure 83: Energy levels per year recorded by the Climate Recording Checkpoints for the whole simulation period. Blue line is for Checkpoint 1, Red for 6. ..... 223
Figure 84: Recorded values for reconstructed temperatures per Climate Recording Checkpoint. 224
Figure 85: Recorded values for reconstructed precipitation values per Climate Recording Checkpoint. ..... 224
Figure 86: Correlations between all visit counting output variables for scenario Habitat-A. For instance, the correlation between matchedIntervalCoverage (simulation score) and matchedIntervals (count of all matched intervals) is 0.98, and very significant (three red stars, as have all correlations). ..... 228
Figure 87: Illustrating the bias in the checkpoint dataset: overview of the used intervals in the simulations. X-axis gives Time, with each interval defined by starting point and end point in years ago. Vertical axis is the latitude of the site. ..... 229
Figure 88: Mapping models resulting from mutation (in green) and combination (in red). Horizontal axis the simulation number (from 1500 onwards), the Y-axis presents the simulation score. Data from the Evolved set in scenario Energy-B. ..... 231
Figure 89: Illustrating the effects of using coastal resources. Left simulation 1384 with coastal resources turned on (scenario Habitat-J). On the right the same simulation with coastal resources disabled (Habitat-I). ..... 240
Figure 90: The attractiveness of coastal resources. Simulation 1365 from Scenario Energy-B but with dynamic hominins. ..... 241
Figure 91: Overview of the average foraging range for part of the simulation period. Largest value around 9.3, minimum value around 7.0, standard deviation around 2. Figure created in HomininSpace, by calculating the average foraging range of all present groups for each time step. Simulation 1384, scenario Habitat-I, the maximum value for the foraging range is 11. ..... 243
Figure 92: Simulation score versus maximum foraging range. For low values of the foraging range good scores are difficult to obtain. Increasing the maximum beyond 9 does not influence the score much. Values are from scenario Habitat-C. ..... 244

Figure 93: Illustrating the ability to cross larger waterbodies to some of the islands off the coast of France. Produced in Scenario Energy-D with dynamic hominins, simulation 1365..... 245
Figure 94: Overview of the results for the main biological values for CCSM (left) and MIROC climate models with unified colour coding schemes..... 305

# LIST OF EQUATIONS

Equation 1: Relation between Temperature and Primary Productivity. .....54
Equation 2: The effect of Precipitation on Primary Productivity. .....55
Equation 3: Relation between Secondary Biomass and Net Primary Production. .....57
Equation 4: Calculation of Effective Temperature using the mean extreme temperatures. .....59
Equation 5: Computing Secondary Biomass from Primary Productivity (a, b from Table 6). .....60
Equation 6: Computing the available Secondary Biomass for hominin consumption. .....62
Equation 7: Calculating the result for Primary Productivity at time step t+1. .....63
Equation 8: Calculating the number of hominins per segment for the next time step. .....70
Equation 9: Combining interval scores for all checkpoints to compute the simulation score. .....108

# 1. INTRODUCTION

> *"... all archaeological reasoning includes the use of some type of analogy or model, applying knowledge learned in the present to the past, whether explicitly or implicitly"*
>
> (Yellen 1977, 50)

## 1.1 Background

This study presents an agent-based simulation model exploring the patterns of presence and absence of Late Pleistocene Neanderthals in western Europe. Computer based simulation techniques are well suited to explore explicit models of our ideas about the past. They can play an important role in understanding key events in the evolution of our species. The usage of models in archaeological studies has increased due to the relatively recent acceptance of the method by archaeological researchers (McGlade 2014), especially by those interested in human evolution (Lake 2014; Cegielski and Rogers 2016). Used for hypotheses exploration, theory building and method testing (Lake 2014), models have been constructed in the investigation of such wide ranging subjects as the evolution of cooperation (Pereda *et al.* 2017), surface deposit formation (Davies *et al.* 2016), or Neanderthal cannibalism (Agustí and Rubio-Campillo 2017).

Dispersal is one of the key factors in understanding the spread of a species and a fundamental component of paleoanthropological research (Gamble 1993; Dennell 2003; Roebroeks 2006). Computational modelling is increasingly used to address issues in past hunter-gatherer dispersals and distributions (Banks, d'Errico, Peterson, Kageyama, *et al.* 2008; Benito *et al.* 2016; Gilpin *et al.* 2016; Kolodny and Feldman 2017; Melchionna *et al.* 2018). Computational techniques addressing dispersals require explicit models on demography and on spatial behaviour of individuals and groups. The way for the demographic success of the global dispersal of *Homo sapiens* was paved by several Out-Of-Africa migrations, by more than one species of the genus *Homo* (Stringer 2003; Derricourt 2005; Klein 2008; Stringer 2016). They and their descendants managed to colonize huge areas far into Eurasia. Patterns of dispersal can offer possible explanations for the spatiotemporal patterning found in the archaeological record. But *Homo erectus*,

*Homo neanderthalensis*, the Denisovans and ultimately modern humans disperse through the landscape in a process for which many details are still unknown and thus difficult to model (Dennell and Roebroeks 2005; Holmes 2007; Verpoorte 2009; Reyes-Centeno *et al.* 2014).

Better understanding of the underlying dispersal principles including chosen routes, resource acquisition strategies and the influence of population density is needed, as well as insight into demographic fluctuations in response to environmental change. The study of the dispersal characteristics of the Neanderthals in the landscape can help to understand the essentials in the human colonisation success (Lieberman 2002). Insight into the why, how and when of Late Pleistocene hominin dispersals through the employment of computational modelling techniques can open a window into deep human population history.

## 1.2 Study objectives

This study attempts for the first time to create a hominin model while avoiding parameter value biases and explicitly modelled handicaps. This can be achieved by assigning and varying parameter values automatically while optimizing the match with the relevant archaeology. Therefore, the main research aim addressed in this thesis is how to find parameter values for a generic hominin dispersal model through space and time without using values from ethnographically derived data nor from the paleoanthropological and archaeological literature and further to apply this technique in the exploration of a Neanderthal case study.

Any computer model is finalized by assigning values to the model parameters. There are at least two issues with selecting parameter values. Firstly, earlier research that formed the basis of my master thesis illustrated that a bias in selecting parameter values from the literature would lead to unwarranted expectations about the truth value of the modelling results (Scherjon 2015a). Well informed parameter values, assigned using paleoanthropological literature, comparative research, or ethnographic analogies can only produce results that compare well to the archaeology of the real world if all relevant elements from that world have been included in the model, something which is difficult or maybe impossible to realize. It was shown that running simulations only using such informed values will miss parameter values that produce more matches of the modelled hominin world with the archaeologically attested presence (Scherjon 2016). When

assuming that more matches means a better hominin model, the best models were *not* found using user selected, pre-determined parameter values.

Secondly, when implementing a hominin model this is often done to compare indirectly or directly against a model for modern humans (Kolodny and Feldman 2017). Neanderthals are the most modelled non-sapiens hominin species. In general, Neanderthals are constructed as a less capable and inferior hominin, with for instance a lower birth rate (Zubrow 1989) or higher mortality rates (Flores 1998; Sørensen 2011). These and other studies have, not surprisingly, confirmed that if such superiority is assumed and implemented using selected model parameter values (Zubrow 1989; Flores 1998; Flores 2011; Gilpin *et al.* 2016; Horan *et al.* 2005) Neanderthals would indeed have been likely to go extinct. Such results do not prove that Neanderthals were less capable than modern humans, merely the obvious that if you model a handicap the impeded species will suffer.

## 1.3 Methods and approaches

All models are based on assumptions. Any assumption is false to some extent (Wimsatt 1987). Good modelling finds those assumptions that matter most and are least biased. In this study I present an agent-based simulation system called HominínSpace that implements a generic demographic and social parameterized model of hominin dispersal. Agent-Based Modelling (ABM) is a style of modelling where individuals and their interactions with other individuals and the environment are explicitly programmed (Grimm *et al.* 2005) and where their repertoire of actions can cause the emergence of phenomena representing (past) human behaviour (Bonabeau 2002, 7280, Epstein 2006, Gilbert and Ahrweiler 2009). Models are simulated through time within a reconstructed high-resolution environment. HominínSpace implements a set of explicit and implicit assumptions, the cataloguing of which is an integral part of this research.

The approach in this thesis entails two lines of research. Firstly, to augment the existing HominínSpace system in such a way that parameter values are assigned autonomously, that is without intervention by the user. The system will be able to traverse the parameter space that is created by the complete set of possible parameter value combinations in an attempt to find those parameter values that will result in a simulation that matches well with the archaeological data. If the output of a simulation reproduces the phenomena of the real system well, the model is a possible explanation for the occurrence of those phenomena (Axelrod 1997). To this end the system uses an underlying data set with archaeological

presence data to compare against. Simulations executed in the HomininSpace modelling system are used to assess the character of past hominin dispersals. To address questions about past hominin dispersal in a changing environment the model integrates paleoclimate data with archaeological data (Nakazawa and Bae 2017).

Secondly, this study collects, evaluates, and describes the available archaeological data for a Neanderthal case study using the patterns of presence (and absence) of Neanderthals in western Europe in the Late Pleistocene to illustrate the applicability and usefulness of the developed tool, and to assess the validity of the method. The area and time frame are well delimited and intensely researched. An area where climatic changes cause major spatial changes due to emerging and subsiding landmasses which have influenced hominin pathways (Guiot *et al.* 1989). A large dataset is compiled with evidence of Neanderthal presence attested by fossil and/or archaeological finds and constrained by radiometric dates. The database that results from this effort is interrogated in the HomininSpace system to determine promising parameter values and to validate simulation results.

This study further implements model elements to address several questions that are unresolved in the current debate about Neanderthals. The HomininSpace simulation tool is used to answer these in the context of the Neanderthal case study. During each simulation the modelled presence of hominins is compared to the archaeology by computing matches between modelled and archaeologically documented presence of Neanderthals (cf. Janssen 2009). Comparing the results for the different questions will add to the discussion about and inform on specific details of Neanderthal dispersal.

Many hypotheses about Neanderthal behaviour in the landscape exist (for examples see Shea (1998)). Seven questions were selected based on their relevance in the current debate on Neanderthals or because they were of interest to close colleagues as well as their suitability for inclusion in the model. Every simulation will explore a combination of one or more of these questions by including or excluding relevant model elements. The questions concern a variety of behaviours and are detailed in the Neanderthal case study. The questions are grouped into three major categories, supplemented by three neutral model[1] questions that are used for system testing:

[1] A neutral model explicitly omits or sets to zero the effects of one factor or process, while retaining all others.

- Environment reconstruction:
    1. How does Energy reconstruction of the environment (McNaughton *et al.* 1989) compare to Habitat reconstruction (Binford 2001; Kelly 1983) when matching the archaeology with modelled hominins?
    2. What is the effect of adding coastal resources to the model (Fa *et al.* 2016)?
- Hominin behaviour:
    3. How does the addition of an ebb and flow dispersal with the available resources influences the match with the archaeology (Hublin and Roebroeks 2009; Dennell *et al.* 2011)?
    4. What is an optimal maximum foraging range or would an unlimited range be better when matching archaeology (Benito *et al.* 2016)?
    5. What happens if we assume that Neanderthals were able to cross large open water systems (Leppard and Runnels 2017)?
- Population distribution:
    6. Absence data indicates locations where Neanderthals were NOT present for some period of time. What influence has adding such data when comparing simulation results with only archaeology attested presence (Jiménez-Valverde and Lobo 2007)?
    7. What happens when population core areas are implemented that produce new Neanderthal groups under certain conditions (Jennings *et al.* 2011)?
- Neutral models for system verification:
    8. How does the reconstructed topography influence the dispersal characteristics of Neanderthal groups?
    9. How does the reconstructed energy distribution in the simulation area influence the dispersal characteristics of Neanderthal groups?
    10. What are the results if the Neanderthals are implemented with random movement instead of going for the maximum amount of available resources?

The model underlying the HomininSpace system is a collection of parameters that describe modelled hominin demography, social and cultural life with the associated source code that describes how this parameter set is implemented. Due to the wide range of possible values for each model parameter in HomininSpace the total parameter value space is huge and the sheer number of possible unique models is enormous[2]. It is shown that it is computationally impossible to run simulations for all possible models in order to find the best match preventing an exhaustive exploration of the parameter space. To solve this optimization problem HomininSpace implements a genetic algorithm inspired by the processes of natural selection (Forrest 1993). It is a domain-independent and thus meta-heuristic way of traversing the parameter space searching for the best matching simulations (Goldberg 1989).

[2] Almost 400 million if we conservatively assume that each parameter can take only three values out of its possible range: the minimum, medium and maximum value.

## 1.4 Thesis structure

The HomininSpace modelling and simulation system is the complete software architecture to run the simulations in this study. That includes a computational and explicit model (in source code) and computing infrastructure including modelling software tools and libraries, input files, batch scripts, post processing modules, etc. The design, implementation and usage of the HomininSpace modelling system are presented and discussed in four parts that together form this thesis:

**Part One: The Lay of the Land**

Part One gives a general background to the main topics. This part starts with an introduction into modelling and simulation in general (Chapter 2). This chapter puts forward and justifies simulation techniques as an important tool to explore questions relating to research of the past. Chapter 1 introduces the Neanderthal case study. It describes the environment of Late Pleistocene western Europe and the hominins that lived there during this period.

**Part Two: Creating the Actors**

Part Two describes the model and its components. First, in Chapter 1, the modelled environment is discussed. In HomininSpace an energy landscape is reconstructed and two different approaches to create such landscapes are described. The changing topography is discussed with rising and falling sea levels that are included in the model. Chapter 1 creates a model for hominin dispersals in general and introduces specifics for Neanderthals. The final chapter of this part describes the archaeology as it is used and included in the underlying database of the HomininSpace system (Chapter 6).

**Part Three: Setting the Stage**

The implementation of the model in HomininSpace is the core of Part Three. Chapter 7 discusses the design of HomininSpace following the Overview, Design and Details (ODD) protocol. Section 7.7 explains the model parameter space and how to explore this using a Genetic Algorithm. Chapter 1 describes how the system is built and includes details about the input data and the user interface. The possibilities and consequences of the implementation of specific questions on hominin behaviour are described in Chapter 9.

**Part Four: And Action!**

The final part of this thesis presents the simulation experiments done in this research. Chapter 1 describes the different scenarios that are explored with a summary of the

simulation results for each of them. This is followed by the two chapters that analyse (Chapter 1) and discuss (Chapter 1) the simulation results. Chapter 1 concludes the thesis with the remit of this research and an evaluation of the HominінSpace environment, a discussion of possible extensions, and gives recommendations on how to use the tool in further research. Appendices include abbreviations and supplementary materials.

# PART ONE: THE LAY OF THE LAND

"*…the explosion in the number of simulation models, which taken together with their output utility, suggests that simulation really has finally come of age as part of the archaeological toolkit*."

(Lake 2014, 277-278)

# 2. AGENT BASED MODELLING

## 2.1 Introduction

The use of computer systems is today well established for solving problems of interest in archaeology. Increases in computer power have added to administrative facilities like database, spreadsheet and word processing capabilities, to the application of Geographical Information System (GIS) techniques, and the creation of large simulations implementing dynamic models and solving huge numbers of complex calculations in real-time. This chapter introduces HomininSpace; an agent based modelling and simulation system to explore hominin dispersal in a landscape through time. Modelling involves abstraction, simplification, and formalization in order to better understand how the world works. As such each model is a simplified reflection or representation of reality (Castle and Crooks 2006). Building and disputing models is fundamental to science and scientists often debate which one is the more accurate model. Models are typically used when it is either impossible or impractical to create measurable experiments (Section 2.2).

Simulations are run to model and reproduce parts of some natural, ethological, social or conceptual process (Drogoul and Ferber 1994). The design of a model usually starts with a problem definition, then the identification of the desired features of the system to be modelled, followed by a definition of possible system representations. Then the model is implemented within a modelling environment according to the design specifications and at the required abstraction levels. This also involves selecting the best simulation tools for the needed features and conversion of data into usable formats. The simulations are run using selected input data in defined scenarios. Section 2.3 will discuss different modelling techniques and motivates the selection of a particular tool.

Section 2.4 gives an overview of the development history of HomininSpace. The implementation of the underlying model in HomininSpace combines results from many different individual disciplines including anthropology, archaeology, geology, and climatology. This provides an increased understanding of the fundamental elements of the behaviour of past hominins and offers a better insight into the subdisciplines, the data they create and how they are connected (Shaman *et al.* 2013). HomininSpace is implemented as an agent based model combining the representation of dynamic processes with geographical and real-world data allowing modelling of both processes (dynamic) and patterns (Mathur 2007).

## 2.2 Why model?

"Because we all do! But with simulation we can do it better!"[3] Even though people are taken aback by the assumed difficulty of implementing a computational model we are by all means a modelling species (Holland and Quinn 1987). The power of modelling is in part derived from the fact that one needs to explicitly specify all the details of the model that is to be used, and thus to engage the underlying assumptions. This forces the researcher to consider all needed elements, and to quantify them. Simulation, the execution of a model through time, and other process-based modelling techniques require an explicit description of the processes that are relevant to the system of interest. Such processes can be constrained by archaeological data that has timing information included to provide spatially explicit representations of patterns through time.

Modelling and simulating are inherent features of our cognition (Niazi and Temkin 2017). Abstraction may well be an indispensable element in any attempt to formalize real world systems. In this sense modelling is used more often than is realized. Some even go as far as to state that "*science in general is impossible without model[l]ing*" (Dershowitz and Gurevich 2008). To stress this point, it might be adequate to think of the following lines which belong to A.M. Turing, one of the founding fathers of computer algorithms: "*...if one wants to treat the problem seriously and systematically one has to replace the physical puzzle by its mathematical equivalent*" (Turing 1954, 11). In the context of this thesis, the puzzle is the complex way hunter-gatherers might find their way around the Palaeolithic landscapes and the mathematical equivalent is an abstraction in the agent-oriented modelling methodology.

Using computers, simulations become more feasible and allow more complex setups with the almost continual increase in personal computer power (Clarke 1973; Moore 1965; Brooks 2001). A model must be explicit in all details before a computer can work with it. A computer does what it is told, and does not do anything that is not explicitly requested[4]. Simply put, the model must be unambiguously explicit in all its details. Explicit models allow discussion with and replication by other researchers and thus reproduction of research. It is also possible to test the effect of large numbers of model parameters and wide ranges of parameter values over time on predefined starting scenarios (Bonabeau

[3] Private conversation with Iza Romanowska, 2017.
[4] Note that observed irregularities with computer usage always reflect errors by either user or software developer.

2002). With the general availability of powerful computers simulation has become possible and more fashionable in archaeological research (Lake 2014).

Students today are much more "computer fähig" then previous generations (Marc 2001; Kay and van Harmelen 2012). Computer hardware is more than sufficient where even cellular phones today have more built-in processing power than some desktop computers, especially since the introduction of the iPhone by Apple which reversed the traditional design of 'phone first, computer second' (Linge and Sutton 2015). Development tools become very sophisticated and some are dedicated towards simulation building (Nikolai and Madey 2009). It is possible to write meaningful simulations with less than one A4 of source code (Schmitz 2018).

Computer simulations have been used in archaeological research for almost 50 years (Lake 2014). Although initially hampered by insufficient hardware in the 1970s, the theoretical background in complex systems theory and significant improvement in computing power allowed many researchers to include more and more computational modelling in their research agenda (Kohler and Van der Leeuw 2007). Agent-based modelling especially offers the possibility of addressing individuality and emerging phenomena in social interactions in complex systems (Premo *et al.* 2005). Agent based models are now recognized as a powerful means to explore the relations between individual actions and larger social structures for any time scale (Rogers and Cegielski 2017). The capability to model change and emerging phenomena makes the method well suited to analyse certain categories of archaeological hypotheses (Cegielski and Rogers 2016).The method is quite appealing to archaeologists due to the analogy between agents producing simulation results and the humans in the past creating an archaeological record. And a significant number of archaeologists today has experience with writing, replicating and reviewing simulations as is illustrated by the five sessions (out of 42) at the Computer Application & Quantitative Methods in Archaeology (CAA) 2018 conference that were dedicated to some aspect of modelling and simulation[5].

Simulation as a technique is used for a wide range of purposes that include prediction, performance, training, entertainment, education, proof and discovery (Axelrod 1997). As a methodology it adds a novel way of doing science: like deduction it starts with explicit assumptions but then generates new data that can be analysed inductively. But unlike

[5] http://2018.caaconference.org/sessions/, accessed 11 December 2017. Sessions include S3, S9, S10, S17, S19, and S22.

standard induction, this data is not derived from measurements in the real world, but extracted from explicitly stated model environments (Axelrod 1997, 25).

### 2.2.1 In or out: what to include in a model

Archaeological models necessarily (over)simplify the past. A model is thus inherently incomplete[6]. And if any factor is not included in the model, it will not influence the results and might not be included in the discussion (nor in the reconstruction of that past). This realization has resulted in more complex modelling with positive feedbacks (Kohler 2012). Elements of a model can be parameterized, where different values for a parameter create different models. The addition of model elements or parameters can result in over fitting, where the exclusion of details can lead to under-fitting.

The selection of parameters is ultimately guided by the principle of parsimony (Forster 2000). Adding more parameters will capture noise in the data, and it is therefore essential to keep the number of parameters low, or to prune the model when possible. By adding parameters it is possible to fit all data (it takes about 30 parameters to fit an elephant according to Burnham and Anderson (2003, 30)). Under-fitted models will capture little structural information and are equally ill fitted for inferences about novel data sets (Wagenmakers 2003). A model is constructed by people, and therefore implements the biases of the author(s). It is very difficult to avoid such biases and effort must be invested in mitigating the steering effects. Minimally, the biases must be stated as explicit as possible. Ultimately it is the creator of a model who can and must rationalize what is included and what is not.

### 2.2.2 Model credibility

Modelling and simulating past hominins in a reconstructed environment using demographic parameters that are documented ethnographically (like birth rates, mortality figures, procreation, and consumption) allows representation of human behaviours in a reproducible manner. Actors who interact according to individual level behaviours create a system that is characterized by aggregate patterns which are quite comparable in scale to those observed in the archaeological record (White 2013, 124). Reality is modelled to some degree with the hominins behaving according to the rules as put forward by the model.

[6] Cliff Notes version: All models are wrong, cited by John Hawks: http://johnhawks.net/weblog/reviews/neandertals/neandertal_dna/neandertal-ancestry-iced-2012.html, verified 8 okt 2013.

The validity of such a model, that is 'how well the model represents hominin behaviour' can be addressed by comparing the simulation results of the model with the known material remains that are the results of hominin behaviour in the past (Gilbert 2008). A prerequisite for a valid comparison is that the reflected behaviours are not directly programmed into the model, and that these behaviours are not the emerging phenomena that are being investigated. As such, comparison must be made with elements of the archaeological record and the corresponding aspects of the model that are the result of the dynamics of the simulated system. The elements of the archaeological record may not have been used to construct the model. These comparisons allow assessment of the degree to which the internal dynamics of the model match those of the system it represents (White 2013, 137-138).

Since a model is always by definition a simplification of reality, the results from simulating that model reflect some elements (but not all!) of the real world. Statements based upon modelling results thus have an inherent believability problem. Within archaeology, a model should be able to explain past observations, predict future findings, should be refutable, should enable an estimation of the level of confidence in the model and should be simple (Marwick 2017). One of the most important gains is that one unrepeatable process (the behaviour of hominins in the past) can be imitated by another process (the simulation), one that can be replicated and parameterized at will (Hartmann 1996). Sometimes simulated experimental data is even preferential to compare theories against, since for this data input values are known (Konigsberg and Frankenberg 2013, 295).

The evaluation of simulation results can be quantified using a utility function, the design of which is a key element in the modelling effort. Simulating models does not duplicate reality and correctness of a model cannot be formally proven. It is up to the designer of a model to ensure that the model contains the key elements of the modelled process, and that therefore simulation results actually represent what could have happened in the real world.

An important issue with models when used as an explorative tool in archaeological reconstructions or interpretations is the stochastic nature of the input and the results (Peeters and Romeijn 2016). Generally, only after a multitude of simulation runs certain patterns are certified to occur and not due to chance events, and thus by definition have a statistical likelihood and comparative reality value. This diminishes the credibility of the

model and the simulation results. Thus, a sound theoretical basis is required beforehand and a statistical significant validation of the results afterwards.

A fundamental aspect of the scientific method in general is that experiments can be reproduced and replicated by other researchers, who can sub sequentially build upon these results to advance the field of research (Wilson *et al.* 2014). Credibility, that is the acceptance of model and results as correct, increases with the quality of the development process (Law and McComas 1991). Computational reproduction of simulation results and replication by different researchers will strengthen the model as a true approximation of past behaviour and will increase the credibility of the research (Marwick 2017). To allow proper reproduction, certain basic requirements must be met. At the very least these include making publicly available the raw input data, source code and output data (Peng 2011, Sandve *et al.* 2013).

## 2.3 Techniques for modelling dispersal in landscapes

Neanderthals no longer exist and thus cannot be observed nor can they be subjected to experiments to test hypotheses. Ethnographic data has limited applicability, allows little experimentation as well, is generally difficult to reproduce, has limited control group facilities and has limited time depth. Computational modelling and complex systems theory can offer remedies to all of these issues (Gilbert 2008; Kohler and Van der Leeuw 2007; Miller and Page 2007b; White 2013, 131). For exploring dispersal hypotheses in a constructed landscape there are basically three different techniques available: Mathematical Modelling (MM), Cellular Automation (CA) and implementing an Agent Based Model (ABM). Each of these methods has its specific advantages and disadvantages.

A well-known example in MM is the "wave of advance" model for population expansion (Ammerman and Cavalli-Sforza 1984; Fisher 1937). Fisher (1937) defined the constant rate of advance ($r$) in his model of the spread of an advantageous gene as follows: $r = \sqrt{(2 * g * m)}$, with $g$ the growth rate and $m$ the migration rate per unit of time and space. This has been applied to many other diffusions, for instance to the spread of agriculture in Europe by Ammerman and Cavalli-Sforza (1984) and Davison *et al.* (2006), or Paleoindian dispersals in North America using environmental carrying capacity and estimated diffusion coefficients by Steele *et al.* (1998). A disadvantage of the very abstract mathematical models is that often restrictive or unrealistic assumptions are imposed that limit their

applicability. Examples are linearity, homogeneity (identical individuals) and normality (Bankes 2002, 7199).

A CA is a dynamic system in which space and time is discrete and where the system is specified through a regular matrix of cells, their boundary conditions and their relations with other cells, a finite set of states for each cell, and rules that determine the dynamics of the cells (Wolfram 1983, 602). In CA the basic unit is the cell, and each cell is updated by a state transition function synchronously in discrete time steps (see for instance Pfeifer *et al.* (2008)). Good examples are by Surovell (2003), who studied the colonization of the Americas with a grid based simulation using the carrying capacity of the coastal areas, or SteppingOut implementing Out-of-Africa dispersal by Steven Mithen and Melissa Reed (Mithen and Reed 2002). One of the more restrictive elements of grid-based modelling is the memory requirements when modelling large areas using small grid sizes. Using large or equally-sized grid cells gives the simulation an abstract character and limits free movement through space. Modelling individual entities is cumbersome in CA models. ABM addresses some of the limitations of both MM and CA.

### 2.3.1 Agent Based Modelling

Agent-based Modelling (ABM) [7] focuses on the behaviour of individual entities who act according to certain rules (Abar *et al.* 2017). It enables in a bottom-up approach the study of how aggregate system-level and individual-level patterns emerge without a central controller (Bonabeau 2002; Bankes 2002, 7200; Young 2002, 138). This means that individual activities (interactions, processes) produce system level dynamics (like cultural norms or institutions) which are not visible (or even known to be present) at the individual level. This is contrary to more traditional models where the characteristics of a population are averaged and changes in these characteristics are simulated for the whole population. In ABM the characteristics and actions of each autonomous individual are tracked through time and systematically the aggregated consequences are established. Due to these properties heterogeneous agents in such a system form a natural specification of the concepts in many social problems (Bankes 2002, 7199; Macal and North 2009) for which empirical methods of analysis are often the only available alternative (Edmonds and Bryson 2004 ).

[7] Referred to as Individual-based Modelling (IBM) in ecology, and conceptually the same as in Multi-agent Systems (MAS) where the agents are complete systems like the computers in a network.

Detecting the emergence of such phenomena often relies on graphical output in which the operator is needed for identification of (newly emerged) features. Micro motives can result in macro behaviours (Schelling 1978, with an illustrative example in the analysis of non-malicious segregation in black and white neighbourhoods resulting from only minor preference differences). Defining measurable macroscopic behaviour across multiple simulations can give rigor to the analysis. Non-linearity of interactions between individuals and the unpredictability of the effects of individual behaviour on social dynamics and structures complicates understanding of macro scale implications (Miller and Page 2007a).

Simulations in ABM progress through time in discrete time steps, where all actions for one time step are completed before the next time step is started. Agents in ABM interact in rule-based and goal-oriented ways. An agent perceives (senses) and interacts with its environment and with other agents. Its behaviour depends at least partially on its own state and experience. The agent decides what actions are needed to satisfy its objectives. Table 1 describes the general characteristics of agents in ABM (Macal and North 2005). ABM has been extensively used in ecology, biology, and the social sciences (Abar *et al.* 2017).

**Table 1: Common characteristics of agents in Agent Based Modelling.**

| ***Characteristic*** | ***Description*** |
|---|---|
| **Individuality** | Each agent is uniquely identifiable, with individual characteristics in the form of a distinct set of attributes. The social organization is decentralized. |
| **Environment** | There may or may not be influence from the agent on the environment, but there is always influence from the environment on the agent. |
| **Autonomous** | Each agent is goal directed and independently decides what action to take to achieve its goals, according to certain rules. |
| **Flexibility** | An agent has the ability to learn and adapt its behaviour over time based on its experience. |
| **Local view** | The individual agents can only access that part of the environment which they can perceive. Not one has a global view of the complete system. |

An important concept of ABM is the environment in which the agents operate. Often the agents have some influence on the (local) environment, but they are always influenced by that environment. Adding data from a Geographical Information System (GIS) to an ABM system maps the characteristics of the real world onto the space of the agents and greatly enhances the representation possibilities to visualize patterns (Mathur 2007). Information in a GIS tends to be static, whereas the dynamic environment in an ABM generally tends to be schematic (instead of real-world data). A combination yields the benefits of both approaches. A small change in the simulated environment of an ABM can change the nature, and even the occurrence, of high-level behaviours (Polack *et al.* 2010). As a result,

simulated environments are as important as agents in order to reach the purpose of the simulation study.

The ABM methodology shares with all other modelling approaches the difficult and creative process of deciding what to include (more realistic) and what to exclude (more abstract) from the model (cf. Janssen 2009). Ultimately it is the modeller that decides, but via peer-review processes and other techniques it can and must be ensured that the model actually represents those parts of the real world that the modeller intended to model. Agent based models are calibrated using data on individual behaviours, which can include archaeological data.

Validation of modelling results is one of the most important methodological challenges in ABM (Gilbert 2008). This can be approached in two ways: first one can look at the process and quantify the way the model represents elements of the real process, or one can compare the modelling results (aggregated patterns) against real, empirical data (and implicitly assume or explicitly define that a good match means that the model represents the causal process well). The independence of the calibration process with the validation step provides ABM its explanatory power (but as noted by Angus and Hassani-Mahmooei (2015) this is surprisingly often lacking in publication efforts). It is essential to address such critical elements of the method documenting the modelling process (Marwick 2017).

Humans and human groups are highly individual, have territories and live in diverse dynamic social structures. Analytical models are less suited to model these characteristics (Pitt *et al.* 2003, 110). In archaeology ABM has been used to investigate a number of research questions. Good examples are the MAGICAL project with agents harvesting resources (Lake 2000), ENKIMDU simulating Mesopotamian settlement systems (Christiansen and Altaweel 2005) and the Prehistoric Patagonia model where initial hunter-gatherers become farmers (Barceló *et al.* 2008). For a good overview of the use of ABM in archaeology, see Barceló *et al.* (2008) and Lake 2014. ABM has been chosen as the modelling technique to implement HomininSpace.

## 2.4 HomininSpace – a model of the past world of hunter-gatherers

The core of my research and the main topic in this thesis is the HomininSpace modelling and simulation system. HomininSpace builds upon the experience gained when implementing SteppingIn, a simulation tool I implemented in order to compare alternative scenarios about how Europe was populated by modern humans (Scherjon 2011). The

current thesis draws further upon the development of an early version of HomininSpace (version 1.0, Scherjon 2015a). Underlying research question in that study was the identification and qualification of key hominin characteristics (parameters) needed to model dispersal in the landscape and to assign values to these key parameters from biological, archaeological, paleoanthropological and ethnographic data and literature. Simulated presence of modelled hominins is then compared against archaeological data from specific sites to assess the quality of the model, where more matches means a better model. A case study of different mobility types for Neanderthals in western Europe served to assess the usability of the tool and explored whether distributions of Neanderthals in the landscape through time were caused by continuously tracking preferred habitats (Hublin 2009, Roebroeks *et al.* 2011).

The tracking of favourable habitats has also been described as the "ebb and flow" of populations (e.g. Hublin and Roebroeks 2009), and involves individuals or groups of individuals moving to areas where the most favourable circumstances are found. When conditions worsen, populations would retreat into refugia with more benign environments (Stewart and Lister 2001). The dynamic "ebb and flow" of constantly moving populations has often been opposed to a more static "sources and sinks" model where local populations must adapt behaviourally and/or genetically to cope with the changing climate or become (locally) extinct when conditions become less favourable (Pulliam 1988; Pulliam 1996). They are replenished from more productive areas when the situation improves (Dennell *et al.* 2011; MacDonald *et al.* 2012). Obvious examples of this "sources and sinks" model are most species of flora. Since individuals cannot move by themselves they invariable die when the climate deteriorates sufficiently. For species to live there again when conditions improve the area must be re-colonized from other source areas.

Analysis of the simulations in HomininSpace 1.0 implementing a static versus a dynamic mobility suggested that the archaeology of Late Pleistocene Neanderthals best matches hominin groups following a static strategy where they occupy an area and stay there even if the environment becomes less favourable. Note that static here does not mean sedentary. Static hominins in HomininSpace still move around collecting resources but they do this in a confined local area from which they cannot leave. It was concluded that Neanderthals would have followed at least partly this more static strategy and were thus maybe less mobile staying in more confined areas than previously thought.

Results and details from the HomininSpace 1.0 program were presented at international conferences including the Computer Applications and Quantitative Methods in Archaeology (CAA) and the European Society for the study of Human Evolution (ESHE) meetings (Scherjon 2015b, 2015c). The enthusiasm that was expressed by peer researchers illustrated the interest in this topic. Due to the importance of peer review and replication in modelling studies a replication of some of the core elements of HomininSpace was implemented in 2016 by Iza Romanowska, a researcher trained in Palaeolithic Archaeology and computer modelling.

This replication in the NetLogo simulation environment failed to reproduce the patterns observed in the simulation results. In the original model a 'source and sink' mobility pattern consistently outperformed the 'ebb and flow' hominins for those simulations that have a good match with the archaeology. Doubts imprinted by the replication efforts inspired more simulations in HomininSpace, and then the identification of a pattern of peak simulation results. Many parameter value combinations lead to good matches with the archaeology and indeed, the 'source and sink' implementation scored higher on these than the 'ebb and flow' mobility type for the same parameter values. But next to many low scoring simulations there were occasionally some parameter value combinations that produced high scores in which the reverse was observed: 'ebb and flow' scored higher than 'source and sink' for the same parameter values (Figure 1).

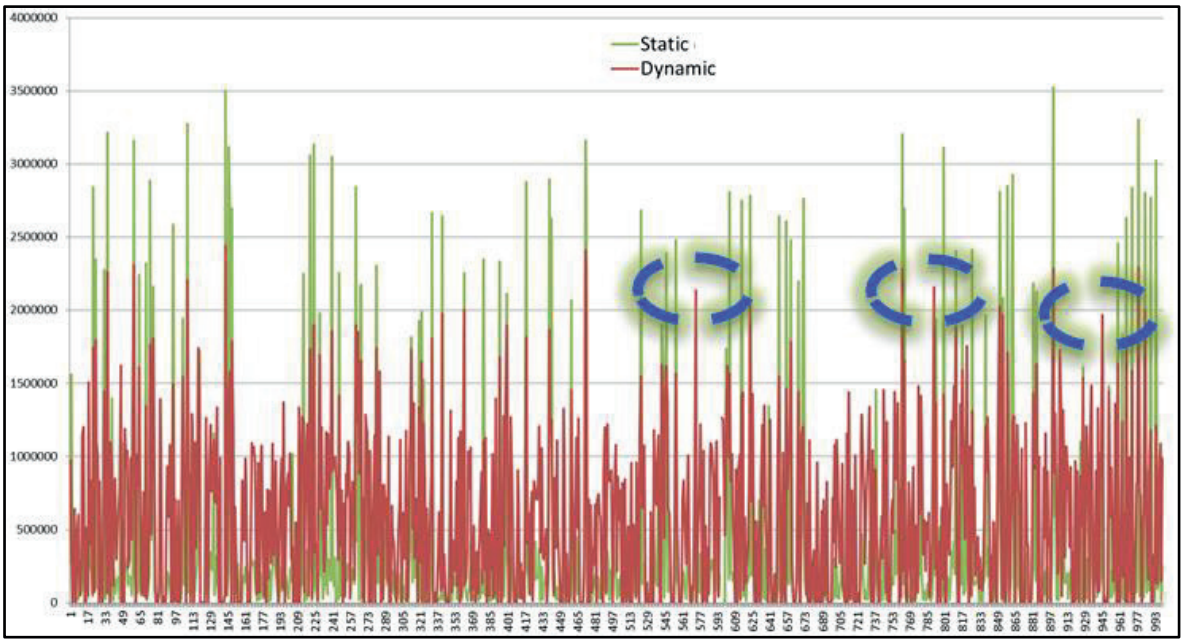



**Figure 1: Simulation scores in HomininSpace 1.0. Blue circles indicate some local maxima where the dynamic or 'ebb and flow' hominin implementation (in red) score higher than the static or 'source and sinks' hominins (in green).**

This important find made clear that the bias in value selection from a researcher should not be used when creating hominin parameter values in this model. Instead in an exhaustive search many different simulations must be executed to find the best simulation results. The new system, HomininSpace 2.0, is able to vary autonomously all parameter values, run simulations, collect and analyse results. The system then identifies promising parameter value combinations and creates new ones in an attempt to improve the match with the archaeology. This is an intelligent and autonomous implementation that can operate without intervention and more important without guidance by a human researcher. Therefore a dedicated implementation of a Genetic Algorithm model (cf. Calvez and Hutzler 2006) is developed that optimizes parameter values while matching archaeology (Scherjon 2016). This automated exploration also allows more hypotheses to be explored simultaneously. Analysis finally identifies those parameters that are relevant for answering the different questions, and the parameter values that give for each question the best match with the archaeological data.

Computer based simulations in archaeological research are mostly related to the testing of hypotheses, to theory building, or to the development of new methodologies (Lake 2014, 260). HomininSpace was developed as a tool to be used to explore the effect of implemented model elements on the behaviour of hominins in a reconstructed environment, with matching archaeological data as the main indicator of success. It is specifically designed to allow easy introduction of new functionality that can be combined with already implemented model elements to analyse more complex (and maybe more realistic) scenario's. The functionality in the model that is associated with certain questions can be activated at will and is used to project the consequences of the alternatives for the behaviour of the system (Nichols 2001).

The model is constructed following a bottom-up, pattern-oriented strategy (Grimm 1994; Grimm *et al.* 1996). In this approach, a pattern in the real world is observed and within the model variables and processes are included to enable (but not force) a pattern to emerge in simulations (Grimm *et al.* 2005, 987). Today one can see hunter-gatherers move through their environment according to certain preferences, and in the archaeology archaeologists find (in)direct traces of past mobility through time in a changing environment. HomininSpace implements a changing environment and a parameterized behavioural repertoire for modelled hominins. This technique is bottom-up in the sense that at the individual level the hominins (agents) are given certain characteristics and are then allowed to interact amongst each other and with the environment. The emerging system level

patterns in the simulation output are analysed and compared against the archaeological data. It then becomes possible to systematically explore how changes at the lower level affect the patterns that emerge at the system level.

The environment in which the hominins in HomininSpace live uses modern day topographical data for land and sea masses, augmented for reconstructed sea levels. Climate parameters that are most influential on the environmental circumstances (temperature and precipitation) are reconstructed for the whole simulation period. The hominins are modelled with parameters inspired by ethnographic data. The use of ethnographic analogies for modelling the hominins in HomininSpace may seem an obvious choice, but the Neanderthals from the Late Pleistocene are not modern humans and they lived in an environment that is largely non-existent today. If they are used such analogies should be well motivated (Wylie 2002, 147-153).

Within the model underlying HomininSpace ethnographic data is used to (1) identify the physiological variables and cultural behaviours that influence dispersal, and (2) illustrate the regularities and the range of variability of these behaviours and constraints thereof. Within the ethnographic record there is not one single culture that forms a direct analogy to the Neanderthal way of life. Instead this record is used to identify the parameters and possible value ranges thereof to be used in the design and implementation of the hominins in the model. HomininSpace is a spatially explicit model that attempts to quantify relevant variables that are involved in hominin dispersal.

# 3. CASE STUDY: THE LATE PLEISTOCENE IN WESTERN EUROPE

## 3.1 Introduction

This chapter provides a background to the core elements of the case study implemented in this research. Introduced are the topography of the research area, the environment during the simulated period in that area and the Neanderthals that inhabited Late Pleistocene Europe and whose presence is attested by the archaeology. Details of the model are given in the chapters that follow. The topographical area in this case study covers the part of Western Europe between latitudes 41.3 and 51.5 and between longitudes -6.3 and 8.5. This area includes more or less the area currently known as France. This area is well suited for agent-based modelling because it is enclosed by natural borders: in the south by the Pyrenees mountains and the Mediterranean Sea, in the west by the Atlantic Ocean and the Channel, in the north by the Ardennes hills and in the east by the Rhine river and the Alpine mountain ranges. The area includes different landscape elements like mountains, hills, plains and coastal areas.

Neanderthals lived here in the Late Pleistocene and archaeology attests to their occupational history. The area is selected also because archaeological research intensity is probably highest here if compared to the wider Neanderthal range, with the greatest density of observation points. Another reason why this area is particularly useful in this study is that according to the archaeology there are parts with a more or less continuous occupation in the south whereas in the north large areas seem to become depopulated when the climate deteriorated. The variation in population density through time and space combined with the relatively high number of archaeological sites provides a unique opportunity to investigate responses of the local hominin population to a changing environment.

HomininSpace simulates hominins moving through a reconstructed and dynamic landscape for an extended period of time. The components and the data flow in the simulation system are visualized in the schema presented in Figure 2. During execution output data is generated for further analysis and illustrative purposes. These include topographical (GIS) display of the simulation and simulation results. Archaeological data resulting from past hominin behaviours is used to compare simulation results against in order to explore hypotheses on those behaviours.

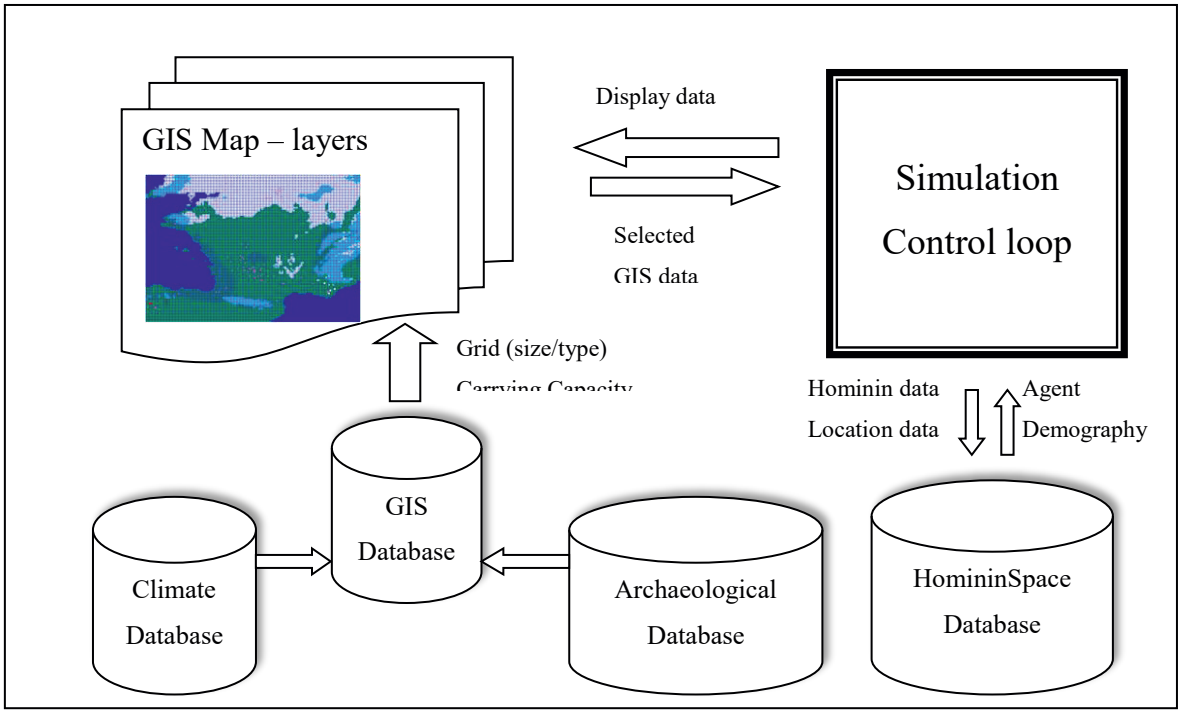



**Figure 2: Components and data flow in the HominінSpace simulation system.**

The simulation period for all experiments in the case study starts at 131,000 years ago and ends at 50,000 ago. Each thus starts at the end of Marine Isotope Stage (MIS) 6 and encompasses MIS 5, MIS 4, and ends in MIS 3[9]. This time frame is selected to ensure that the archaeology used in the simulations is exclusively produced by Neanderthals (Higham *et al.* 2014; Tsanova 2006; Hublin 2015; Roussel *et al.* 2016), which is important for the interpretation of the simulation results. The environment for the selected area and time period is discussed in Section 3.2. Not only the varying climate and the effect thereof on flora and fauna in the simulation area are illustrated but also the effects of advancing and retreating glaciers and associated fluctuating sea levels on the topography itself. The hominins that populated this changing environment are presented om more detail in Section 3.3.

## 3.2 The environment: climate and topography

One of the most important parameters defining the natural environment is the local climate. Climate is the long-term pattern of weather conditions in a particular area, and includes temperature, air pressure, humidity, precipitation, sunshine, cloudiness and wind fluctuations[10]. Climate in the research area is influenced to a large extent by the fact that the Atlantic Ocean, limiting the area to the west, provides a relatively warm sea climate.

[9] MIS stages after Lisiecki and Raymo (2005) (cf. Richter *et al.* 2013).
[10] Definition from http://www.dictionary.com/browse/climate, accessed 16 January 2018.

Further to the east a land climate prevails while in the south the Mediterranean Sea produces warmer climates year round. The mountain ranges which include the Pyrenees and the Alps have their influence on the climate also and to the north glaciers advance when the temperature drops.

In temperate western Europe the fluctuating climate during the latter part of the Quaternary (Cohen and Gibbard 2011) may have had major effects on the spread of flora and fauna. The cyclic changes in the global climate of the Quaternary have been described as dramatic (Guiot *et al.* 1989). Ice and deep sea cores give compelling evidence for sometimes rapid climate change (Sanchez Goñi and Harrison 2010). During glacial and interglacial cycles climate change causes geographic distributions of plants and animals to shift and for all species the changing environment influenced the use of refugia, speciation, range sizes and latitudinal patterns (Dynesius and Jansson 2000; Hewitt 1996). These shifts are caused by differences in the energy levels received and stored in the landscape. The general effect of these changes on the flora is that forest is replaced by woodland and, with increasing aridity and seasonality, in turn by grassland in northwest Europe (Janis *et al.* 2000). Cold and dry conditions in colder periods like the Last Glacial Maximum (LGM) favoured extensive tundra and steppe. Northern hemisphere forest biomes were displaced southward. Boreal evergreen forests (taiga) and temperate deciduous forests were fragmented, while European and East Asian steppes were greatly extended (Prentice *et al.* 2000).

The Late Pleistocene is generally seen as a colder period. But it is initiated by a warm stage, the Eemian. The Eemian Interglacial in Europe is characterized by high pollen abundances of oak (*Quercus*), hazel (*Corylus*), alder (*Alnus*), beech (*Fagus*), hornbeam (*Carpinus*), and other temperate woody taxa with differences among sites likely reflecting interregional differences in Eemian climates as well as variation in substrate. After the last interglacial, the proportion of treeless vegetation increased, with high pollen abundances of grass (*Poaceae*), sagebrush (*Artemisia*), chenopods (*Chenopodiaceae*), pine (*Pinus*), and juniper (*Juniperus*) reported from records spanning the last glacial period (Allen *et al.* 1999). This suggests that summer precipitation, at least, was insufficient to sustain forest (Prentice *et al.* 1992; Prentice *et al.* 2000). What came into existence in large parts of continental Eurasia is referred to as the Mammoth Steppe.

The Mammoth Steppe is a landscape remarkable for its unusual combination of mammal species: woolly mammoths, bison, horses, lions, camels and many more (Guthrie 1982, 309; Guthrie 1990; Stewart 2007). The environment was very productive and characterized

by a very diverse flora and fauna with a large variety in species[11]. The ecosystem of the Mammoth Steppe disappeared during the Pleistocene – Holocene transition and was replaced by the modern tundra, taiga, and steppe environments.

In the Mediterranean area of France the climate was completely different, with a humid and warm climate at the beginning of the Late Pleistocene. When it becomes colder in the north, the south turns more arid (and only slightly colder), causing the landscape to open (Montuire and Desclaux 1997). These phases alternated until the end of the Late Pleistocene when also the southern parts become more colder and more open.

While climate has changed the landscape drastically for people living during the Quaternary, the topography itself changed relatively little (van Gijssel 2006). The landscape in western Europe was shaped in the last 200,000 years by two major processes: glaciation and sea level change. Other processes changed the nature of the landscape and include earthquakes, erosion and sedimentation systems, and volcanic and tectonic activities. Events with lasting impact are diversions of major river systems (for instance the Thames) and creation of volcanic plateaus and mountains (like the Chaîne des Puys in the Massif Central region, around 70 ka).

The area of this case study was hardly affected directly by glaciations, with the Weichselian arctic glaciers reaching no further south than Denmark and covering part of England (see Figure 3). Only the Alpine glaciers (and to a lesser extent those from the Pyrenees) have left their traces but the extent of their glaciers is limited and in the grid scale of the simulations their effects are ignored. The effects of a glaciation extend however much further than the area of the main body of ice. Besides influencing the local climate, glaciers can physically block and divert existing river systems and create new ones due to melt. The weight of the ice depresses the land and can aggravate the above mentioned processes. Rebounding surfaces effect the area long after the ice has gone, influencing local sea levels and again diverting river systems.

[11] See for instance the Stage 3 database on identified faunal remains, on-line available through http://www.esc.cam.ac.uk/research/research-groups/oistage3, verified 26 April 2011.

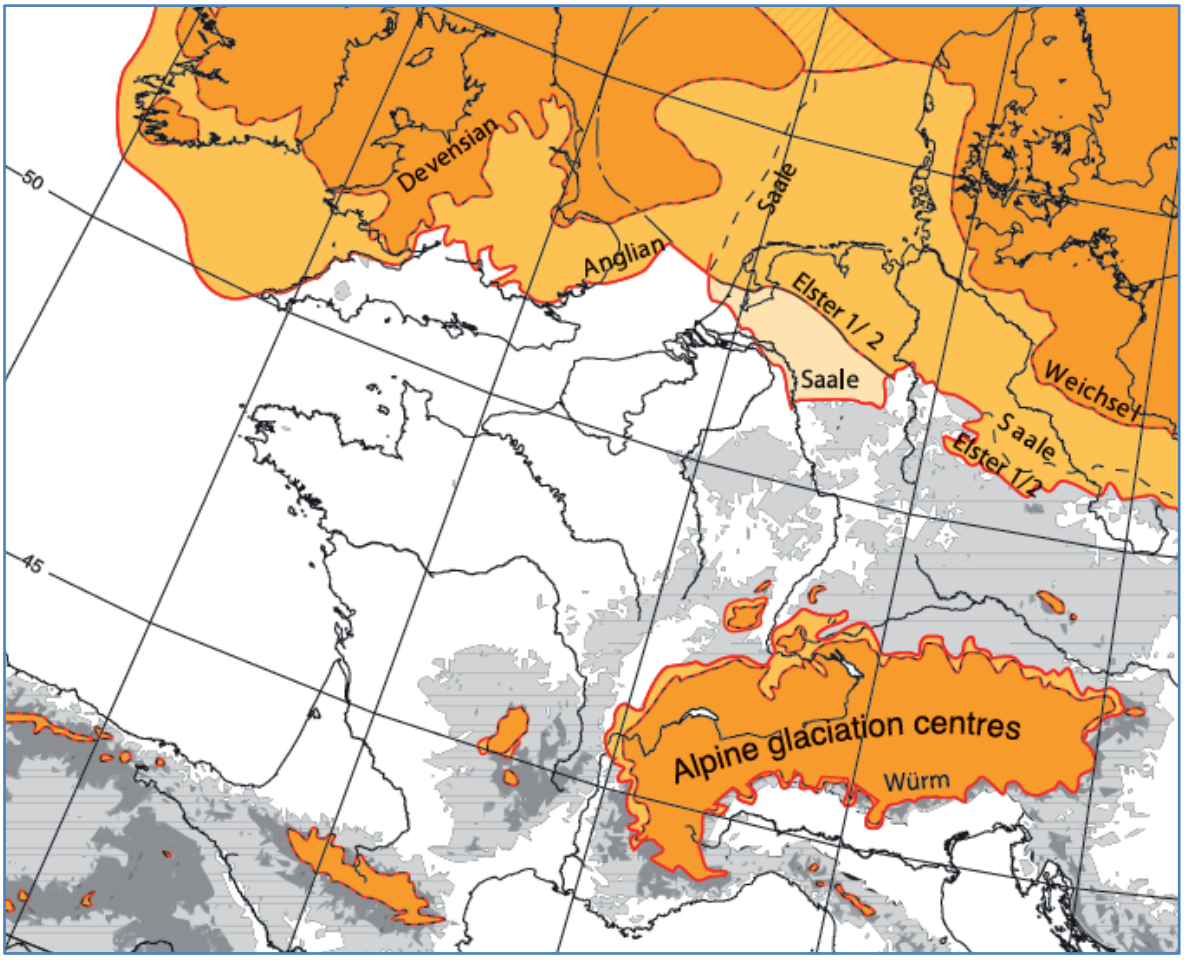



**Figure 3: Maximum ice sheet coverages (adapted from van Gijssel 2006, Figure 3.4). The solid red line follows the maximum limits of the Pleistocene glaciations.**

A good example illustrating the influence of ice sheets is the Channel area between modern day England and France during the severe cold stages of the Quaternary. Here, the accumulation of sweet water in glaciers encroaching from the North made the sea level drop as much as 130 meter, opening a land bridge from main land Europe to the British Islands (Gupta *et al.* 2007). The main river systems in the larger area, including the Thames, the Rhine and the Meuse were diverted into these newly exposed land areas. The huge river system, additionally fuelled with melting waters from the ice, might have posed a formidable obstacle for any hominin trying to reach across (Gibbard 1988). But nevertheless Neanderthals reached the shores of England later during the early phases of MIS stage 3 (White 2006).

Due to the importance of coastal areas (Cohen *et al.* 2012) reconstructing past sea levels is essential for understanding dispersals of species and interpretation of the stratigraphy of geological sequences (Cohen *et al.* 2012). Kelly (1983) separates production environments with many marine resources from those with only terrestrial influences. Sea level changes influence the landscape in a more or less dramatic way. Sea levels on a global scale are

influenced mainly by the presence of glaciers, where more water stored as ice in the glaciers means a lower sea level (Grant *et al.* 2012).

Sea level rise and fall had significant effects on people in the past as well. Due to the diversity of available resources, occupation densities near water bodies were generally high. Lowering of the sea level by growing ice-sheets exposed new lands, while former beach areas became part of the inland ecology. Climate associated with the land-sea border will shift and communities will have to respond. Accordingly, sea level rise forced populations to relocate, causing local population densities and movement which had to be resolved. In some cases rising local sea levels might have been rapid and unexpected, removing communities in the process. Rates of mean sea-level rise reached at least 1.2 meter per century during all major episodes of ice-volume reduction (Grant *et al.* 2012).

Figure 4 presents the bathymetry of the European tectonic plate with in pink those landmasses that emerge when the global sea level drops 120 meters, the assumed global level drop for the LGM (Fairbanks 1989). This map is constructed by Smith and Sandwell (1997) and illustrates what landmass might have been accessible during the simulations.



**Figure 4: Topography for western Europe, sea level lowered with 120 meters. Pink areas are now submerged. Adapted from http://topex.ucsd.edu/marine_topo/mar_topo.html, accessed 8 April 2013.**

## 3.3 The hominins: Neanderthals

The Neanderthal[12] phylogenetic lineage arises somewhere around 500 ka (Meyer *et al.* 2016). In Europe these hominins were typically sturdily built, with short but robust limbs and a specific cranial shape (Weaver *et al.* 2016). Through time they displayed more and more typical Neanderthal features (Hublin 2009; Klein 2009a). The species disappeared around the time that *Homo sapiens* (modern humans) appeared in Europe in the areas previously occupied by Neanderthals (Hublin 2017). The possible replacement of Neanderthals by modern humans is one of the most important questions of palaeoanthropology (Hublin 2017;Hajdinjak *et al.* 2018). They are unarguably the best researched extinct hominin but many aspects of Neanderthal daily life remain poorly known (Anwar *et al.* 2007, 250).

To obtain food Neanderthals frequently hunted some of the largest and most dangerous prey animals of their time (Bocherens *et al.* 2005; Hoffecker and Cleghorn 2000; Roebroeks 2006). These include large herbivores like horse (*Equus* sp.) red deer (*Cervus elaphus*), woolly rhinoceros (*Coelodonta antiquitatis*), woolly mammoth (*Mammuthus primigenius*) and bison (*Bison priscus*). Even when smaller animals are more abundant, it is rather more cost effective to hunt these larger animals in terms of return rates for procurement effort (Kelly 1995, 86-87). Most dietary protein (up to 80%) were obtained from terrestrial animals (Naito et al 2016). Animal species were locally hunted both selectively (i.e. equids) as well as unselectively (i.e. cervids) (Marín *et al.* 2017). Research suggests that Neanderthals were quite capable and successful in obtaining and using biogeographical information on hunted resources (Adler *et al.* 2006; White *et al.* 2016).

Mobility, the sum of small scale movements through larger geographic and temporal scales, enables hunter gatherers to survive (Kuhn *et al.* 2016). Binford (1980) demonstrated that the character of mobility depends on the environment. He used effective temperature to illustrate the relation between environment and different settlement strategies. He reduced earlier schemes (e.g. Murdock 1967) into two main types of mobility: collectors and foragers, where foragers would move the group around looking for resources and collectors would leave the group at one specific location and collect resources on individual or task-specific forays to bring them back to the group. For a

[12] Throughout this work the terms *Homo neanderthalensis* and Neanderthal are synonyms and taken to denote the same species. They are used interchangeably. The species name is written with an 'h' as in the original spelling. As much as possible individual Neanderthals are consequently referred to as 'he'. Wherever applicable this should be taken to mean both male and female hominins of any age.

discussion see Kelly (1995) who found that larger distances were covered in colder climates(117). Environmental productivity steers hominin residential moves (Kelly 1983, 292) and influences population density with most productive and thus most habitable areas being colonized first (Codding and Jones 2013). Occasionally, hunter-gatherer groups become more sedentary, and cease to move residentially as a group while increasing individual logistic mobility (Kelly 1995, 148).

Individual mobility in the form of migration can be tested with strontium analysis (Bentley 2006). Such analysis suggests that at least some Neanderthals were very mobile, albeit not over very large distances (Richards *et al.* 2008). If relative tibial rigidity reflects adaptation to terrestrial mobility Neanderthals were walking or running to a far greater degree than any modern human of today (Shaw and Stock 2013). Used raw material sources and transport may also provide a proxy for past mobility and seem to suggest that Neanderthals in only few instances moved material over large distances (Adler *et al.* 2006; Féblot-Augustins 1999; Féblot-Augustins 2009; Soressi and Hays 2003; Slimak and Giraud 2007).

Gamble and Steele (1999) calculated that areas used by Neanderthal groups could be relatively large (2025 $km^2$ for Grotte Vaufrey VIII). But daily travelled distances for Neanderthals are likely to have been relatively short (Verpoorte 2006), supported by the fact that more than 90% of the raw material was transferred less than 20 km from its source (Féblot-Augustins 1999). This could suggest that most groups were not migrating over larger distances or that lithic artefacts were not taken along, and that the exceptional artefacts were either exchanged between Neanderthal groups or transported over longer distances by some individuals or groups.

Neanderthals disappeared from the archaeological record around the time modern humans arrived in Europe. It is currently assumed that modern humans arrived in western Europe not before 50,000 years ago (Appenzeller 2013; Higham *et al.* 2014; Tsanova 2006; Hublin 2015; Roussel *et al.* 2016; Teyssandier 2008). The presence of the first modern humans is attested by physical remains and via association with sites and lithic assemblages attributed to the Aurignacian culture family. The distribution of the different Aurignacian technologies appears to trigger the transition from Middle to Upper Palaeolithic. This process often includes transitional industries and seems to temporally coincide with the eventual disappearance of the Neanderthals.

The earliest appearance of these archaeological cultures is around 43-42 ka (Higham *et al.* 2011). For some, most notably the Châtelperronian, a Neanderthal origin is inferred

(Hublin *et al.* 2012; Roussel *et al.* 2016; Ruebens *et al.* 2015). The presence of modern humans is likely a direct (Mellars 1992) or indirect (through stimulus diffusion of culture) cause for the Neanderthal demise (Roussel *et al.* 2016). Due to geological processes (like low sedimentation rates), periglacial processes (like cryoturbation), and due to poorly recorded excavations there are quite a few interpretive problems for sites dating to the transition period around 50-35 ka (Ruebens *et al.* 2015). To prevent incorrect assessment of the Neanderthal record in this study by potential mixing of early modern human influences the upper boundary of the simulations is set to 50 ka.

# PART TWO: CREATING THE ACTORS

*"Van Helmont, in the midst of odd experiments to obtain mice from junk and sawdust, made one rather intelligent experiment. He grew a willow twig weighing 5 lb in a large clay pot containing 300 lb of soil and irrigated it with rainwater. After 5 years, he harvested a willow tree of 164 lb, with a loss of only 2 oz of soil. Van Helmont concluded from this that water was condensed to form plants."*

(Lieth 1973, 305)

# 4. THE MODELLED ENVIRONMENT

## 4.1 Introduction

HomininSpace models the environment of the Late Pleistocene for part of western Europe for simulations with modelled hominins. The topography is reconstructed for a delimited area using current day topographical and bathymetry data taking a fluctuating sea level into account (Section 4.2). HomininSpace further uses values for past temperatures and precipitation levels (described in Section 4.3) in order to compute how much energy there is in the landscape available for hominins who sustain themselves mostly with the consumption of large ungulates (Section 4.4). Ultimately this gives a carrying capacity for every grid cell for every time step in a simulation that will guide mode and speed of hominin dispersal.

Within the HomininSpace simulation system climate and topography are modelled to study the effect that temperature and precipitation can have on the behaviour of Late Pleistocene hominins in the landscape (Banks, d'Errico, Peterson, Kageyama, *et al.* 2008; cf. Eriksson *et al.* 2012). It is suggested that these are probably the two most important environmental variables that influenced changes in the abundance of fauna and flora through time (Kelly 1983, 279; Krohne 2001, 47-51; Monserud and Leemans 1992; Thackeray 2013). These variations drive availability and prevalence of food resources and were among the most important factors determining whether hominins could survive and thrive in any particular area (Hamilton *et al.* 2007; Freeman and Anderies 2015).

It is however notoriously difficult to reconstruct Pleistocene environments and to isolate individual parameters that would have attracted or deterred hominins, if only because ancient hominins lived in a wide diversity of landscapes and were more or less successful in each one of them (Roebroeks *et al.* 1992; Zuk 2013). Using only a few variables to model biomes as different as those of the British Isles, the western European mainland, the Mediterranean coast or the diverse mountainous areas of the Pyrenees and Alps invariably leads to simplification. HomininSpace makes use of the fact that all these environments have in common that they can exist only by virtue of both the influx of energy and the amount of water supplied to each square meter (Guthrie 1990, 214), which are in turn influenced by topography (Boivin *et al.* 2013, 38).

Hominins, like all animals, fulfil their energy requirements by extracting it from the landscape around them. Energy in general is provided by the sun delivered onto the earth in the form of heat and light, and converted into primary biomass using water and nutrients from soil and air. This process is referred to as primary production and is mainly done by plants and marine algae (some bacteria also contribute). All energy in the environment is ultimately generated by the available sunlight[13]. The amount of solar energy is described by the Milankovitch cycles, the orbital forcing which is the combined effects of obliquity (axis tilt), eccentricity of the orbit of the earth around the sun and the precession of the orbit (Hays *et al.* 1976). This fluctuating solar radiation energy or insolation is further subject to atmospheric losses due to scattering and absorption and latitudinal differences between absorbed and emitted radiation and therefore varies throughout the days and months.

Insolation is the most important source of energy in primary productivity, the production of organic materials from atmospheric carbon dioxide. The produced vegetation forms the base of every food chain. The most important factor defining the net production is solar energy but nutrient availability, water management in the area and local temperature also play a part (Hawkins *et al.* 2003). Primary production decreases with latitude (cf. Damuth 1991; Silva *et al.* 2001) and the amount of evaporation and variation in geography cause most deviation from expected solar radiation production (Binford 2001, 82).

Herbivores or primary consumers directly consume plant material and carnivores, situated at higher levels of the food chain, consume other animals. This accumulation of living secondary biomass by heterotrophs is referred to as secondary production (Allaby 2005; Benke and Huryn 2006). There is a direct relation between terrestrial primary productivity and the availability of ungulates (Coe *et al.* 1976; Janis *et al.* 2000; Oesterheld *et al.* 1992), which are the prime targets for Pleistocene hominins. The productivity hypothesis states that the primary production will limit the secondary production of herbivores, and the amount of herbivores (secondary biomass or SB) will limit the number of predators (Hawkins *et al.* 2003, 3106).

Hominins can sustain themselves with both plant materials and meat. The archaeological record is so far not sufficiently detailed to allow determination of exact proportions (Kaplan *et al.* 2000, 180). The underlying model in this thesis however assumes that

[13] The contribution of internal heat from the earth's core is relatively low (less than one percent) and mainly visible in the movement of the continents.

hominins primarily feed on large terrestrial game to satisfy their energy needs (Salazar-García *et al.* 2013; Snodgrass and Leonard (2009) but see Power *et al.* (2018)), as summarized in the schema below.

| *Primary Production* ==➔ | *Secondary Production* ==➔ | *Neanderthals* |
|---|---|---|
| (Vegetation) | (Herbivores) | (Predators) |
| Limited by solar energy and water availability. | Limited by Primary Production. | Limited by Secondary Production. |

The amount of energy available as herbivores is often predicted using empirical data from different climates and vegetation types to infer a formula with one or more climatic parameters. Examples include Coe *et al.* (1976) who are modelling large herbivores on the east African grasslands using rainfall, or Eisenberg and McKay (1974) who also used rainfall to infer the amount of herbivores in tropical forests. To model secondary biomass in different biomes a single climatic parameter is however generally not sufficient (Binford 2001, 106). HomininSpace uses precipitation as well as temperature and implements two different model variants to calculate the available energy in the landscape as the availability of large edible ungulates as food for hominins in the area of western Europe during the Late Pleistocene (Section 4.4).

## 4.2 Reconstructing topography with fluctuating global sea levels

Paleobathymetric variation can influence early hominin dispersals by exposing or hindering likely routes (Nakazawa and Bae 2017). The model underlying HomininSpace uses reconstructed sea levels which are compared to modern day topography data for land masses and upon current day bathymetry (bed rock data) for marine areas. This latter is rationalized by the fact that the bed-rock level of the sea bottom represents the maximum depth where erosion and sedimentation of sand layers constantly change the local conditions. The model ignores further the effects of suppression of land surfaces by glaciers, rebounce when the ice retreats (glacio-hydro-isostatic contributions to local sea-levels (Lambeck and Chappell 2001)), and the formation of natural sea defences like dunes and swamp marshes that might influence or hinder (local) sea level rise or fall.

The simulation area is divided into grid cells with the Spatial Analysis in Macroecology tool (SAM) (Rangel *et al.* 2006; Rangel *et al.* 2010). The area used in the Neanderthal case

study 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). This area includes the South of England but excludes Sardinia and little of Spain is included. The area encompasses all of the Pyrenees mountain range and the western part of the Alpine regions.

A grid cell is defined as 0.1 x 0.1 degrees (roughly 10 x 10 kilometers, depending on the latitude) with the area composed of 14.948 grid cells (Figure 5). The surface area of 10 x 10km has been selected after extensive performance testing, and represents a compromise between detail of climate and topographical data and simulation duration. The area of a single grid cell is generally insufficient to support a hominin group. Foraging radii of modern hunter gatherers hardly ever fall below ten kilometer. For the enviromental settings of today, 4394 of the cells in the simulation area are under water. This number reduces considerable when the sea level drops in colder periods and more land emerges. The dimensions of a grid cell can mask the real reasons why Neanderthals were somewhere in the past. They select places to stay but since a modelled group only forages in an area and varying motivations for specific site selection are not implemented (for instance a preference for locations near other-than-food-resources).

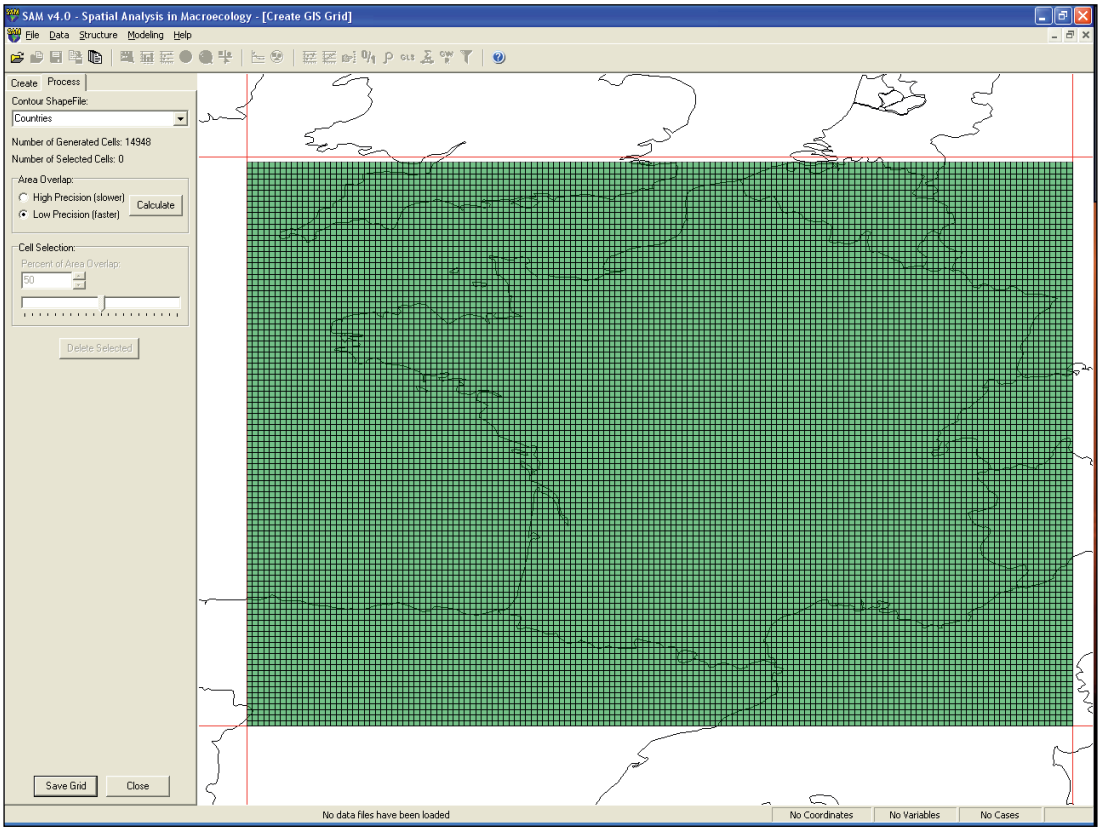



**Figure 5: The simulation area divided into 14.948 grid cells in the SAM tool.**

SAM is also used to distribute topographical and climate parameters onto these grid cells. Next to the topographical data (altitude or elevation) these parameters include 19 bio-values from http://www.worldclim.org (both present and past climate data). In each grid cell all the datapoints with information (there can be more than one datapoint in a grid cell depending on the underlying dataset) for each parameter are collected, and they are converted into aggregrate values per grid cell: mean, min, max, std, etc. For most variables HomininSpace uses the mean value. However, for bedrock the maximum value is used (see below). When data is missing the values are automatically interpolated from neighbouring cells by SAM.

Altitude information is derived from the modern day dataset downloaded from http://www.worldclim.org (verified 16 April 2013). Unfortunately this dataset does not contain depth information for areas currently located under sea level but which during the simulations can become dry land. The dataset contains only elevation data for land areas. For the sea grid cells bedrock data from a different dataset, the ETOPO1 dataset is used (https://www.ngdc.noaa.gov/mgg/global/global.html, accessed November 2016). Bedrock data details the altitude information for current day submerged landmass. The model assumes that whenever a single bedrock datapoint in a grid cell emerges (because the water level drops below this height), the total grid cell becomes useable land with the altitude of that point. Thus, for the bathymetry data the aggregated maximum is used (which is the minimum depth). Emerged land cells then take missing values for temperature and precipitation extrapolated from nearby coastal areas (the worldclim current day data does not contain values for water cells).

Altitude data in WorldClim.org is derived from the Shuttle Radar Topography Mission (SRTM) dataset, which contains elevation (above sea level) data on a near global scale in very high resolution. The SRTM radar system flew onboard the Space Shuttle Endeavour in February 2000 (http://www2.jpl.nasa.gov/srtm/, address verified 1 October 2013 with links to the original datasets). WorldClim.org provides datasets with aggregrated data in different spatial resolutions. HomininSpace uses the 2.5 arc minutes distribution (5-8 km), for modern-day data and for past climate data sets. This is the highest uniform resolution for which all datasets are available. The data can be downloaded in ZIP file format (compressed data set) with one data layer for the altitude besides the bioclimate variables.

Bathymetry data for parts of the Channel area and Atlantic Ocean and other sea grid cells were obtained from the National Geophysical Data Center (NOAA). Among the data that

this organization makes available is the ETOPO1 dataset (Amante and Eakins 2009). Figure 6 presents the data taken from the website (http://maps.ngdc.noaa.gov/viewers/wcs-client/, accessed 13 April 2013). This data is retrieved in XYZ format (geocentric coordinates) (Frakes 2003). SAM can read a matrix of XYZ formatted data (Rangel *et al.* 2010), and because these files are georeferenced it can map the data into the grid cells. Those grid cells that have no altitude value from WorldClim (sea and lake cells) take the altitude from the ETOPO1 dataset. The ETOPO1 dataset compares well against other datasets (see http://vterrain.org/Elevation/Bathy/bathy_comp.html, with more bathymetry data available at http://vterrain.org/Elevation/Bathy/, both accessed 12 April 2013).



**Figure 6: Bathymetry data for the simulation area (the red square), taken from the ETOPO1 dataset.**

Mean global sea level values are taken from Bintanja and van de Wal (2008) (Figure 7). This data matches rather well the more recent sea level reconstructions made by Grant *et al.* (2012) (Figure 8). The latter argue that ice-core data chronology as used by Bintanja and van de Wal (2008) is not fit for sea level reconstruction since this assumes that the ice volume is in a systematic phase with either Antarctic or Greenland climate, which may not be the case (Grant *et al.* 2012, 744). They connect dated cave speleotherms from the eastern Mediterranean to a high resolution foraminiferal record using the Red Sea basin isolation concept.

However, although Grant *et al.* (2012) claim a high resolution continuous record, their data for the LGM for instance contains only four values from 13,857 to 26,055, a period of 13,000 years, with a minimum value for Red Sea level of -115.74m. Bintanja and van de Wal (2008) provide a continuous record with values every 100 years and a minimum sea level value for 19,800 years ago of -123.41m. They also provide their level as a relative global sea level, where Grant *et al.* (2012) provide a Red Sea relative level. The results from Bintanja and van de Wal (2008) closely match the benthic foraminiferal d$^{18}$O curve (Waelbroeck *et al.* 2002). For further comparison and an overview of reconstruction methods see Medina-Elizalde (2013).

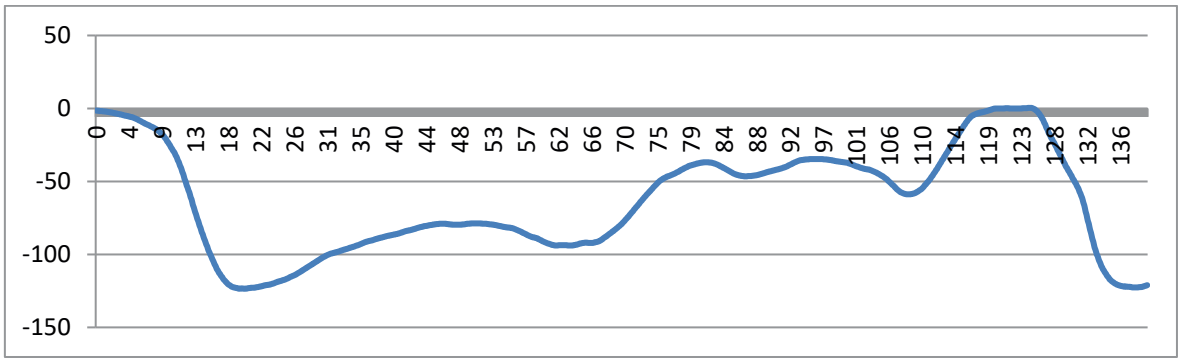



**Figure 7: Reconstructed global sea level in meters (Y-axis), from today to 140 ka (from Bintanja and van de Wal (2008)).**

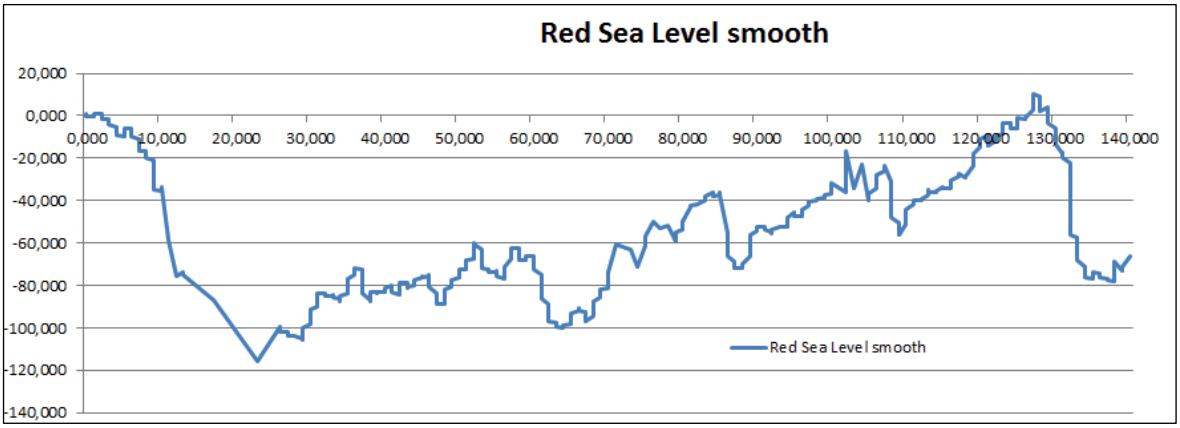



**Figure 8: Red Sea level until 140 ka adapted from the supplementary materials from Grant *et al.* (2012) converted to equidistant horizontal axis.**

Global sea level values taken from Bintanja and van de Wal (2008) are interpolated between time steps where needed. In each grid cell the height (altitude or depth) is stored where 0 (zero) is the modern day sea level. In each time step the reconstructed value of the sea level is added to all height levels. A resulting value of -10 meters then exposes all land mass with a depth between 0 and 10 meters. A value of +2 meters sea level rise will drown all current day landmasses with an altitude below two meters.

## 4.3 Distribution of reconstructed temperature and precipitation

The spatial and temporal distribution of the temperature and precipitation in HomininSpace is derived from two different datasets: (1) the current distribution of climatic variables over the world and (2) their distribution as reconstructed in high resolution for the Last Glacial Maximum (LGM, ~21 ka). The LGM is a very cold period with enlarged ice sheets and low atmospheric $CO_2$ levels. Modern day data is considered to represent a very warm period. The aim for HomininSpace is to have reconstructed environmental data for each time step for the simulation period (roughly 131-50 ka, with each time step representing one year in real time). Such continuous data is not available and expensive to compute. Time series for HomininSpace must be obtained through other means. Local conditions are estimated using a reconstructed global yearly mean temperature record as an index for interpolation between the possible local climatic extremes.

Since past climate variables cannot be measured directly, local paleoclimates are generally reconstructed through interpretation of a diverse range of proxy indicator patterns (Isarin and Bohncke (1999); Lowe and Walker 2014). These include data from lake sediments, tree rings, ice cores and peat deposits. Most of the terrestrial archives are however short and/or discontinuous which is why instead often the longer and more continuous marine data sets are used (Sanchez Goñi *et al.* 2008; Rasmussen *et al.* 2014). Most bioclimatic indicators respond to more than just climate and must be used with care when reconstructing and interpreting past conditions to compare simulation results against (Kageyama *et al.* 2013). Correlation between marine and terrestrial sequences are notoriously difficult but occasionally possible, for instance by using common event markers (Sier *et al.* 2011). The European continental climate records for instance correlate rather well with the global oxygen isotope record from marine environments (Mosbrugger *et al.* 2005).

Temperatures reconstructed from ice volumes are yearly global mean values. Local temperatures (and precipitation values) can differ enormously with strong regional expression, and between locations and continents the yearly means can be far apart (Consortium 2013). As such, these nearly continuous data sets from the stable isotope data analysis are very good *relative* indicators of the temperature fluctuations through time but less suitable for *absolute* local reconstructions. Within the HomininSpace modelling system such a record is therefore used to scale between high resolution local spatial reconstructions for two extreme climatic phases (Lawing and Polly 2011).

To reconstruct the yearly mean temperature Bintanja and van de Wal (2008) use the integrated LR04 marine stack of benthic $\delta^{18}O$ from 57 globally distributed sediment cores (Lisiecki and Raymo 2005). They quantify and separate the two main components affecting the $\delta^{18}O$ values using models that describe the relationship between surface temperature and deep-water temperature and between surface temperature and the storage in ice sheets (see also Moum *et al.* (2013) or Shen *et al.* (2005), 379). They then build a continuous reconstruction of atmospheric temperature, ice volume and global sea level for the last three million years. Figure 9 presents their data for the period of interest in the Neanderthal case study. In their supplementary information section Bintanja and van de Wal (2008) compare their results favourably against different and independent proxy records of variable length. Reconstructed values become more uncertain further back in time, especially before 400 ka (Bintanja *et al.* 2005).

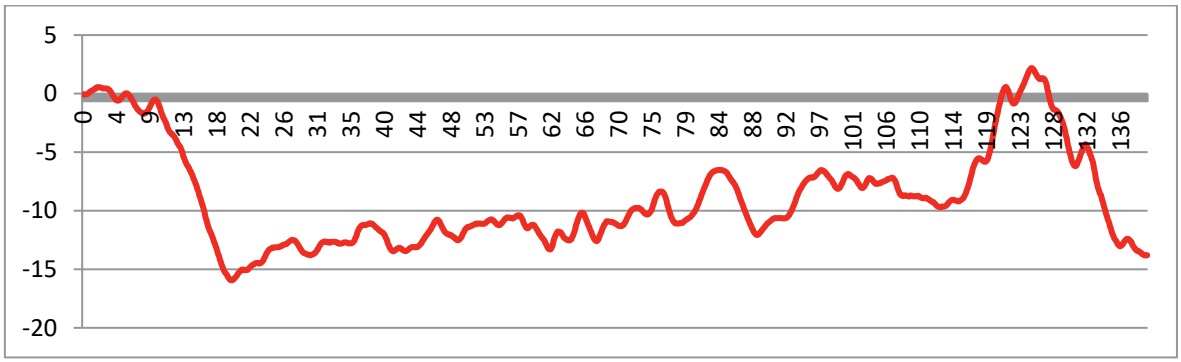



**Figure 9: Overview of reconstructed global mean temperatures (Y-axis) for the last 140ky, drawn according to Bintanja and van de Wal (2008).**

In the temperature record global and many local minima and maxima can be found, with some of the more prominent ones given in Table 2 (these points in time will be used to monitor environment reconstructions). Dating of the LGM climatic minimum matches estimates from other areas (Barrows *et al.* 2002; Yokoyama *et al.* 2000). A yearly temperature of minus 15.9 ºCelsius is an absolute minimum for the period considered. It is a mean value, suggesting that colder periods throughout the year were present. The yearly temperature during the Mid Holocene is hardly higher than modern day temperatures, albeit differently distributed throughout the year, which is why this study uses current day data as maximum. The reconstructed data[14] is available as an Excel database in the supplementary information of Bintanja and van de Wal (2008) (see subsection 8.5.2 for a detailed description of this input file). This high resolution temperature record is used in

[14] Global 3Ma Temperature, Sea Level, and Ice Volume Reconstructions collected by R. Bintanja and R.S.W. van de Wal. *IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series* # 2011-119. NOAA/NCDC Paleoclimatology Program, Boulder CO, USA.

this study to interpolate between climatic extremes since it represents the climate more accurately than the isotope data itself and since it is a very good relative indicator of global conditions through time.

**Table 2: Local and global minimum and maximum temperatures from Bintanja and van de Wal (2008), supplementary information. Temperatures are offset in degrees Celsius relative to modern day yearly mean values.**

| *Description* | *Date (years ago)* | *Temperature (in Celsius)* |
|---|---|---|
| A local maximum around 2 ka | 2,100 | 0.56 |
| A local maximum around 6 ka (Mid Holocene) | 5,800 | 0.03 |
| A global minimum around 19 ka (LGM) | 19,600 | -15.94 |
| A local minimum around 61 ka | 61,500 | -13.31 |
| A local minimum around 78 ka | 78,100 | -11.85 |
| A local maximum around 83 ka | 83,800 | -6.51 |
| A local minimum around 88 ka | 88,700 | -12.07 |
| A local maximum around 97 ka | 97,200 | -6.53 |
| A local minimum around 112 ka | 112,800 | -9.70 |
| A local maximum around 120 ka | 121,400 | 0.57 |
| A global maximum around 124 ka | 124,800 | 2.19 |

Data on climatic conditions is provided by WorldClim, one of many sources of climate data and a database of global climate layers with a very high resolution (Hijmans *et al.* 2005). It was developed at the University of California, Berkeley and is freely available through http://worldclim.org (first accessed 25 July 2013). The data for modern day climate parameters are collected from weather stations and other sources from around the world and originally sampled in 30 second resolution. Coarser datafiles have been created by calculating the mean of the original variable over larger grid cells (Hijmans *et al.* 2005). Bioclimatic variables are derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables. The bioclimatic variables represent annual trends (e.g., mean annual temperature, annual precipitation), seasonality (e.g., annual range in temperature and precipitation) and extreme or limiting environmental factors (e.g., temperature of the coldest and warmest month, and precipitation of the wet and dry quarters) and are used in numerous biogeographical studies (Elith *et al.* 2006; Broennimann *et al.* 2012) and widely recognized for their practical value (Fick and Hijmans 2017). A list of all variables is given in Appendix 1.

The used data for LGM conditions at WorldClim is calculated using a coupled general circulation models: Community Climate System Model (CCSM), available from http://www.worldclim.org/past, verified 10 January 2018 (Collins *et al.* 2006). It is

favourably compared against other modelling results in Appendix 2. Such models integrate mathematical functions describing atmospheric and oceanic circulation, sea ice, land surface characteristics and properties of the atmosphere (Hijmans *et al.* 2005). Running such simulations is time consuming, computationally intensive and expensive. Data is only available for a limited number of time steps. That is why this study interpolates between high resolution extreme local conditions. Results farther back in time are progressively more uncertain. Data from the WorldClim database was selected due to its availability of similar data sets for current and past periods, the easily usable format in high resolution (30 arc second or 1-km resolution), and the included level of detail (19 bioclimatic parameters for all datasets).

### 4.3.1 Reconstruction via interpolation between climate extremes

The WorldClim database contains data on modern day environmental conditions, together with future projections and past climate reconstructions. However, intermediate results between reconstructed points in time are not available, especially not at a high resolution temporal scale. The model underlying the current research uses a linear interpolation between two known data points of extreme climates to generate continuous values for climatic parameters in the past.

A straight linear interpolation between current climate and a general circulation model of the LGM using stable oxygen isotope ratios provides a good estimate of paleoclimate parameters for other time periods, including extrapolation for points in time outside the range of the two measuring points (Lawing and Polly 2011). Lawing and Polly (2011, 8) compare the results of two general circulation models with their interpolated dataset of temperature and precipitation. They find that the differences between their data and each of these modelled reconstructions are less than or equal to the differences between these two models. This suggests that their interpolated model is of equal quality as the general circulation models.

Lawing and Polly use an interpolation that is governed by stable oxygen isotope ratios which unfortunately are not only influenced by oceanic $\delta^{18}O$ but also by local deep-water temperatures (Imbrie *et al.* 1984). The underlying research therefore instead uses reconstructed surface temperatures in which those effects are compensated (Bintanja and van de Wal 2008). HomininSpace uses an similar interpolation based on isotopic values measured from within ice cores, but with the isotope ratios converted into a mean global temperature reconstruction that compensates for effects from upwelling deep water that

represent earlier climatic influences (Bintanja and van de Wal 2008). It is expected that this compensation provides more realistic results for reconstructing temperature and precipitation at any point in time.

For each grid cell values for LGM (global temperature offset at -16) and for current day (global temperature offset is 0) local precipitation and temperature are available[15]. The reconstructed global temperature from Bintanja and van de Wal (2008) is used as an index in each step to interpolate in each grid cell between these two values to obtain interpolated values for precipitation and temperature. This interpolation method is illustrated in Figure 10. For a hypothetical location two temperature values are given: the value for the LGM which is 5, and for today a value of 12. Horizontally in Figure 10 the reconstructed yearly temperature (a global mean value) is used to index the interpolated value. Thus, when at any time during the simulation this yearly temperature takes the value -8, the interpolated temperature will be calculated at 8.5. In a similar manner this global mean temperature is used to index the line between precipitation values for LGM and today.

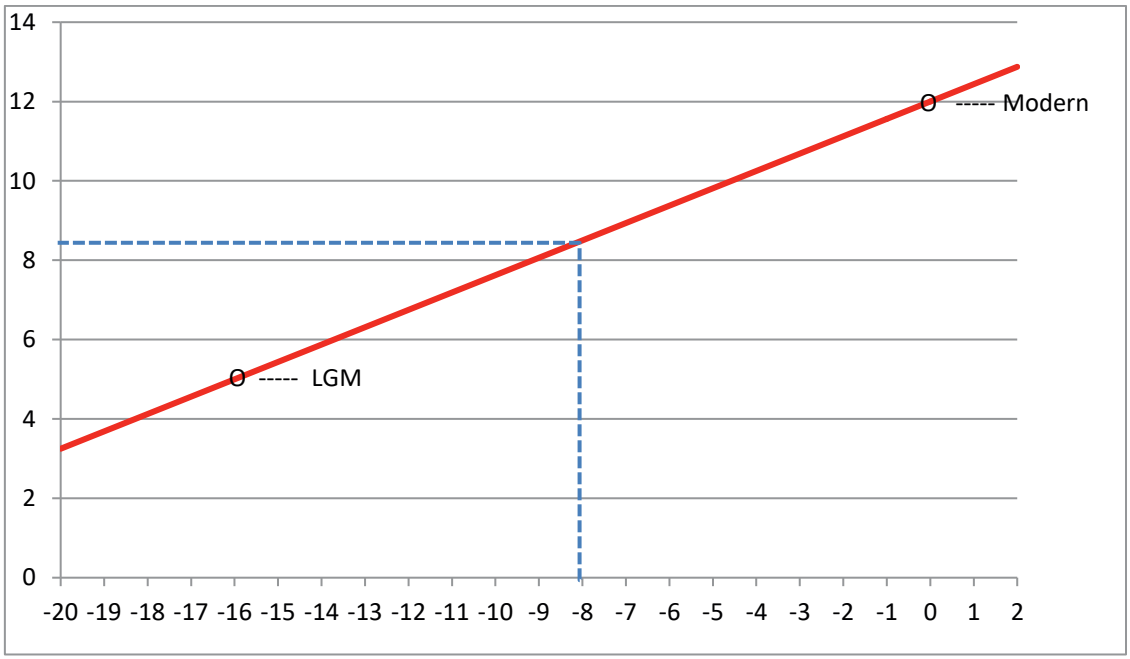



**Figure 10: Interpolation of temperature values between two given data points: LGM and modern day values, with horizontally the reconstructed global mean temperature used as index for the linear interpolation.**

The distribution of the climatic values for the interpolation for mean annual temperature and precipitation are given in , Figure 12, Figure 13 and Figure 14. For the warmest and coldest months Figure 15 presents the values for modern day and LGM. A frequency

[15] For some grid cells, notably current day sea areas, values are calculated after interpolation from nearby land cells.

distribution for three LGM parameters is given in Figure 16. Temperatures are in degrees Celsius, rainfall is in mm per year. For those grid cells where current day data is not available (for instance for exposed shelves in the Channel area) values for temperature and precipitation are extrapolated from nearby grid cells (compare with the similar extrapolation of soil properties in Hoogakker *et al.* (2016, 55)). Color coding of the grid cells in these and other figures is according to the values in Table 3.

**Table 3: Color coding of grid cells for temperature and precipitation.**

| Color coding for temperature (top) in Celsius, precipitation (bottom) in mm per year. | | | | | |
|---|---|---|---|---|---|
| -30 | -15 | 0 | 15 | 30 | 40 |
| 300 | | 1400 | | | 2400 |

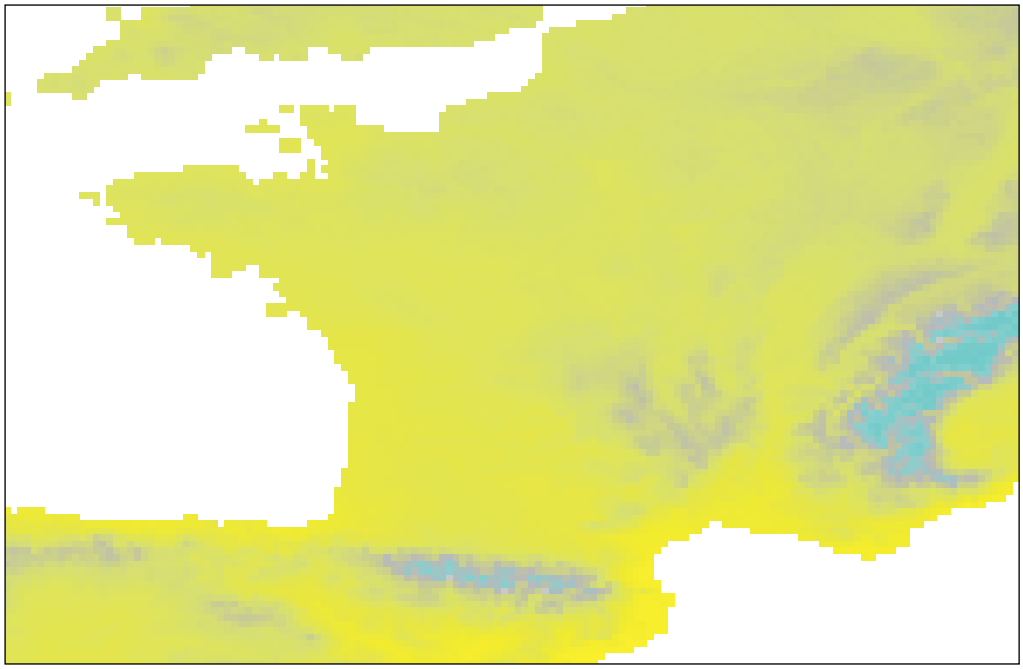


**Figure 11: Mean annual temperature distribution for modern day conditions.**

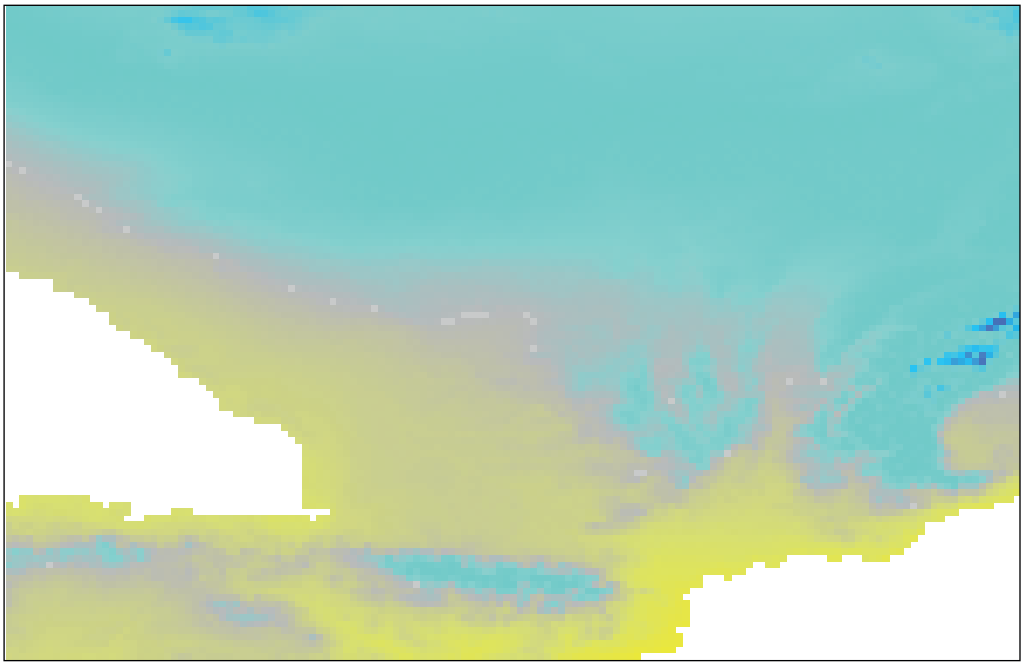


**Figure 12: Mean annual temperature distribution during the LGM (source CCSM).**

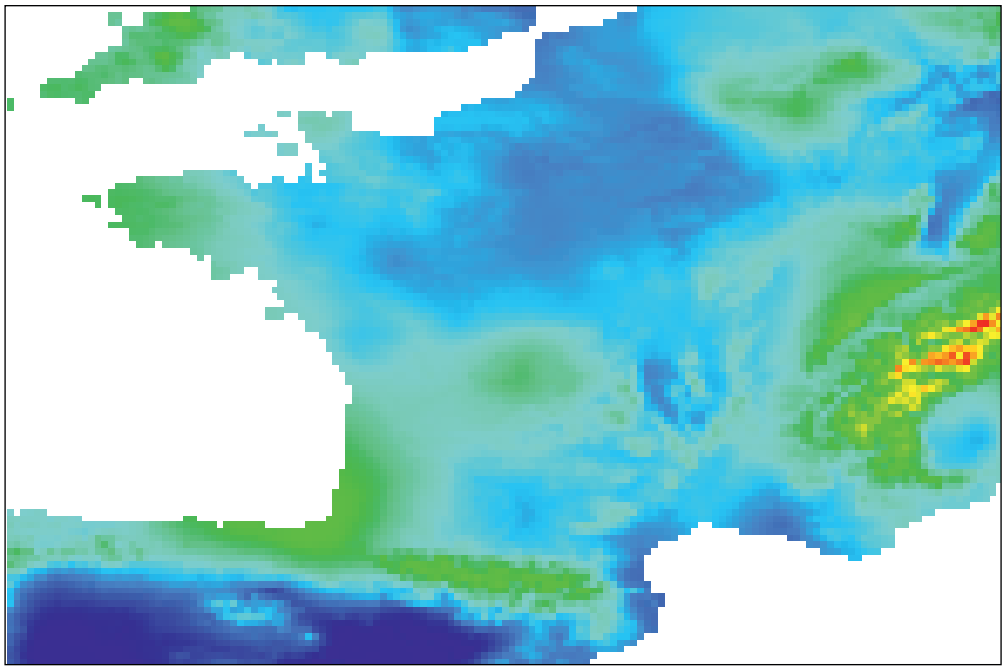


**Figure 13: Annual precipitation distribution for the modern day conditions.**

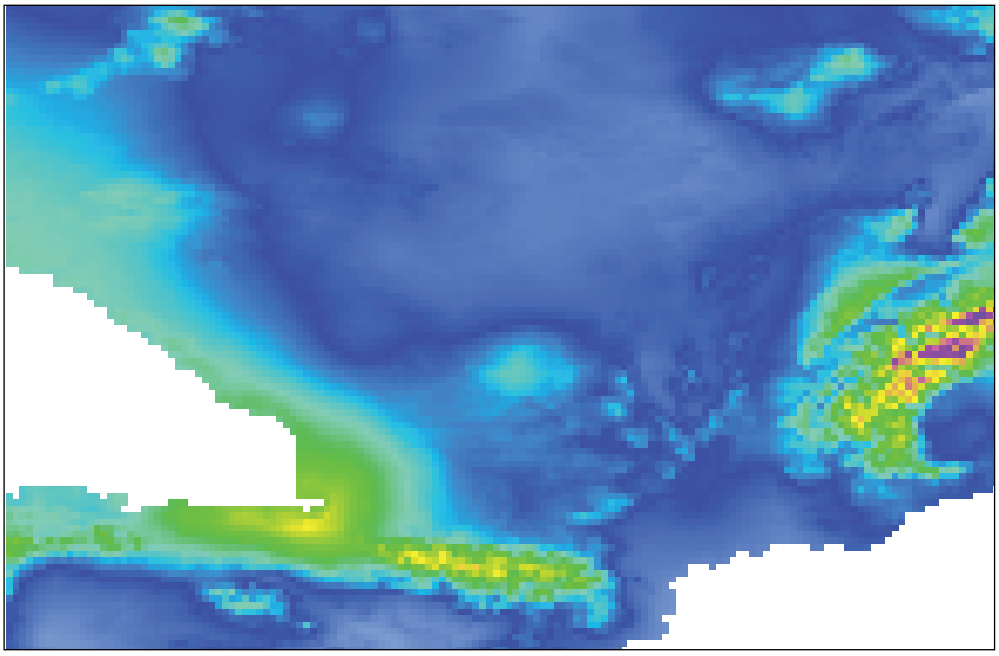


**Figure 14: Annual precipitation distribution during the LGM (source CCSM).**

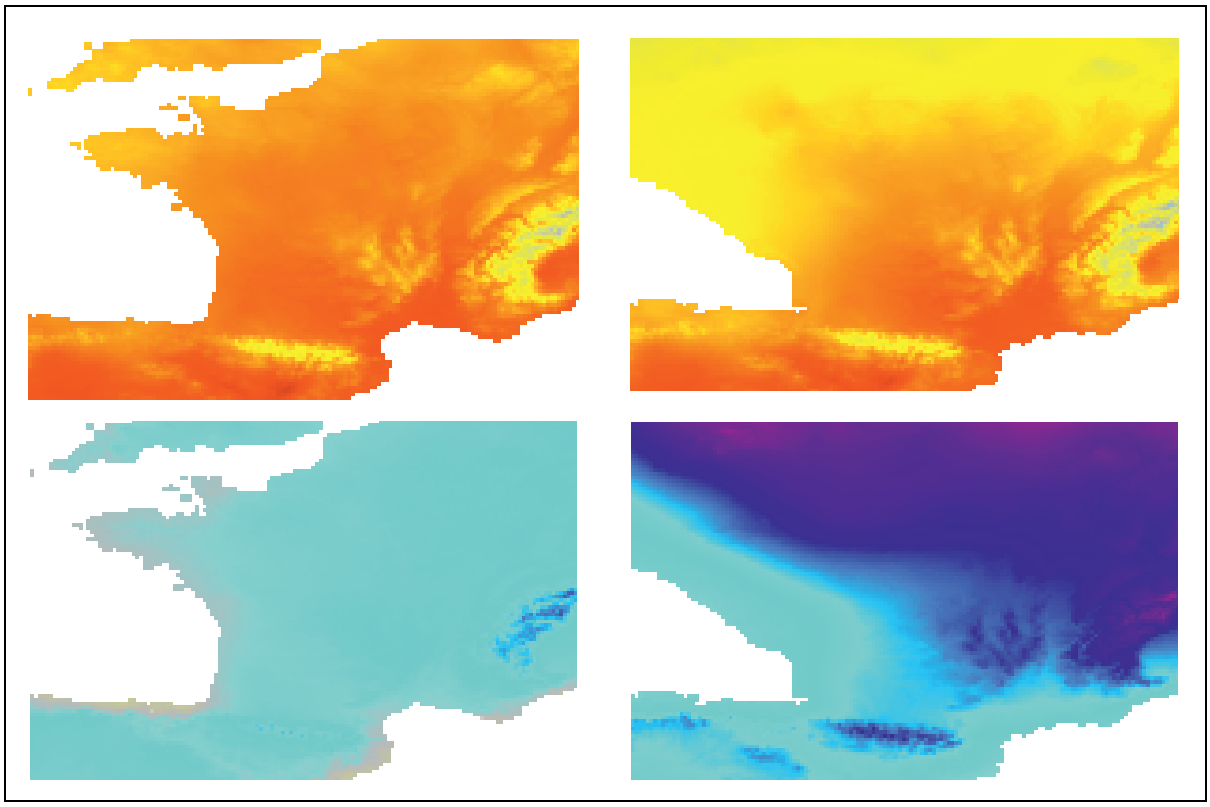
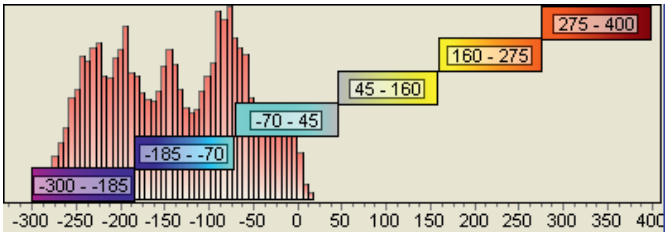


**Figure 15: Overview of modern day (left) and LGM temperatures of the warmest month (top) and coldest month (bottom).**



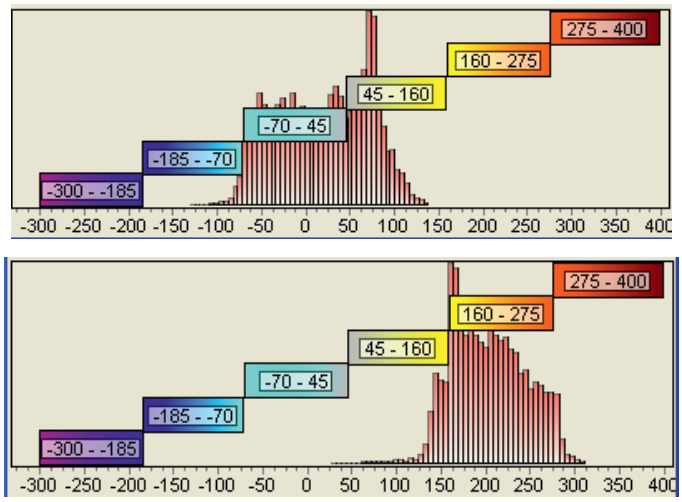



**Figure 16: Relative frequency distribution for the three LGM temperature climate parameters: coldest month (top), yearly mean, and hottest month (bottom).**

## 4.4 An energy landscape in two forms

HomininSpace uses the two most important climatic parameters, temperature and precipitation, as defining factors for the reconstructed climate in the simulations, and computes from these the amount of energy in the landscape. Reconstructing the available energy levels within the HomininSpace modelling system involves several steps. The aim and final prediction of these calculations is an amount of energy available in the form of edible secondary biomass (large ungulates). This is recalculated for every year, with results stored in each grid cell, and forms the basis of the carrying capacity per cell, taking fluctuating sea levels into account.

HomininSpace explores and compares two different methods to calculate secondary biomass given the primary production. The first one (Subsection 4.4.2), henceforth referred to as the *continuous energy model*, uses data on primary production and available secondary biomass from a wide range of climates and vegetation types to infer a single direct relationship between the two important climatic parameters (precipitation and temperature) and herbivore availability across different biomes (McNaughton *et al.* 1989). The second variant (Subsection 4.4.3) will use these same variables to first identify the prevailing climate type, then determine the biome type and only then use empirical data for each biome to calculate available secondary biomass (Binford 2001; Kelly 1983). Since this last method first models the habitat for an area this method is referred to as the *habitat model* in this work.

Both approaches start with primary productivity and use data from linear regression on empirical data to infer relationships between climate parameters and biomass. In every time step during a simulation several reconstructions are done in each grid cell, where reconstruction of the secondary biomass depends on the chosen variant. The schema in Figure 17 indicates the pathways for the two variants with steps one and four the same in both approaches.

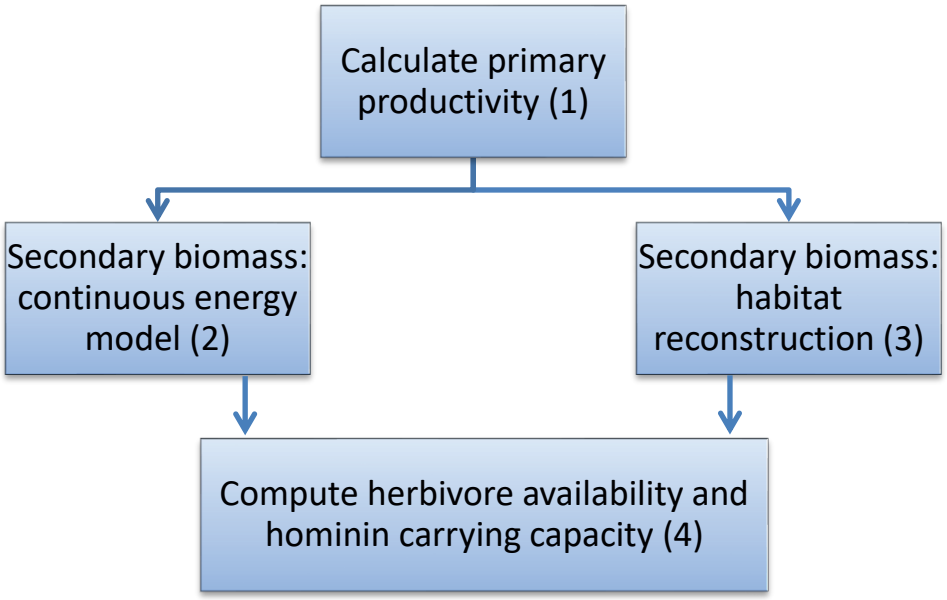



**Figure 17: Schema illustrating both pathways to compute carrying capacity.**

Each of these steps are summarized here and detailed in the following sections.

1. Primary Productivity (PP) is based on mean yearly temperature and precipitation using formulas designed by Lieth (1973). This gives the PP mass in kg per grid cell per year.
2. In the *continuous energy* variant the secondary biomass that is required in the final step 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 the HominінSpace simulation system.
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, a composite variable using temperature of the coldest and hottest months of the year. This value specifies which kind of climate (polar, boreal, cool temperate, etc.) is realized. Then the type of vegetation is considered per kind of environment. For several types of environment more than one type of vegetation can be present, depending on the amount of precipitation. For instance, cool temperate areas can be covered by deciduous forest or open woodland. 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 amount of secondary

biomass that can be sustained by the computed PB can now be calculated and is based on empirical key values from Kelly (1983), the reconstructed vegetation type and the calculated PB. Using empirical data takes into account the changing flow of energy between trophic levels.

4. The herbivores (SB) consuming the primary production are not prefabricated meat slices, they are not all ungulates and certainly do not all have a useful size fit to be hunted or consumed by hominins. To convert the SB values into available energy several modifiers are applied, effectively reducing the amount of secondary biomass present in the landscape into packages of meat available to hominins.

### 4.4.1 Calculating primary productivity

The amount of energy remaining in vegetation from the process of photosynthesis after respiration is defined as the (net) primary production. This is the plant material (the primary producers) that is potentially available to secondary consumers, the herbivores. Both temperature and precipitation govern primary productivity. The relation between temperature (T) and primary productivity (PP), derived from empirical data, is given in Equation 1, with T in degrees Celsius[16] (Lieth 1973, 325):

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

**Equation 1: Relation between Temperature and Primary Productivity.**

The constant *e* equals 2.71828182845904, the base of the natural logarithm. The formula is derived from meteorological data taken from the Climate Diagram World Atlas and assumes that the net productivity does not exceed three kg per square meter per year and that the relation between temperature and productivity is sigmoid shaped (a logistic growth equation). An example of the calculations for different mean annual temperatures is shown in Table 4, where a grid cell is defined as a square area of 10 x 10 kilometres.

**Table 4: Example values for primary productivity for given temperatures.**

| *Temperature* | *Denominator in Formula (1)* | *Primary Productivity in g/m²/year* | *Primary Productivity in kg/gridcell/year* |
|---|---|---|---|
| | | | |
| 5 | 3,054433211 | 982,1789 | 98.217.894,88 |
| 10 | 2,133148453 | 1406,372 | 140.637.187,99 |
| 15 | 1,625002268 | 1846,151 | 184.615.126,92 |
| 20 | 1,344727855 | 2230,935 | 223.093.467,53 |
| 25 | 1,19013898 | 2520,714 | 252.071.400,92 |

[16] Throughout this thesis all temperatures are given in degrees Celsius.

Lieth (1973) found that precipitation in a similar fashion steers productivity. The formula for this process integrates an empirically found relationship for more arid climates, a saturation curve for yield factors and an exclusion threshold for lower temperatures (Equation 2 with P in mm per year). Again it is assumed that productivity is less than three kg per square meter per year and the graphic representation of the formula has a saturation curve form. Figure 18 presents both datasets and their derived formula in one graph.

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

**Equation 2: The effect of Precipitation on Primary Productivity.**

Both formulas assume an exclusion line for low precipitation and low temperature levels respectively. For any combination of temperature value and precipitation level the minimum factor controls production (according to von Liebig, referenced in Lieth (1973, 326)). Thus, for each grid cell in each turn both values are computed and the smaller of the two is used. It is further assumed that these formulas are valid for all values occurring in the simulations. Alternative formulas have been derived from empirical data (Wisiol 1984, 472) but they often include the parameter evapotranspiration which is unavailable for the simulation period, or they are based on data for a single climate type (only for desert, etc.). For problems with modelling NPP in open vegetation areas versus forested areas see Del Grosso et al. (2008).

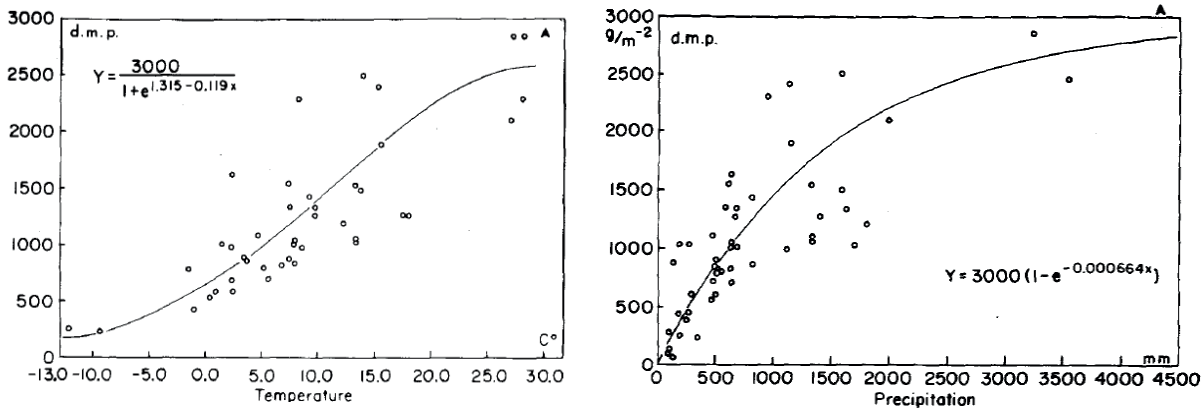



**Figure 18: Datasets and derived equations for productivity based on temperature (left) and precipitation values. Figures reproduced from Lieth (1973), Figures 5 and 6.**

### 4.4.2 Secondary biomass: the continuous energy model

This first variant to calculate herbivore availability is based on empirical data collected by McNaughton *et al.* (1989). In landscape ecology is has been recognised for some time that environments are not very homogeneous, and form patchworks of concentrations of different plant (and animal) species (Hansson *et al.* 1995, preface). The resulting complexity in the landscape is almost impossible to reconstruct and deviates from an even distribution of stable resources (Grøn 2018) as created by the habitat reconstruction method of the next subsection. Within the continuous energy approach in every time step during the simulation the following calculations are done in each grid cell:

1. Primary Productivity (PP) is based on mean yearly temperature and precipitation using formulas designed by Lieth (1973). This gives the PP mass in kg per grid cell per year;
2. The amount of herbivore availability or secondary biomass that can be sustained by the computed PP can now be calculated and is based on empirical key values from McNaughton *et al.* (1989).

The relationship between primary productivity and secondary biomass as collected by McNaughton *et al.* (1989) from different environments is depicted in Figure 19. The dataset consists of 51 data points from desert (labelled 1 in Figure 19), tundra (2), temperate grassland (3), temperate vegetation succession of old fields (4), unmanaged tropical grassland (5), temperate forest (6), tropical forest (7), salt marsh (8) and agricultural tropical grassland (9). Few points are from anthropogenic landscapes (4, 9). NAP is the net above-ground primary productivity, here simplified to be equal to PP.

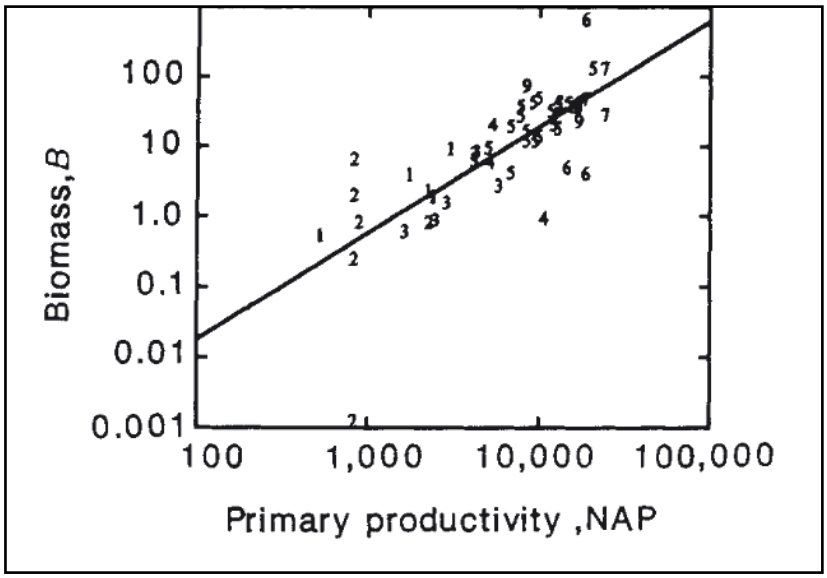



**Figure 19: Relationship between herbivore biomass (B) and net primary productivity (NAP) for different types of ecosystem (from McNaughton *et al.* (1989, Figure 1)).**

Regression analysis produced the following formula (Equation 3) to calculate the Secondary Biomass (SB), with log indicating common logarithms and NAP equated to PP, as implemented in HominínSpace:

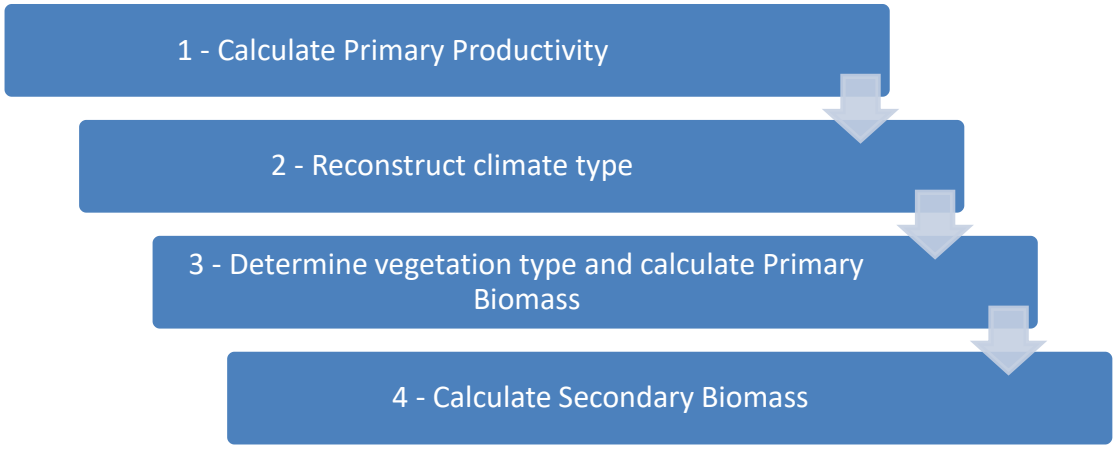
$$\log SB = 1.52 * (\log NAP) - 4.79 \ \ in\ kJm^{-2}$$

**Equation 3: Relation between Secondary Biomass and Net Primary Production.**

This gives a value for secondary biomass independent of biome type and is used to construct a continuous energy landscape based on temperature and precipitation.

### 4.4.3 Secondary biomass: the reconstructed habitat model

Most global models of primary productivity depend on stratifying a geographical area into a number of functional units, land cover types, biomes, or vegetation classes (Ellis and Ramankutty 2008; Feddema *et al.* 2005; Haxeltine and Prentice 1996; Thomas *et al.* 2004). In such models the major biome types (desert, tundra, grasslands, savannas and the tropical, temperate and boreal forests, each identifiable primarily by their predominant vegetation) are predicted from climate parameters like temperature, precipitation and correlation with the water balance (Bond *et al.* 2005, 525; Prentice *et al.* 1992). See Appendix 1 for a list of major (mega) biome types and what they encompass. Within the habitat model of the HominínSpace simulation system identification of biome has been deferred to the second stage of the production calculation, with the computation of secondary productivity. In each time step during a simulation the following reconstruction steps are executed in each grid cell:




1. Primary Productivity (PP) is based on mean yearly temperature and precipitation using formulas designed by Lieth (1973). This gives the PP mass in kg per grid cell per year;
2. The type of climate is reconstructed using temperature ranges based on data collected by Binford (2001). Calculated is Effective Temperature (ET), a composite variable

using temperature of the coldest and hottest months of the year. This value specifies which kind of environment (polar, boreal, cool temperate, etc.) is realized;
3. Then the type of vegetation is considered per kind of environment. For several types of environment more than one type of vegetation can be present, depending on the amount of precipitation. For instance, cool temperate areas can be covered by deciduous forest or open woodland. 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;
4. The amount of Secondary Biomass (SB) that can be sustained by the computed PB can now be calculated and is based on empirical key values from Kelly (1983), the reconstructed vegetation type and the calculated PB.

Temperature and precipitation values used in the preceding steps are calculated per grid cell using linear interpolation between the modern day and LGM temperature and precipitation distributions (see Section 4.3). Within a single reconstructed habitat there are differences between cells due to the fact that precipitation and temperature are used to calculate productivity, and these parameters vary throughout the total simulation area.
To calculate the secondary biomass available for hunter-gatherers, first the Primary Biomass must be estimated. This is dependent on the type of biome in which the primary production is realized. The major biome types (desert, tundra, grasslands, savannas and the tropical forests, temperate forests and boreal forests) can be predicted from temperature and precipitation and correlate with the water balance (Bond *et al.* 2005, 525). First, the general type of climate for which the biome has to be selected is determined based on the given temperature. For this there are a few rules of thumb using a variable called Effective Temperature (ET). This calculated value was proposed by Bailey (1960) to refer to the amount of solar energy that is available at a certain location. Its usage in selecting a type of climate is based on three empirically found constants:

- 18° C - the minimal mean temperature of the coldest month of the year that will sustain tropical plant communities;
- 10° C - the minimal mean temperature expected at the beginning of the growing season along the boundary between polar and boreal environments;
- 8° C - the minimal mean temperature at the beginning and the end of the growing season;

From these constants effective temperature values can be calculated (Equation 4) relating directly to empirically found biological boundaries separating different zones of biological activity. For instance, an ET of 18° C or higher indicates that there is no killing frost[17] all

[17] Many plants can be damaged by freezing temperatures, depending on type of plant and temperatures reached (https://en.wikipedia.org/wiki/Frost, accessed December 2018).

year round, and an ET of less than 10° C suggests fewer than 30 days in the year without a killing frost. Effective Temperature is calculated using the mean temperature for the warmest month (MWM) and the mean temperature for the coldest month (MCM) (Binford 2001, 59):

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

**Equation 4: Calculation of Effective Temperature using the mean extreme temperatures.**

With the ET the type of climate for each grid cell can be determined. Boundary values for different types of climate are given in Table 5.

**Table 5: Different types of climate and their defining boundaries, derived from Binford (2001), Table 4.04, page 70.**

| # | 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 | ---- |

Both the Cool temperate as well as the Warm temperate climates can produce two different biome types (types of vegetation) (Kelly 1983). For both climates this depends on different amounts of rainfall. For Warm temperate environments it is possible to get Temperate Deciduous Forest (relatively humid) or Woodland/Scrubland (more arid), and for Cool temperate climates this would be Temperate Evergreen forest (humid) versus Temperate Grassland (arid). The more arid biomes are characterized by continuous grass layers with trees in a climatic regime of distinct wet and dry seasons. Besides rainfall the main additional forces maintaining such biome types are herbivory, fire and frost (Du Toit and Cumming 1999). In the grasslands, most of the biomass is reproductive tissue (blades of grass) and as such, production is relatively high and biomass low (Binford 2001, 82).

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 (page 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.

Not all energy stored in PP is equally available to other trophic levels. The ratio of Primary Biomass (PB) (the producers) and Secondary Biomass (SB) (the consumers) indicates both how much herbivores are available as prey as well as the availability of PB itself to herbivores. A high level indicates more potential prey as well as relatively more available plant material in the form of foliage, fruits and seeds (as opposed to wood). 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 6. 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, specifically the herbivores that will provide the meat that supplies the energy for the hominins in the underlying model.

**Table 6: Calculating PB and SB from PP. Values derived from Table 3, page 284 (Kelly 1983, 284, Table 3). Type climate from Table 5 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 6 are combined. These are derived from Kelly (1983, Table 3, 284) where a linear relationship is found between PP and PB. This translates into Equation 5 for estimating SB (PP is per cell and in gram, hence the division by 1000 to get to kilograms):

$$SB = \frac{PP * b}{a} / 1000 \quad \text{in kg/cell/year}$$

**Equation 5: Computing Secondary Biomass from Primary Productivity (a, b from Table 6).**

That SB indeed varies with climatic variables is also empirically attested (Binford 2001, 101). Examples with calculated values according to Equation 5 for different biomes are given in Table 7.

**Table 7: Calculated secondary biomass per grid cell in kilograms per year for different mean yearly temperatures (theoretical values) in different biomes[18].**

| *Temp* | *Biome1* | *Biome2* | *Biome3* | *Biome4* | *Biome5* | *Biome6* | *Biome7* | *Biome8* | *Biome9* |
|---|---|---|---|---|---|---|---|---|---|
| ***-10*** | 64.73 | 14.16 | 174.44 | 260.80 | 302.01 | 157.30 | 377.54 | 169.97 | 195.58 |
| **-5** | 110.58 | 24.18 | 297.98 | 445.49 | 515.89 | 268.70 | 644.90 | 290.33 | 334.08 |
| **0** | 181.45 | 39.68 | 488.97 | 731.03 | 846.55 | 440.92 | 1,058.26 | 476.42 | 548.22 |
| **5** | 280.68 | 61.39 | 756.36 | 1,130.80 | 1,309.49 | 682.04 | 1,636.96 | 736.96 | 848.01 |
| **10** | 401.90 | 87.90 | 1,083.02 | 1,619.18 | 1,875.05 | 976.61 | 2,343.95 | 1,055.24 | 1,214.25 |
| **15** | 527.57 | 115.38 | 1,421.69 | 2,125.50 | 2,461.38 | 1,282.00 | 3,076.92 | 1,385.22 | 1,593.96 |
| **20** | 637.53 | 139.43 | 1,718.00 | 2,568.51 | 2,974.39 | 1,549.20 | 3,718.22 | 1,673.93 | 1,926.18 |
| **25** | 720.34 | 157.54 | 1,941.15 | 2,902.14 | 3,360.74 | 1,750.43 | 4,201.19 | 1,891.36 | 2,176.37 |

Since the formulas and categories are based on extant biomes and actual measured data, they are applicable only in environments that exist today. Types of landscape that no longer exist (like the Mammoth Steppe) are not given by Kelly since actual values for secondary biomass or MWM/MCM cannot be measured and productivity must therefore be modelled. It has been shown that especially the cool temperate grassland was very productive through a combination of intense sunlight and fertile loess providing mosses, lichens, grasses and shrubs feeding Mammoth Steppe herbivores (Guthrie 1990). This biome is best implemented by applying the figures from the African savannah.

Terrestrial biomes are distinguished primarily by their predominant vegetation, and are mainly determined by temperature and rainfall (Forseth 2012). The model in HominínSpace assumes a direct relation between temperature, precipitation and vegetation type, and acknowledges but ignores other factors that (could) play an important part in the resulting biome, most importantly soil type and groundwater levels (Fan *et al.* 2013). The limiting effects of temperature and precipitation levels however surpass the influence of other parameters and their availability significantly correlates with primary productivity, today (Li *et al.* 2013) and in the past (Janis *et al.* 2000).

### 4.4.4 The number of edible ungulates: carrying capacity

Carrying capacity in the underlying research is defined as the amount of kilocalories in the form of a number of edible ungulates within a certain area that are available for

[18] See also Supplementary Materials for the spreadsheet “Climate Calculations.xlsx” to compute these values.

consumption by the modelled hominins at a given time. The carrying capacity is a variable that is used in the hominin demographic calculations (see Chapter 1). Research has shown that in some of the most relevant biome types about half of the plant production is eaten by large (> 5 kg) mammalian herbivores (Du Toit and Cumming 1999). The remaining yearly production is eaten by small mammals, insects, decomposed by fungi, etc. The model in HomininSpace therefore assumes that 50% of the herbivore mass in any area consists of individuals from species that can be hunted and consumed.

It is further estimated that from every kilogram of secondary ungulate biomass only about 60% is consumable by hominins. The remaining 40% are bones, skin and assorted inedible items including hooves and horns. These can be useful raw materials but are not usable for caloric consumption (White 1953; White 2006, 11). It has been shown that Neanderthals can switch between different types of prey species, depending on the availability in the landscape (Hodgkins *et al.* 2016). The available secondary biomass for consumption is given by the equation below.

$$SB_{avail} = \frac{1}{2} * \frac{6}{10} * SB \quad \text{in kg}$$

**Equation 6: Computing the available Secondary Biomass for hominin consumption.**

The linear relationship in this study between primary production and edible large ungulate biomass is but a model of reality in which the interaction between the different components is an important element. See for other examples the work from Augustine and McNaughton (2006) or Hobbs (1996). The underlying model has to necessarily simplify and generalize matters and states that from a hypothetical secondary biomass of 100 kg, the following remains available for consumption:

- 50% is large ungulates = 50kg;
- 60% of 50kg = 30kg edible material.

When prey animals are harvested from the environment, prey populations can recover. Generally there is an optimal harvest size which still allows full recovery within the hunting cycle. Since HomininSpace calculates the available prey species based on primary productivity, harvesting actually reduces primary biomass and thus recovery of prey populations and secondary biomass must be translated into recovered primary biomass. Ecosystem responses to disturbances are complex due to species – species interactions and their relationships with the environment (Ives and Carpenter 2007). Full recovery of certain forest type environments to re-establish after disturbances for instance can take

very long (Cole *et al.* 2014). Therefore recovery of the environment from harvesting by hominins is modelled after larger ungulate recovery mechanisms, and very conservatively and rather arbitrarily implemented as a 20% yearly rate of recovery and prey populations increase (Steinmetz *et al.* 2010, 44), summarized in the equation below.

$$PP_{t+1} = PP_{P,T} - PP_{consumed} + \frac{1}{5} * PP_t \text{ , in kcal}$$

**Equation 7: Calculating the result for Primary Productivity at time step t+1.**

# 5. THE MODELLED HOMININS

## 5.1 Introduction

Developing a model is a trade-off between different interests. Predictive accuracy, computational needs and realism compete in the process of finding an acceptable balance when building models that are valid descriptions of the actual processes (Levins 1966). The main factor steering dispersal is the resulting population sum of births, deaths, immigration and emigration through time (assuming no transport technology is available). The parameters concerning hominin behaviour in HomininSpace are qualified in this chapter and include demographic parameters, energy needs, group dynamics and mobility factors. Together they implement a generic hunter gatherer model with parameters selected from ethnographical literature. Finding specific values for these parameters is the main aim of this study. The use of data on modern human hunter-gather societies as a model for Neanderthal parameterization in this study is based on the following:

- Similar **anatomies**. Both hominin species are bipedal and anatomically similar, with similar body proportions, size, weight and stature (Trinkaus 1986);
- Matching **behaviour** (Villa and Roebroeks 2014) with both species implementing hunting and gathering as a main subsistence strategy which is core to the underlying model of this thesis;
- **Compatible DNA**. Remaining Neanderthal DNA in modern human DNA suggests that offspring of mixed parents is viable and that offspring of mixed parenthood is able to reproduce (Green *et al.* 2010). This suggests that *Homo sapiens* could be a good model for and even be the same species as *Homo neanderthalensis* when this criterion forms a major part of the definition of species (Klein 2009b; Mayr 1950);
- Comparable **life histories** (Smith *et al.* 2007; Weaver and Hublin 2009; Smith *et al.* 2010) and mortality patterns (Trinkaus 2011);
- There appears to be no significant **ecological differences** between Neanderthals and early modern humans (Chase 1989; Marean and Kim 1998);
- Neanderthals shared symbolic thinking with early modern humans and are **cognitively indistinguishable**, based on retrieved material culture as discussed by Hoffmann *et al.* (2018);

In short, both species are generally considered ‘sufficiently’ similar (O’Connell 2006). It is assumed that “after having assessed similarities and differences further similarities can be inferred” (Bernbeck 1997, 86, Kowarik *et al.* 2012). Of course there are many differences between the two species. Where many Neanderthals probably got most of their energy from large herbivores, many modern humans likely extended their diet to include more aquatic resources (Richards and Trinkaus 2009), Neanderthals probably used more energy

in their daily activities (Verpoorte 2006; Churchill 2009 but see Heyes and MacDonald (2015)), etc. These differences are acknowledged, but assumed to be differences in degree and not in kind. Nevertheless, material evidence should be used to test or support the validity of a model based on ethnographic data (Isaac 1968). Within HomininSpace, archaeological data is used to verify the proposed models by comparing simulation results to material culture remains.

The archaeological record for the Neanderthals never revealed evidence for the use of the wheel, for boats, for the use of draft animals, for sleds, etc. (Wobst 1974, 152). These hominins had to walk and carry everything they owned with them. Planning for the future would not include stocking much food surplus, and distance covered and speed of movement would be limited (Lee and DeVore 1968, 12). Also, modern hunting and gathering populations are invariably influenced by the modern cultural environments around them with imported diseases, negative migration balances, access to domesticated flora and fauna, contacts with sedentary populations, etc. (Wobst 1974). For this and other reasons cultural uniformitarianism from the present to the past is doubtful in itself (Martelle Hayter 2011). The cultural criteria used by any anthropologist further colour any interpretation of the ethnographic record (Marks 2012). And even those hunter and gatherer populations that still exist today show great diversity in many aspects of life (Binford 2001; Kelly 1995; Lee 1968). As such, even when sharing a similar environment two cultures can be a world apart.

This chapter identifies the model parameters that describe the variability in the hominins in HomininSpace. Parameters are model elements that can be changed by assigning different values. A unique model is created by assigning a specific value combination to the parameters. The parameters form the hominin model of HomininSpace together with the variables and logic that is embedded in the source code of HomininSpace and that describes how the parameters are used. The parameters are summarized with their selection motivated and with more details provided in the remaining sections of this chapter.

Demographic model parameters are discussed in the Section 5.2. Energy is a basic requirement for any living individual, and Section 5.3 will discuss the modelled energy requirements of Neanderthals. Within the HomininSpace modelling system the energy is used for two major action categories: it steers hominin movements through the landscape in which they aim to maximize possible intake, and it enables procreation. The basic unit in HomininSpace is the hominin group. These groups of hominins can be created and can

also be removed from the simulation, by leaving the simulation area, through extinction or when merging with another group. Group interaction and mobility is discussed in section 5.4. Section 5.5 summarizes all parameters and other model elements of the model.

## 5.2 Demography

Demographic processes can enable dispersal. If sufficient new individuals are born, members of a species can disperse into new territories. The demographic process in HomininSpace is implemented as a statistical procedure. This means that the system does not track individuals, but accounts for demographics with births and deaths recorded for the group as a whole. This implementation is possible due to the fact that the basic agent in HomininSpace is the group (not the individual hominin) and prevents random effects from demographic stochasticity. This process mimics population statistics, but is implemented at the group level.

In every group there are three age categories that individuals can belong to: `pre-fertile`, `fertile` and `post-fertile` (a common biological categorization, see for instance an application in the modelling of *Lynx pardinus* by Gaona *et al.* (1998)). This limited categorization ignores any differences between infancy, childhood, juvenile and adolescence stages in the pre-fertile category (Bogin 1999), since the interest of this research is neither in growth nor in foraging efficiency, and also lumps energy requirements for this phase into one (average) constant. Gurven and Kaplan (2007, 322) have identified different death profiles for these three categories, and they found that from birth through childhood, for adults and for surviving individuals after 40 years of age a different average mortality rate can be calculated supporting this separation into three segments (compare however the four categories identified by White (2013)). HomininSpace implements a different mortality rate for each of three age categories.

There is no separation between males and females in HomininSpace and for each category 50% of the hominins are assumed to be female, with birth and death rates being applied to both sexes equally. This is rationalized by the fact that, especially in the fertile category, males will most likely run many risks when hunting whereas females have a high death rate when giving birth or from complications thereafter. Other causes of death include disease and illnesses, inter-group and intra group violence, and accidents. For different types of societies, including extant hunter-gatherers, Gurven and Kaplan (2007, 348) find an adult mortality rate of 1%. Hill *et al.* (2007) observe a higher percentage of adult

mortality rates of 2% for the Hiwi hunter-gatherers of Venezuela. The Hiwi experience more violence related deaths than other groups who mainly suffer from diseases and the authors suggest that the Hiwi might reflect Palaeolithic cultures better. Only about 50%-60% of all hunter-gatherer infants reach adulthood (Gurven and Kaplan 2007; Kelly 1995). This means an overall mortality rate of 4%, see Table 8.

Although empirical data (Gurven and Kaplan 2007, 325) suggests a decreasing mortality rate for the pre-fertile category and an exponentially increasing rate for post-fertile individuals (referred to as the 'Siler' model), the HominínSpace model implements a flat parameterized death rate for each age segment since there are no aging individuals in HominínSpace, only groups. As with all parameters this natural death rate is constant throughout the simulation period[20]. One of the disadvantages of working with a mortality rate is that the segment will never become empty since there is always a remainder. For instance, with a mortality rate of 15% almost 5% is still alive after 20 years. This unrealistic situation is solved via group extinction (see subsection 5.4.1).

**Table 8: Different mortality rates (in %) and the compounded effects on a starting population of 100 with no replenishment. Bold rows are results after 15 and 20 years.**

| *Perc* / *Year* | *1%* | *2%* | *3%* | *4%* | *5%* | *6%* | *7%* | *8%* | *15%* |
|---|---|---|---|---|---|---|---|---|---|
| ***1*** | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| ***2*** | 99.00 | 98.00 | 97.00 | 96.00 | 95.00 | 94.00 | 93.00 | 92.00 | 85.00 |
| ***3*** | 98.01 | 96.04 | 94.09 | 92.16 | 90.25 | 88.36 | 86.49 | 84.64 | 72.25 |
| ***4*** | 97.03 | 94.12 | 91.27 | 88.47 | 85.74 | 83.06 | 80.44 | 77.87 | 61.41 |
| ***5*** | 96.06 | 92.24 | 88.53 | 84.93 | 81.45 | 78.07 | 74.81 | 71.64 | 52.20 |
| ***6*** | 95.10 | 90.39 | 85.87 | 81.54 | 77.38 | 73.39 | 69.57 | 65.91 | 44.37 |
| ***7*** | 94.15 | 88.58 | 83.30 | 78.28 | 73.51 | 68.99 | 64.70 | 60.64 | 37.71 |
| ***8*** | 93.21 | 86.81 | 80.80 | 75.14 | 69.83 | 64.85 | 60.17 | 55.78 | 32.06 |
| ***9*** | 92.27 | 85.08 | 78.37 | 72.14 | 66.34 | 60.96 | 55.96 | 51.32 | 27.25 |
| ***10*** | 91.35 | 83.37 | 76.02 | 69.25 | 63.02 | 57.30 | 52.04 | 47.22 | 23.16 |
| ***11*** | 90.44 | 81.71 | 73.74 | 66.48 | 59.87 | 53.86 | 48.40 | 43.44 | 19.69 |
| ***12*** | 89.53 | 80.07 | 71.53 | 63.82 | 56.88 | 50.63 | 45.01 | 39.96 | 16.73 |
| ***13*** | 88.64 | 78.47 | 69.38 | 61.27 | 54.04 | 47.59 | 41.86 | 36.77 | 14.22 |
| ***14*** | 87.75 | 76.90 | 67.30 | 58.82 | 51.33 | 44.74 | 38.93 | 33.83 | 12.09 |
| ***15*** | **86.87** | **75.36** | **65.28** | **56.47** | **48.77** | **42.05** | **36.20** | **31.12** | **10.28** |
| ***16*** | 86.01 | 73.86 | 63.33 | 54.21 | 46.33 | 39.53 | 33.67 | 28.63 | 8.74 |

[20] In times of resource insufficiencies or extreme cold rates are doubled or more, up to a maximum of 100%.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ***17*** | 85.15 | 72.38 | 61.43 | 52.04 | 44.01 | 37.16 | 31.31 | 26.34 | 7.43 |
| ***18*** | 84.29 | 70.93 | 59.58 | 49.96 | 41.81 | 34.93 | 29.12 | 24.23 | 6.31 |
| ***19*** | 83.45 | 69.51 | 57.80 | 47.96 | 39.72 | 32.83 | 27.08 | 22.29 | 5.36 |
| ***20*** | **82.62** | **68.12** | **56.06** | **46.04** | **37.74** | **30.86** | **25.19** | **20.51** | **4.56** |
| ***21*** | 81.79 | 66.76 | 54.38 | 44.20 | 35.85 | 29.01 | 23.42 | 18.87 | 3.88 |
| ***22*** | 80.97 | 65.43 | 52.75 | 42.43 | 34.06 | 27.27 | 21.78 | 17.36 | 3.29 |

Three years is the average length of time between births for hunter-gatherers (Kelly 1995, Table 6-7; Pennington 2001, 184). The HominінSpace model assumes that all fertile women bear one child every $n$ year. This is implemented with a birth-rate of $\frac{100}{n}$% of all female adults giving birth to a single new-born[21]. Ethnographic data suggests that in general within a few years of reaching adulthood and after being married a first baby is born (Kelly 1995, 245-246; White 2013, 126). Population growth is exponential (Kremer 1993), and continues until carrying capacity levels of the environment are reached. In HominінSpace, for each year that population size exceeds carrying capacity, penalties are applied: growth rates are halved and death toll is doubled (cf. White 2013, 137).

When anthropologists compare nomadic populations with sedentary groups the birth interval for the nomadic way of life is longer and birth intervals are prolonged to at least four years (Short 1987). It appears that these intervals mainly depend on energy recovery after birth, which in turn depends on food quality and food availability (Valeggia and Ellison 2009). Fine-tuned birth and death rates are important elements of a realistic simulation for hominins. It might be possible to obtain more realistic birth rates from morphological features from hominin skeletons (e.g. Gunz *et al.* 2010), or from genetic profiles (Arenas *et al.* 2013). Since these rates are important for population distribution, comparing these rates to other species, modern humans in particular, is highly recommended (as done by Zubrow (1989), Skinner (1997) or Flores (1997)).

In most hunter–gatherer populations, fertility has a parabolic character with few births occurring after reaching an age of 40 (Pennington 2001, 175). During the fecundity period the average hunter-gatherer woman produces five to six children, but a mean of eight is not exceptional (Hewlett 1991, Table 2; Kelly 1995, 244; Pennington 2001, Table 7.2). Note that in HominінSpace it is assumed that hominin females give birth to one new born per delivery. A variety of factors may influence fertility (the number of children born),

[21] Twins are a practical issue with HGs. Australian aboriginals immediately killed one of any new-born twins, since it is impossible to carry two infants while foraging. HominінSpace assumes only one new-born per pregnancy.

reproductive span (the period in which reproduction occurs) and birth-intervals, including ovarian function, contraception and abortion, infanticide, coital frequency, nutritional stress, and lactation (e.g., Hill *et al.* 2007; Kelly 1995; Pennington 2001; White 2013, 125). Most of these factors are culturally defined and thus impossible to quantify for the past while others did not exist for prehistoric hominins. Only nutritional stress is quantifiable and responded to as discussed above.

Population growth in the HomininSpace model depends on mortality and birth rates that are included as parameters in the model. They are applied at the end of every year. Distribution within the population segments is assumed to be uniform. Thus in each time step for each age segment the number of people in that age segment divided by the length of the segment move up to the next segment (with any deaths subtracted). Newborns, to be added to the first segment, are calculated using women of the `fertile` age segment. The mathematical model is given below (Equation 8). Base mortality rates and birth percentages are constant throughout the simulation period, assuming that health and fertility did not substantially improve or decline in the Late Pleistocene, although local resource shortages impact both. In the calculations that follow *Pre(t)* is the number of pre-fertile hominins at time *t*. *Fertile(t)* is the number of fertile hominins and *Post(t)* the number of post-fertile individuals. For each segment there is a specific mortality rate *dPre, dFertile* and *dPost.* The length of the pre-fertile segment is 15 and the length of the fertile segment in this example is 25. New population numbers for time *t+1* are calculated as follows, with *b* the birthrate in percentages[22]:

```
Pre(t+1)     = Pre(t)-dPre*Pre(t)-Pre(t)/15+(b/100)*(Fertile(t)/2);
Fertile(t+1) = Fertile(t)-dFertile*Fertile(t)-Fertile(t)/25+Pre(t)/15;
Post(t+1)    = Post(t)-dPost*Post(t)+Fertile(t)/25;
```

**Equation 8: Calculating the number of hominins per segment for the next time step.**

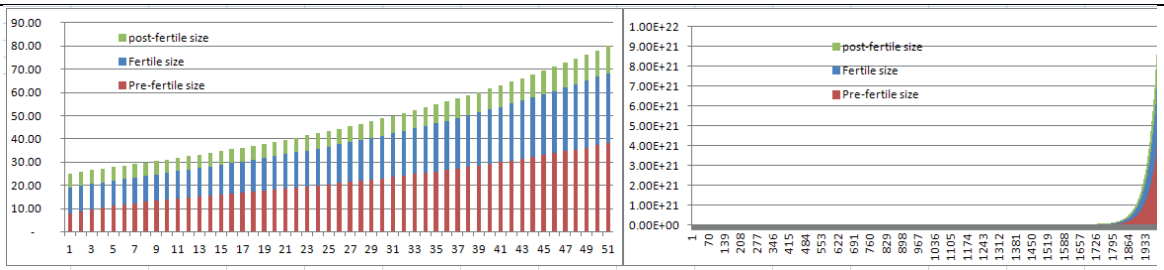
An example of results from these calculations is shown in Figure 20. These are taken from the spreadsheet '*HomininSpace Population Calculator.xlsx*' which implements the demographic model mathematically (see Supplementary Materials). Figure 21 contains seven graphs presenting population growth (or reduction) for different illustrative values for the demographic parameters. This figure uses the same initial sizes as in the spreadsheet in Figure 20 and serves to illustrate the effect of subtle changes to these parameters. Values for the selected parameters define the population through time and are

[22] Based on the calculations shown in http://www.di2.nu/pop_sim.html, accessed June 2012.

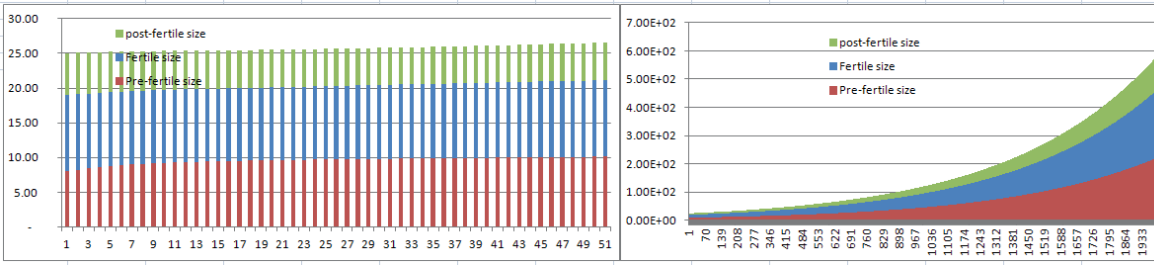
therefore included as parameters in the HominінSpace model (birthrate, segment sizes and mortality rates per segment).

| Year | Newborns | Pre-fertile size | Pre-fertile deaths | Becoming fertile | Fertile size | Fertile deaths | Becoming post-fertile | post-fertile size | Post-fertile deaths | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.82 | 8.00 | 0.32 | 0.53 | 11.00 | 0.22 | 0.44 | 6.00 | 0.48 | 25.00 |
| 1 | 1.79 | 8.96 | 0.36 | 0.60 | 10.87 | 0.22 | 0.43 | 5.96 | 0.48 | 25.80 |
| 2 | 1.79 | 9.80 | 0.39 | 0.65 | 10.82 | 0.22 | 0.43 | 5.92 | 0.47 | 26.54 |
| 3 | 1.79 | 10.54 | 0.42 | 0.70 | 10.82 | 0.22 | 0.43 | 5.88 | 0.47 | 27.24 |
| 4 | 1.79 | 11.20 | 0.45 | 0.75 | 10.88 | 0.22 | 0.44 | 5.84 | 0.47 | 27.92 |
| 5 | 1.81 | 11.80 | 0.47 | 0.79 | 10.97 | 0.22 | 0.44 | 5.81 | 0.46 | 28.58 |
| 6 | 1.83 | 12.35 | 0.49 | 0.82 | 11.10 | 0.22 | 0.44 | 5.78 | 0.46 | 29.23 |
| 7 | 1.86 | 12.87 | 0.51 | 0.86 | 11.26 | 0.23 | 0.45 | 5.76 | 0.46 | 29.89 |
| 8 | 1.89 | 13.35 | 0.53 | 0.89 | 11.44 | 0.23 | 0.46 | 5.75 | 0.46 | 30.54 |
| 9 | 1.92 | 13.81 | 0.55 | 0.92 | 11.64 | 0.23 | 0.47 | 5.75 | 0.46 | 31.21 |
| 10 | 1.96 | 14.26 | 0.57 | 0.95 | 11.86 | 0.24 | 0.47 | 5.76 | 0.46 | 31.88 |
| 11 | 2.00 | 14.70 | 0.59 | 0.98 | 12.10 | 0.24 | 0.48 | 5.77 | 0.46 | 32.57 |
| 12 | 2.04 | 15.13 | 0.61 | 1.01 | 12.36 | 0.25 | 0.49 | 5.79 | 0.46 | 33.28 |
| 13 | 2.08 | 15.55 | 0.62 | 1.04 | 12.62 | 0.25 | 0.50 | 5.82 | 0.47 | 34.00 |
| 14 | 2.13 | 15.98 | 0.64 | 1.07 | 12.90 | 0.26 | 0.52 | 5.86 | 0.47 | 34.74 |
| 15 | 2.18 | 16.40 | 0.66 | 1.09 | 13.19 | 0.26 | 0.53 | 5.91 | 0.47 | 35.51 |
| 16 | 2.23 | 16.83 | 0.67 | 1.12 | 13.50 | 0.27 | 0.54 | 5.96 | 0.48 | 36.29 |
| 17 | 2.28 | 17.26 | 0.69 | 1.15 | 13.81 | 0.28 | 0.55 | 6.03 | 0.48 | 37.10 |
| 18 | 2.33 | 17.70 | 0.71 | 1.18 | 14.13 | 0.28 | 0.57 | 6.10 | 0.49 | 37.93 |
| 19 | 2.39 | 18.14 | 0.73 | 1.21 | 14.46 | 0.29 | 0.58 | 6.17 | 0.49 | 38.78 |
| 20 | 2.44 | 18.59 | 0.74 | 1.24 | 14.80 | 0.30 | 0.59 | 6.26 | 0.50 | 39.66 |

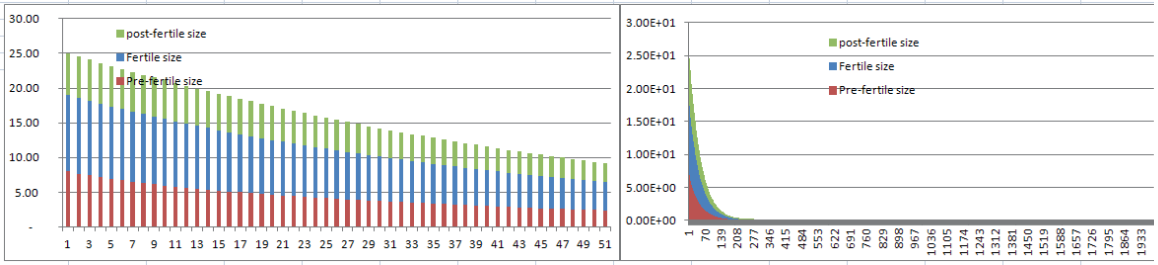
**Figure 20: Population calculations for a group with initial size 25: Pre=8, Fertile=11 and Post = 6. Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.33.**



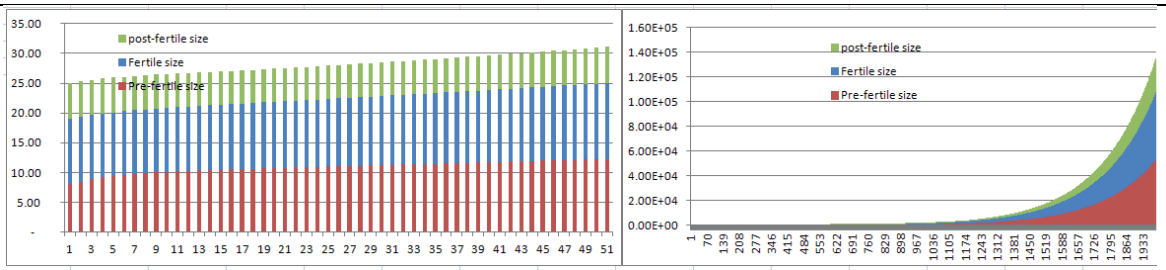

**Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.33. Default configuration, with progresively increasing numbers. Effective population explosion after 1800 years.**



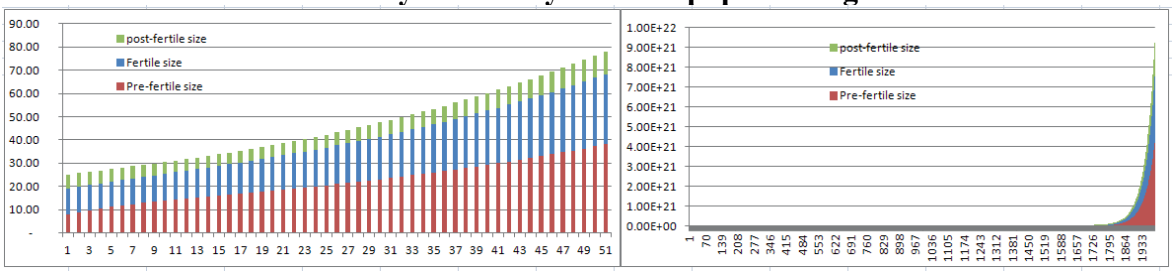

**Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate 0.2. Illustrates the effect of a lower birthrate with same mortality rates (compare vertical axis with the first graph).**



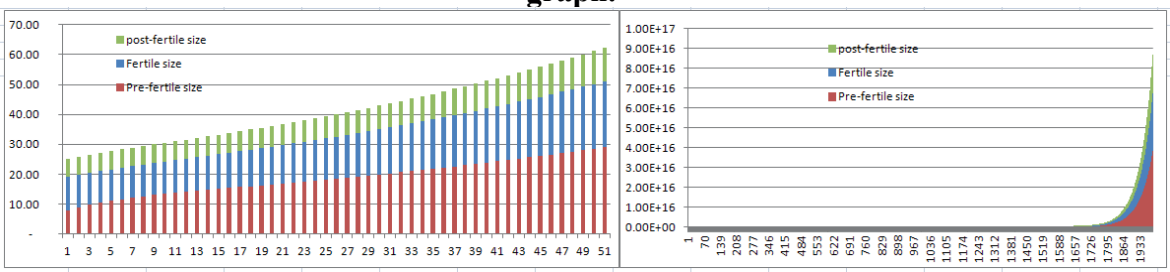

**Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.1. Extinction due to very low birthrate.**



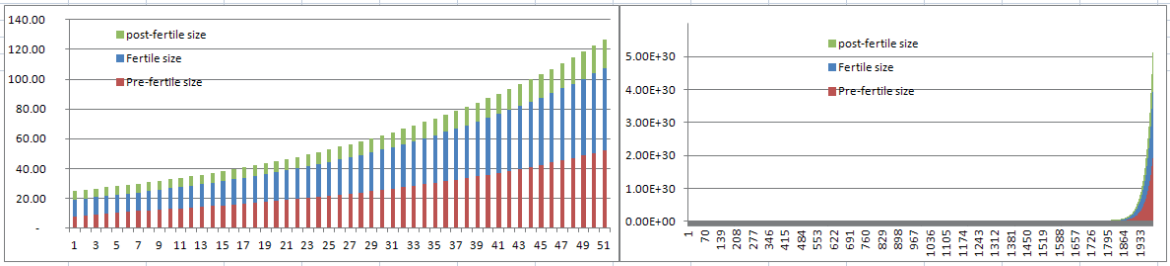

**Mortality rates are 0.1, 0.02 and 0.08 respectively, the birthrate = 0.33. Very high infant mortality rate. Very reduced population growth.**




**Mortality rates are 0.04, 0.02 and 0.1 respectively, the birthrate = 0.33. Very high post-fertile mortality rate. Effect of this change is negligible, as opposed to the first graph.**




**Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.33. The length of the fertile period is reduced to 20 (from 25). This effectively delays the population buildup, but does not change the shape of the curve.**




**Mortality rates are 0.04, 0.02 and 0.08 respectively, the birthrate = 0.33. The length of the pre-fertile period is reduced to 10 implementing early maturation. This enormously boosts the population buildup, already clearly visible after 50 years.**

**Figure 21: Population composition after 50 years (left) and after 2000 years (right) with unlimited carrying capacity. Graphs visualize typical exponential growth and are created with "HomininSpace Population Calculator.xlsx", see Supplementary Materials.**

## 5.3 What does it take to be a Neanderthal - energy requirements

For practical reasons the model in HomininSpace parameterizes overall mean subsistence values differing per age segment, constant through time and space and averaged over sex. In the model a carrying capacity of X indicates how many kilograms edible meat are

available for consumption. It is assumed that hominins are capable of extracting that amount from the environment. These kilograms meat are converted into kcal and that gives a maximum to the amount of available energy. This does not mean the population will always be at maximum level (X*kcal/hominins). Instead mortality and birth rates for hominins are positively or negatively adjusted based on whether the population level and their energy usages are above or below the local carrying capacity (White 2013, 137). For every year with resource shortages a penalty is applied to birthrate (that is halved) and mortality rates (which are doubled). These rates return to normal when sufficient resources are found. When there is less than 10% of the required resources available, a group is immediately removed.

Energy requirements for modern humans are suggested to average 2000 kcal, but this number varies according to activity, sex, age, size, etc. (Kelly 1995, 101). Energy, obtained from digested food, is spent on movement, resting, sleeping, thermoregulation including diet-induced thermogenesis, daily activities, reproduction, digesting food, and maintenance and growth of the body (Snodgrass and Leonard 2009). Isotopic analysis of Neanderthal bones suggests that most of their protein intake came from animal protein and fat (Richards *et al.* 2000; Snodgrass and Leonard 2009) which does not exclude however consumption of additional types of food that are not reflected in the isotope data.

Neanderthal energy intake is estimated to be higher than that of modern day modern humans based on a larger body mass and more muscles, high levels of physical activity and intake of high protein diets, mostly meat (Aiello and Wheeler 2003; Churchill 2005; Klein 2009b; Snodgrass and Leonard 2009; Steegmann *et al.* 2002). Comparative research amongst modern day northern populations with similar lifestyles suggests an elevated Basal Metabolic Rate (BMR) as a response to colder climates. Higher energy requirements (intake and increased travel costs due to physiological properties) can influence mobility and will influence the effective foraging radius, the frequency that campsites are moved and investments in furniture and technology (MacDonald *et al.* 2009; Verpoorte 2006).

Energy supply is limited and must be extracted from the local environment (Bird *et al.* 2009). Some researchers suggest that an average Neanderthal would expend between 3000 and 5500 kcal/day, depending on physique, activities, age and sex (Churchill 2005; Churchill 2006; Sorensen and Leonard 2001). The mobility cost for reproductive women would be in the high end (Wall-Scheffler and Myers 2013). Sorensen and Leonard (2001) estimated energy requirements of 4000-7000 kcal for male Neanderthals, and 3000-5000

kcal for women. Snodgrass and Leonard (2009) suggested a similar amount for women but 4000-6000 for men. Steegmann *et al.* (2002) estimated 3360-4480 kcal were needed for male Neanderthals. And Froehle and Churchill (2009) distinguished cold and temperate climates, with energy use in temperate zones 3227-3527 for males, and 2297-2547 kcal for females. In colder periods energy use would amount to 4469-4877 kcal per day for men, and 3180-3190 for females. MacDonald *et al.* (2009) found 5020 kcal per day for male and 4590 kcal per day for female Neanderthals, based on BMR estimations from skin surface area. From body mass calculations they found values of 4450 kcal/day and 4230 kcal per day. They used the maximum relative amount of energy expenditure above basal needs for hunter-gatherers (MacDonald *et al.* 2009, 215). However, the assumed higher body mass of Neanderthals from which these estimates are reconstructed is contested (Heyes and MacDonald 2015). The exact energy needs are thus disputed but that Neanderthals need energy and extract it from the environment is a given fact.

Energy is derived from the sun and becomes available in a landscape as flora and fauna. Isotopic analysis of Neanderthal fossil material suggests that most of their food intake came from animal protein and fat (Richards *et al.* 2000; Snodgrass and Leonard 2009). Research has shown that Neanderthal diet in different periods and environmental conditions focused on medium to large terrestrial game (Hublin 2009; Salazar-García *et al.* 2013). Fish (Bocherens *et al.* 2013), small terrestrial and aquatic game (Archer and Braun 2013; Richards and Trinkaus 2009), birds (Fiore *et al.* 2016 but see Rufà *et al.* 2015) and vegetables (Henry *et al.* 2011) were all consumed by Neanderthals but meat was likely the main component of their diet (Hardy *et al.* 2015), maybe comparably limited in diet breath to some earlier hominins in Wallacea (O'Connor *et al.* 2017). Meat contains high quality protein and the essential nine amino acids that humans cannot make themselves and is one of the most successful ways of coping with energy needs in cooler environments (Steegmann *et al.* 2002, 571). Meat is denser than plant materials and contains some essential minerals, including iron, glucose and some vitamins (Kelly 1995, 105).

Meat can provide around 3000 kcal/kg (Diem and Lentner 1962) and is included as a parameter in the HominínSpace model. The calculations further assume a 60% available meat return on live animal weight (White 2006, 11), but see Lyman (1979) or Stewart and Stahl (1977) with lower percentages). Today butchery results deliver similar meat percentages. See for instance http://lovelivegrow.com/2011/01/how-much-meat-from-a-pig/, accessed 2 August 2013. Here a return percentage of 48% is mentioned, but larger animals tend to have larger percentages of usable meat and since blood and marrow in the

past probably had greater value than today the 60% quoted is likely not far off the useable percentage . In particular, in order to prevent protein poisoning from eating too much lean meat, maintain vitamin C levels, and to get essential carbohydrates and fats, a hunter could eat raw meat, plus internal organs like the liver and brain or extract marrow for the additional materials (see Speth (2010), also for other risks and limitations of high-protein intake). In certain cultures hunting large and dangerous animals serves social or symbolic purposes (Speth (2013) and references therein) but in HomininSpace all hunting is assumed to be subsistence related. This assumes that time and effort are optimally spent in hunting and gathering (see for example Dusseldorp (2009)).

### 5.3.1 Temperature tolerance

The Late Pleistocene is characterized by a generally lower annual mean temperature distribution compared to today. Exposure to extreme cold could have increased the energy requirements of Neanderthals compared to ethnographic data even further. The specific physique of the Neanderthals including shorter limbs and high muscularity offered only limited advantages for coping with cold temperatures (Aiello and Wheeler 2003; Snodgrass and Leonard 2009; Sørensen 2009). It is suggested that an elevated BMR might be related to colder climates (Churchill 2009).

Some of the energy requirements for preventing heat loss from the bodies of Middle and Late Pleistocene hominins must have been generated or sustained using external means which could include domestic fire and/or clothing. It is estimated that about 5477 Watt per month would be needed to warm and provide cooking fuel for a mixed group of about 25 Neanderthals (Sørensen 2009). This is about 1.5 kilogram of wood per day per individual, providing 4.8W/kg or 1136 kg wood per month for the group. Although wood supply can be limited in certain environments like grasslands and cold rugged areas (Pryor *et al.* 2016), the model in HomininSpace assumes that alternative fuels in the form of animal dung or bones were available and that heating through fire or furs was no limiting factor to the modelled hominins. Internal resistance to cold and cultural means to mitigate climate effects are combined into one model parameter, the *temperature tolerance*. This threshold is the minimum temperature where hominins still survive. If at any time the coldest reconstructed temperature falls below this threshold, the hominins will die. Benito *et al.* (2016) show that the coldest temperature is a good predictor of habitat suitability for Neanderthals.

## 5.4 Are Hominins social animals? - hominin groups

Hunting animals in groups requires cooperation and coordination of efforts (Marín *et al.* 2017). Living and hunting in social groups facilitates hunts with game driving, the final kill and the distribution of the meat afterwards. Most authors therefore assume that Neanderthals moved and hunted in groups (Radloff and Du Toit 2004; Kaplan *et al.* 2000 and the papers in the special edition of Quaternary International, Volume 297, 2013: "Worldwide Large-Scale Trapping and Hunting of Ungulates in Past Societies"). In these groups non-breeding females might have assisted male hunters (Biesele and Barclay 2001; Kuhn and Stiner 2006). Other foraging activities are also often more effectively done in groups (for instance bee honey extraction, see Mulder *et al.* (2000)).

However, there is hardly archaeological data supporting the theory that Neanderthals actually lived and foraged in groups, but extended family structures have been suggested for the genetically related individuals identified in El Sidron (Lalueza-Fox *et al.* 2011. Teamwork at collaborative tasks has been attested with chimpanzees suggesting a deep origin within the hominids (Melis and Tomasello 2013). For an evolutionary pathway to cooperative social behaviour in hominins see for instance Shultz *et al.* (2011) and Lehmann *et al.* (2007). The underlying model in this study assumes that hominins live in groups, comprised of individuals of different age categories. Maximum group size is a parameter in the model.

The model in HominínSpace assumes that whatever the composition in age and sex when sufficient resources are available in the area, each group is capable of finding and retrieving all resources necessary to meet its subsistence requirements. Differences in hunting and gathering behaviour between males and females (e.g. Bird and Bliege Bird 2005), or between adults and children (MacDonald 2007; White 2013) are of no influence on the foraging efficiency of the modelled group as a whole. With insufficient resources groups will suffer as described in the next subsection. Groups, when sufficiently repopulated, have the potential to live forever.

### 5.4.1 Group demographics - new groups, group extinction, merging of groups

A group moves through the landscape and interacts with the environment and other groups as a whole (cf. Agustí and Rubio-Campillo (2017)). Groups can procreate by splitting a new group from the mother group. When that happens a group is created consisting of fertile adults of both sexes to ensure healthy growth potential. Groups look for

opportunities to split when the total group size exceeds a given, parameterized value. The default for this number is 50, the largest size for what Dunbar (1993) refers to as the overnight camp or band size, derived from ethnographical observations of traditional hunter-gatherers, but different group sizes have been observed (Hill and Dunbar 2003; Zhou *et al.* 2005; Hamilton *et al.* 2007; Layton and O'Hara 2010, 101). The model in HomininSpace assumes that groups exceeding the maximum size will split. In the HomininSpace model new groups can also be created in *core areas*. These are locations on the map in which, conditions permitting, new groups are created (see subsection 9.4.2). A newly created hominin group consists of 25 individuals, with 15 being of pre-fertile age (cf. Sørensen (2009)).

Groups can join other groups (merge). Especially when the group size is below a minimum threshold (this is a parameter) they will send a join request to other groups in the area around them aiming to avoid extinction (Soltis *et al.* 1995). 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 for their own number in the area, or when the group size already exceeds the maximum size[23].

Ethnographic research has shown that group extinction is a common process in hunter-gatherer societies, where groups are eliminated for a variety of reasons (Soltis *et al.* 1995, 477). The percentage of eliminated groups varies wildly across cultures and therefore the model in HomininSpace has adopted a viability criterion. Since death is a stochastic process in HomininSpace, and during each time step that a group is hungry only a percentage of a group dies, a minimum threshold has been established after which a group is considered extinct. The group cannot be sustained when the number of individuals falls below a parameterized minimum (Birdsell 1968; Wobst 1974). The default for this threshold is three fertile adults, considered the minimum size critical for regeneration and survival (Nabel *et al.* 2013). This number is however parameterized and if a group contains this number or less reproductive adults it is considered not viable and if it does not join another group that year it is removed from the simulation. Groups are removed from a simulation when they move out of the simulation area.

Newly created groups are not able to interact with other groups for several time steps after creation. Since a new group is always created within the foraging range of the parent group

[23] Other reasons are easily found but have not been implemented, like different cultural affiliations or just chance.

any new group is always in direct competition for resources with other groups, and will most likely become "hungry" (see below). Such new groups are allowed to merge with other groups only after a certain number of years, to prevent an intense creation-hungry-merge cycle. The number of these years is included as a parameter. Although exact ranges are unknown, this modelled behaviour is partly based on the intrinsic mammal dispersal pattern that drives away all or some of one's offspring, to avoid inbreeding or competition for resources or members of the opposite sex (Dobson 1982). In humans most often young males are expelled, sometimes accompanied by fertile female companions. With mammals in areas with high population densities emigration rates increase, forcing groups to move away (Baker 1978). Such expansive behaviour is sometimes even actively promoted genetically, especially in frontier regions (Phillips, Brown, *et al.* 2006).

### 5.4.2 Social interaction - home range and foraging range

Each grid cell in HomininSpace can contain one group only. When a group is present this cell is referred to as the *home range* of the group and will be defended against other groups at all costs. Small groups defending their own home range stand a good chance even against much larger groups (Crofoot and Gilby 2012). Therefore, the model in HomininSpace prevents other groups from occupying an already claimed home range or even exploiting it in their foraging activities. Each grid cell, and thus also the home range, is implemented as an area of 10x10km or 100km$^2$. Such a limited area hardly ever produces enough resources to meet the subsistence needs of a group for longer periods. Members of the group will have to constantly leave the home range and forage outside the protected area. They will have to collect resources within the *foraging range* (FR) for that group. Since a time step in HomininSpace is one year, this is actually the *annual range* as generally referred to in the literature.

The FR is variable and recalculated for each individual group based on the characteristics of that group and the environment each year. The range is implemented as a square area around the home grid cell (not a circle!). FR gives the number of grid cells from the center to any side. The length of one side is two times the range plus one (see Figure 22 and below). The flexibility of the range reflects the ability of hominins to obtain resources from any environment. Foraging ranges can optionally be limited by a maximum size, included as a parameter in the HomininSpace model. For an example of modern day hunter-gatherer ranges see the values collected by Kelly (1995), table 4.1, or by Weaver and Steudel-Numbers (2005).

At the beginning of each simulated year the estimated required FRs for all groups are computed in random order. These ranges are based on the subsistence needs of the members of the group, the remaining available resources in the area (defined by the climate parameters and previous consumption), and the presence and subsistence needs of other groups that occupy nearby regions. To calculate the FR the number of individuals in each age segment is multiplied by the energy needs for members of that segment and added together. Then the available energy (usable meat) from the home range is subtracted. If there are still resources needed the foraging range is increased by one grid cell. The energy in each cell within the new foraging range is added together. If this is not enough to satisfy the needs of the group the FR is again increased and this process is repeated until the energy requirements are met or the maximum FR is reached (see Figure 22).

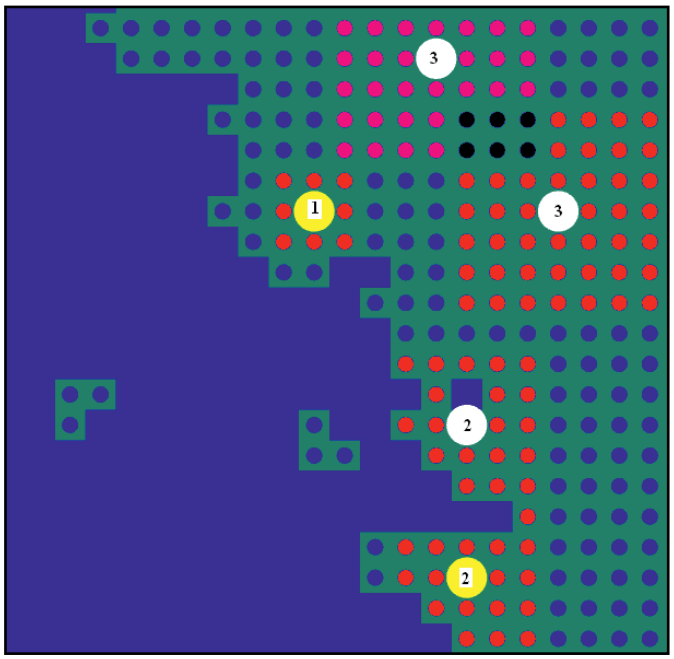



**Figure 22: Foraging ranges and resource acquisition. Larger circles are home ranges with the value of the foraging range of the groups in the centre. A value of '2' means two grid cells in all directions (horizontal, vertical and diagonal) are included when foraging for resources. Black are shared grid cells.**

When a FR is calculated that supplies enough energy all grid cells within that foraging range are 'claimed' by this group. That is, all grid cells have a counter for the number of claims, and within the foraging range all counters are increased by one. All grid cells that have received two or more claims can provide only that fraction of the available energy to each claiming group. The number of cells in a foraging range = $(2*range+1)^2$. So for range=0 the number of cells is 1, for range=1 the number of cells = 9, for range = 2

cells=25, for range=3 cells = 49, etc. If no maximum foraging range is set, the foraging area can theoretically include the whole map area.

Subsequently all groups consume the energy they need, in random order. Groups with not enough resources will consume all available energy in their foraging range. With insufficient resources (for instance because another group has consumed part of the claimed shared resources or because the foraging range is too small) a group is considered *hungry*. Hungry groups have penalties on birth rates and suffer from increased death rates. Groups are allowed to move through another foraging range in search for food (for a comparative example with hyena's see Hofer and East (1993)). The most likely situation in which a group A is deprived of resources is when group B claims cells in the FR *after* the computation of the foraging range for A.

### 5.4.3 Mobility

If the local climate deteriorates this will affect resource availability and influence hominin presence in the area. Generally, populations will decline, either through migration following preferred habitats or through local extinction of individual groups (Bradtmöller *et al.* 2012; Dalén *et al.* 2012; Hublin and Roebroeks 2009; Roebroeks *et al.* 2011). Net population dynamics in general are steered by four processes: immigration into the area, emigration out of that area, the rate of reproduction of the species under the given circumstances and the number of individuals that die per time unit. Climate change will cause geographical ranges for species to contract, expand, shift or stay the same depending on the climatic tolerances and environmental conditions (MacDonald *et al.* 2012; Stewart *et al.* 2010, 661).

Groups move through the landscape and available energy steers their direction. In the model of HomininSpace there are two types of hominin mobility strategies, referred to as *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 are depleted (with insufficient resources groups will shrink and subsequently need fewer resources). There is one exception where static groups that have decided to stay can move, 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. Dynamic mobility is characterized by constant movement, always directed towards the most favourable area within range. As groups tend to deplete the resources in the area

where they stay they 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. This factor two is derived from the radius leap-frog pattern described by Binford (1982) where residential moves between home bases are twice the foraging distances of a group. Habitat patches are defined as the sum of the available energy for each cell combined with its surrounding adjacent eight neighbours. When implementing movement it was found that hominin groups should not be attracted by the high productivity of single grid cells to avoid so called *death traps*: a grid cell with high productivity surrounded by cells with very low productivity. Groups that move onto such attractive cells will have problems sustaining themselves and quickly vanish. Examples of these death traps are small islands and promontories, both surrounded by mostly zero productivity sea cells.

When the cell with the highest productivity for this group is 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 (claim on resources). Thus other groups will have to share the surplus of resources of those cells with this group. There are no separate locomotion costs (Alexander 2002) associated with these residential moves since they are included in the energy budgets for each individual. The single move in each time step is in effect the net displacement of all residential moves that year. Thus an effect of energy requirements on frequency of displacement or duration of stay (Verpoorte 2006) is not included in the HomininSpace model.

Groups will never move onto sea or lake patches, and they will not move onto a cell that is already occupied by another group. When the borders around the simulation area are activated, they also never move to a grid cell located near any border. If these borders are not activated groups can leave the simulation completely.

## 5.5 Overview: building blocks for the underlying model

Table 9 summarizes the model elements in the HomininSpace model as identified in this chapter. The default values are provided for illustrative purposes and have been discussed elsewhere (Scherjon 2015a). HomininSpace identifies three age categories for hominin species: each individual is either pre-fertile, fertile or post-fertile. Model variables included in this table can be one of two types:

- *Parameters*: variables that can be given a value and that control the behaviour of the agents or the system in a certain way. These are read from the parameter input file, the naming of the variables in the table below is consistent with that file;
- *Constants*: constant values, the same for all simulations; these can be modified only by changing the source code and recompiling the executable that runs the simulations.

**Table 9: Variables on hominin behaviour included in the HominSpace model.**

| ***Name*** | ***Type*** | ***Description*** | ***Default value*** |
|---|---|---|---|
| ***CohortSize_PreFertile*** | parameter | The size of the pre-fertile segment. | 15 |
| ***CohortSize_Fertile*** | parameter | The size of the fertile segment (the length of the post-fertile segment is not limited). | 25 |
| ***DeathRate_PreFertileCohort*** | parameter | Mortality rate for the pre-fertile segment per group. | 4 |
| ***DeathRate_FertileCohort*** | parameter | Mortality for the fertile segment per group. | 2 |
| ***DeathRate_PostFertileCohort*** | parameter | Mortality for the post-fertile segment per group. | 8 |
| ***BirthRate*** | parameter | Birth-rate in a percentage of the fertile category per group (a value of 33% means one child per three years for each female). | 33 |
| ***Subsistence_PreFertileCohort*** | parameter | Energy needs for an individual in the pre-fertile segment in kcal. | 3,000 |
| ***Subsistence_FertileCohort*** | parameter | Energy needs for an individual in the fertile segment in kcal. | 4,000 |
| ***Subsistence_PostFertileCohort*** | parameter | Energy needs for an individual in the post-fertile segment (in kcal). | 3,500 |
| ***Years_Before_Group_Maturity*** | parameter | Period in years before newly created groups can interact with other groups (for instance to join). | 5 |
| ***GroupSize_BeforeMerge*** | parameter | Minimum size (threshold) for the group. If the size of the groups falls below this number the group will attempt to merge, if possible. | 3 |
| ***GroupSizeFertile_BeforeMerge*** | parameter | Minimum size for the fertile category. If this segment contains fewer individuals, the group will also attempt to merge. | 3 |
| ***GroupSize_BeforeSplit*** | parameter | Maximum size of a group. If the number of individuals exceeds this number the group will split into two groups. | 50 |
| ***Calories_Per_Kg_Meat*** | parameter | The amount of kcal that can be extracted from one kilogram of meat. | 3,000 |
| ***Temperature_Tolerance*** | parameter | Minimum temperature that can be sustained by the hominins. If the coldest temperature during a year falls below this value the hominins die. | -18 |
| ***Max_ForagingRange*** | parameter | Maximum annual foraging range. Groups cannot forage outside this range. | 15 |
| | | | |
| ***Penalty growthrate*** | constant | Penalty applied to the growth rate when subsistence requirement are not met. | 0.5 |

| | | | |
|---|---|---|---|
| ***Penalty mortalityrate*** | constant | Penalty applied to all mortality rates when subsistence requirement are not met. | 2 |
| ***FemaleMale*** | constant | The percentage of females in the population. | 50 |
| ***usableMeat*** | constant | The percentage of one kilogram of animal that is usable meat for meeting subsistence needs. | 60 |
| ***maxGroup size*** | constant | The maximum size for a group before offspring is created, in individual members. | 50 |
| ***newGroups prefertile*** | constant | The size of the pre-fertile segment in newly created groups (from core areas). | 8 |
| ***newGroups fertile*** | constant | The size of the fertile segment in newly created groups (from core areas). | 11 |
| ***newGroups postfertile*** | constant | The size of the post-fertile segment in newly created groups (from core areas). | 6 |
| ***Homerange*** | constant | The number of grid cells that are considered by a group as their territory. | 1 |
| ***ViableGroup-size*** | constant | If the number of individuals in the group falls below this number the group is removed. | 3 |
| ***Scanning modifier*** | constant | The scanning range is defined by default as two times the annual foraging range. Within scanning range the hominins are aware of resource availability. | 2 |
| ***Factory ScanningRange*** | constant | The radius (in grid cells) of the area that a factory must monitor when creating new groups. | 30 |
| ***openBorders*** | constant | Defines if borders are open or closed. If True (open), hominins can leave the simulation area. | False |

# 6. THE ARCHAEOLOGICAL DATA IN THE MODEL

## 6.1 Introduction

Archaeological data is often used to validate the results of simulations (Lake 2014). HomininSpace takes data points from the archaeological record from which we know that hominins *were there* at a specific point in time. This presence information is the main core of the archaeological database in HomininSpace (Section 6.2). The archaeological record of western Europe is probably the area with most detail in both temporal scale and geographical distribution of Neanderthal sites, specifically for the Late Pleistocene (Locht *et al.* 2010).

For HomininSpace a chrono-archaeological database (compare van Andel (2002, 6)) has been built to store Checkpoints in Space and Time (CSTs) (Scherjon 2012). Each CST has a list of radiometric dating results (Section 6.3) that indicates in what time frames hominins were at that location. The date with the associated 1-δ uncertainty values form an interval of presence during which hominins were at that site with 68.3% certainty. Thus a CST consists of X and Y coordinates (a latitude/longitude location) and a list of chronological intervals, compiled according to specific criteria. In HomininSpace the CSTs are implemented as active agents storing data about the hominin agents that are present in that area (Section 6.4). Checkpoint locations and associated dates have been selected when they fulfil the following requirements:

- The location is an archaeological site featuring in peer reviewed scientific publications;
- Methods, techniques and results of sampling and dating are described in detail;
- Stratigraphy is clearly described and position of the samples is unambiguous or from a closed archaeological context;
- Samples have a good association with other finds and are properly described;
- Results of the laboratory are discussed;
- There is no dispute in the literature on the nature of the dates;
- Calculated dates are earlier than 50 ka;
- Standard deviation of a date is not larger than half the date itself (dates like 130 ka +/- 90,000 are rejected). Informative value of measurements violating this rule is considered low;
- Results are actual chronometric dates (not given in MIS periods or relative to other localities).

Some included CSTs do not fit all the criteria. Any deviation from these quality standard selection criteria is indicated and discussed. The database file with all the checkpoints is included in the Supplementary Materials[24].

## 6.2 The archaeology of presence and absence

The use of large datasets in a spatial context to validate the model of past hominin behaviour is an attractive approach to assess the quality of a simulation model. Related work to this study is done in Germany on modelling the earliest hominins in Europe and in the Role of Culture in Early Expansions of Humans (ROCEEH) project (Timm *et al.* 2014), and with a similar focus in Italy (Muttoni *et al.* 2018), in Japan on Initial Upper Palaeolithic populations in Eurasia in the Replacement of Neanderthals by Modern Humans (RNMH) project (Kondo *et al.* 2018). Previously the Stage 3 project in the UK collected an extensive database on European Palaeolithic sites (van Andel and Davies 2003) which was updated in the Préhistore à l'Actuel: Culture, Environments et Anthropologie (PACEA) database (D'Errico *et al.* 2011); both were used in many studies. None of these projects however attempts to combine a detailed environment reconstruction with a generic hominin model and an evolutionary algorithm to find parameter values validated against a database of hominin presence and absence.

When using archaeological data it must be acknowledged that such data becomes available only through many biasing filtering processes (Clarke 1973, 16-17; Schiffer 1985). These include excavation biases, where some areas are more scrutinized than others (Dennell and Roebroeks 2005, 1102; Surovell and Brantingham 2007), biases introduced by the geography of the region with accessibility of relevant deposits (Bynoe *et al.* 2016), a research bias (Roebroeks *et al.* 2011), a bias created by the archaeological methods used (Discamps and Faivre 2017; Wheatley 2004), or a bias created by taphonomy (Soressi 2016) where younger sites are over-represented due to the loss of older sites to environmental and climatic factors, to mention just a few.

A bias can also be introduced through differences in the site function type. The archaeological materials in caves (living areas) might be different from those found at special activity sites like kill/butchery sites. Certain sites might not be recognized due to the fact the typical artefacts are not present there because they were not needed for the

[24] Named "Neandertal sites - north west Europe.xlsx".

specific task at that site. A (small) bias is also imposed by the limitations of this author since only English and French literature was studied.

A comprehensive overview of all Neanderthal sites is currently missing in the literature, although some attempt to collect as many as possible (for instance the database constructed in the Replacement of Neanderthals by Modern Humans (RNMH) project (Akazawa *et al.* 2013)). Some collection biases can be accommodated or corrected by statistical techniques (Surovell *et al.* 2009) but in this study the individual radiometric dates are used as they are reported, and as such no corrections were applied to this data to compensate any collection bias. The literature was searched for radiometric dates obtained for Palaeolithic sites in the case study area. Undated archaeology is ignored for all purposes. The actual archaeology, its inferred function, or cultural affiliation is ignored. For individual sites only presence information is extracted and used. The traces of hominin activity in the past rarely survive sufficiently to support inferences about the behavioural patterns of past hominins but HomininSpace will use as many securely dated time points as possible from a large area to attempt such inferences.

True absence information, the fact that hominins were *not* somewhere at certain periods, is at least equally valuable as presence information. However there are two issues with absence information: absence of evidence is not evidence of absence and the inherent limited applicability of any absence data. Absence of information can have many reasons including inaccessibility of sites, unavailability of research resources, inhospitable climate, natural barriers, or removal of evidence by taphonomy. Excavation preferences can also induce preferential site selection, suggesting absence caused by simply not looking.

Any cause of absence can have a local or wider effect. When for instance continent wide glaciation occurs larger areas can become devoid of life, by restricting access to all resources. However, a clear local absence in one site caused by a rock fall blocking access to a cave does not mean that hominins were not present in any other location nearby. For individual sites it is very difficult to ascertain that given absence truly means that hominins were not in the area and that the evidence can be extended to a wider area around the site (Phillips, Anderson, *et al.* 2006).

## 6.3 What dating information to use?

When modelling past hominins faith increases in models that match the archaeology well and decreases for models that predict distribution patterns poorly. For HomininSpace, only

absolutely dated archaeology is used as a reference against which simulation results will be matched. This study uses the radiometric dates associated with archaeological deposits resulting from visits by past hominins. Only dates that indicate a presence before 50 ka are used to validate the simulations. Computational resources have further limited the lower boundary to 131 ka, resulting in a simulation period of 81 kyr.

The chronological framework for the area is mainly based on a number of stratigraphic sequences with mostly luminescence dating of heated flint and sediments, added to U-Th/ESR dates, measured on bones clearly associated with human occupation. Dating information derived from stratigraphic, archaeostratigraphic or bio-climate correlations are not included in the database (e.g. dating using the occurrence of early Weichselian "sols steppiques" in Goval (2008)). Identification of exact dates is difficult for relative dates and the database in HominінSpace is therefore constructed without qualifiers like 'before' or 'after'. It is acknowledged that by applying such rigorous selection standards many potential positive presence data points are rejected, but priority has been given to an radiometrically dated framework to validate against.

Radiometric dates are included to construct presence intervals with 1-δ uncertainty (within one standard deviation). The probability of a measurement dating actual presence on the site for 1-δ is only 68.3%. 2-δ could be used, which would give a 95.5% certainty (Richter *et al.* 2013). But 1-δ is chosen since this is the de facto standard in reporting archaeological dates. This follows Guibert *et al.* (2008), a major collection of dates for the area and period of study. The individual ages are combined per layer where possible. A one sigma interval in dating accuracy is deemed sufficient in this research and this limited precision is taken into account in the discussion.

### 6.3.1 Database structure

All collected data is stored in an Excel database file. The name of the file is "Neandertal sites - north west Europe.xlsx". The structure of this database is given in Table 10. Included for each site is dating information. Each radiometric date represents an individual moment in time, defined by the calculated date plus and minus one standard deviation given in kilo years (ka). As such, when a bovid tooth from La Ferrassie is dated using the isochron ESR technique with a resulting date range of 61 +/- 5 ka (Blackwell *et al.* 2007), a checkpoint for this site is created that checks for presence of hominins from 66 to 56 ka inclusive. In total there are 470 dates in the database. Dates from the same locality (archaeological site) are combined into a single Checkpoint object with multiple ranges.

The date ranges are henceforth referred to as Presence Intervals, both in the source code as well as in this thesis. Other checkpoints may overlap the area (when two sites are in the same grid cell) but the check for presence is done for each checkpoint separately. Table 11 lists all the included checkpoints, with all the intervals in Table 12.

**Table 10: Data fields in the archaeological data database.**

| ***Field*** | ***Name*** | ***Remarks*** |
|---|---|---|
| **1** | # | Unique number |
| **2** | Locality | Site name |
| **3** | Latitude | Position: latitude (decimal format) |
| **4** | Longitude | Position: longitude (decimal format) |
| **5** | Age (ka) | Radiometric age |
| **6** | Plus (kyr) | One sigma (standard deviation) plus |
| **7** | Minus (kyr) | One sigma (standard deviation) minus |
| **8** | Dating method | Take from the literary source |
| **9** | Year | Identification of measurement: year |
| **10** | Laboratory | Identification of measurement: laboratory where measurement was done |
| **11** | Sample | Identification of measurement: sample identification with number given by the laboratory |
| **12** | Layer | Identification of measurement: layer (also used to combine multiple measurements into one single interval) |
| **13** | Associated material/culture | Cultural interpretation or association of dated object (illustrative purposes only) |
| **14** | Type of site | Open air or cave (this value is not used in this study) |
| **15** | Confidence | 1, 2, or 3, with a 1 indicating the lowest confidence level |
| **16** | Remarks | Any additional information |
| **17** | Reference (found in) | Literature source from which measurement was taken |
| **18** | Primary source | Original source where the date was first published (if different from the Reference) |

When two or more intervals are within the same archaeological *layer*, they are combined into one single interval using the minimum and maximum interval values. This process is illustrated here with data depicted in Figure 23, where multiple TL dates are present in two layers for the open air site Port Racine (Cliquet *et al.* 2003, 53). Combining these dates with their standard deviations per layer results in two intervals for Port Racine: (1) Port Racine 62-81 ka for “secteur 4” and (2) Port Racine 55-88 ka for “secteur 1”. This assumes that the layers were correctly identified by the original researchers, that these layers separate different periods with visits by past hominins, and that these dates do not contain anomalous dating results.

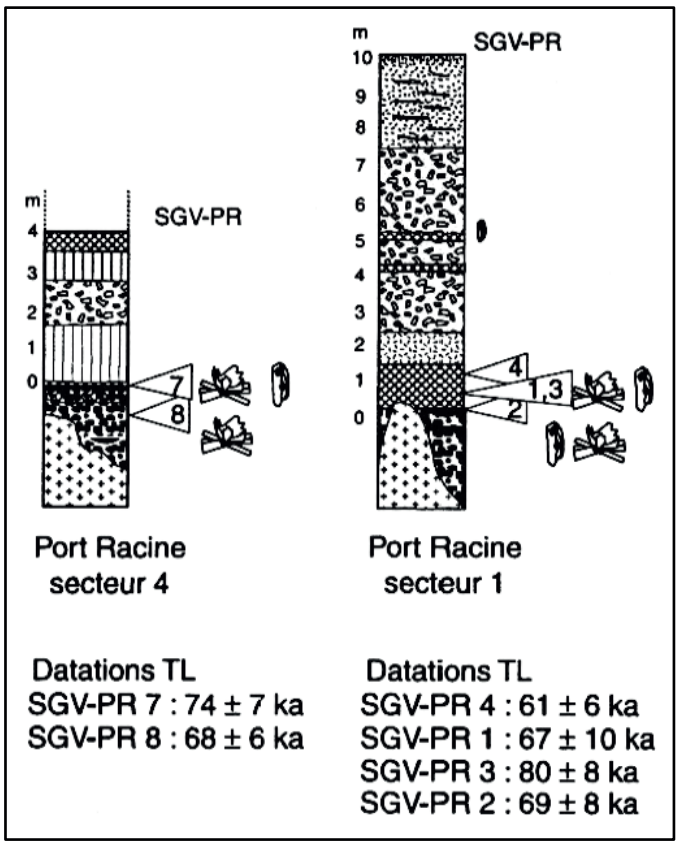



**Figure 23: Example illustrating the interval construction. Port Racine, two layers with dates, taken from Cliquet *et al.* (2003).**

Geographical locations are taken from the Radiocarbon Palaeolithic Database Europe (RPDE), which can be downloaded in ".kmz" format, the native Google Earth specification language. The RPDE is a collection of Lower, Middle and Upper Palaeolithic sites with environmental conditions and available chronometric dating. Version 11 supplied most locations and was downloaded November 2012[25]. Version 20 supplied more locations and was downloaded in August 2016, and was renamed to RPED v20. The Access database was constituted by the INQUA-Commission on Palaeoecology and Human Evolution and can be freely downloaded and used.

Although much effort has been put into collecting as many published dates as possible and to include when available unpublished archaeological data for the research area, the publication archive is huge and added to on a daily basis. By the nature of such an archive completeness is unattainable and as such, the database used by the simulation system is implemented to be easily updatable when new dates are published. Simulations are reproducible and can be rerun with the same parameter values but with new checkpoint datasets when needed.

[25] From http://ees.kuleuven.be/geography/projects/14c-palaeolithic/index.html, accessed 2012.

## 6.4 Checkpoints in Space and Time (CSTs)

The regular Checkpoints in Space and Time (CSTs) are simulation objects that register the presence of hominin groups that move through the area within defined presence intervals based on archaeological information. There are other types of (check)points that need a location in HomininSpace (Subsection 8.5.3): monitoring checkpoints that register presence data without intervals, starting points where initial hominin groups are created, and climate monitoring points that record the local climate. See Figure 24 for an overview of all checkpoints. Checkpoints are defined in the checkpoint input file.

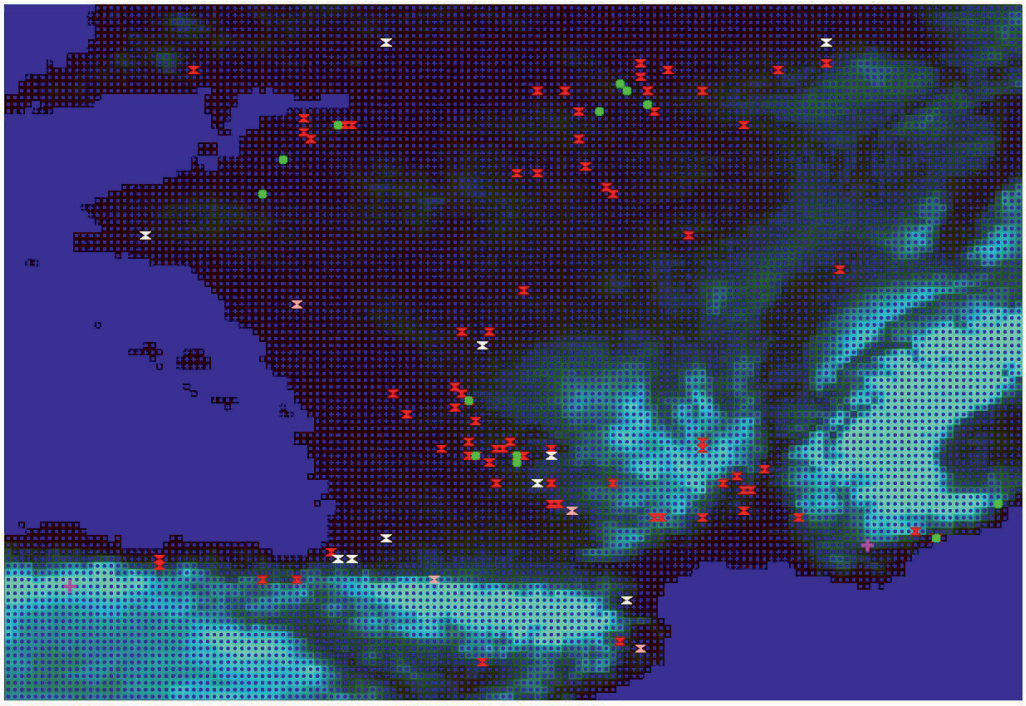


**Figure 24: Overview of all checkpoints. Red are regular checkpoints, white monitoring checkpoints, green starting locations, pink climate monitoring points and purple plus signs are core areas.**

All CSTs are associated with an archaeological site and have one or more intervals defined. A hominin group will be recorded by a CST whenever the gridcell containing the CST is visited in the foraging phase of that group. Information stored includes visiting group type, visit count, and whether the visit is within any of the presence intervals. Table 11 gives geographical position of all regular checkpoints. The size of this collection is limited (n=83), illustrating the deficiency in accurate chronological information for western Europe which is the most intensely studied Late Pleistocene area (Vieillevigne *et al.* 2008). In total 470 intervals are associated with these checkpoints. This provides only a snapshot view on the behaviours of hominins living in this huge area and during this large time span.

**Table 11: Regular Checkpoints in Space and Time with geographical position.**

| # | *Name* | *Latitude* | *Longitude* | # | *Name* | *Latitude* | *Longitude* |
|---|---|---|---|---|---|---|---|
| **1** | **Abauntz** | 43.02722 | -2.04166 | **43** | **Igue des Rameaux** | 44.1524 | 1.75329 |
| **2** | **Abri Bourgeois-Delaunay** | 45.66667 | 0.4667 | **44** | **Jupiter** | 43.47889 | -1.45306 |
| **3** | **Abri des Canalettes** | 43.993 | 3.257 | **45** | **Kent's Cavern** | 50.46987 | -3.53 |
| **4** | **Abri des Pêcheurs** | 44.40806 | 4.20722 | **46** | **La Butte d'Arvigny** | 48.63278 | 2.61778 |
| **5** | **Abri du Brugas** | 44.05889 | 4.48389 | **47** | **La Chapelle aux Saints** | 44.9833 | 1.7167 |
| **6** | **Abri du Maras** | 44.31694 | 4.55528 | **48** | **La Ferrassie** | 44.9667 | 0.9333 |
| **7** | **Abri Suard** | 45.66667 | 0.46667 | **49** | **La Folie** | 46.61904 | 0.35471 |
| **8** | **Angé** | 47.29444 | 1.251944 | **50** | **La Quina** | 45.504 | 0.303 |
| **9** | **Anse de Query** | 49.65167 | -1.23972 | **51** | **La Roche à Pierrot** | 45.75389 | -0.63833 |
| **10** | **Artenac** | 45.85 | 0.3333 | **52** | **La Rochette** | 45.01667 | 1.10167 |
| **11** | **Ault** | 50.1075 | 1.45167 | **53** | **La Roquette II** | 43.94278 | 3.90278 |
| **12** | **Barbas I** | 44.8499 | 0.5667 | **54** | **Le Moustier** | 45.00198 | 1.07 |
| **13** | **Baume Vallée** | 44.96667 | 3.88306 | **55** | **Le Prissé** | 43.48139 | -1.45833 |
| **14** | **Beauvais 1** | 49.4325 | 2.13278 | **56** | **Le Rescoundudou**[26] | 44.40833 | 2.575 |
| **15** | **Bérigoule** | 43.98425 | 5.2547 | **57** | **Le Rozel** | 49.47111 | -1.8425 |
| **16** | **Biache-Saint-Vaast** | 50.3 | 2.95 | **58** | **Les Canalettes** | 43.95 | 3.1833 |
| **17** | **Boxgrove** | 50.86901 | -0.69028 | **59** | **Les Cottés** | 46.67333 | 0.83333 |
| **18** | **Cantalouette II** | 44.864 | 0.54793 | **60** | **Les Forêts** | 44.98804 | 0.11429 |
| **19** | **Caours** | 50.13028 | 1.87833 | **61** | **Les Pradelles** | 45.733 | 0.4333 |
| **20** | **Champlost** | 48.02417 | 3.67111 | **62** | **Lezetxiki** | 43.08 | -2.5 |
| **21** | **Chez-Pinaud Jonzac** | 45.43774 | -0.41926 | **63** | **Mauquenchy** | 49.60389 | 4.46778 |
| **22** | **Combe Brune** | 44.87167 | 0.50444 | **64** | **Ormesson** | 48.78583 | 2.54083 |
| **23** | **Combe Grenal** | 44.8084 | 1.223 | **65** | **Payre** | 44.68 | 4.75 |
| **24** | **Combe-Capelle Bas** | 44.76619 | 0.81344 | **66** | **Pech de l'Azé I** | 44.85883 | 1.25331 |
| **25** | **Coudoulous I** | 44.16864 | 1.69023 | **67** | **Pech de l'Azé II** | 44.85937 | 1.25264 |
| **26** | **Cova de l'Arbreda** | 42.16111 | 2.74694 | **68** | **Pech de l'Azé IV** | 44.85941 | 1.25259 |
| **27** | **Covalejos Cave** | 43.38639 | -3.95722 | **69** | **Pié-Lombard** | 43.71667 | 7.04972 |
| **28** | **Croix du Canard** | 45.09174 | 0.47448 | **70** | **Pont-des-Planches** | 47.53556 | 5.92167 |
| **29** | **El Castillo** | 43.29 | -3.965 | **71** | **Port Racine** | 49.53333 | -1.9 |
| **30** | **Estret de Trago** | 41.87 | 0.72 | **72** | **Roc de Marsal** | 44.90636 | 0.96705 |
| **31** | **Fermanville-La Mondrée** | 49.68306 | -1.45 | **73** | **Saint Germain-des-Vaux** | 49.72639 | -1.92278 |
| **32** | **Fonseigner** | 45.3167 | 0.6095 | **74** | **Saint-Amand-les-Eaux** | 50.44806 | 3.42722 |
| **33** | **Fresnoy-au-Val** | 49.83974 | 2.06103 | **75** | **Sainte-Anne** | 45.06667 | 3.86667 |
| **34** | **Gouberville** | 49.69 | -1.32 | **76** | **Saint-Hilaire-sur-Helpe** | 50.13222 | 3.90417 |

[26] In the RDPE database this is not "Le Rescoundudou" but "Rescoundudou" (without Le), to the south.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **35** | **Grossoeuvre** | 48.94111 | 1.19638 | **77** | **Saint-Illiers-la-Ville** | 48.97778 | 1.54278 |
| **36** | **Grotte de Coudoulous II** | 44.46864 | 1.69023 | **78** | **Savy** | 49.84278 | 3.20861 |
| **37** | **Grotte des Barasses II** | 44.51167 | 4.36472 | **79** | **Scladina** | 50.48528 | 5.02583 |
| **38** | **Grotte du Figuier** | 44.32333 | 4.54889 | **80** | **Seclin** | 50.55 | 3.03333 |
| **39** | **Grotte du Lazaret** | 43.69094 | 7.29499 | **81** | **Sous les Vignes** | 44.48332 | 0.93333 |
| **40** | **Grotte Vaufrey** | 44.80023 | 1.19922 | **82** | **Villiers-Adam** | 49.07083 | 2.22083 |
| **41** | **Grotte XVI** | 44.80348 | 1.18453 | **83** | **Walou Cave** | 50.59056 | 5.69389 |
| **42** | **Havrincourt 1** | 50.11722 | 3.0825 | | | | |

In names, dashes are optional. For example “Sous les Vignes” is also known as “Sous-les-Vignes”.

In the database, a Confidence level is used with the following meaning:

1. unsure, data missing, disputed in the literature, only available through secondary sources (n=28)[27];
2. reasonable, well dated element but with incomplete description or it is disputed (n=204);
3. very confident, well described, undisputed, complete results in literature (n=237).

In the database is indicated upon what criteria the confidence level is based. Where reasonable the level is based on an evaluation by the original source (mainly Guibert *et al.* (2008, 20)) with the criteria described by the authors (all the zero level or very bad dating results have already been dropped since these are considered unusable in the context of this research). Note that when intervals are combined the confidence level of the resulting interval is the lowest level of the original intervals. The combined presence intervals per checkpoint are listed in Table 12.

**Table 12: Checkpoints with x and y grid cell locations and their interval ranges.**

| Checkpoint | X | Y | Intervals |
|---|---|---|---|
| **Abauntz** | 42 | 17 | 40000 - 54000 |
| **Abri Bourgeois-Delaunay** | 67 | 43 | 130000 - 166000, 91000 - 116000, 111800 - 128670, 77060 - 79280, 107000 - 117000, 65000 - 121000 |
| **Abri des Canalettes** | 95 | 26 | 68500 - 78500 |
| **Abri des Pecheurs** | 104 | 31 | 49000 - 59000 |
| **Abri du Brugas** | 107 | 27 | 57200 - 68800 |
| **Abri du Maras** | 108 | 30 | 81000 - 99000, 53000 - 57000, 40000 - 52000, 69000 - 95000 |
| **Abri Suard** | 67 | 43 | 111000 - 141000, 217000 - 287000, 45000 - 57000, 94000 - 108000 |
| **Ange** | 75 | 59 | 59700 - 161800 |
| **Anse de Query** | 50 | 83 | 114720 - 136360 |
| **Artenac** | 65 | 45 | 96650 - 119000, 122200 - 123200, 64000 - 70000 |

[27] Included in the database to increase geographical spread.

| | | | |
|---|---|---|---|
| **Ault** | 77 | 88 | 45000 - 65000 |
| **Barbas I** | 68 | 35 | 117000 - 175000 |
| **Baume Vallee** | 101 | 36 | 71000 - 86000 |
| **Beauvais 1** | 83 | 81 | 51600 - 59600 |
| **Berigoule** | 115 | 26 | 61300 - 122100, 49800 - 97600 |
| **Biache-Saint-Vaast** | 92 | 90 | 162000 - 188000, 112000 - 166000 |
| **Boxgrove** | 55 | 95 | 72000 - 142000 |
| **Cantalouette II** | 67 | 35 | 55000 - 66870 |
| **Caours** | 81 | 88 | 109000 - 139000, 108700 - 124500, 105000 - 131000 |
| **Champlost** | 99 | 67 | 52500 - 60900 |
| **Chez-Pinaud Jonzac** | 58 | 41 | 55300 - 80700, 67000 - 95600, 59500 - 82700, 40200 - 75300, 56100 - 79200, 56000 - 75200, 36400 - 79600, 47900 - 89300 |
| **Combe Brun** | 67 | 35 | 56600 - 69600, 105000 - 131000 |
| **Combe Grenal** | 74 | 35 | 40000 - 48000, 54000 - 68000, 55000 - 69000, 61000 - 75000, 91000 - 126000 |
| **Combe-Capelle Bas** | 70 | 34 | 48000 - 57000, 48300 - 57500, 49100 - 64700, 33900 - 61600 |
| **Coudoulous I** | 79 | 28 | 88200 - 112600, 110100 - 146200, 127700 - 152000, 119000 - 161000, 40300 - 65300, 61800 - 67600, 99600 - 138300 |
| **Cova de l-Arbreda** | 89 | 8 | 74300 - 93700 |
| **Covalejos Cave** | 22 | 20 | 87857 - 95857 |
| **Croix du Canard** | 67 | 37 | 74900 - 83700 |
| **El Castillo** | 22 | 19 | 58500 - 79900, 54300 - 84400 |
| **Estret de Trago** | 69 | 5 | 65600 - 141000, 67500 - 122000, 37200 - 46200, 42900 - 58800, 38400 - 47600, 41500 - 50700 |
| **Fermanville-La Mondree** | 48 | 83 | 63700 - 75700 |
| **Fonseigner** | 68 | 40 | 44900 - 55500, 47300 - 58300, 49600 - 63200 |
| **Fresnoy-au-Val** | 83 | 85 | 99300 - 114300 |
| **Gouberville** | 49 | 83 | 108000 - 148000 |
| **Grossoeuvre** | 74 | 76 | 122000 - 138000 |
| **Grotte de Coudoulous II** | 79 | 31 | 39000 - 51000, 34600 - 85700, 23300 - 56600, 79100 - 97400, 78000 - 102600, 86900 - 147500 |
| **Grotte des Barasses II** | 106 | 32 | 98000 - 124000, 43000 - 53000, 53000 - 66000 |
| **Grotte du Figuier** | 107 | 30 | 43000 - 61000 |
| **Grotte du Lazaret** | 135 | 23 | 53000 - 175000 |
| **Grotte Vaufrey** | 74 | 35 | 110000 - 140000, 102000 - 116000 |
| **Grotte XVI** | 74 | 35 | 53600 - 74500 |
| **Havrincourt 1** | 93 | 88 | 61200 - 71500, 50800 - 65200 |
| **Igue des Rameaux** | 80 | 28 | 34800 - 53500, 34500 - 69200, 129300 - 158600, 75600 - 102200 |
| **Jupiter** | 47 | 21 | 45000 - 55000, 58000 - 78000, 90000 - 132000 |
| **Kents Cavern** | 27 | 91 | 72000 - 142000 |
| **La Butte d-Arvigny** | 88 | 73 | 97000 - 141000, 56000 - 116000, 59000 - 87000, 42000 - 94000, 75000 - 125000 |
| **La Chapelle aux Saints** | 79 | 36 | 44000 - 60000 |
| **La Ferrassie** | 71 | 36 | 56000 - 66000, 53000 - 57000, 41800 - 57800, 41800 - 59600, 58500 - 78800, 79500 - 100300 |
| **La Folie** | 66 | 53 | 55300 - 60100 |
| **La Quina** | 65 | 42 | 39400 - 46600, 40300 - 58000 |
| **La Roche a Pierrot** | 56 | 44 | 29700 - 52000, 33100 - 43600, 38100 - 46700 |
| **La Rochette** | 73 | 37 | 49100 - 55900 |
| **La Roquette II** | 101 | 26 | 52900 - 61500 |

| | | | |
|---|---|---|---|
| **Le Moustier** | 73 | 37 | 32400 - 45200, 35500 - 53900, 36300 - 53800, 37700 - 42900, 36500 - 45300, 36600 - 51000, 38000 - 45200, 37300 - 44500, 38700 - 53300, 38200 - 48200, 44800 - 55800, 50800 - 60800, 40700 - 49500 |
| **Le Prisse** | 47 | 21 | 78700 - 89900 |
| **Le Rescoundudou** | 88 | 31 | 62500 - 201000 |
| **Le Rozel** | 44 | 81 | 92000 - 112000, 102000 - 124000, 97000 - 117000, 104000 - 126000 |
| **Les Canalettes** | 94 | 26 | 60800 - 88400 |
| **Les Cottes** | 70 | 53 | 46100 - 55300 |
| **Les Forets** | 63 | 36 | 85300 - 103900 |
| **Les Pradelles** | 66 | 44 | 53000 - 62200 |
| **Lezetxiki** | 37 | 17 | 109000 - 153000 |
| **Mauquenchy** | 107 | 83 | 69800 - 84200, 75400 - 90600 |
| **Ormesson** | 87 | 74 | 46800 - 53000, 41400 - 48200 |
| **Payre** | 110 | 33 | 117000 - 147000 |
| **Pech de l-Aze I** | 75 | 35 | 24000 - 60000, 66000 - 196000 |
| **Pech de l-Aze II** | 75 | 35 | 60000 - 162000, 102000 - 158000, 120000 - 155000, 113000 - 161000, 37300 - 133000, 47800 - 63000, 51000 - 66300, 35900 - 101400, 40700 - 84700, 49900 - 85000, 63200 - 90400, 105000 - 133000 |
| **Pech de l-Aze IV** | 75 | 35 | 40000 - 59600, 41000 - 57000, 61900 - 78500, 61900 - 97800, 79600 - 118000, 46000 - 76000 |
| **Pie-Lombard** | 132 | 24 | 62300 - 77700 |
| **Pont-des-Planches** | 121 | 62 | 45000 - 60900, 41100 - 54800 |
| **Port Racine** | 43 | 82 | 62000 - 81000, 55000 - 88000 |
| **Roc de Marsal** | 72 | 36 | 45000 - 52200, 48900 - 57100, 46900 - 53900, 42600 - 60700, 52000 - 82000 |
| **Saint Germain-des-Vaux** | 43 | 84 | 55100 - 87600 |
| **Saint-Amand-les-Eaux** | 96 | 91 | 45860 - 52540 |
| **Sainte-Anne** | 101 | 37 | 96200 - 129700, 82700 - 97500, 80800 - 102300 |
| **Saint-Hilaire-sur-Helpe** | 101 | 88 | 89600 - 108200 |
| **Saint-Illiers-la-Ville** | 77 | 76 | 97200 - 112800 |
| **Savy** | 94 | 85 | 48000 - 54000 |
| **Scladina** | 112 | 91 | 110000 - 150000 |
| **Seclin** | 92 | 92 | 80000 - 106000 |
| **Sous les Vignes** | 71 | 31 | 40300 - 47800 |
| **Villiers-Adam** | 84 | 77 | 99000 - 121000 |
| **Walou Cave** | 119 | 92 | 78500 - 105800, 38000 - 57000 |

# PART THREE: SETTING THE STAGE

*"On two occasions I have been asked, - 'Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?' In one case a member of the Upper, and in the other a member of the Lower House put this question. I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question."*

(Babbage 1864)

# 7. OVERVIEW, DESIGN CONCEPTS AND DETAILS (ODD) OF HOMININSPACE

## 7.1 Introduction

This chapter provides elements of the design of the HomininSpace modelling system and follows the Overview, Design concepts and Details protocol (ODD) (Grimm *et al.* 2006). 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 (OpenABM, Janssen *et al.* 2008) includes the ODD protocol as a requirement in their Model Review process[28]. HomininSpace was documented in such a manner that it passed this review process allowing incorporation of HomininSpace into the OpenABM repository allow open access to code and facilitating replication and re-use by other researchers.

This chapter includes several sections selectively adapted from the HomininSpace ODD document, which is in original form included in the Supplementary Materials. The ODD protocol prescribes several elements including purpose of the model, the process overview and scheduling (Section 7.2), the simulation period and stochasticity in the model (Section 7.3), and the model parameters (Section 7.4). The simulation results will be compared against archaeological data to assess the validity of the implemented model (Section 7.5). To enable autonomous modification of model parameter values a Genetic Algorithm has been implemented (Section 7.6 and Section 7.7). These last two sections are largely based on a previously published paper (Scherjon 2016). The model with documentation and simulation results is available online from the OpenABM website[29].

## 7.2 Process overview and scheduling

A simulation in HomininSpace always starts with an initialization phase (detailed in the flowchart in Figure 25). The system is initialized by reading configuration and data files, either from file or as entered by the user in the User Interface (UI). The configuration consists of a simulation number, a random seed, a set of parameter values and the model settings that define what model elements participate in this simulation. Data files contain

[28] See https://www.openabm.org/page/model-review-overview, accessed 27 October 2014.
[29] See https://www.openabm.org/model/5294/version/1/view, accessed December 2017.

grid data, climate data for modern day and past environments, and archaeological data. The simulation grid is created and filled with the environment information. Factories, single grid cells that represent population core areas where new hominin groups can be produced, are created if included in the model. Data for the reconstructed global sea levels and yearly mean temperatures for the simulation period are read from file and stored in memory.

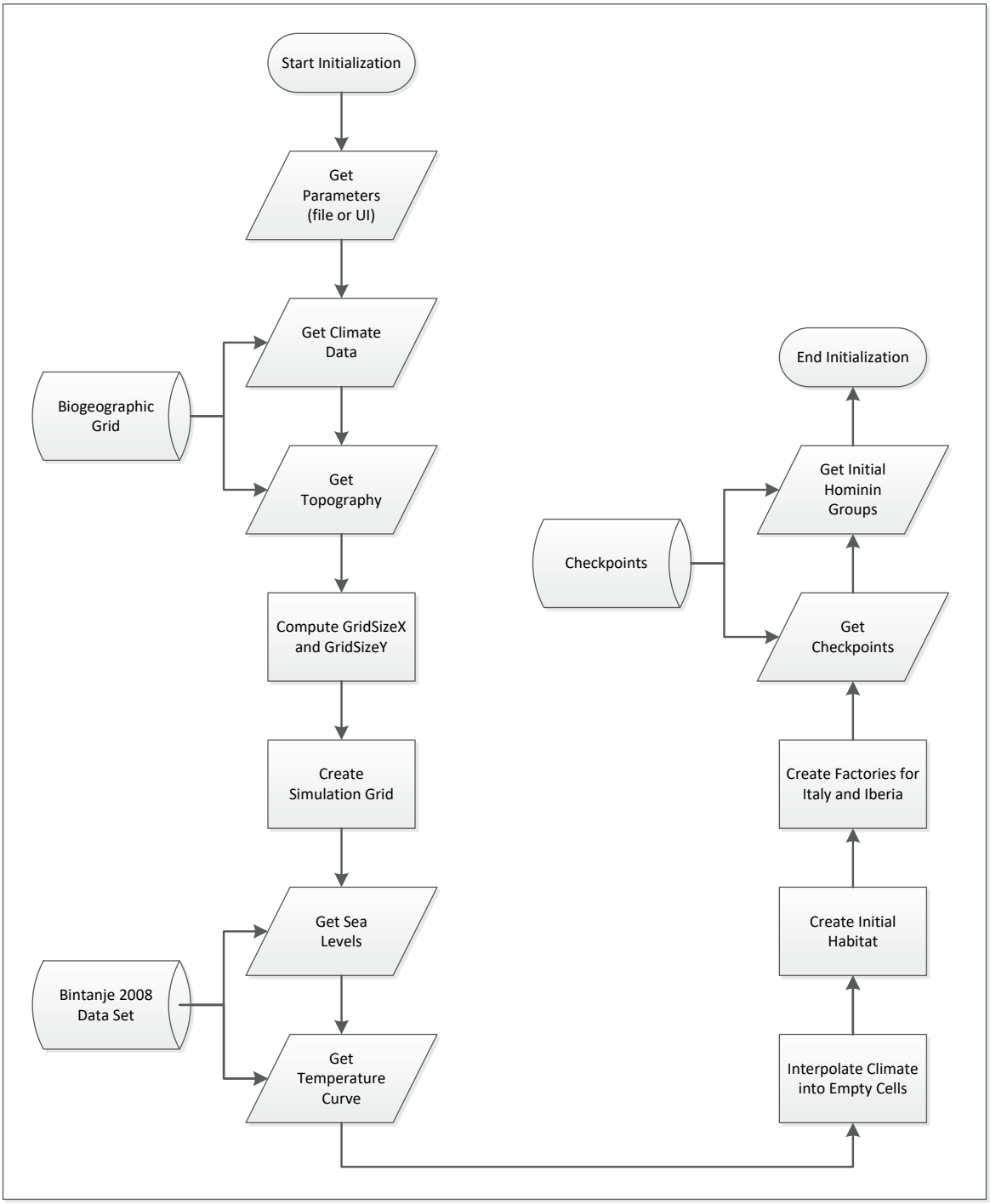



**Figure 25: Flowchart presenting the Initialization process in HomininSpace.**

After initialization of the system the main simulation run routine is executed. Within this loop the simulation executes the model, time step after time step. The sequence of events

for a time step is given in Figure 26. The first action in each time step is updating the environment with new climatic data and topographical information (incorporating sea level changes). Ultimately, the edible biomass is calculated for each grid cell.

After updating all grid cells each hominin group that can move will scan the environment delimited by two times the foraging range (the radius leap-frog pattern as described by Binford (1982)). The location which is most attractive to the group because it has the highest energy levels becomes the destination for this group. Upon moving to this new location the foraging range for this group is recalculated, and presence is registered with the grid cells and resources are claimed. The groups will then consume the resources they require (if available) and the hominins in the group will either produce offspring or they will die, depending on the obtained resources.

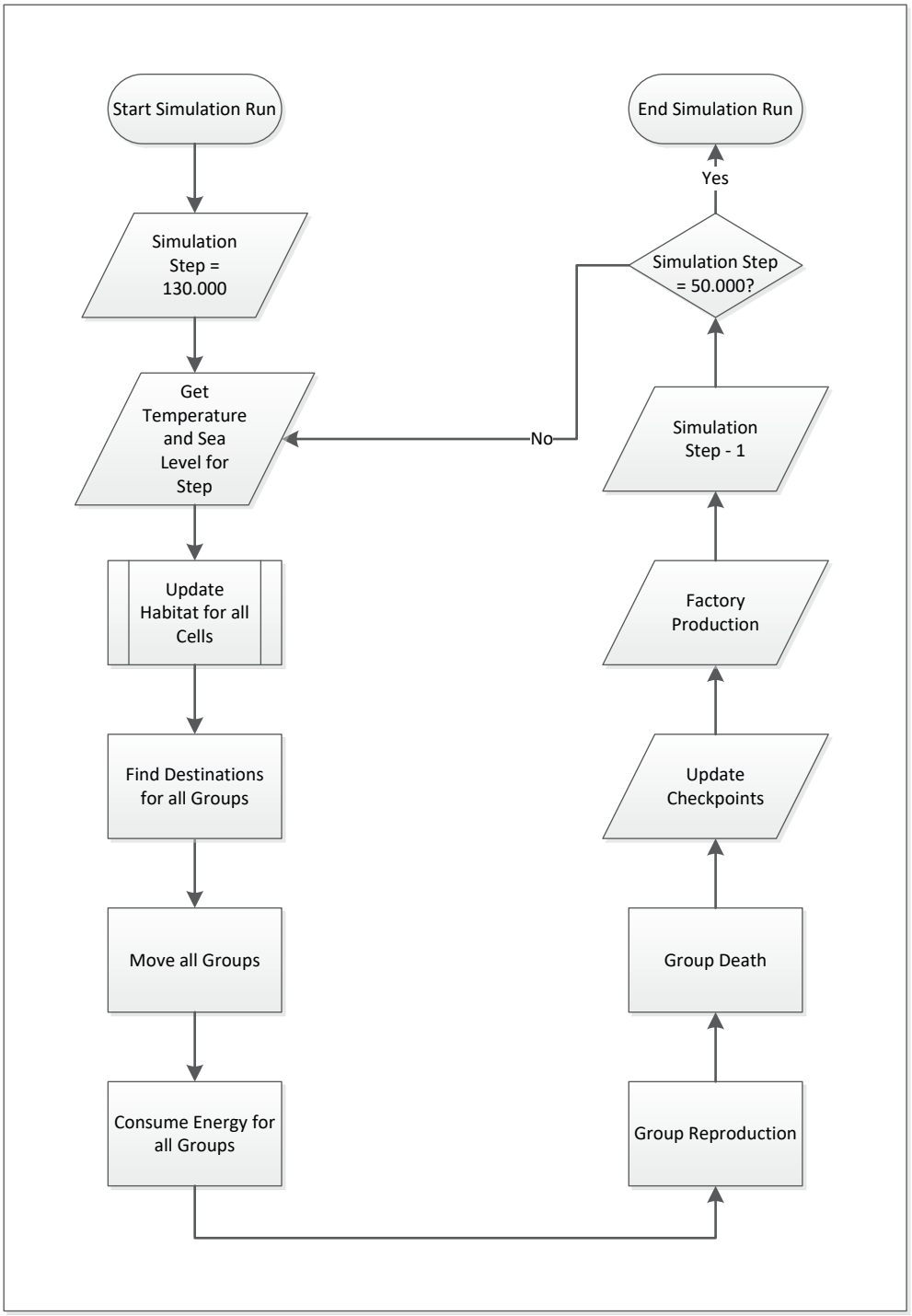



**Figure 26: Flowchart of the Simulation process.**

## 7.3 Simulation time and stochasticity

Each simulation time step is defined to be one year. One year is the most natural complete set of environmental events, including a full seasonal cycle that includes summer and winter and dry and wet periods. It is also a unit that is readily recognized in the geological and biological records allowing reconstructions of environmental variables to some extent. Each simulation has a starting time expressed in years ago and a number of time steps (years) to execute. In the experiments performed for this research all simulations start at

131 ka and end at 50 ka. That means that every simulation run has 81,000 time steps. In each time step the following actions are executed in this order:

1. Increment timer, get new global sea level and global mean temperature;
2. Update all habitat cells by recalculating available energy;
3. Move all hominin groups and update foraging ranges;
4. For all groups: consume required energy, grow if there is sufficient energy, check for merge and split possibilities, perform reproduction and then check for death conditions;
5. Check production conditions for each factory (if present). If favourable, produce one new group per factory.

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. Note that energy levels are computed in integers. This added random value prevents groups always selecting for instance the left path when presented equal route choices. The random seed, an integer value driving the random number generator, is included as an identifying element for the simulations. Each time a simulation is run with the same parameter values and the same random seed value, identical results are produced.

Parameter values for the Standard parameter set were generated randomly, specifically without bias (that means no normal distribution for any of the values). Selection of the parameter values for improvement in the subsequent parameter value space exploration is random, but organized in tournaments (see Section 7.6) to reduce the random character in the final distribution of parameter values.

## 7.4 Model parameters and the `ViabilityIndex`

In the preceding chapters the building blocks for the underlying model in HomininSpace have been identified. Some of the variables of the model have been marked as parameter, indicating that these can be changed. Table 13 lists and describes all 16 available parameters. There are numerous other model elements but these cannot be simply altered by the user, since they are defined as constants or otherwise implemented in the source

code. Parameters were selected as candidates for variation because they are disputed in the literature (birth rates, energy requirements, temperature tolerances, foraging range), because they are directly related to other modifiable parameters (mortality rates) or because they were found important in agent dynamics and interaction (cohort sizes, number of years before group maturity). A simulation is uniquely defined by the combination of the parameter values and the seed value for the random number generator.

**Table 13: The 16 parameters that can be changed in the HominінSpace model.**

| *Name* | *Description* |
|---|---|
| ***1. BirthRate*** | The number of females that conceive this year (a value of 33% means one child per three females, or one child every three years for a female). |
| ***2. DeathRate_PreFertileCohort*** | Death rate for the pre-fertile segment per group, the percentage that does not survive. |
| ***3. DeathRate_FertileCohort*** | Death rate for the fertile segment per group, the percentage that does not survive. |
| ***4. DeathRate_PostFertileCohort*** | Death rate for the post-fertile segment, the percentage that does not survive. |
| ***5. Subsistence_PreFertileCohort*** | Energy needs for an individual in the pre-fertile segment of a group in kcal. |
| ***6. Subsistence_FertileCohort*** | Energy needs for an individual in the fertile segment of a group in kcal. |
| ***7. Subsistence_PostFertileCohort*** | Energy needs for an individual in the post-fertile segment of a group in kcal. |
| ***8. Years_Before_Group_Maturity*** | Period in years before a newly created group can interact with other groups (join) or settle. |
| ***9. GroupSize_BeforeMerge*** | Groups can merge with other groups if their total size is smaller than this value. |
| ***10. GroupSizeFertile_BeforeMerge*** | Groups can merge with other groups if their size of the fertile segment is smaller than this value. |
| ***11. GroupSize_BeforeSplit*** | A group can procreate by splitting in two, thereby generating a new group. This value is the minimum size before a group can split. |
| ***12. Temperature_Tolerance*** | Minimum temperature that can be sustained by the hominins. If the coldest temp in a year falls below this value the group of hominins dies. |
| ***13. CohortSize_PreFertile*** | The size of the pre-fertile segment in a group, with age boundaries in years. |
| ***14. CohortSize_Fertile*** | The size of the fertile segment in a group in years[30]. |
| ***15. Calories_Per_Kg_Meat*** | The number of kilocalories that can be extracted from one kilogram of meat. |
| ***16. Max_ForagingRange*** | Maximum annual foraging range from the current location. Groups cannot forage outside this range (in grid cells). |

The parameters are used in a demographic model that computes new population numbers based on the given values. The `ViabilityIndex` is the theoretical number of individuals remaining when a default hominin group (size 25) is followed for 100 years, given birth- and death rate values. For instance, for parameter value set 5 (see Table 21)

[30] Note that the length of the post-fertile segment is not limited, but chances of survival decrease progressively with age.

the `ViabilityIndex` = 548, meaning that the population of 25 individuals has grown to 548 after 100 years. The algorithm to calculate this number is given in Figure 27. The `ViabilityIndex` is used in the analysis of the simulation results and gives a rough indication of the viability of a given population without taking resource availability into account. It is passed as a parameter for presentation purposes only and not included in the model.

```
# calculate viability index by simulating 100 years and counting survivors
  sizePreFertile = 8
  sizeFertile = 12
  sizePostFertile = 5

  # check if viable population by running for 100 years, assuming sufficient
    resources
  for (year in 1:100 ) {

    # size of age category
    becomingFertile = sizePreFertile / sSizePreFertileSegment
    # size of fertile age category
    becomingPostFertile =  sizeFertile / sSizeFertileSegment;
    newBorns =  (sizeFertile / 2.0) * (sBirthRate / 100.0);

    dPre = sizePreFertile * (sDeathRate_PreFertile/100.0);
    dFertile = sizeFertile * (sDeathRate_Fertile/100.0);
    dPost = sizePostFertile * (sDeathRate_PostFertile/100.0);

    sizePreFertile =  sizePreFertile + newBorns;
    sizePreFertile = sizePreFertile  - becomingFertile;
    sizeFertile = sizeFertile  + becomingFertile;
    sizeFertile = sizeFertile  - becomingPostFertile;
    sizePostFertile = sizePostFertile  + becomingPostFertile;
    sizePreFertile = sizePreFertile  - dPre;
    sizeFertile = sizeFertile  - dFertile;
    sizePostFertile = sizePostFertile  - dPost;

  }

  totalPopulationSize = sizePreFertile + sizeFertile + sizePostFertile
  cat (",")
  # there is always a fraction left, even if unviable population
  cat (floor(totalPopulationSize))
```

**Figure 27: Calculating the ViabilityIndex for a group with 25 individuals. Code fragment is taken from the "CreateHSParameters.R" source file.**

## 7.5 How to match the archaeology

HomininSpace uses an extensive collection of archaeological presence data: the CSTs with associated intervals. Simulations are scored against this archaeological data set. When comparing simulation results against the archaeological attested presence it is essential to assume that this record is representative for the actual presence in the past (Wheatley 2004). When using presence only data those areas where nothing has been found (or where nobody has searched) cannot influence the results. The effect of using presence-only data is clearly illustrated with the theoretical example of the omni-present hominin that will

automatically be the best matching Neanderthal. To avoid this one can introduce *absence areas*. These are periods in time and space where the model states that hominins were *not* present, a fact that can be used in model validation (Jiménez-Valverde and Lobo 2007) (see Subsection 9.4.1). Unfortunately, only limited absence data on Neanderthals is available. On the other hand, when the simulation results match the archaeology, it fits the data regardless of how this data was collected. It then becomes important to define what constitutes a 'fit' with the archaeological data.

### 7.5.1 Translating simulation results into numbers: `MatchedIntervalCoverage`

This study assumes that an archaeological site that presents evidence for the presence of Neanderthals (for instance stone tools) is prove of the past presence of at least one Neanderthal group (of whatever size). Since a single stone knapping individual can produce hundreds of stone tools, it is very difficult to assess the number of presence events that created any archaeological assembly (see for instance the supplementary information in Kolodny and Feldman (2017)). In HomininSpace a database is constructed of presence intervals constructed using radiometric dates with one standard deviation (SD). Each interval is from a date minus one SD to that date plus one SD. If intervals overlap, or if two intervals are from the same layer they are combined into one interval using the minimum and maximum values. Each checkpoint has one or more intervals.

During the simulations, modelled presence of agents is compared against these intervals to compute the score or *precision* of each simulation (Crooks *et al.* 2015). If an agent visits a checkpoint (either by foraging in the same grid cell or by making that grid cell its home range) the simulation time step is compared against the interval(s) for that location. If that specific time step falls inside an interval, one point is added to the score for that interval. This is done for each time step that any group visits the checkpoint. A time step can be scored only once (if multiple groups forage in the same grid cell, this still count as one archaeology producing visit). If the time step does not fall inside an interval this is also registered, but it does not add to the simulation score.

The score for a single interval is the number of visits divided by the length of the interval. Thus the maximum score for each interval is 1, when visited by as many visits as the length of the interval in years. HomininSpace thus assumes that more visits to the same location in a single interval make for a better match. HomininSpace does not interpret the interval structure: if two different archaeological layers have overlapping dating intervals, there are two intervals to be matched (if these two intervals are however in one layer they

are combined into one interval). The scoring procedure is illustrated in Figure 28 and Figure 29. With a more precise date (smaller sigma) the maximum score is easier to achieve.

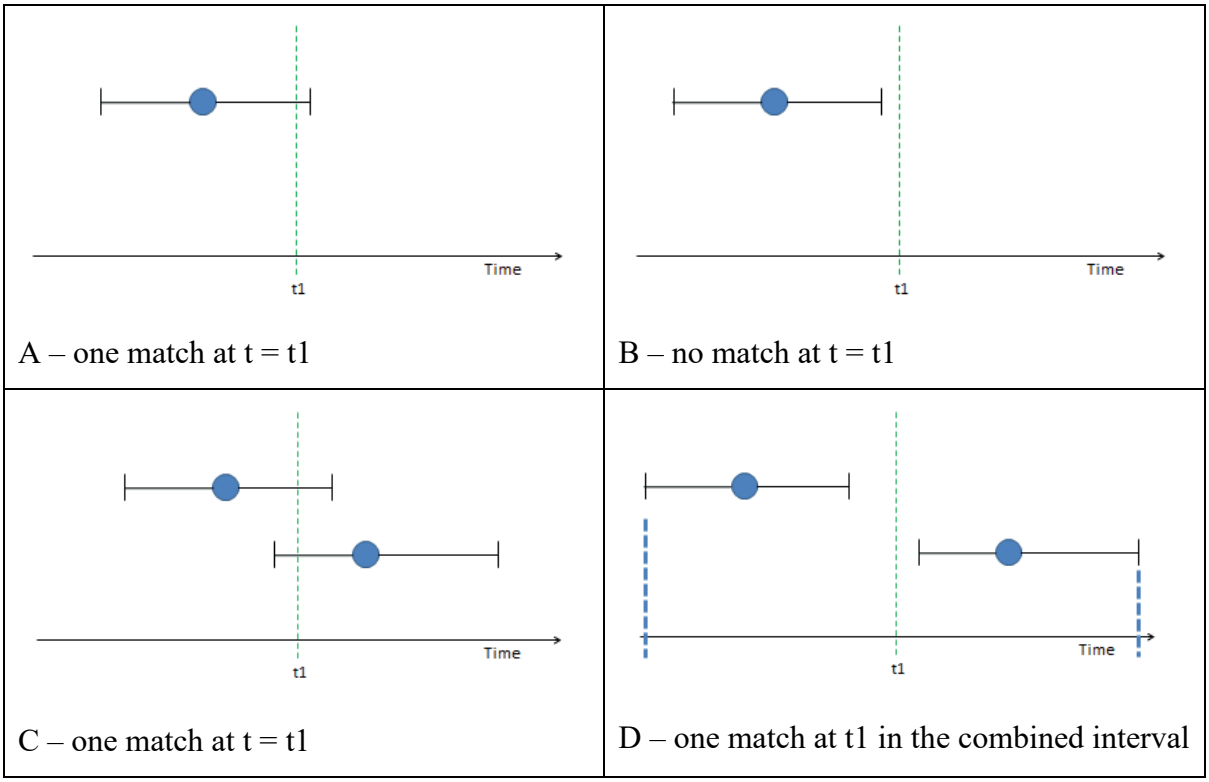



**Figure 28: Counting matching visits. A visit at simulation time *t1* matches a dated interval for a single archaeological layer (A, C and D). Only in B there is no match.**

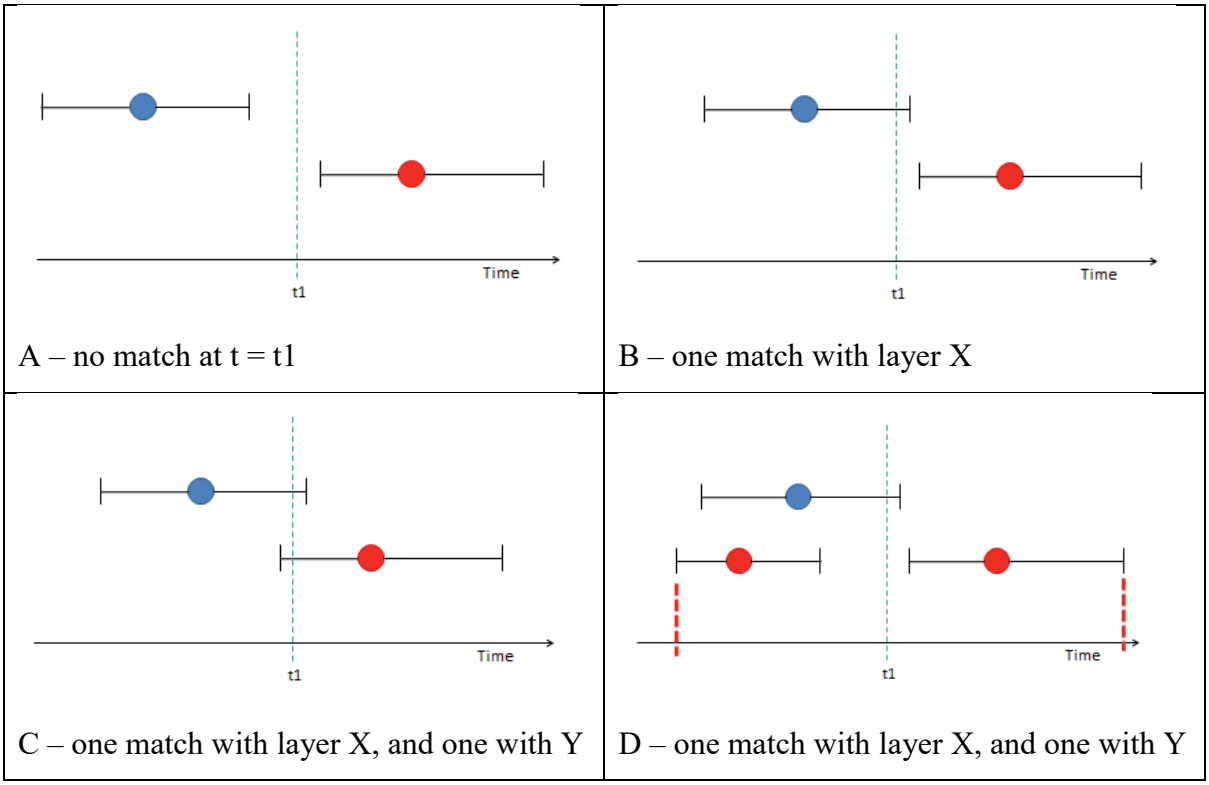



**Figure 29: Counting matching visits. A visit at time *t1* can match two dated intervals that represent two archaeological layers (blue dots for layer X, red for layer Y).**

In HominінSpace a simulated visit at a checkpoint in a time step does not mean hominins were there only one time. It means that within one year a group of hominins visited the site at least once and could have left archaeology to be dated as part of the archaeological record of the site. Each time step in the simulation represents one year, and thus a single group can visit a checkpoint only once per time step. Other groups that live in the area can visit the checkpoint in the same time step, but only one visit per time step is actually counted. A checkpoint can be re-visited in next time steps. It is assumed that one or more visits can create the archaeological layer as it is identified by the archaeologists, but that more visits give a better chance of producing the archaeology. That means that a higher score implies a better model, or more plausible parameter values.

The fitness function in the Genetic Algorithm (the next subsection) that finds new parameter value combinations requires a single number that indicates how successful a simulation is: the *simulation score*. This number is used to compare one simulation against another. A larger number suggest a better match with the archaeology and thus a better model. To compute the total simulation score an interval score is calculated for each interval. This is the actual number of visits divided by the length of the interval. The division by the length ensures that intervals that are more accurate and thus shorter have more weight in the total simulation score. The simulation score is calculated by adding all interval scores together for all checkpoints (Equation 9). This number is referred to as `MatchedIntervalCoverage`. The maximum possible value or theoretical maximum of this variable in the configuration as used for the simulations in this research is a score of 39,200. In this thesis any general reference to simulation score or simulation results refers to this number.

$$\boldsymbol{MatchedIntervalCoverage} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{CSTs}} \sum_{\boldsymbol{n=0}}^{\boldsymbol{intervals}} \frac{\boldsymbol{visitCount}}{\boldsymbol{length}}$$

**Equation 9: Combining interval scores for all checkpoints to compute the simulation score.**

## 7.6 Genetic Algorithms - searching through N-dimensions

Optimization of agent-based models is a complex problem. Computation of a solution for the model involves the execution of a simulation in all but the most trivial cases. The search for optimal solutions is therefore computationally expensive, especially with an extensive environment and when the number of parameter combinations is large. An

exhaustive algorithm like a Monte-Carlo search implementation would require too many computational resources to efficiently find optimal solutions. Instead this research applies Genetic Algorithm (GA) based search and optimization techniques in an attempt to characterize Neanderthal models that match the archaeology well while exploring the parameter space (Grimm and Railsback 2005; Revay and Cioffi-Revilla 2018).

In a typical GA, a population of potential solutions is evolved in the solution space towards higher values of a fitness function (Forrest 1993). Genetic Algorithms form, together with Evolutionary Strategies and Genetic Programming, the Evolutionary Algorithm programming paradigm (Coello Coello 2002). Differences between these techniques are fading (Michalewicz 1992, 132), and while the chosen terminology no longer reflects our current understanding of genetics, they all aim to simulate an evolutionary process in the computer (Coello Coello 2002). These are nature-inspired numerical optimization techniques that operate in a virtual laboratory. Basic elements in evolutionary algorithms are a population of individuals to work with, a string with values which can be manipulated that define individuals (referred to as the genes or the chromosome of that individual), a fitness function that calculates how well adapted an individual is within the modelled environment (Michalewicz 1992), and an optimization technique targeting optimal solutions by modifying individuals. The values for model parameters are adjusted until experimental optima are reached with no substantial improvement in the simulation results for any newly generated models (Calvez and Hutzler 2006). Compared to random search methods GA implement a stochastic search with directed randomness.

The term Genetic Algorithm was coined by John H. Holland (1975) but the general principle was already recognized by Alan Mathison Turing in the 1948 essay 'Intelligent Machinery': '*There is the genetical or evolutionary search by which a combination of genes is looked for, the criterion being the survival value'* (Turing 1948, 16). The idea is that an individual represents a single solution to the problem at hand; that individuals vary from one to the other; that an evaluation function represents the environment in calculating the fitness of each individual; and that the computer searches for the individual with the highest fitness value (the optimal solution) in the problem space, without actually programming that solution, but instead by manipulating the genes of the individuals mimicking the processes of variation and natural selection (Coello Coello 2002). The technique is often referred to as meta-heuristic due to its domain independence.

GAs are very well suited for exploring nonlinear search spaces, with potentially multiple locations present that can yield good solutions, for instance in social agent-based modelling studies (Revay and Cioffi-Revilla 2018). A non-traditional methodology is required when traversing the larger resulting parameter spaces searching for optimal solutions, where exhaustive search algorithms are replaced by heuristic procedures. A GA differs from more traditional optimization and search algorithms in at least four ways (Goldberg 1989):

- A GA searches in a population of points, not just in a single point;
- GAs operate with probabilistic, unbiased transition rules, not deterministic ones;
- A GA operates on a set of parameter values, not individual parameters themselves, and does not require any explicit knowledge of the actual structure of the solution space (what these parameters are about);
- A direct, explicit fitness function is used.

In the structure of a GA, a probabilistic selection is made from the population based on some measurement of the individual's adaptation to the environment or its fitness within it. The design of the fitness function is at the discretion of the modeler. After the selection of one or more individuals, operators are applied on the selection. These operators are inspired by the early perception of how the genomes of offspring were created in nature. They include Mutation, where random changes are made to the genes of an individual, and Crossover (a special kind of Recombination), where the genes of two individuals are mixed in sequence. Then the fitness function is calculated for the generated offspring, and the new individuals are inserted into the population, generally replacing other individuals. The process is then restarted, until some stop criteria are met. Generally the stop criterion is a certain value for the fitness function being attained, or no measurable improvement of the calculated fitness values after a certain number of generations.

An Agent-based Model (ABM) is a complex system aiming to reproduce the dynamics of the real world, and generally one that cannot be solved mathematically. Therefore GAs form a perfect exploration and tuning method for such models (Calvez and Hutzler 2006). Emerging properties, high parameter sensitivity, and non-linear solutions often characterize the solution space for such models. There is no analytical description, and simulation results cannot easily be reduced to the input data. There is development and emergence in a non-trivial way. As such they are well suited to modelling complex systems and to explore this complexity. Archaeological research questions connected to such models often incorporate the resultant output of the (inter)actions of many individuals

through time, and as such can be targeted by ABM techniques (Lake 2014). However, there have been few attempts to apply GAs to archaeological simulations, although they are implicitly used in the more widely applied Genetic Algorithm for Rule-Set Prediction tool (GARP) for biogeographical niche modelling (Banks, d'Errico, Peterson, Vanhaeren, *et al.* 2008).

Every simulation in HomininSpace is characterized by a combination of parameter values and the random seed. This combination of parameter values is referred to as the *chromosome* for that individual simulation. In HomininSpace, a GA is then used to search for the parameter value combination that when simulated optimally matches the archaeology. First, an initial solution population is constructed by varying parameter values randomly, and then running all the simulations with these parameter sets. Subsequently, simulation results are used to select promising parameter combinations, generate new individuals (new parameter combinations), and running simulations for that evolved offspring in a search for a better solution.

## 7.7 Traversing the parameter space

Values chosen for the parameters will decide how well the simulation matches the archaeological presence data, and thus score the simulation results into the `matchedIntervalCoverage` variable. Maximizing this variable is one of the aims of this study. All the value combinations for all parameters form a 16-dimensional parameter space, with each parameter adding one dimension in which variation is possible. The search traverses this parameter space. Previous research has shown that human insight and analysis falls short in identifying the best matches, and some form of machine learning is needed where no explicit instructions nor biases steer the system (Scherjon 2015a).

### 7.7.1 Generating the Standard parameter value set

The *Standard* parameter value set is a collection of 1500 generated parameter value combinations which are used as the starting point for every exploration. Simulations for these 1500 parameter combinations are executed first, after which a genetic algorithm will attempt to optimize individuals from this set. The parameter values for the Standard set have been generated randomly, with values taken indiscriminately from an interval defined by minimum and maximum values (see Table 14). I specifically use a random probability distribution for each parameter (Iman and Conover 1980) and do not construct a (optimized) Latin Hypercube sample with equal distances or attempt any other structuring

of the search space (Bates *et al.* 2003). Instead I use the genetic algorithm for the optimization of the sampling (Liefvendahl and Stocki 2006).

Each value from this interval has an equal chance of being selected (specifically *not* implementing a normal distribution of values for parameters like birth or mortality rates). The generation of this Standard set is done with an R script named "CreateHSParameters.R" and is referred to as spanning the parameter space. The minimum and maximum possible values are constructed as a wide interval around the default value, which has been taken from the literature (Scherjon 2015a). This avoids the (researcher) bias in selecting values for these parameters. For example, 1500 times a random value is selected for the BirthRate parameter from the interval [1..50], one for each parameter value combination in the Standard set.

**Table 14: Minimum, maximum, step size, and default values for the Standard parameter set.**

| *Name* | *Minimum* | *Maximum* | *Step* | *Default* |
|---|---|---|---|---|
| ***1. BirthRate*** | 1 | 50 | 1 | 33 |
| ***2. DeathRate_PreFertileCohort*** | 1 | 15 | 1 | 4 |
| ***3. DeathRate_FertileCohort*** | 1 | 15 | 1 | 2 |
| ***4. DeathRate_PostFertileCohort*** | 1 | 15 | 1 | 8 |
| ***5. Subsistence_PreFertileCohort*** | 2000 | 5000 | 250 | 3000 |
| ***6. Subsistence_FertileCohort*** | 2000 | 5000 | 250 | 4000 |
| ***7. Subsistence_PostFertileCohort*** | 2000 | 5000 | 250 | 3500 |
| ***8. Years_Before_Group_Maturity*** | 1 | 10 | 1 | 5 |
| ***9. GroupSize_BeforeMerge*** | 1 | 5 | 1 | - |
| ***10. GroupSizeFertile_BeforeMerge*** | 1 | 5 | 1 | - |
| ***11. GroupSize_BeforeSplit*** | 6 | 150 | 10 | - |
| ***12. Temperature_Tolerance*** | -30 | -5 | 1 | -18 |
| ***13. CohortSize_PreFertile*** | 8 | 20 | 1 | 15 |
| ***14. CohortSize_Fertile*** | 15 | 40 | 1 | 25 |
| ***15. Calories_Per_Kg_Meat*** | 2500 | 3500 | 50 | 3000 |
| ***16. Max_ForagingRange*** | 1 | 15 | 1 | 15 |

### 7.7.2 Searching the highest scoring parameter value set with Genetic Algorithms

The parameters that are used in the model underlying the HominínSpace simulation system are described in Table 13. Parameter values are based on ethnographic data with wide plausible extended value ranges from which values are randomly selected. Systematical exploration of the total parameter space employing these 16 parameters would require an exponential number of simulations. For example, if one takes only three values per

parameter (a minimum, a maximum and a value in between), this would require the simulation of a 3^16 parameter value sets (or almost 300 million unique simulations). Therefore the combinatorial explosion with so many parameters requires a non-exhaustive exploration. Nevertheless, due to the non-linear character of the solution space, a systematic exploration of the parameter space is needed (Calvez and Hutzler 2006). A Genetic Algorithm (GA) implements such a search.

The implemented GA in HomininSpace selects individuals from the population of solutions using a tournament selection procedure (Miller and Goldberg 1995). For the selection of each individual, a selection round or tournament is organized in which *n* individuals are randomly chosen from the total population. From this subset the highest scoring (most optimal) solution is declared winner. This procedure ensures that even mediocre solutions can (occasionally) produce offspring and prevents swamping of the population by the offspring of a few extraordinarily good performing individuals.

The selected individuals participate in the creation of the next generation new individuals through the application of several operators. HomininSpace implements a integer-coded GA where each individual is represented by a string of 16 integer values (the chromosome), one integer for each model parameter (Herrera *et al.* 1998). The operators that are applied are *mutation* and *crossover*. Then additional simulations for the newly generated parameter combinations are executed, the fitness value computed, and the offspring is added to the general population. This process is repeated until the end of each experiment (Figure 30).

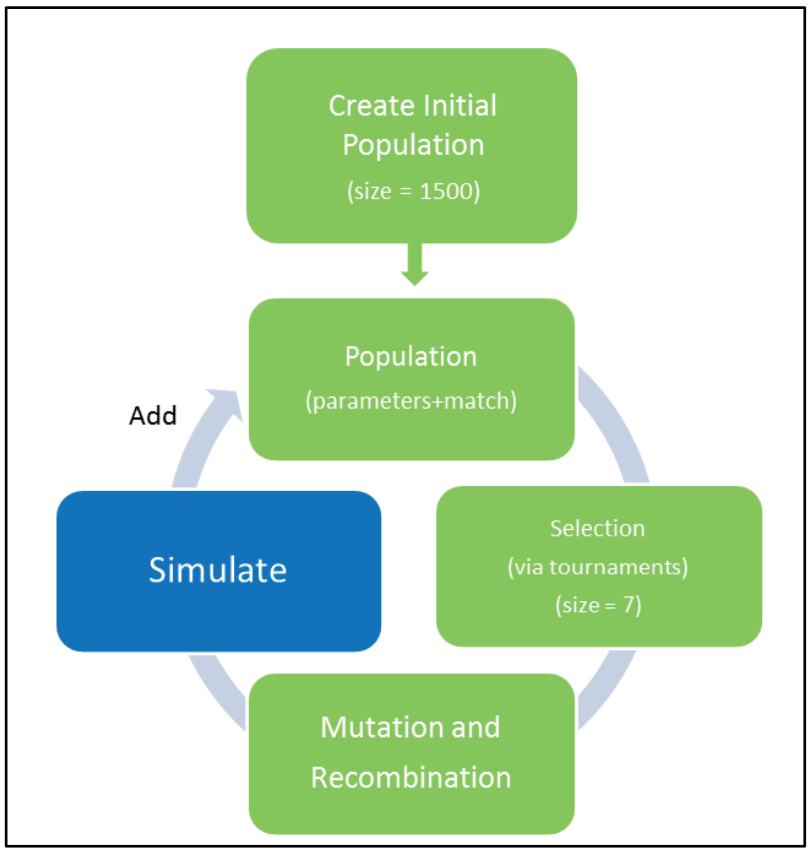



**Figure 30: Functional representation of the implemented GA.**

The mutation operator is implemented as the random modification of one single parameter value by plus or minus 10%. This implements a low rate mutation mechanism to increase coverage of the search space and to prevent convergence to a local optimum (Yao 1993). An additional advantage of the chosen mutation operator is that when needed, the search space can be expanded even beyond the original chosen random value domain borders. In other words, values can be created that are not present in the (initial) population (Aleksandra *et al.* 1997). Crossover aims to recombine two good parent solutions into potentially even better offspring solutions. Implemented is a uniform multi-point crossover, where the offspring is a single individual that is a stochastic mix of the parameter values from both parents (Syswerda 1989). This combines very well with tournament selection (Aleksandra *et al.* 1997, 7860). The algorithm ensures that both parents contribute the same amount of genetic data. Seven individuals (four crossover results, three point mutations) are created in each generation, because the hardware running the experiments has seven parallel processors available for executing simulations.

For HominínSpace, the generation size has been defined as seven new individuals for the very practical reason that in each computer that participated in the experiments seven

computing cores were available for execution of the simulations. Thus seven simulations were run simultaneously. Since offspring of each generation is added to the population without replacement of parents or other less performing individuals, no lineage is terminated prematurely, and each individual competes until the very end of each experiment. This ensures that the best performing individuals survive intact through each generation. However, successful lineages still might tend to dominate the tournaments, swamping more successful but unfortunate individuals. This must be taken into account when reporting successful individuals and when defining the stop criterion of the algorithm.

Since the GA methodology cannot guarantee that the optimum solution has been found, the algorithm must have a stop criterion that allows a sufficient number of iterations before halting the search. Since swamping by individual lineages does not always occur (especially when many similar scoring individuals are present in the solution base) the stop criterion for HomininSpace has been defined as follows:

- A minimum of 500 new individuals must have been generated, that is more than one third of the original population size;
- If more than 100 new individuals have been generated without any improvement in the simulation score the search ends.

# 8. BUILDING THE SIMULATION SYSTEM

## 8.1 Introduction

The core of the HomininSpace modelling and simulation system is an agent-based model. To turn that into a running simulation the model has to be programmed and executed in a computer simulation environment. Many such environments exist, and Repast Simphony has been selected for implementation of the model (Section 8.2). A development environment generally consists of more than just a compiler for the source code in which the model is implemented, and can include additional third party tools as well (for an overview see Section 8.3). The model is then executed within the structure of the HomininSpace simulation system (Section 8.4), using the input files that define the case study with topography, climate data, information about the archaeology, and the model (Section 8.5 and 8.6) and creating output data during the simulations (Section 8.7).

Parts of this chapter have been extracted from the "HomininSpace – user manual" document that is included in the Supplementary Materials which includes further details, formal descriptions and data file examples. HomininSpace, including source code, data files, and examples was accepted into the OpenABM model repository (Janssen *et al.* 2008). After an anonymous peer review it was included in the Computational Model Library of the Computational Modeling in Social and Ecological Sciences Network (CoMSES Net) under model number 5294[31].

## 8.2 Agent-based modelling tools

Many agent-based modelling toolkits for building simulations are available (e.g., Abar *et al.* 2017; Gilbert and Bankes 2002; Nikolai and Madey 2009; North *et al.* 2006; North *et al.* 2013; Railsback *et al.* 2006; Tobias and Hofmann 2004). Nikolai and Madey (2009) alone review 53 toolkits and developing environments. Out of the existing tool sets REcursive Porous Agent Simulation Toolkit (Repast Simphony or in short Repast) was selected for implementing HomininSpace (Crooks 2007; Macal and North 2009). Repast is a turn based simulation system for developing and executing agent-based models (Tatara *et al.* 2006). It is freely available as open source software (North *et al.* 2013).

[31] https://www.comses.net/codebases/5294/releases/1.0.0/, accessed 8 March 2018·

Repast has been selected for the Google Summer of Code 2011 project, ensuring publicity, additional example programs and a larger user base (see http://www.cscs.umich.edu/~rlr/SoC/index.php?n=Main.SoC2011, accessed 28 April 2011, or http://www.google-melange.com/gsoc/homepage/google/gsoc2011, accessed 22 March 2013). Repast has for instance been used to analyse urban segregation and urban growth (Mathur 2007), to model crime and identify burglary risks (Malleson 2006), to model pedestrians and evacuation routes (Castle and Crooks 2006; Usman *et al.* 2017), to model the spread of diseases (Pais *et al.* 2017), to build FamilyNet2 (White 2013) which is part of the development of the ForagerNet2 model, a complex model that encompasses representations of space, social learning, and social networks (White 2012), the modelling of noise in the genes of growing bacteria populations (Stiegelmeyer and Giddings 2013), and for large scale emergency response simulations (Hawe *et al.* 2012).

Repast was selected to implement HomininSpace because it was positively reviewed, has an active user base, has good quality documentation, there are many example programs, it is described and used in the literature, and has built-in support for many standard libraries. The model can be programmed in Java, a multi-purpose and platform independent programming language and runtime environment that this author is proficient in. In addition, the predecessor of HomininSpace, the agent based modelling system SteppingIn (Scherjon 2012) was successfully implemented by this author using a prior version of Repast.

### 8.2.1 Repast Simphony version 2.2

Repast Simphony consists of two major components: a development and a runtime or execution environment. The runtime part will execute model runs and visualize the simulation results. The development module is used to maintain the source code of the simulation and to manage the program structure. Figure 31 identifies the main elements of the user interface. The Repast environment includes point-and-click model configuration and operation (execution, enquiries), integrated two or three dimensional GIS and other model views, connection to standard data sources including database support, and automated connections to external programs for statistical analysis and model visualisations (Tatara *et al.* 2006).

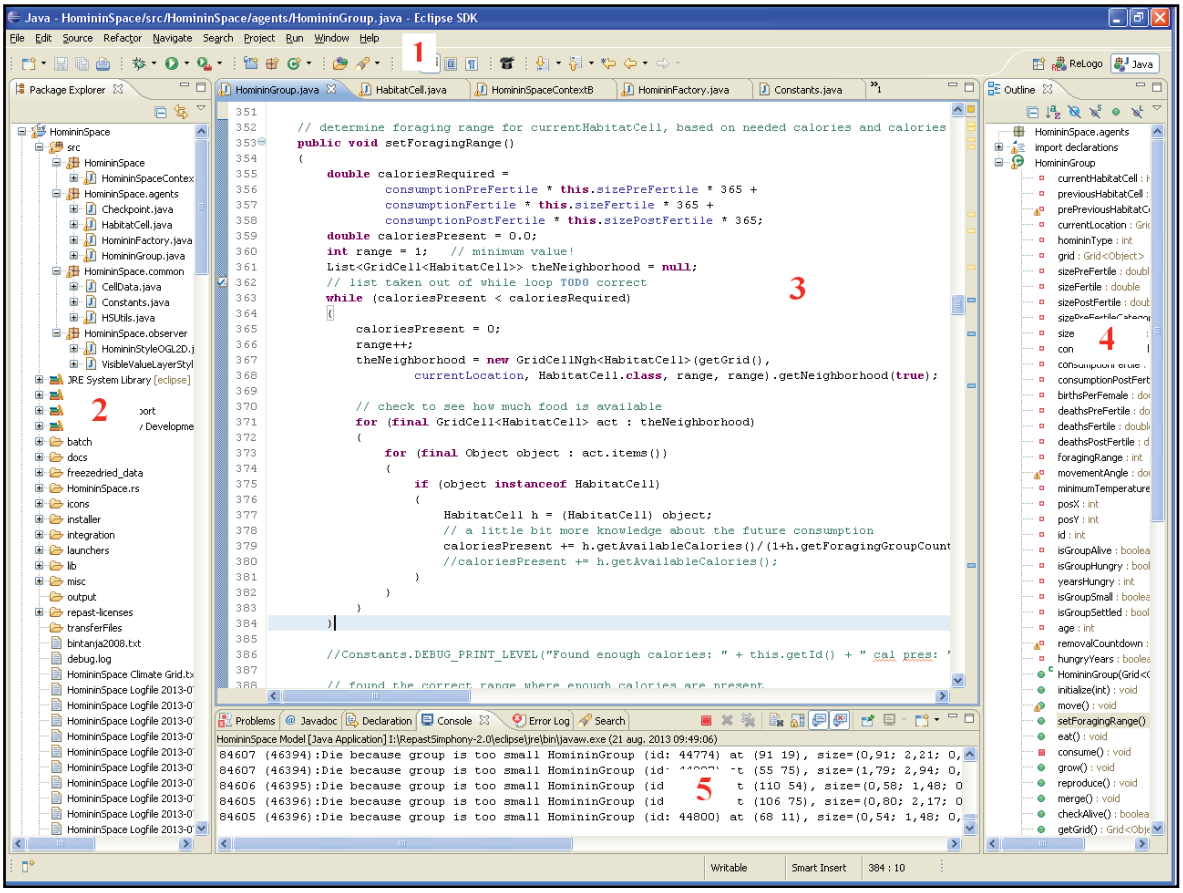



**Figure 31: The Repast Simphony development environment. 1 = main menu; 2 = project directory structure; 3 = source editor; 4 = file structure; 5 = log information window.**

Software development in Repast Simphony can be done in the Object Oriented Programming (OOP) language Java (Gosling *et al.* 2005). The advantage of OOP in agent based modelling is that objects or classes in the programming language naturally translate into agents in the model. Within OOP, methods and attributes (data) for one object are kept together, in one entity (the object as instantiation of its class). For agents to be autonomous and individual, one needs the individual characteristics and processes together in one single instance. Although other types of development environments can and have been used, agents and OOP fit together intuitively, even as new initiatives to build more socially tailored development environments are under way (Borrill and Tesfatsion 2010, 5).

The Java Runtime Environment (JRE) which executes the java code contains next to libraries and core classes also the Java Virtual Machine (JVM). Java classes are compiled into bytecode that runs in any JVM. There are JVMs for almost any environment, from computers to telephones, making the sales pitch ‘write once, run everywhere’ literally true (Gosling and McGilton 1996). This allows upscaling of the computing environment up to certain super computer configurations if such computing power is needed. It currently is

the most popular programming language (TIOBE Index for November 2017, https://www.tiobe.com/tiobe-index//, retrieved 13 November 2017), and as such a good choice for any software development effort.

Repast itself has predefined classes for building agent-based simulations and to access the simulation infrastructure during runtime. The central object in a Repast simulation is the `Context` class, which stores agents and maintains the relationships between agents (called `projections`). The `ContextBuilder` stores the information on the agents in the Context. RepastS uses configuration files ("model.score" and "scenario.xml") to specify the roles of the classes in a simulation model, to identify the `Contextbuilder` class and to define the default context (Collier and North 2010). Repast uses the `geotools` library to display geographical information (www.geotools.org, accessed march 2011). This open source java library is used to manipulate geospatial data and to display the `Projections` objects and thus the agents with a geospatial location (Railsback *et al.* 2006).

## 8.3 The development environment

In the Repast Java development environment the modeller writes source code that creates model pieces (agents, grid) in the form of Java objects which are passed to the runtime system using a declarative configuration (North *et al.* 2005). The runtime system must be told how to instantiate and connect the model components. Models are managed based on interactive user input (graphical interface) and declarative and imperative requests from the components themselves.

### 8.3.1 The source code

When programming in an object oriented language like Java all entities are objects, with associated data and methods described by their class definition. In repast all model components are plain Java objects. There are several classes included in the HomininSpace source code. The context creator class which constructs the main context, implements the main loop and returns it to the Repast run environment is the `HomininSpaceContextBuilder` class. Part of the source code for this class is shown as an illustrative example in Figure 32. The required agents and projections (grid) are added to this context.

```
/**
 * Custom @link ContextBuilder implementation for the HomininSpace
modelling environment.
 *
 */
public class HomininSpaceContextBuilder extends
DefaultContext<Object> implements   ContextBuilder<Object> {

	/*
	 * Constructor, build simulation environment
	 */
	@Override
	public Context<Object> build(final Context<Object> aContext)
	{
		// schedule start and end of simulation
		ISchedule schedule =
			RunEnvironment.getInstance().getCurrentSchedule();
		// schedule last routines to be executed
		ScheduleParameters endParams =
			ScheduleParameters.createAtEnd(ScheduleParameters.
		LAST_PRIORITY);
		schedule.schedule(endParams, this, "end");
		// schedule start
		ScheduleParameters startParams =
			ScheduleParameters.createOneTime(1);
		schedule.schedule(startParams, this, "start");
		// schedule end
		RunEnvironment.getInstance().endAt(tickCount);

		// schedule main loop
		ScheduleParameters loopParams =
			ScheduleParameters.createRepeating(1, 1,
			ScheduleParameters.LAST_PRIORITY);
		schedule.schedule(loopParams, this, "activateAgents");

		return theContext;
	}
}
```

**Figure 32: Source code excerpt from the "HomininSpaceContextBuilder.java" class.**

The grid based environment in which the other agents reside is a collection of objects accessible through the `Grid` object. The grid matrix is constructed from `HabitatCell` objects, each containing a `CellData` object. The interacting agents in the modelling system are the `HomininGroup`, `HomininFactory` and `Checkpoint` classes. There are several support classes, data definitions, and interface specifications, including the sources `HSUtils` and `Constants` (see Figure 33 which lists all the Java source files included in this project). Verbatim code and interface specifications are included in the Supplementary Materials.

```
/HomininSpace/src
   HomininSpace
      /HomininSpace/src/HomininSpace/HomininSpaceContextBuilder.java
   HomininSpace.agents
      /HomininSpace/src/HomininSpace/agents/Checkpoint.java
      /HomininSpace/src/HomininSpace/agents/HabitatCell.java
      /HomininSpace/src/HomininSpace/agents/HomininFactory.java
      /HomininSpace/src/HomininSpace/agents/HomininGroup.java
   HomininSpace.common
      /HomininSpace/src/HomininSpace/common/CellData.java
      /HomininSpace/src/HomininSpace/common/Constants.java
      /HomininSpace/src/HomininSpace/common/HSUtils.java
   HomininSpace.observer
      /HomininSpace/src/HomininSpace/observer/CheckpointStyleOGL2D.java
      /HomininSpace/src/HomininSpace/observer/HomininStyleOGL2D.java
      /HomininSpace/src/HomininSpace/observer/VisibleValueLayerStyleOGL.java
```

**Figure 33: The Java source files in the HomininSpace development environment.**

### 8.3.2 Software development tools

For the development of HomininSpace and execution of the model and its components several tools have been used, and some of these are mentioned throughout this thesis. These tools are listed below, with in bold the abbreviation that is used to reference that tool where appropriate:

- REcursive Porous Agent Simulation Toolkit (**Repast Simphony** or in short **Repast**) (Repast version 2.2, https://repast.github.io/download.html, downloaded 15 September 2014). This is the main simulation development environment. HomininSpace has been programmed in the Java computer coding language environment.
- The **Yourkit** Java profiler tool to identify bottle necks in execution time (https://www.yourkit.com/java/profiler/, verified 28 November 2017).
    - Identified several hotspots where much execution time is spend, with the implemented solution:
        - For `Factories`, no longer calculate all available resources, just search one valid spot for the new group;
        - For `SetForaging` range, limit the calculations to ranges < `maxForagingRange` (was endless!) Added `maxForagingRange` to the local variables for groups;
        - Some debug statements (that are not printed but do require string operations) are commented out.
        - For performance reasons groups that become extinct are removed from the simulation but are not physically removed from memory but stored in a separate list, ready to be re-initialized when needed. In some simulations more than 60,000 groups are simulated, activating and

deactivating these from a list consumes considerably less memory and processing power than instantiating and destroying them as objects.

- Some scripts that are used in HomininSpace have been written in **R**, a system for statistical computation and graphics (R Core Team 2017). Scripts were developed in the R Studio environment (RStudio 2017). Downloaded from (https://www.rstudio.com/products/rstudio/download/). R can be found at https://cran.rstudio.com/, both addresses verified 28 November 2017.
- For exploratory statistical analysis the Statistical Package for the Social Sciences (**SPSS,** today part of IBM Business Analytics (IBM Corp. 2015)) was used. https://www.ibm.com/analytics/nl/nl/technology/spss/, verified 28 November 2017; However, all numbered statistical output was generated with R scripts, to allow proper replication of these results (Marwick 2017).
- Converter for degrees to decimal http://andrew.hedges.name/experiments/convert_lat_long/ (possibility to indicate W/E), accessed 2 August 2016. Via Excel sheet: http://support.microsoft.com/kb/213449, accessed 28 May 2013.
- Spatial Analysis in Macroecology, version 4.0 (**SAM**) (http://www.ecoevol.ufg.br/sam/, accessed 14 February 2012 (Rangel *et al.* 2006; Rangel *et al.* 2010) is used to select grid cells and to distribute GIS data sets (for instance the WorldClim datasets) over these grid cells.

## 8.4 The runtime environment

Repast offers simulation execution and facilities for agent storage, display, interaction, data analysis and data presentation (North *et al.* 2005). Common tasks like scheduling, object activation, simulation termination, data acquisition and storage, object communication, etc. are supplied by the Repast system. There is a strict separation between model, data storage and visualisation.

Simulations can be started in the Repast runtime user interface (Figure 34). A user interface is no more than a representation of internal data structures, so it is also possible to execute simulations without user interface. Running simulations with the user interface requires installation of the Repast system on the target computer. For the execution of batches the runtime system is prepackaged in an archive file. Practical details including a list of Repast system files are given in the HomininSpace User Manual, included in the Supplementary Materials.

At the start of a HomininSpace simulation the system is initialized by reading the configuration and data files, including grid data, climate data for modern day and the past, and archaeological data with the checkpoints. The simulation grid data structure is created

and filled with the environment information. Factories if desired are created to implement core areas. Data for the reconstructed global sea level and the yearly mean temperatures for the simulation period are read from file databases and stored in memory.

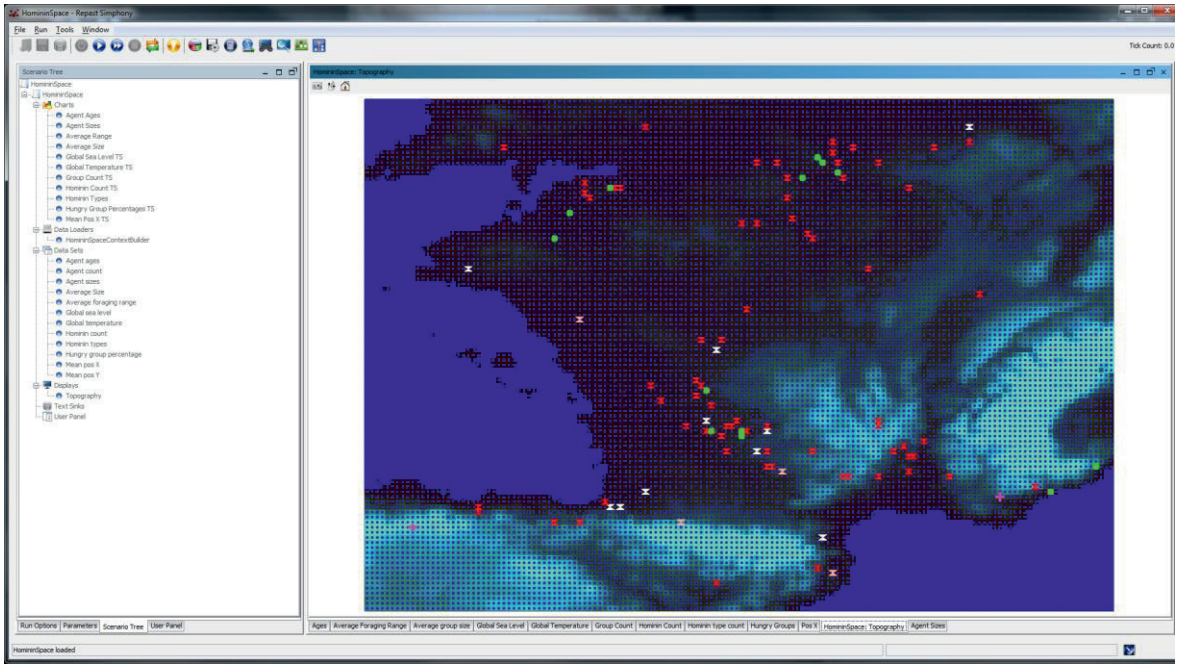


**Figure 34: User interface of HomininSpace after initialization. On the right the Topography layer for the simulation area, on the left the scenario tree.**

### 8.4.1 Batch execution of simulations

In batch mode simulations are executed without user interface. Multiple simulations can be run in parallel and all results are collected and combined after a batch is completed. An interactive user interface is not needed and not present with batch execution. It is possible to set model parameter values for each run. For batch processing the batch variant of the Repast launch scripts must be started. A dialogue window will ask for a file with batch parameters for each simulation in the batch (Figure 35). The user must also specify the number of processors of the target computer (this defines the number of parallel simulation executions).

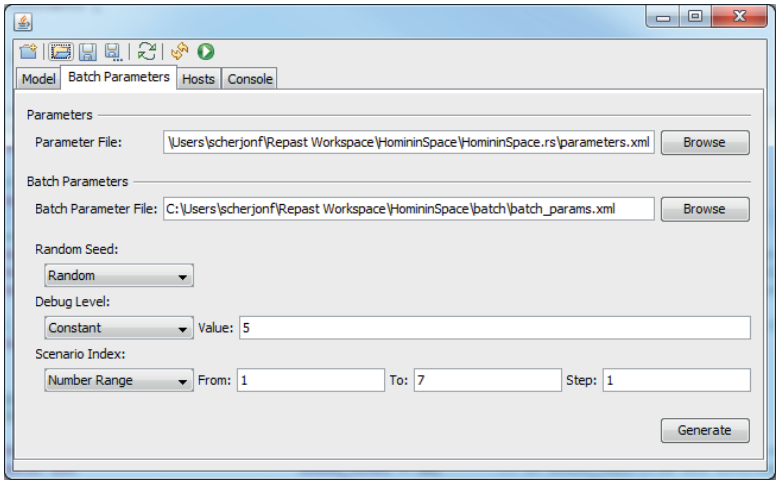



**Figure 35: Values for the batch parameters provided in the Batch Parameters tab.**

To facilitate the execution of large numbers of simulations it is possible to start batch simulations outside of the development environment. To prepare this the complete model has to be available in a single archive batch file named “complete_model.jar”. This file is created by executing the batch Model command in the development environment. Configuration of the batch can be provided as required. The following steps describe standalone batch execution in HomininSpace:

1. Create a directory for execution: copy the created “complete_model.jar” file into this directory and create the files “local_batch_run.properties” and “batch_unrolled_params.txt”;
2. Start execution with: “`java -cp "./lib/*" repast.simphony.batch.LocalDriver local_batch_run.properties`” which reads the appropriate parameter values from the file “HomininSpace Parameters.txt” and the settings for this simulation from “HomininSpace Settings.txt”;
3. Collect results via: “`java -Xmx512m -cp "lib/*" repast.simphony.batch.ClusterOutputCombiner . combined_data`”;
4. Combine results with previous batch executions: merge data from each "combined_data\Batchresults.txt";
5. Ensure that the input file is adjusted: remove those simulations that were executed from the parameter values file “HomininSpace Parameters.txt” and return to step 2 for further batch processing. If the Genetic Algorithm is used it will update the parameter values file with new models and then return to step 2.

The content of an example parameter value file is presented verbatim below. The first line contains the headers for each column, separated by commas. Each subsequent line presents a model with associated parameter values, starting with the unique simulation number (see also Table 21).

```
ScenarioNumber,RandomSeed,BirthRate,DeathRate_PreFertileCohort,DeathRate_FertileCoho
rt,DeathRate_PostFertileCohort,Subsistence_PreFertileCohort,Subsistence_FertileCohor
t,Subsistence_PostFertileCohort,Years_Before_Group_Maturity,GroupSize_BeforeMerge,Gr
oupSizeFertile_BeforeMerge,GroupSize_BeforeSplit,Temperature_Tolerance,CohortSize_Pr
eFertile,CohortSize_Fertile,Calories_Per_Kg_Meat,Max_ForagingRange,Parent_Scenario_1
,Parent_Scenario_2,ViabilityIndex
2306,7355,50,3,1,2,5000,2500,3250,7,4,0,26,-21,9,38,3350,10,1774,2039,92976
2307,6971,43,3,1,3,3375,2200,3000,1,4,4,16,-29,8,38,3150,10,1993,2266,45951
2308,5432,43,1,1,3,2500,2750,5000,1,4,0,16,-28,8,40,3465,10,2220,2255,92541
2309,5432,43,3,1,4,3750,2500,4250,1,4,1,16,-29,9,38,3150,13,2209,2193,31274
2310,8477,43,3,1,12,3575,4750,3250,1,5,2,16,-29,9,40,3150,14,2143,0,32354
2311,8477,43,2,1,2,2750,4750,4250,1,6,1,16,-29,8,38,3150,10,2259,0,63218
2312,7355,48,1,1,12,3250,2250,3250,1,4,1,16,-32,8,40,2900,13,2208,0,201902
```

## 8.5 Data files defining a case study

The HomininSpace system is designed as a general hominin simulation model, with explicitly no hominin species specific elements in the source code or the model structure. For instance, the birth-rate which could be different between species is a model parameter, not a fixed value. The simulated environment and time frame are all taken from input files, as well as the archaeology that simulation results are compared against. That means that any hominin species from any region from any time frame can be analysed with this tool by simply providing the appropriate input files.

During initialisation of a simulation the input files for the model elements are read from several input files. These files are located in the model directory of the HomininSpace system which is defined in the `Constants` object in the source code. HomininSpace has been designed with flexibility in mind and all specifics for a certain case study are read from file. Replacing input files creates a different case study. Files are provided for the topography and climate data, for past sea levels and global temperatures, and with archaeological data in the form of checkpoints.

### 8.5.1 Simulation grid, topography and climate data

The simulation area is divided into grid cells with the Spatial Analysis in Macroecology tool (SAM). SAM is used to distribute topographical and climate parameters onto these grid cells. Next to the topographical data (elevation) these parameters include 19 bio-values from http://www.worldclim.org (both for present and past data sets) and the bed-rock values from the ETOPO1 dataset. The values for the global mean temperature in the past are used as an index for a linear interpolation between the extremes of modern day

climate and LGM climate data. After processing by SAM a selected group of variables are stored in the climate data file called "HomininSpace Climate Grid.txt". See Appendix 1 for a complete list of all variables. It contains one line for comments and 14948 data lines. Each line contains the bioblimatic values for one grid cell. The structure of the file is described in Table 15. The file is read by the following routine in "HSUtils.java":

```
public static List<CellData> readClimateDataFile(final String
cellDataFileName)
```

The file is read into the HomininSpace model at startup. Interpolation of climate data for submerged land is done by HomininSpace. On reading the data a list of `CellData` objects is returned, each containg the relevant climate data for one grid cell.

**Table 15: Structure of the climate data file after preprocessing by SAM.**

| *Field* | *Description* |
|---|---|
| **1** | Longitude (X Centroid) |
| **2** | Latitude (Y Centroid) |
| **3** | n_alt, the number of altitude points in this cell |
| **4** | Mean_alt, the mean altitude (current day) |
| **5** | Mean_bio1, annual mean temperature (current day) |
| **6** | Mean_bio5, max temperature of the warmest month (current day) |
| **7** | Mean_bio6, min temperature of the coldest month (current day) |
| **8** | Mean_bio12, annual precipitation (current day) |
| **9** | Mean_bio13, precipitation of the wettest month (current day) |
| **10** | Mean_bio14, precipitation of the driest month (current day) |
| **11** | Mean_wc_2_5m_CCSM_21k_bio_14, CCSM value for bio14 (LGM) |
| **12** | Mean_wc_2_5m_CCSM_21k_bio_1, CCSM value for bio1 (LGM) |
| **13** | Mean_wc_2_5m_CCSM_21k_bio_5, CCSM value for bio5 (LGM) |
| **14** | Mean_wc_2_5m_CCSM_21k_bio_6, CCSM value for bio6 (LGM) |
| **15** | Mean_wc_2_5m_CCSM_21k_bio_12, CCSM value for bio12 (LGM) |
| **16** | Mean_wc_2_5m_CCSM_21k_bio_13, CCSM value for bio13 (LGM) |
| **17** | Max_etopo1_bedrock, the maximum bedrock value (minimum depth) |

### 8.5.2 Reconstructed global mean temperature and sea level

Reconstructed temperatures and sea levels for the simulation period are read by the HomininSpace system from the input file "Bintanja2008.txt", taken verbatim from Bintanja and van de Wal (2008). A description of the file is given in Table 16. From this file only the data for fields one (timestamp), five (surface temperature) and nine (global sea

level) are used. The routine that reads the file is provided by the "HSUtils.java" class and has the following signature:

```
public static void readBintanjeDataFile(final String cellDataFileName)
```

**Table 16: Structure of input file "Bintanja2008.txt". Timestamp (1), reconstructed temperatures (5) and sea levels (9) are used by HominінSpace.**

| *Field* | *Description* |
|---|---|
| **1** | Time (years ago) |
| **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) |

On reading the data three private array variables in the `HSUtils` class are constructed: `timeStamps`, `seaLevels`, and `temperatures`. Values from these variables are available for each tick (year) through the routines `getSeaLevel(tick)` and `getTemperature(tick)`. When there is no value for a given year, the closest value is returned. The values of the Neanderthal case study are depicted in Figure 36 (global mean temperatures) and Figure 37 (global sea level).

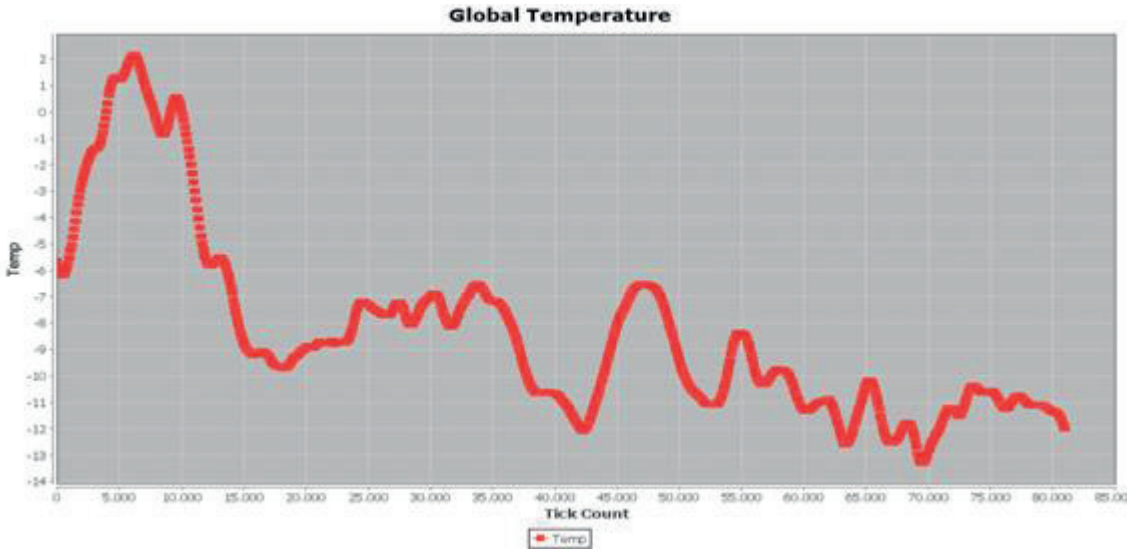



**Figure 36: Reconstructed global mean temperature relative to modern day temperatures. Tick 0 is the start of the simulations (131 ka).**

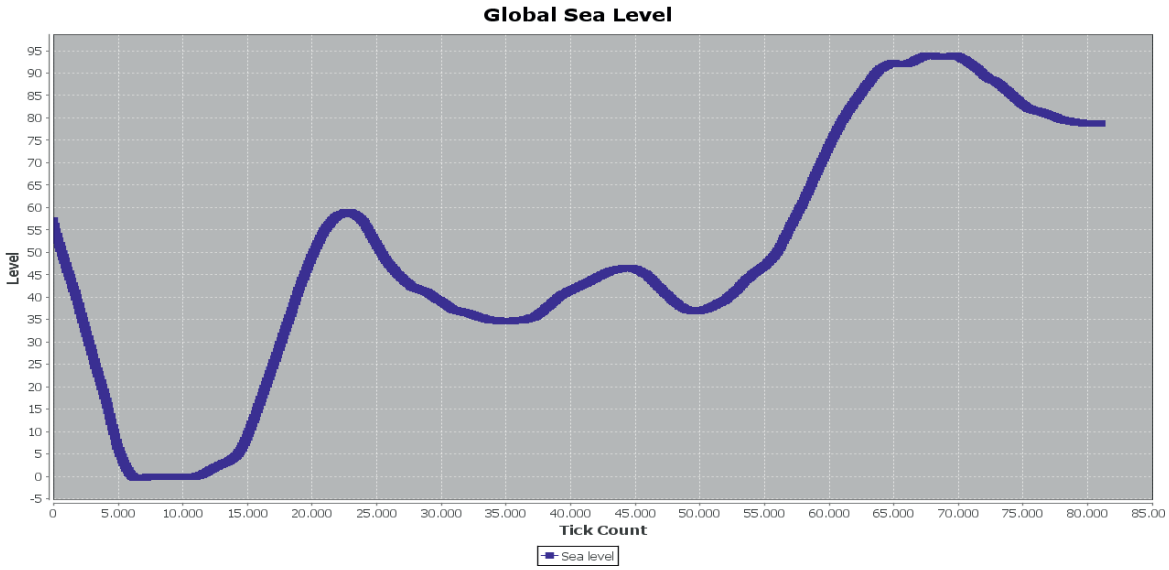



**Figure 37: Global sea level relative to the modern day sea levels Values in meters *below* modern day level. Tick 0 is the starting point of the simulations (131 ka) with a global sea level of almost 60 meters below the level of today (0).**

### 8.5.3 Checkpoints

Checkpoints are objects located in a grid cell with a specific task. There are four different types of checkpoints in a simulation: regular Checkpoints in Space and Time (CSTs), initial starting locations for agents, monitoring checkpoints, and climate monitoring checkpoints. Checkpoints all have a geographical location and are stored as checkpoints with a specific type in one single input file named “HominінSpace Checkpoint List.txt”. The routine that reads this file is provided in the “HSUtils.java” source file and has the following signature:

```
public static List<Checkpoint> readCheckpointDataFile(final String fileName)
```

After successfully reading the file a list with `Checkpoints` is returned. More than one checkpoint can be in the same location (each grid cell is roughly ten square kilometers, and some archaeological sites are located very close together).

The checkpoint file starts with a line that can contain comments, followed by any number of checkpoints, one per line. The four checkpoint types are indicated by the value of the first field in each line. A value of one (1) indicates a regular checkpoint, a value of two (2) is a monitoring checkpoint, three (3) is a starting location for local groups that are present when the simulation starts, and four (4) indicates a climate monitoring checkpoint. For regular checkpoints all fields described in Table 17 must have a value, for the others only latitude and longitude needs to be provided. If the checkpoint location falls outside the grid, the checkpoint is ignored. Fields are separated by tabs.

For the regular CSTs each line is one interval. Intervals for the same locality (site) are stored in one CST. I have assigned a confidence level indicating the estimated reliability of the date provided. It is used in calculating scores for the visits. Preliminary analysis suggests that using only reliable dates seems to make no difference in matching archaeology when compared to using all the dates.

**Table 17: Layout of the checkpoint input file. Fields are separated by tabs.**

| *Field* | *Description* | *Value type* | *Description / remarks* |
|---|---|---|---|
| **1** | Type checkpoint | Integer value | 1 = regular checkpoint<br>2 = monitoring checkpoint<br>3 = starting location<br>4 = climate monitoring checkpoint |
| **2** | Locality (site name) | Character string | Any characters (but no tabs) |
| **3** | Latitude | Floating point value | Example (Artenac):<br>45.85, 0.3333. |
| **4** | Longitude | Floating point value | |
| **5** | Age (ka) | Floating point value | Example: Artenac 110 +9/-7 |
| **6** | One sigma plus (ka) | Integer value | |
| **7** | One sigma minus (ka) | Integer value | |
| **8** | Confidence level | Integer value | 1 = unreliable date<br>2 = modest confidence<br>3 = high confidence in measurement |
| **9** | Type interval | Character string | Can be: absence/presence for indicating an absence or presence interval. |
| **10** | Layer name | Character string | Name of the archaeological layer with which the date is associated. |

In total, the default input file contains 4 climate monitoring points, 13 monitoring points, 13 starting points, and 83 checkpoints with 470 dated intervals. When two or more intervals refer to the same archaeological layer they are joined together to form a new interval with as starting point in time the minimum of all starting points, and as the end the maximum of all ending points. After reading the checkpoints and joining any overlapping intervals a total of 198 intervals are included in the simulations. For easy reference a list of all checkpoints and joined intervals is included in the Supplementary Materials.

### Initial population starting points

When a simulation is started from the Last Interglacial onwards (all simulated scenarios), a number of hominin groups is assumed to be already present in the simulation area. This is the initial population. This is modelled as groups of hominins located at known penultimate glaciation archaeological presence sites (Bocquet-Appel and Tuffreau 2009)[32]. For each group in the initial population a location is provided in the checkpoint input file as a starting location. During initialization of the simulation at each of these starting points

[32] From the sites by Bocquet-Appel and Tuffreau (2009) three are not in France: Pietraszyn, Ripiceni Izvor and Külna.

a group of 25 individuals is created[33]. An overview for these points is given in Table 18. For each site the decimal version of the location is given. An geographical view of where these starting points are is given in Figure 38.

**Table 18: Default starting locations for groups of the initial population.**

| # | *Name* | *Latitude (dec)* | *Longitude (dec)* |
|---|---|---|---|
| **1** | **Abri Suard (La Chaise)** | 45.66667 | 0.46667 |
| **2** | **Bapaume les Oisiers** | 50.14600 | 2.83700 |
| **3** | **Barbas I** | 44.84990 | 0.56670 |
| **4** | **Beaumetz-les-Loges** | 50.24238 | 2.65206 |
| **5** | **Combe-Grenal** | 44.80840 | 1.22300 |
| **6*** | **Mont-de-l' Evangile** | 49.85463 | 2.42174 |
| **7** | **Le Lazaret** | 43.69028 | 7.29500 |
| **8*** | **Longavesnes** | 49.97088 | 3.05989 |
| **9** | **Piégu** | 48.60000 | -2.55000 |
| **10[1]** | **Port Pignot** | 49.68330 | -1.45000 |
| **11** | **Vaufrey** | 44.79997 | 1.20000 |
| **12[2]** | **El Colombo** | 44.14400 | 8.19090 |
| **13[3]** | **La Cotte-de-St-Brelade** | 49.18500 | -2.20200 |

(*) = not in the RPDE database, location by approximation
(1) = This is also known as Fermanville-Port-Pignot
(2) = Site is called Colombo Cave
(3) = This site is located on the island of Jersey




**Figure 38: Google Earth view of the starting locations for the initial population.**

[33] Note that at a starting point no checkpoint is created, only a group of hominins.

### Monitoring checkpoints

Monitoring checkpoints have no presence interval list. These checkpoints are included to record hominin presence during the simulation and can be used for verification and validation purposes. Included are those sites with known hominin presence but where dates fall outside the simulation period, well known hominin sites without exact dates, locations that can be used to detect population expansion, etc. (). These sites include Maastricht-Belvédère, Kent's Cavern, and more.

**Table 19: Monitoring checkpoints located at example positions.**

| # | *Name* | *Latitude (dec)* | *Longitude (dec)* |
|---|---|---|---|
| **1** | **Les Rochers de Villeneuve** | 46.41528 | 0.74083 |
| **2** | **Les Bosses** | 44.46535 | 1.50190 |
| **3** | **Grotte des Fieux** | 44.85000 | 1.70000 |
| **4** | **Petit Bost** | 45.08130 | 0.46083 |
| **5** | **La Micoque** | 44.93300 | 1.01667 |
| **6** | **Maastricht-Belvédère** | 50.85000 | 5.68330 |
| **7** | **Boxgrove** | 50.86901 | -0.69028 |
| **8** | **Kent's Cavern** | 50.46987 | -3.53000 |
| **9** | **Grotte de Saint-Marcel** | 44.33080 | 4.54200 |
| **10** | **La Grotte du Pape** | 43.63000 | -0.70000 |
| **11** | **El Castillo** | 43.28960 | -3.96499 |
| **12** | **Isturitz** | 43.36670 | -1.19611 |
| **13** | **Abri Olha** | 43.36206 | -1.38676 |
| **14** | **Montou-la-Joliette** | 42.78328 | 2.83330 |
| **15** | **Kervouster** | 48.05000 | -4.23330 |

Many known Middle and Late Pleistocene sites have not been included in this list. Included monitoring checkpoints are just to illustrate the use of such sites and for system tests. If any area is to be monitored in greater detail, more checkpoints surrounding that area can be included.

### Climate recording checkpoints

These points register every time step for the grid cell they are located in the two most important climate parameters: precipitation and temperature. They also store the type of reconstructed habitat, the type of topography (which can change due to fluctuating sea levels) and the amount of available energy through time. Climate recording checkpoints are provided for verification and validation purposes[34]. They have been used for instance to create a graph of local habitat type through time for comparison with reconstructed past

[34] Note that for performance reasons these checkpoints are not installed by default. If the user wants to use this functionality it can be activated by setting in the source code the boolean `USE_CLIMATE_RECORDING_CST` to true.

climate graphs from other sources. This can also be used to verify that the grid cell does become a water cell when the sea level rises. These checkpoints do not store absence and presence data for hominin groups. Table 20 lists the climate recording checkpoints that are installed by default. The user can add more as needed. The names of the checkpoints here are indicative of their geographical location (Figure 39). Locations have been selected for their informative value.

**Table 20: The climate monitoring checkpoints with their locations.**

| # | *Name* | *Latitude (dec)* | *Longitude (dec)* |
|---|---|---|---|
| **1** | Climate Monitor 1 Atlantic | 47.0 | -2.0 |
| **2** | Climate Monitor 2 Occitanie | 44.4 | 1.3 |
| **3** | Climate Monitor 3 Mediterranean | 42.0 | 3.2 |
| **4** | Climate Monitor 4 Pyrenees | 43.0 | 0.5 |
| **5** | Climate Monitor 5 Scladina | 50.5 | 5.0 |
| **6** | Climate Monitor 6 Channel | 50.0 | -1.1 |

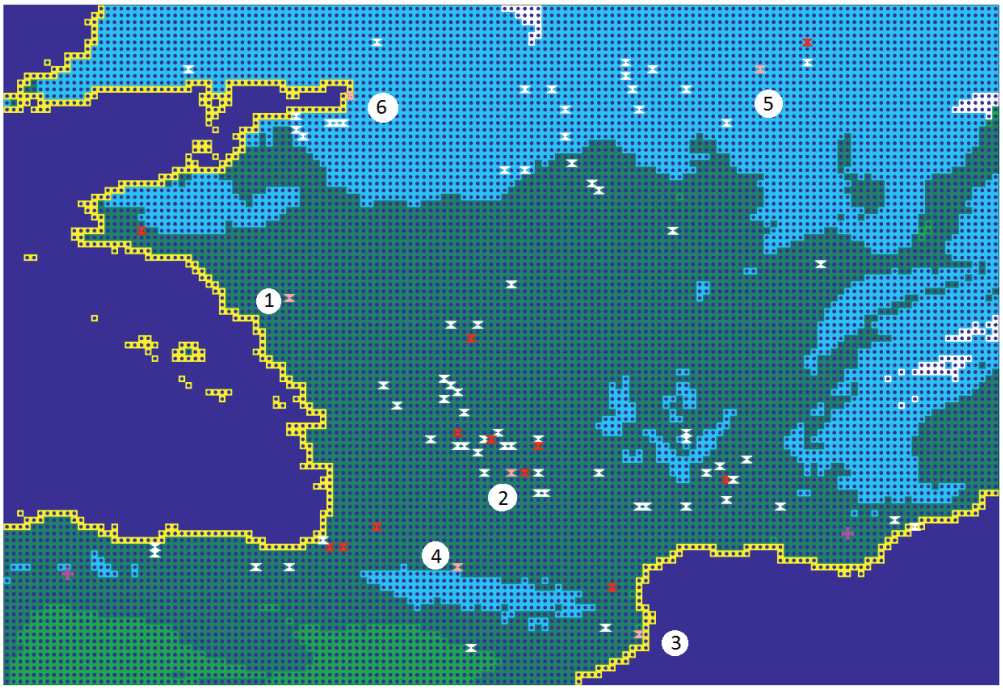



**Figure 39: Climate recording checkpoint locations (pink crosses numbered 1-6).**

## 8.6 Input files defining a simulation

There are two files defining what simulation the HomininSpace simulation system actually executes: a file with the values for the model parameters including the value for the random seed, and a file which defines what model elements participate and thus which questions are being answered (see Chapter 9).

### 8.6.1 Model parameter values

Parameter values for all simulations are stored in a file: “HominінSpace Parameters.txt”. This comma separated file contains a header with the parameter names and then for each line a (different) set of values. When running a single simulation from the user interface it is possible to select the line number with the values to be used. The 16 parameters have been described in Table 13. An example from the Standard parameter values list file is given in Table 21 with the header line with the names of the parameters included.

**Table 21: The header and the first ten parameter value sets (comma separated) from the Standard parameter file.**

| 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,ViabilityIndex |
|---|
| 1,30573,10,4,9,6,2750,4000,4750,6,5,5,46,-28,16,16,3300,10,0,0,0 |
| 2,31589,7,9,9,12,2500,5000,2500,1,2,2,16,-20,13,19,3350,8,0,0,0 |
| 3,6218,36,5,6,8,4000,3750,3000,3,2,4,76,-29,20,37,2850,3,0,0,30 |
| 4,16673,46,6,8,11,4000,2250,2250,8,2,2,96,-15,9,23,2600,5,0,0,103 |
| 5,2875,43,6,1,7,5000,3000,3000,8,2,5,146,-6,18,24,3250,3,0,0,548 |
| 6,20649,30,7,14,14,4750,3000,2250,10,1,1,26,-16,9,21,2850,12,0,0,0 |
| 7,8368,22,4,4,6,4000,4750,4250,3,2,1,116,-26,12,34,3250,4,0,0,34 |
| 8,20427,32,10,5,11,4250,5000,2250,3,3,3,126,-5,12,38,3150,10,0,0,18 |
| 9,9611,27,9,8,11,3000,2500,2750,6,4,5,56,-11,10,33,2750,11,0,0,1 |
| 10,23136,6,2,9,11,2000,2250,2500,8,4,4,6,-27,20,23,3200,12,0,0,0 |

Note that `ScenarioNumber`, `RandomSeed`, `Parent_Scenario_1`, `Parent_Scenario_2`, and `ViabilityIndex` are variables used for administrative and analytical purposes: `ScenarioNumber` gives each simulation an unique number, `RandomSeed` provides the seed for the random number generator, `Parent_Scenario_1` and `Parent_Scenario_2` are the parents for these parameter values if this set is generated by the Genetic Algorithm, and the `ViabilityIndex` is a pre-calculated measure of the effectiveness of the demographic sub-model.

### 8.6.2 Simulation settings

For hypotheses testing purposes and to answer specific questions some model elements can be activated or deactivated. Such model elements include the use of coastal resources, and the maximization of a foraging range. A detailed description is given in Chapter 9. A setting is a Boolean variable that defines if a certain model element is included in the simulations or not. There are ten different settings. Each setting is a Boolean variable that can have a value of either True or False, where True stands for activated. The model

settings are read from a file that is located in the model data directory with the name: "HominінSpace Settings.txt". This file contains two lines: one with a header with setting names, and one line with values for the settings.

The names of the settings, in the same order as they appear in the input file are: `ENERGY_CONTINUOUS, STATIC_DISPERSAL, USE_BEACHES, USE_MAXIMUM_FORAGING_RANGE, GROUPS_CAN_CROSS_WATER, USE_FACTORIES, DEATH_PENALTY_FOR_ABSENCE, ZERO_MODEL_DISABLE_WATER, ZERO_MODEL_DISABLE_ENERGY, ZERO_MODEL_RANDOM_WALK`. An example line with values is: "`false, false, true, false, false, false, false, false, false, false`" with all settings activating their false value but the coastal resources (#3) which is activated. See Table 23 (Chapter 9) for a description of what model elements can be activated for what purposes.

## 8.7 Simulation output

Output of a simulation run consists of a log file with logging and debug information, simulation data collected in maps, time series and charts (user interface only), and a results file with data collected during the simulation. Log information can be written at the end of each time step (optional). Visual output in maps and time series are created interactively and continuously in the user interface. When the simulation terminates, the file with results is written. Figure 40 presents the flow chart describing this process, without user interface.

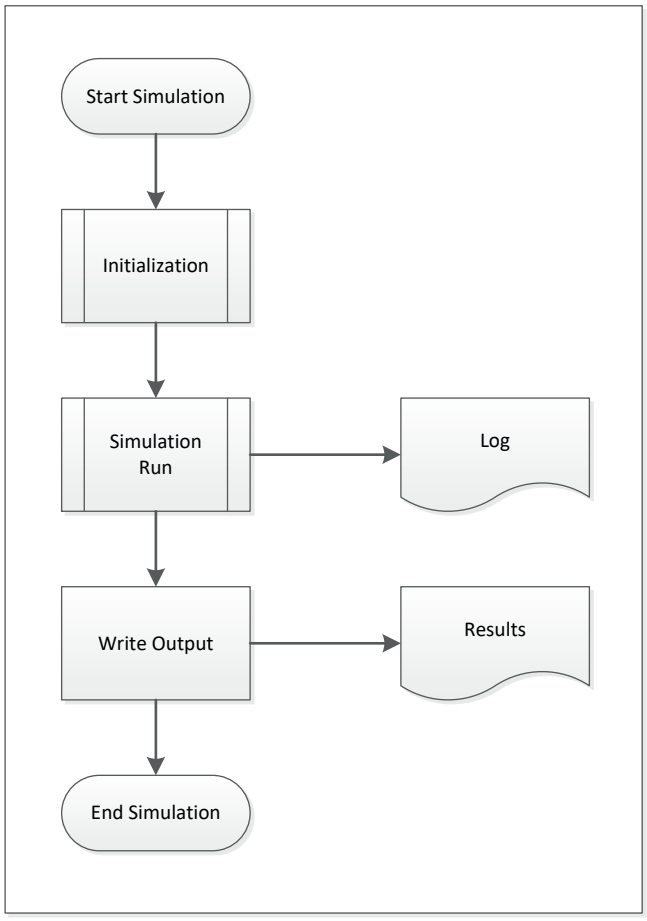



**Figure 40: Flowchart - main simulation routine.**

### 8.7.1 Logging and debug information

The log file is written in the data directory of HomininSpace. The name of this file is "HS Logfile [date][timestamp].txt". For example the file "HS Logfile 2017-11-27 1144.txt" is the logfile written on 27 November 2017, where the simulation started at 11:44 in the morning. The content of the file is variable and is controlled by the Debug Level parameter[35]. With each logging statement in the source code a level is given, and if that level is less than the Debug Level, the information is written into the file. So if the user asks for level 0, no information is written to the file, and if the user selects 1000, all information is written (which can be costly in execution time). The default value is 5, where only information about the model construction and simulation results is written to the file. The main purpose of the debug log file is verification of the system. Debug information is also written to the debug console if the simulation is executed with user interface. Fragments of an example log file are given verbatim below, the full file is included in the Supplementary Materials ( "HS Logfile 2018-11-29 0045.txt").

[35] Either provided by the user via the user interface or in the "local_batch_run.properties" parameter file.

```
HomininSpace Simulation Thu Nov 29 00:45:24 CET 2018
HomininSpace Modelling and Simulation system, version 1.9.98
Reading run settings: HomininSpace Settings.txt
ParameterIndex: 1384
Parameters: 1395,5759,38,4,3,4,2750,5000,4000,6,1,4,16,-
26,14,17,3300,15,0,0,169
SimulationId: 1395
HomininSpace system initialized.
Scenario index: 1384
startingTimeStep: 131000
timeStepIncrement: 1
DEBUG_LEVEL: 10
tickCount: 81000
minimumRequiredTemperature: -26.0
birthsPercentage: 38
sizePreFertileGroup: 14
sizeFertileGroup: 17
…
useFactories: false
deathPenalty: false
intialIberiaGroups: 0
intialItalyGroups: 0
ZeroWater: false
ZeroEnergy: false
ZeroWalk: false
Model directory: D:\Repast Workspace\HomininSpace\
Ordering: Local, Iberia, Italy.
Habitat model (Kelly/Binford)
Random seed: 5759
Random seed: 5759
Reading Climate data file: HomininSpace Climate Grid.txt
Reading sealevel file: bintanja2008.txt
Creating grid: 148x101
Starting year = 131000 years BP; sealevel: 57.121; temperature
offset: -5.6483
Creating factories for Iberia (6,9) and Italy (125,20)
Reading checkpoints: HomininSpace Checkpoint List.txt
Added checkpoint Abauntz [x=42, y=17, 1 intervals][40000 - 54000]
Added checkpoint Abri Bourgeois-Delaunay [x=67, y=43, 6
intervals][130000 - 166000, 91000 - 116000, 111800 - 128670, 77060
- 79280, 107000 - 117000, 65000 - 121000]
```

### 8.7.2 Maps, bar charts and time series

Data from a running simulation in the user interface of HomininSpace can be displayed as maps, as bar charts and as time series. At the end of the simulation aggregated overviews of the data are created. These visualization aids are only available from within the user interface. Unless stated differently, all example output in the description below is created

for simulation number 21, a model parameter value set with a high viability index[36]. Hominin groups are represented in maps by coloured circles, occupying one single grid cell regardless of number of individuals. Red dots are checkpoints, pink crosses are core areas. In figures with habitat reconstructed simulation results, the following colours are used for the habitats, with yellow (not shown) being used for beach areas (Table 22):

**Table 22: Color coding of reconstructed habitats in simulation maps.**

| Tundra | Boreal forest | Evergreen forest | Grassland | Deciduous forest | Woodland |
|---|---|---|---|---|---|
| [white] | | | | | |

Maps are used to show information on the simulation grid itself: topography (height levels and sea grid cells) (Figure 41a), reconstructed environment (in two variants: energy level (Figure 41b) and habitat reconstruction (Figure 41c-d)), visit density (count of all visits to any cell in the simulation this far), and death density (count of all groups that died for each cell) (Figure 42). By default topography is shown at initialization, habitat reconstruction is visible during the simulation, and visit densities are presented at the end of the simulation. What is displayed can be configured in the code, by changing in the "Constants.java" source file the values of `VISIBLE_LAYER_WHILE_RUNNING` and `VISIBLE_LAYER_AT_END` accordingly. The different map types are used to verify simulations and to identify global patterns through time.

[36] Parameter values are: 21,1420,50,9,1,12,3500,3250,5000,7,3,5,26,-24,16,33,2700,9,0,0,1367. All settings (see Chapter 9) are set to False.

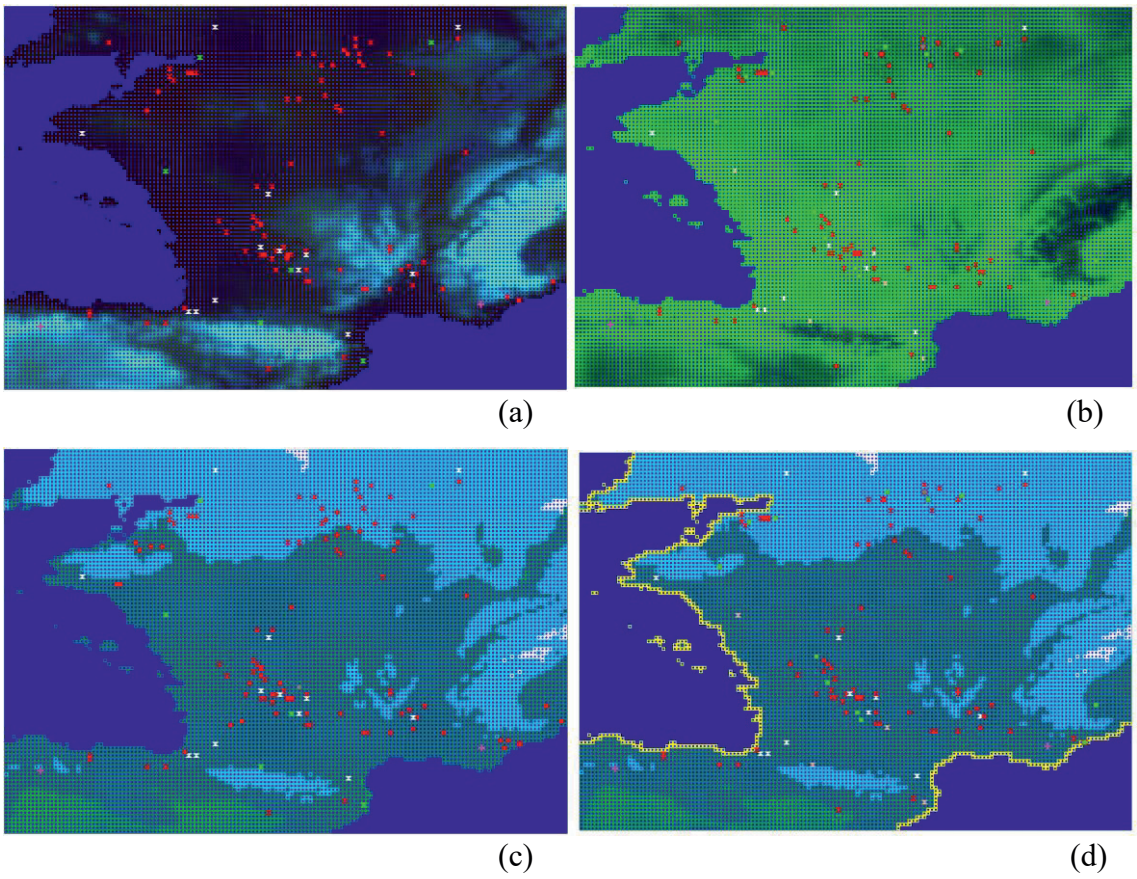



**Figure 41: Different map types at the start of a simulation: topography (a), energy levels (b), habitat reconstruction (c) and habitat reconstruction with beaches (d).**

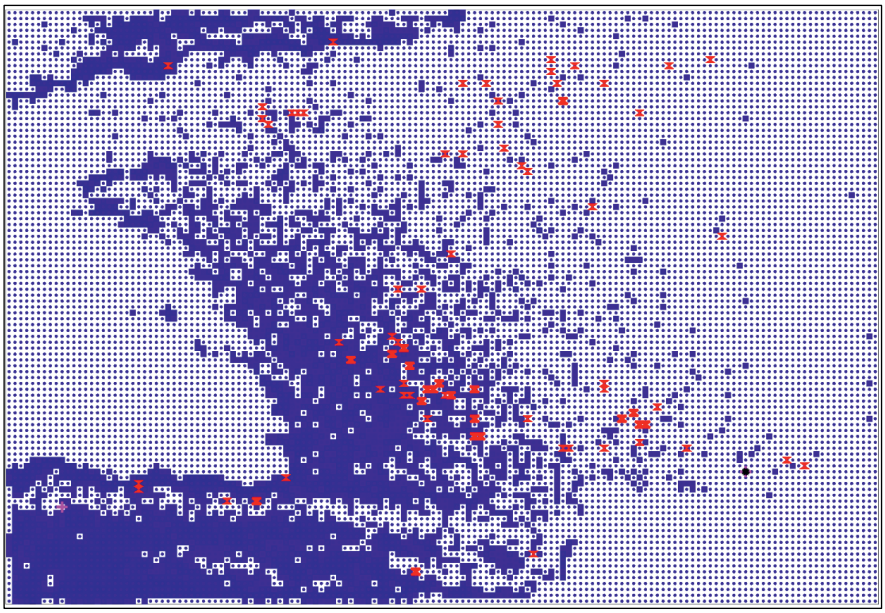


**Figure 42: Death density map at the end of a simulation. Darker colors indicate more perished hominins.**

Bar charts are graphical representations of data in the form of bars. In HominínSpace they present data for one point in time, updated at each time step during a simulation. Two

charts are available: group ages and agent (group) sizes. Group age indicates how long a group survives. Agent size is the total group size, adding together all three age cohorts. These charts are used to display the character of the hominin population at any point in time.

Time series are graphs that can display values or statistical information like average, standard deviation, minimum, and maximum value per time step for variables while running the simulation. The following time series are implemented: global temperature, global sea level, position of all groups (the latitude), hominin count (total and per age cohort, see Figure 43), hominin type count (for three hominin types), group count (Figure 44), hungry group percentage, average foraging range, and average group size. Each of these visualization aids are fed by a data stream defined in the user interface. To enable output through time the data set must have the schedule parameter Frequency set to the value REPEAT, feeding values every time step.

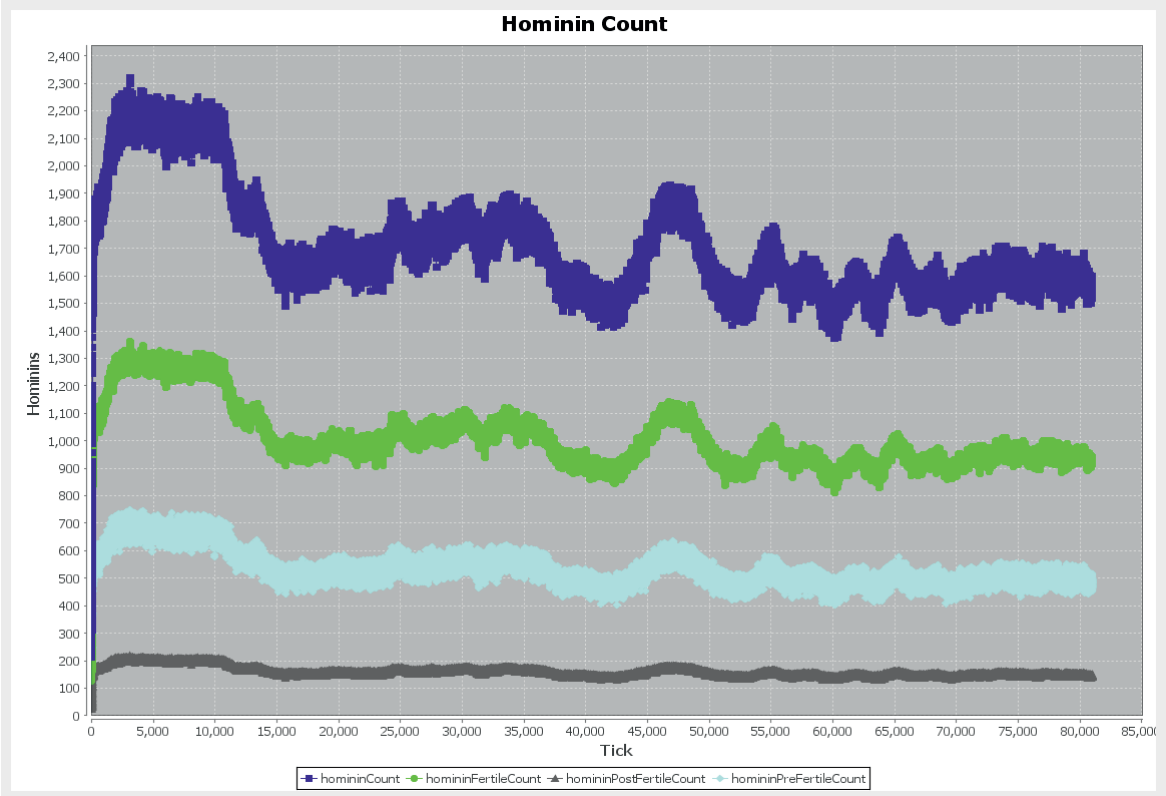



**Figure 43: Example output at the end of a simulation: hominin count (totals in dark blue, different cohorts underneath) through time.**

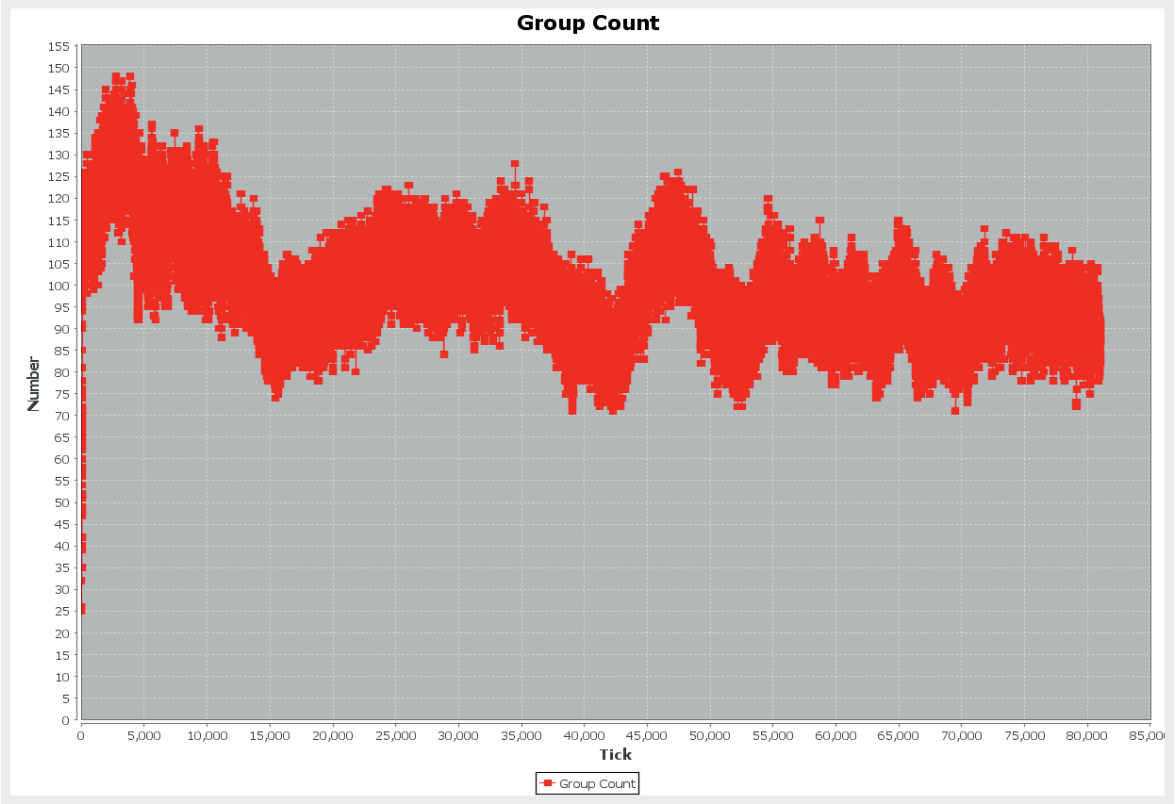



**Figure 44: Example output at the end of a simulation: group count through time.**

8.7.3

### 8.7.3 Simulation statistics

Besides a simulation score, simulation results including statistics on hominin groups and visits are written at the end of each simulation into the file "HS Results [parameter index] [date] [timestamp].txt". For example, the file "HS Results 21 2017-11-27 1147.txt" contains the results for the simulation that used the values for parameters from (Standard) set 21 and ended on 27th November 2017 at 11:47 am. Every results file contains the following data: all parameter values and settings that were used, a number of simulation totals (explained below), the output totals for the climate monitoring checkpoints (each of them can also write an individual output file with data per time step), the output totals for the monitoring checkpoints and finally the output totals for all the regular checkpoints.

The following simulation totals are included in the results file and available for analysis:

`created_groups`, `created_hominins`, `averageGroupAge`, `totalCSTVisits`, `matchingVisits`, `matchedIntervals`, `matchedIntervalCoverage`, `totalVisitsAsHomerangeInsideInterval`, `totalVisitsAsHomerangeOutsideInterval`, `matchingVisitsWithConfidence`, `absenceVisits`, `deathsCrossing`, `deathsCrossingIberia`, `deathsCold`, `deathsFlooding`, `deathsHunger`, `deathsMerging`, `deathsTooSmall`, `Iberia`, `Italy`, and `Offspring`.

For each checkpoint the following totals are written in the results file:

`totalCSTVisits`, `matchingVisits`, `matchedIntervals`, `matchedIntervalCoverage`, `totalVisitsAsHomerangeInsideInterval`, `totalVisitsAsHomerangeOutsideInterval`, `matchingVisitsWithConfidence`, and `absenceVisits`.

Furthermore, each climate monitoring checkpoint writes at the end of a simulation a results file with the following name: “HS Climate [name] [x=[xpos], y=[ypos], [n] intervals][simulation id][date][time].txt”. For example: “HS Climate Climate monitoring station 3 [x=92, y=7, 0 intervals]21 2017-11-27 1300.txt”, with information from simulation 21 for climate monitoring checkpoint “Climate monitoring station 3” with 0 associated intervals written at 27 November 2017 at 13:00. These files contain the following information for each time step: year, temperature, precipitation, energy, topographyType, and climateType. An example is given in Figure 45, illustrating the change in climateType at 127901 years ago. These files can be used to verify reconstructions, analyse local conditions and compare against other reconstruction data sets if available.

```
year,temperature,precipitation,energy,topographyType,climateType
...
127905,15.51,589.8,1312379.9,3,4
127904,15.51,589.8,1312379.9,3,4
127903,15.51,589.8,1312379.9,3,4
127902,15.51,589.8,1312379.9,3,4
127901,15.51,589.8,1312379.9,3,4
127900,15.52,589.67,546706.64,3,6
127899,15.52,589.67,546706.64,3,6
127898,15.52,589.67,546706.64,3,6
127897,15.52,589.67,546706.64,3,6
127896,15.52,589.67,546706.64,3,6
```

**Figure 45: Example output from a Climate Monitor object. Taken from “HS Climate Monitor 3 Mediterranean Habitat [x=94, y=7, 0 intervals] 21 2018-03-23 1303.txt”.**

# 9. ACTIVATING PARTS OF THE MODEL VIA SETTINGS

## 9.1 Introduction

HomininSpace is built to be used as a tool to explore hypotheses and answer questions on hominin behaviour in a reconstructed landscape. To this end the implemented model is run in simulations while comparing simulation results against archaeological attested presence data. Questions in HomininSpace are addressed in scenarios. Each scenario contains a different combination of model elements describing aspects of hominin behaviour in the landscape (Groucutt *et al.* 2015). These elements have been designed to address the questions and implement associated hypotheses that extend the basic parameterized model described in the previous chapters. The system attempts to find the best scoring model parameter value sets per scenario since these describe the hominins most likely to have created the archaeological record under the given hypotheses. This chapter describes the hypotheses that were implemented to answer the questions put forward in the Introduction (Section 1.3).

The in- or exclusion of the model elements are controlled by the so-called *settings* in HomininSpace. A setting is a Boolean value that can be True or False, activating or deactivating certain functionality or elements within the model. An example are extra coastal resources that can be made available to the hominins (True) or not (False). Each setting has a unique and descriptive name. There are ten settings in total which must all be set to either True or False. The chosen combination of all ten settings describes the scenario to be explored. Three settings have been used in the development of the system (the zero model settings). They specifically control the implementation of certain key features (for instance the energy distribution in the landscape, or the direction of the mobility of the hominin groups). All settings are listed with a short description in Table 23 and are described in detail in the following sections.

**Table 23: A list of the settings in the system. Each setting is a Boolean value, and must be activated or deactivated by assigning a value of True or False.**

| *Setting name* | *Description* |
|---|---|
| **ENERGY_CONTINUOUS** | The energy in the landscape is distributed via an energy continuum based on climate parameters (True) or via reconstructed habitats (False). |
| **STATIC_DISPERSAL** | The mobility in the landscape has a static character (True) or more dynamic where the groups keep |

| | |
|---|---|
| | moving with the changing environment (False). |
| **USE_BEACHES** | Do beaches have extra resources (True) or not (False). |
| **USE_MAXIMUM_FORAGING_ RANGE** | Is there a maximum foraging range that groups adhere to and that is given in the parameter file (True) or is there no maximum range (False). |
| **GROUPS_CAN_CROSS_WATER** | Groups can cross larger water bodies (True) or not (False). If not, they cannot go to islands. |
| **USE_FACTORIES** | Activate core areas or factories (True) or not (False). If True, a maximum of one new group is generated per year near a factory if there are sufficient resources in the immediate surroundings. |
| **DEATH_PENALTY_FOR_ABSENCE** | Does the system implement the death penalty for simulations where absence intervals are violated (True) or are absence intervals ignored (False). |
| **ZERO_MODEL_DISABLE_WATER** | Are there water grid cells in the North sea or Mediterranean (False) or are these grid cells all landmass (True). This setting has been used during system development. |
| **ZERO_MODEL_DISABLE_ENERGY** | Do grid cells have a reconstructed energy level (False) or do all grid cells have the same amount of energy that replenishes every year (True). This setting has been used during system development. |
| **ZERO_MODEL_RANDOM_WALK** | Groups move randomly (True) or they move to that area which has the most resources (False). |

## 9.2 Model elements of the reconstructed environment

In HomininSpace the underlying assumption is that energy available in the landscape will provide forage for the hominin groups that move through this landscape. For calculating energy levels reconstructed precipitation and temperature values are available. Available energy per grid cell is calculated using a continuous energy level reconstruction or with a habitat reconstruction with associated energy levels per habitat (Subsection 9.2.1). Furthermore, extra resources can be available for coastal areas with possibly great effect on the population distribution (Subsection 9.2.2).

### 9.2.1 Energy versus habitat reconstruction

Population responses to environmental change can be assessed by combining an energy budget and an agent based model build around individual responses to local circumstances (Johnston *et al.* 2014). Energy required by hominins for subsistence is taken from the landscape in the form of hunted or scavenged game, to some extent supplemented by plant foods (Henry *et al.* 2011). Diets for modern human hunter-gatherers are very diverse and difficult to model (Kelly 1995, 65-101). In the model underlying the HomininSpace simulation system the availability of medium to large ungulates is taken as the limiting

resource defining the maximum population size. Individuals obtain energy from ingested food and expend it on movement, maintenance, growth and reproduction (Sibly *et al.* 2013). Exact energy needs by hominins are uncertain (but see Mateos *et al.* (2014)). HomininSpace implements two different approaches to calculate the available energy in the landscape: direct energy computation via extrapolation from the climate parameters, and a habitat reconstruction with for each habitat type an associated energy level.

The difference between a continuous energy landscape and a full habitat reconstruction was already built into HomininSpace 1.0. It is reflected in the way the available energy is calculated, the hominins in the system will not perceive the environment differently. The only thing hominin agents notice is the amount of available energy, whether that is calculated according to a continuous energy model based on data from McNaughton *et al.* (1989), or using reconstructed habitat data from Binford (2001) and Kelly (1983). The location in the source where the secondary biomass is calculated is in "HabitatCell.java", line 233. If the setting ENERGY_CONTINUOUS is True the following function is called: `calculateSecondaryBiomassAccordingToMcNaughton`, otherwise this function is used: `calculateSecondaryBiomassAccordingToKellyAndBinford`. The reader is referred to the source code included in the Supplementary Materials for the actual implementation of both functions, which contain 60-100 source lines each.

In HomininSpace consumption of resources does not change the habitat type. The type is always reconstructed using climate data, not (available or consumed) resources. Consuming resources does change the amount of available kcal. In the user interface it is possible to visualize both alternatives via the `VISIBLE_LAYER_WHILE_RUNNING` constant. This is illustrated in Figure 46, with both figures created just before the simulation is started and before any resources are consumed.

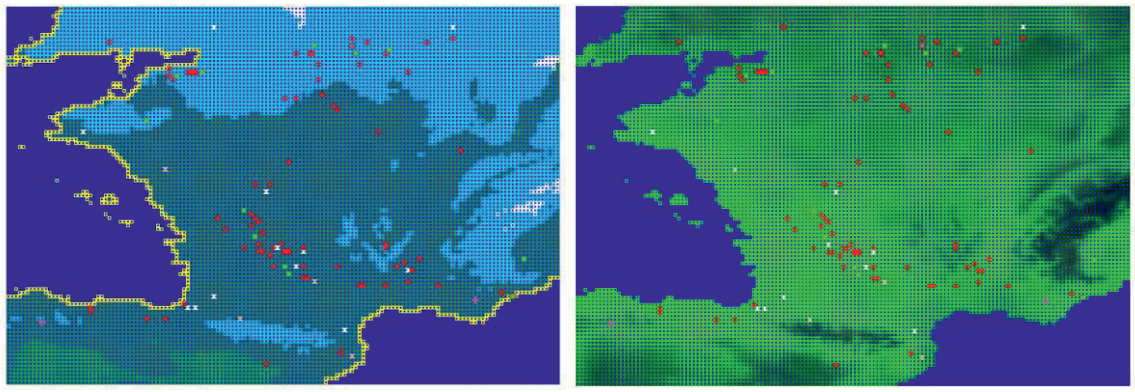


**Figure 46: Habitat reconstruction (left) versus energy level reconstruction.**

### 9.2.2 Coastal resources

Aquatic resources have been exploited throughout the long history of the genus *Homo*. The earliest fresh water fish is attested with *Homo erectus* (Braun *et al.* 2010). Joordens *et al.* (2015) suggest that freshwater molluscs could have been exploited more than 400 ka in Trinil, Indonesia. *Homo sapiens* consumed coastal shellfish from MIS 6 onward in South Africa (Kyriacou *et al.* 2014) and archaeologically retrieved specimens have been used to detect population growth (Klein and Steele 2013). Marine mollusks were exploited in the Mediterranean area by both Neanderthals and modern humans since MIS 5 (Fa *et al.* 2016). In the Levant, data suggest that structural coastal resource exploitation started during the Early Upper Palaeolithic, during which shellfish exploitation became increasingly more frequent and year-round (Bosch *et al.* 2017). The use of molluscs can be traced back to as early as MIS 6 at the Mediterranean coastline of the Southern Spain (Cortés-Sánchez *et al.* 2011). Coastal intertidal foraging return rates can be significant and overall energetic return rates can exceed other foraging activities (De Vynck *et al*. 2016). The question to be explored is if Neanderthals used coastal resources in the simulation area and during the whole simulation period.

Fresh water systems do not play a major role in HominінSpace. Most rivers and lakes are too narrow/small to be included in the topographical map due to gridcell size, and for habitat reconstruction water is assumed to be precipitation only. There is however an extensive coastline, both in the south where the simulation area is bordered by the Mediterranean Sea, as well as the western edge which is formed by the Atlantic and Channel coastal areas including the beaches from southern England. Therefore, the question about the use of aquatic resources focusses on the coastal regions. A value of True for the USE_BEACHES setting activates the coastal resources in a simulation.

In the topography land grid cells have a height greater than zero. If coastal resources are activated, those land grid cells that border at least one water grid cell (sea or lake) are labelled as `TOPOGRAPHY_TYPE_BEACH`. These are the yellow grid cells in Figure 46, the habitat reconstruction. Coastal resources are located on the beach cells (land), not in the water since foraging by hominins occurs only on land grid cells. The additional coastal resources are created by multiplying the regular resource production (based on habitat or energy level) by two. Additionally, coastal resources replenish completely every year, reflecting the seasonal character or the resource and assuming no over-exploitation of coastal resources is possible. Figure 47 illustrates that coastal resources apparently can be

popular with the coastal areas being highly populated[37]. Coastal resources are implemented in the "HabitatCell.java" source file.

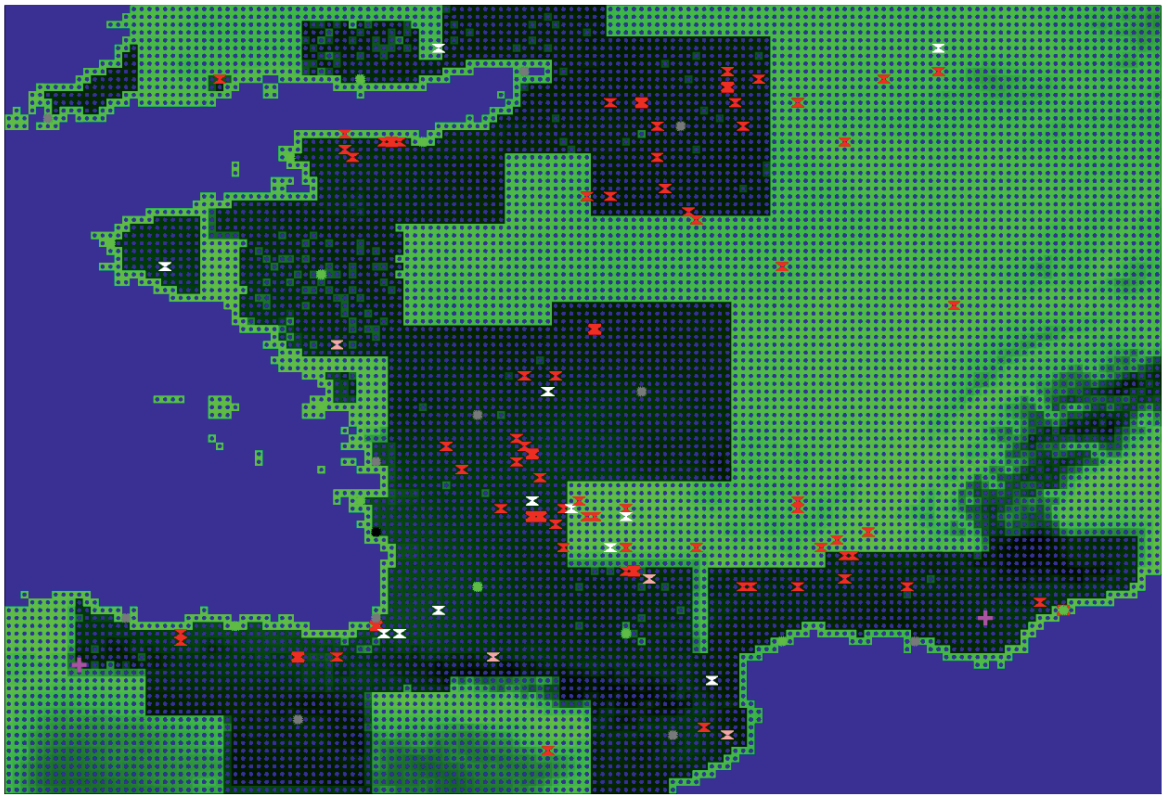


**Figure 47: Illustrating the effect of coastal resources on population distribution.**

## 9.3 Model elements of hominin behaviour

Hominins move through a reconstructed landscape in HominínSpace foraging for food. The user can explore hominin dispersal characteristics and demographical change. Hominin groups can implement a more static mobility versus more dynamic movement (Subsection 9.3.1). When groups forage it feels natural to impose a maximum foraging range (Subsection 9.3.2) but what maximum should that be? Some hunter-gatherers travel huge distances per year. And are hominins capable of crossing large surface waters (see Subsection 9.3.3)?

### 9.3.1 Mobility type: dynamic versus static hominins

Dramatic climatic fluctuations characterize the environment for the Late Pleistocene hominins. It is suggested that many organisms were pushed southwards as the northern areas became too inhospitable (Barnes 2010; Stewart and Lister 2001; Schmitt and Varga 2012). The Neanderthal range responded by contracting and expanding accordingly. Such

[37] Note that coastal resources replenish every year: even if inland areas shows intense foraging (dark colour), beaches remain light green (high energy levels).

processes could be explained by either large scale habitat tracking or regional extinction and subsequent recolonization (Hublin and Roebroeks 2009; Dennell *et al.* 2011). This research contrasts the "ebb and flow" model (habitat tracking) with the "regional extinction" theory. Henceforth the regional extinction model is also referred to as "sources and sinks" where needed to avoid confusion since also with habitat tracking there will invariably be local extinctions(e.g. Riede and Pedersen 2018). Habitat tracking hominins would move towards areas with better environments when the climate deteriorates.

With "sources and sinks" a sink area can have a declining population due to unfavourable environmental conditions but that population is replenished by immigration from different source locations with a net population growth and a surplus of individuals. Local populations are assumed to remain in the area and will become extinct when repopulation levels are below death rates. See for example Finlayson (2009) identifying Iberia as a source for western Europe, or Dennell *et al.* (2011) who identify Southwest Asia as a possible population source for Europe.

Distribution maps of past species are generally however coarse-grained palimpsests of the traces of population expansion and contraction, probably hiding many populating events. There might have been many phases of range expansion and contraction or migration caused by the rhythm of climatic oscillations and these maps most probably understate the full impact and give only a rough approximation of the former distribution of the species (Dennell and Roebroeks 2005; Hublin and Roebroeks 2009). For the colder phases of the glacial-interglacial cycles, there are indications that large parts of the Neanderthal range in northern Europe were deserted in the early Middle Pleistocene, with Neanderthals maintaining core populations in the south (Hublin and Roebroeks 2009). The hominin presence in Pakefield, with a reconstructed near modern climate, is often seen as an example of pre-Neanderthal habitat tracking (MacDonald *et al.* 2012).

Neanderthals are usually seen as a western Eurasian species, but in fact little is known about the limits of their range, both in terms of their former distribution as well as regarding the factors which limit their survival (Roebroeks 2010). A different mobility pattern in response to a changing climate is one of the suggested causes that could have led to the demise of the Neanderthals (Holliday and Falsetti 1995). For the colder phases of the glacial-interglacial cycles, there are indications that large parts of the Neanderthal range in northern Europe were deserted with Neanderthals maintaining core populations in the south (Hublin and Roebroeks 2009). The limits of the Neanderthal geographic range are

usually constructed by drawing lines around the maximum distribution of their fossil remains, but differences in site preservation, reconstructed periods as well as in research intensity and history make such estimates very rough and preliminary (Roebroeks *et al.* 2011; Roebroeks and Soressi 2016). In HomininSpace large mountain ranges and arid areas are passable although distribution is hindered by limited resources in and around such grid cells which makes them less attractive destinations.

The tracking of favourable habitats has been described as the "ebb and flow" of populations (e.g. Hublin and Roebroeks 2009). This is the setting STATIC_DISPERSAL with value True. The "ebb and flow" of moving populations has often been opposed to a "sources and sinks" model where some local populations must adapt behaviourally and/or genetically to cope with the changing climate or become extinct when conditions become less favourable (Pulliam 1988; Pulliam 1996): value of the setting is False. In the source code of "HomininGroup.java", where this is implemented (lines 145-198), this setting allows a group to 'settle', via the boolean `isGroupSettled`. A group can only settle if the setting STATIC_DISPERSAL is True. If a group is settled it can no longer move and must obtain the required resources from the area limited by its foraging range[38]. A group can only settle if the following conditions are met:

- The group is not hungry, that means that there are sufficient resources in the area given the current size and composition of the group;
- The group is mature (that is older than the age indicated by the Years_Before_Group_Maturity parameter value);
- There are no other groups within the foraging range that are settled (this does not prevent two foraging ranges to overlap);
- The current location is not within three grid cells from any border of the simulation area.

### 9.3.2 Imposing a maximum foraging range

Hominin groups in HomininSpace obtain the required energy from the landscape around them. They forage to retrieve resources translated into kilocalories. This is implemented in a two-step process:

1. First each group assesses how many kcal are available within a searching radius of one grid cell from the current position (this encompasses nine grid cells including the current location). If that is insufficient for their needs the radius is increased by one[39] and the total number of available resources calculated for the new radius. The radius is increased until sufficient resources are found or the maximum foraging

[38] Note that a group that is settled and becomes too small can still merge with another group in the foraging range.
[39] A searching radius of 1 encompasses 9 grid cells, 2 covers 25 grid cells, 3 covers 49 grid cells and 4 is 81 grid cells.

range is reached. In the calculations the groups include the presence of other groups that have claimed resources. The final radius is referred to as the current foraging range in which all grid cells are claimed;

2. Then in the second step, after all groups have calculated their current foraging range, resources are consumed by emptying grid cells in a random order from the current foraging range, sharing resources with other groups if their foraging range overlap.

Implementing foraging strategies in this two-step manner will allow groups to co-exist in the same area, and prevents the random order of group activation to unnecessarily cause groups to become hungry (and in extreme cases to die) when one group by chance forages before the other group which is then left with insufficient resources. Note that the foraging event actually represents one full year of resource extraction from the environment within the (annual) foraging range. Not all grid cells within the calculated foraging range will be exploited every year. If sufficient resources are present some cells are not needed (Figure 48).

If maximum foraging ranges are deactivated (setting USE_MAXIMUM_FORAGING_RANGE is False) resources can be obtained for a potentially unlimited foraging range (only limited by the extent of the map and the resource needs of other groups). Large ranges are not impossible since the foraging behaviour encapsulates the period of one year, and there is no data to limit the migration distance of Neanderthals (Benito *et al.* 2016). Cold and arid circumstances seems to increase mobility and range sizes since resources are more sparse and wider distributed (Binford 1991; Kelly 1983). Also the exploitation of meat resources by hunter-gatherers, especially when the prey is migrating, requires large (annual) foraging ranges (Kelly 1983, 296; Kelly 2003). Nunamiut and for example Cheyenne hunter-gatherers hunt large mammals in colder environments and are extensively mobile (Wragg Sykes 2017).

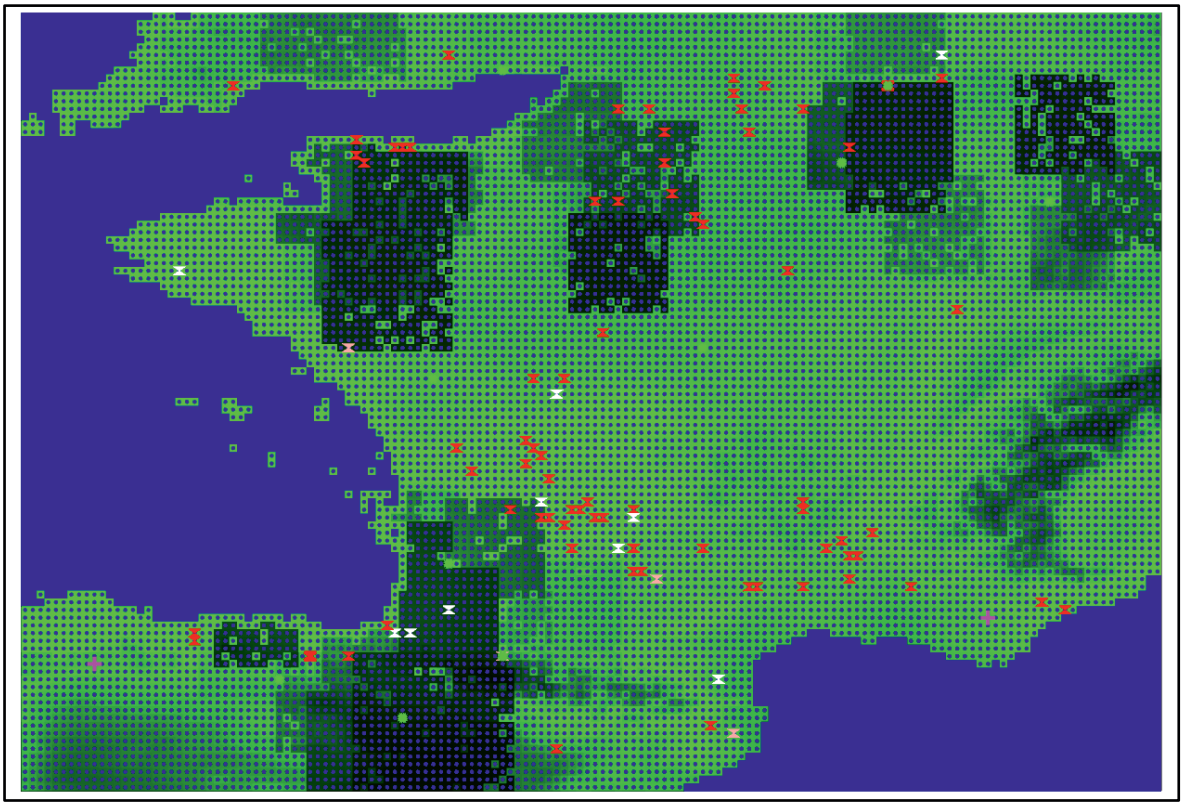


**Figure 48: Screenshot of simulation 1384 illustrating the patch wise foraging behavior with a varying foraging range (different sizes for the foraging squares).**

### 9.3.3 Allowing the crossing of open water systems

Sea barriers can hinder colonisation efforts, particularly difficult crossings like for example the Strait of Gibraltar (Derricourt 2005). Analysis of flora and fauna, including comparisons of genetic material, shows no direct contact between the Maghreb and Iberia (Spain) (Close 2009, 45; Tafelmaier *et al.* 2017). Some similarities have however been observed between Mousterian and other assemblages on both sides (Hublin 2000, 170; Straus 2001) but no Neanderthal remains have been retrieved from African soil yet. This could suggest that Neanderthals were not capable of crossing a relatively small strait of 14 kilometres wide with opposite shores clearly visible. Turbulence makes this a hazardous crossing indeed which could mean that they were not capable of crossing larger water bodies at all. However, very early presence of modern humans has been attested in Northern Africa (Hublin *et al.* 2017), and yet no fossil remains have been found for the simulation period in Spain for that matter. Nevertheless, the ability to travel over large water bodies might have been a specific trait associated with certain modern humans only (Davidson and Noble 1992).

Larger river systems do not seem to be able to stop Neanderthal dispersal, although the Ebro river basin is suggested to separate Neanderthal populations in the south of the Iberian peninsula from more modern northern populations (Zilhão 2000; Vaquero *et al.* 2006; Daura *et al.* 2013) allowing a late survival of Neanderthals in the south. This model (see discussion in Cucart-Mora *et al.* (2018)) and the late survival is however contested (Higham *et al.* 2014).

In the wider Mediterranean area there are suggestions for long-term persistent archaic populations on islands proximate to the Eurasian mainland (Broadfield *et al.* 2001). There is evidence for maritime dispersal in monkeys (Ferràndez-Cañadell *et al.* 2014) and for other large-bodied mammals and this latter could induce short-distance maritime dispersal of local Neanderthal populations (Broodbank 2006; Broodbank 2014). Passive dispersal across larger water bodies has been argued as a possible agent for archaic hominins (Dennell *et al.* 2014; Leppard 2015). This does not necessarily mean that hominins were seafaring but the possibility of crossing is based on the occurrence of rare but suggestive long range dispersal events in the distribution of the genus *Homo* (Strasser *et al.* 2010; Leppard 2015). Thus under certain circumstances water crossing could have been part of the colonisation repertoire of Neanderthals.

The most important area however for which the crossing of larger water bodies influences Neanderthal dispersal in the simulation area is the Channel, where colonisation of England can be prohibited when sea levels are high and England becomes an island (Ashton and Lewis 2002; White 2006; Gilmour *et al.* 2007; Lewis *et al.* 2011; Sier *et al.* 2015; Ashton and Scott 2016). Only a lowered sea level and ameliorated climate allowed over land access to Britain in MIS 3 with hominins possibly crossing the Dogger Plains to avoid the large Channel River system (Wragg Sykes 2010). There is scarce evidence of earlier presence which could only have been made possible by Neanderthals having the ability to cross larger water bodies (Wenban-Smith 2010; Wenban-Smith *et al.* 2010). But this absence could be taphonomical rather than cultural or biological (but see Wragg Sykes 2017).

HomininSpace attempts to explore this issue by in- or excluding the ability to cross larger bodies of water, since it is an important element in hominin dispersal (Leppard and Runnels 2017). In the HomininSpace simulation system, hominins cannot live on and will never move to water grid cells. The fluctuating sea level allows and forbids land access to a small but significant part of the simulation area (the south of Britain). HomininSpace

implements the ability to cross larger water bodies with the `GROUPS_CAN_CROSS_WATER` setting. If this setting is activated (True) any grid cell within dispersal range (two times the active foraging range) can be targeted as a new destination for a group, even if this involves the crossing of (very) large water bodies. If the setting is set to False, islands cannot be reached by the hominin groups.

Since the actual route that an agent follows is never calculated step by step (only the maximum distance is used as a limiting factor) the implementation of this feature requires the identification of all the islands on the map. Only then it can be tested if a new destination is on the same or on a different island. This is not straightforward since the map is quite large (with at least 6000 land grid cells). An intuitive recursive implementation to identify islands requires too many resources often leading to stack overflow. An iterative implementation was made, one that searches for unassigned land cells and when finding such a grid cell, expanding this in all directions. As there are only few islands in the area (generally around nine in total) this is an efficient routine and since it is executed every time step the computational gain in speed opposed to the more elegant recursive implementation is important as well.

The actual code is implemented in “HomininSpaceContextBuilder.java”, routines `assginIslandNumbers`, `assginIslandNumberIter` and `setIslandNumber` (lines 1217-1305).

## 9.4 Model elements of population distribution

There is only one question that is related directly with archaeological data, and that is about absence of evidence of hominin presence (Subsection 9.4.1). The associated hypothesis states that for certain areas it is plausible that hominins were *not* in the area. Such information is very important when attempting to restrict the urge of the system to create an omni-present hominin. Furthermore, new hominins can be created from existing groups or by immigration from more productive areas. HomininSpace can model such areas and offers the possibility to create a core population area anywhere on the map, as described in Subsection 9.4.2.

### 9.4.1 Using hominin absence data

If simulation results have hominins in an area and time frame for which absence has been proven this is clearly an error of the model (Pearson *et al.* 2006). It has been shown that for species distribution models the presence-absence models perform better than models with

presence only data (Elith *et al.* 2006). If simulated hominins are present where the absence is implicit (that is, there is no presence record) this is considered a non-conformity (but not an error!), maybe due to an incomplete model, maybe due to sampling biases (Anderson 2003). A generalised scoring scheme when comparing modelling results and real world presence and absence data is presented in Table 24. In this table modelled absence versus proven presence is less severe than modelled presence with proven absence since a modelled (local) absence could still mean that modelled hominins were nearby (compare archaeological presence). Often, pseudo-absence data is generated to create a counter weight to the presence only information (Benito *et al.* 2016).

If only presence data is used, an always omni-present species would provide maximum match with the data. Excavation results show that, at least in many locations, Neanderthals were not *always* present at *all* sites all the time (for instance illustrated by the overview in Discamps *et al.* (2011), or for a specific location in Bertran *et al.* (2013)), however local such absence must be taken. Modelling efforts and interpretation of simulation results where only presence data is matched must take into account the tendency of the system to 'fill the map', and must try to distinguish between the natural tendency of a species to grow and the intrinsic modelling drive to match all presence points.

**Table 24: Relative scoring schema when matching real world absence and presence versus modelling results. A '+' indicating a positive contribution of the match to an overall modelling score.**

| | Modelled absence | Modelled presence |
|---|---|---|
| Proven absence | ++ | --- |
| Implicit absence | + | - |
| Proven presence | -- | ++ |

The above illustrates the importance of matching absence information as well as the (relatively) limited value of local absence data. If absence is however attested for many sites in a wider, delimited area, with reasons that explain why, and with taphonomy and other potential causes for absence sufficiently countered it can be argued that hominins were indeed *not* present in that area for a given time frame (proven absence). There is one larger area for which absence has been postulated for an extended period of time: the whole of Great Britain during MIS 5 and MIS 4 (Ashton 2002; Ashton and Scott 2016;

White and Pettitt 2011; Wragg Sykes 2017). The initial breach of the chalk barriers formed the Strait of Dover sometime during MIS 12, and since then allowed a rising sea level to isolate Britain from the main land of Europe (for a discussion see Ashton *et al.* (2018)).

For Britain the cold and associated glaciations have been identified as the main cause of absence, with any (larger) waterbody between Britain and the main land of Europe as an additional barrier for hominins (Antoine *et al.* 2003; Scott and Ashton 2011; Sier *et al.* 2015). Two flakes from a questionable late MIS-5 occupation were retrieved (Wenban-Smith *et al.* 2010; White and Pettitt 2011) which are here taken to be insufficient as proof of presence (following Wragg Sykes (2017)). In this study the absence information on Great Britain is used to illustrate the effects of adherence to true absence on modelling efforts, and to explore the hypothesis that there was indeed an absence of Neanderthals in Britain during a significant part of the Late Pleistocene and what that means for a model.

The identification of areas with true absence of hominins is notoriously difficult. Therefore the inclusion of the southern parts of England in the simulation area is key, since here is a larger area for which true absence has been attested (at least with a large degree of certainty, see Section 6.2). From late MIS-7 until the end of MIS-4 or the start of MIS-3 (c.190 – 60 ka), Britain appears to have been "effectively abandoned" (Wragg Sykes 2017). Table 24 indicates that modelled presence in proven absence is the least desirable modelling result.

For evolutionary algorithms penalty functions are the most common method to handle constraints (Yeniay 2005). For genetic algorithms, a popular and simple method is the *Death Penalty* which just rejects unfeasible solutions from the population (Back *et al.* 1991). This states that when a constraint is *not* met, the solution is incorrect and can be removed from the collection. The risk with this method is that most time is spent calculating infeasible solutions that are subsequently rejected.

In the implementation in HomininSpace absence information is taken as an absolute constraint that is punished with the Death Penalty and made explicit by defining absence intervals with checkpoints. These intervals are constructed in a similar fashion as the presence intervals, with the only difference in the input file the type of interval which is 'absence' instead of 'presence'. One of the best dated sites in the UK is Lynford quarry (Ashton 2002), and data for this site suggest an absence sometime before 72 ka (Boismier *et al.* 2012). It also suggests that around 142 ka and earlier Neanderthals might have been

present. For the purpose of this study two checkpoints in England have an (identical) absence interval defined:

1. Boxgrove: 107 ka +/- 35 kyr
2. Kent's Cavern: 107 ka +/- 35 kyr

Whenever a hominin group is near a checkpoint, they are registered in the interval administration. For these routines the type of interval is unimportant. Only when the data is written to the output file the type is included. When the absence hypothesis is tested the genetic algorithm will filter away all solutions with presence registered in the absence intervals before selecting promising parameter value combinations. Only when the setting DEATH_PENALTY_FOR_ABSENCE is True, those simulations are ignored. Note that they are still present in the output file for all other purposes.

### 9.4.2 Using population core areas (hominin factories)

In the HomininSpace model new groups can be created in *core areas*. These are locations on the map from which, conditions permitting, new groups move into the simulation. Two such points have been designated based on suggestions in the literature: the peninsulas of Iberia and Italy. Production from core areas can be activated (True) or deactivated (False) by the user with the setting USE_FACTORIES.

Although hominin occupation in Iberia may have been discontinuous and it is unclear where the western European hominin populations originate from (MacDonald *et al.* 2012), the Iberian peninsula is one of the areas within the HomininSpace modelled area from where new populations can move into 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. The selected locations of these population core areas are based on their southerly location, topographic heterogeneity and ecological diversity, presence of extensive coastlines with associated resource availability, favourable weather patterns and the archaeological record (Carrión *et al.* 2008; Finlayson and Giles Pacheco 2000; Finlayson *et al.* 2006; Jennings *et al.* 2011).

The core area population sources are implemented with *Hominin Factories*. Factories, when conditions are right, will produce as many new groups as the surrounding area will sustain. Factories will check local conditions before production and will make sure that:

- There are no hungry groups in the area around the factory location (this area is based on the Factory *scanning range* and configurable, by default 30 cells);

- There are sufficient resources in the area surrounding the factory for a new group (estimated by adding together all the resources within the scanning range);
- There are empty cells where the new group can be located.

When these conditions are met new groups are created at random empty grid cells in the area. Local populations are thus replenished from more productive areas when the situation improves (MacDonald *et al.* 2012). All new hominin groups are considered to be in optimum condition and consist of the default population structure: 8 pre-fertile size hominins, 11 fertile hominins and 6 post-fertile hominins. These numbers are configurable. New groups immediately start moving and consuming, possibly in direct competition with other groups nearby.

Areas outside the simulation area can accept groups when the option 'Open borders' is selected. 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. However, during the development of HomininSpace it appeared that the open borders of Iberia always absorbed most of the groups present. Especially the dynamic groups, continuously moving towards the highest energy levels, all moved towards the south and disappeared from the simulation. This also occurred when conditions were relatively mild. Therefore, in the simulations the 'Open border' option is not selected. This mimics refugia that continuously contain maximum population densities and only serve as sources. Most competition for resources occurs therefore at the threshold of the border areas.

## 9.5 Neutral models - levelling the playing field

The following model elements implement when applied in succession an increasing more neutral or zero model. In a zero model all elements of the model have equal impact on the results, excluding agent strategies and preferences (Brantingham 2003, 491). Such a model is used to test if there is a bias or error in the implementation of certain model elements like the environment. For instance if all groups move south when there is no energy level differences this could indicate an error in the implementation of the model. A zero model can also be used as the neutral data pattern to be tested against to quantify or to qualify the effect of individual parameters.

When developing a simulation the global patterns that become visible while executing the model (in the case of HomininSpace the distribution of hominins through time and space) should match general expectations derived from reality. For instance, it can be expected

with resource availability in mind that hominins move towards the south when conditions deteriorate, and return when conditions improve. Such expected patterns are used when developing a simulation where deviations from expected behaviour are closely investigated. The following global patterns were observed in developing HomininSpace:

- General movement of the population according to resource availability;
- Patterns of death and visit density to identify hotspots and match these against the available resources (for instance, this identified the death traps, singular localities like islands with high resource availability but hardly any surrounding carrying capacity that attracted many hominin groups that subsequently perished);
- Resource availability and distribution through the landscape through time;
- Birth statistics versus resource availability; in general, if more resources become available it is expected that more hominins are being born.

The environment is for a large part defined by the topography of the area. What would happen if there is no topography at all (Subsection 9.5.1)? That means no water cells and no mountains. It is also possible to create an equal distribution of the energy by simply assigning each grid cell an exact same amount of kcal (Subsection 9.5.2). And suppose the hominins do not search for the best resource patches, but instead move randomly through the environment, foraging where chance brings them? Subsection 9.5.3 implements random movement for groups of hominins.

### 9.5.1 Implementing topographical differences in the landscape

It is possible to turn all water grid cells into land masses with the setting ZERO_MODEL_DISABLE_WATER. A value of True will effectively remove all water from the grid, turning them into accessible landmasses. This levels the playfield by disabling the restrictive element of water cells. Note that height is not a factor of importance for the modelled hominins. This means that hominins will not be restricted in their movement by water bodies like the Mediterranean Sea or the Channel. Temperature and precipitation values will be distributed across the former water grid cells and used in reconstructing energy levels (note that water grid cells if present have zero energy). Without water there are also no beach areas. This setting has only been used to verify that actual water grid cells do limit hominin movement. With this setting activated it can be observed that groups actually move onto former water grid cells. This also verifies the functionality of the lowering sea levels that expose new accessible land.

### 9.5.2 Distributing energy equally throughout the grid

With the setting ZERO_MODEL_DISABLE_ENERGY set to True all energy level reconstructing is turned off. Effectively an equal amount of energy is assigned to all (land) grid cells, and this amount replenishes every year. Energy differences can thus no longer steer hominin groups in the landscape, and movement will resemble random movement. This setting has been used to test if other factors besides resource availability influence the directional decisions of the groups. For instance to check if the shape of the topography (land masses) steer hominins in certain directions.

### 9.5.3 Randomly walking hominins

Food in the landscape can be obtained following an optimal foraging based strategy (Belovsky 1988). Groups head for those areas that can give them the most resources. Opposing this would be a random walk algorithm for hominin groups. Comparison against results from these simulations allows assessment if more complexity in decision making add value to the model. When the setting ZERO_MODEL_RANDOM_WALK is activated (True), hominin group agents select their next destination grid cell randomly from the list with available grid cells they can move to. The choice where to go next is implemented in the HomininGroup.move() routine. Within this routine first a list of grid cells that can be reached is collected, and then they are randomly ordered using the default uniform distribution ("HomininGroup.java", line 209):

```
SimUtilities.shuffle(theNeighborhoodThatCanBeReached,
RandomHelper.getUniform());
```

This is opposed to (setting value False) selecting the grid cell with the most available energy as the best destination, implementing this algorithm ("HomininGroup.java", line 256-266):

```
// potential target area to move to
// to avoid death traps we compare cells and their immediate
environments with each other, not just cells
double cap = h.getAvailableCalories() + h.getEnvironmentCalories();
cap = cap / (1 + h.getForagingGroupCount());
if (cap > bestCapacity)
{
      bestHabitatCell = h;
      bestCapacity = cap;
}
```

# PART FOUR: AND ACTION!

*"... for many years the objective of hunter-gatherer research has been to seek out the essential core of the hunter-gatherer lifeway and consequently to ignore or explain away variability as the product of extraordinary natural environments or particular historical circumstances."*

(Kelly 1995)

# 10. SIMULATION RESULTS

## 10.1 Overview

The HomininSpace simulation system has been designed to answer questions about hominin behaviour in a reconstructed environment while comparing simulation results with archaeological presence (and absence) data. This chapter presents the results for all simulations that were run in this study. The simulations were executed in scenarios, each addressing one or more questions by including or excluding certain model elements. Settings indicate which elements are used. Each scenario first ran the same 1500 simulations, with the same parameter values (including the value for the random seed) in the Standard set. This allows effect assessment of the settings on the simulation score. Then the Genetic Algorithm for each scenario selects individuals and uses these to create new parameter value sets that are subsequently executed until the stop criterion is reached.

A scenario is given a readable name that is used in the analysis and the discussion. These names start with either Habitat or Energy, followed by a single character (for instance, Habitat-A or Energy-C). Scenarios are presented in alphabetical order. Each scenario also has an identifier that lists all the model settings in one string of boolean values (e.g. FTFTFFFFFF). Thus the setting combination FTFTFFFFFF implements a habitat reconstruction with static hominins with a maximized foraging range and with all other settings deactivated. To summarize the contents of each scenario, a descriptive name using the True values of the settings is also provided (for FTFTFFFFFF this is HabitatStaticMaxrange). Table 25 presents an overview of all the scenarios, with identifier, descriptive name, the number of simulations that were run, the maximum simulation score for the Standard set and for all simulations together.

**Table 25: Summary of the simulated scenarios.**

| *Identifier* | *Readable and descriptive names* | *#sims* | *Max Standard* | *Max score* |
|---|---|---|---|---|
| **FFFFFFFFF** | Habitat-A (HabitatDynamic) | 2,148 | 28434 | 30436 |
| **FFFFFTFFF** | Habitat-B (HabitatDynamicAbsence) | 3,102 | 13432 | 13758 |
| **FFFTFFFFF** | Habitat-C (HabitatDynamicMaxrange) | 2,008 | 28532 | 30206 |
| **FFFTFTTFFF** | Habitat-D (HabitatDynamicMaxrangeCoreAbsence) | 2,618 | 18521 | 19168 |
| **FFTTFTFFFF** | Habitat-E (HabitatDynamicCoastalMaxrangeCore | 2,420 | 26370 | 27948 |
| **FFTTFTTFFF** | Habitat-F | 2,240 | 17400 | 17486 |

| | | | | |
|---|---|---|---|---|
| | (HabitatDynamicCoastalMaxrangeCore Absence) | | | |
| **FTFFFFFFFF** | Habitat-G (HabitatStatic) | 2,043 | 30395 | 31016 |
| **FTFFTTFFFF** | Habitat-H (HabitatStaticCrosswaterCore) | 2,192 | 30351 | 31119 |
| **FTFTFFFFFF** | Habitat-I (HabitatStaticMaxrange) | 2,355 | 30123 | 31142 |
| **FTTTFFFFFF** | Habitat-J (HabitatStaticCoastalMaxrange) | 2,210 | 29858 | 30954 |
| **TFFFFFFFFF** | Energy-A (EnergyDynamic) | 2,089 | 28108 | 30522 |
| **TFTTFFFFFF** | Energy-B (EnergyDynamicCoastalMaxrange) | 2,278 | 25249 | 29896 |
| **TFTTFFTFFF** | Energy-C (EnergyDynamicCoastalMaxrangeAbsence) | 2,118 | 4766 | 6273 |
| **TFTTTFFFFF** | Energy-D (EnergyDynamicCoastalMaxrangeCrosswater) | 2,012 | 25249 | 26112 |
| **TTTTFFFFFF** | Energy-E (EnergyStaticCoastalMaxrange) | 2,176 | 30035 | 30651 |
| **TFTTFFFFFT** | Energy-BR (Scenario B + Randomwalk) | 2,127 | 26286 | 28125 |
| **TFTTFFTFFT** | Energy-CR (Scenario C + Randomwalk) | 2,079 | 3301 | 3301 |
| **TTTTFFFFFT** | Energy-ER (Scenario E + Randomwalk) | 2,037 | 29262 | 30439 |
| | | | | |
| 18 scenarios | **Total number of simulations:** | **40,252** | | |

For this study a total of 40,252 simulations were executed. Each run simulated 81,000 years resulting in data for a total of almost 3.3 billion modelled Neanderthal years. And this includes neither the many simulations that were run during development of the tool nor those that were executed in HomininSpace 1.0, for which this study is a successor.

For each scenario one batch of simulations was executed (with an additional duplication batch for scenario Energy-A). Each simulation in a batch has a unique number and a unique set of model parameter values. The results of each simulation are added to a combined data file per batch. A data file with all simulation results for each batch is included in the Supplementary Materials. The first 1500 parameter value combinations for every batch are referred to as the Standard set, and they are the same for all batches. When presented in figures or tables data for this set is always coloured cyan. Data from the Standard set is recombined to create additional parameter value combinations. These subsequent simulations are referred to as the Evolved parameter value set, and in figures and tables coloured in orange.

In the next sections results are presented per scenario, where for each scenario the following information is presented:

- The name of the scenario, the values for the settings and an informal description of the model elements that were included;
- A figure with the matchedIntervalCoverage score for all simulations in the batch, with results from the Standard set coloured cyan and the Evolved set orange;

- A table with the three best scoring parameter value sets from the Standard set in this scenario (again coloured in cyan), followed by the three best scoring parameter value sets from all simulations in the batch (in orange). In general, these latter are from the Evolved set, but occasionally a parameter combination from the Standard set performed so well that even after all improvements by the GA algorithm it still belongs to the top three scores overall.

### 10.1.1 Design of the Experiments (DOE)

There are ten different settings that control what model elements are included in the experiments using Boolean values to include or exclude those parts of the model. Even though three of these were mainly used in system development this leaves seven Boolean values or 128 different possible setting combinations. Since each setting combination requires at least 2000 simulations it is computationally impossible to do all combinations. A selection of settings in numbers and combinations has been made. Table 26 presents an overview of the usage per setting in the selected scenarios. The aim was to include each setting at least once.

**Table 26: Overview of the setting values in all scenarios.**

| *Setting name* | *Setting is True in the following scenarios* |
|---|---|
| **ENERGY_CONTINUOUS** | Energy-A through Energy-E and in Energy-BR, Energy-CR, Energy-ER. |
| **STATIC_DISPERSAL** | Habitat-G through Habitat-J and in Energy-E and Energy-ER. |
| **USE_BEACHES** | Habitat-E, Habitat-F, Habitat-J, Energy-B through Energy-E, and Energy-BR, Energy-CR and Energy-ER. |
| **USE_MAXIMUM_FORAGING_RANGE** | Habitat-C through Habitat-F, Habitat-I and Habitat-J. Energy-B through Energy-E and Energy-BR, Energy-CR and Energy-ER. |
| **GROUPS_CAN_CROSS_WATER** | Habitat-H, Energy-D. |
| **USE_FACTORIES** | Habitat-D, Habitat-E, Habitat-F, Habitat-H. |
| **DEATH_PENALTY_FOR_ABSENCE** | Habitat-B, Habitat-D, Habitat-F, Energy-C. |
| **ZERO_MODEL_DISABLE_WATER** | -- |
| **ZERO_MODEL_DISABLE_ENERGY** | -- |
| **ZERO_MODEL_RANDOM_WALK** | Energy-BR, Energy-CR and Energy-ER. |

## 10.2 Scenario Habitat-A (HabitatDynamic)

This is the baseline simulation scenario against which others can be compared. The settings are all deactivated: FFFFFFFFFF. Therefore it features habitat reconstruction and dynamic hominins. The results for the MatchedIntervalCoverage variable are shown in Figure 49, where for each simulation a dot is shown in the graph, in cyan for the Standard set and orange for the Evolved. As expected the Evolved models score generally high matchedIntervalCoverage values. The top three scores for the Standard (cyan) and for the GA Evolved (orange) parameter sets are given in Table 27. In total there are 2148 simulations that are used from this batch[40]. The maximum score for the Standard set is 28434, for the Evolved set it is 30436, an improvement of 7%. Note that the use of a maximum foraging range is deactivated and all values for the maximum foraging range parameter are thus random and not involved in the GA improvement.

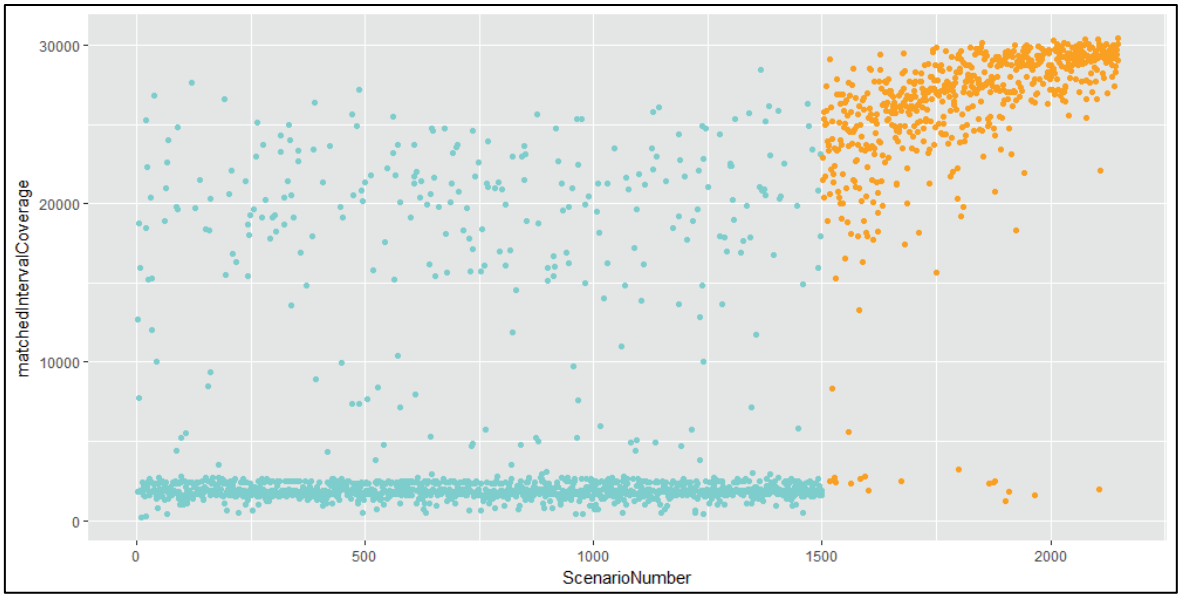



**Figure 49: Simulation results for scenario Habitat-A.**

**Table 27: Best three simulation results and parameter values for scenario Habitat-A.**

| Settings: FFFFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 28434 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 120 | 27636 | 43 | 1 | 1 | 12 | 3750 | 3250 | 5000 | 9 | 1 | 4 | 96 | -29 | 8 | 32 | 2850 | 6 | 62186 |
| Standard 3 | 486 | 27201 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2148 | 30436 | 49 | 1 | 1 | 12 | 2000 | 3000 | 5500 | 9 | 1 | 1 | 16 | -23 | 13 | 41 | 3150 | 6 | 49118 |
| Evolved 2 | 2076 | 30380 | 45 | 1 | 1 | 12 | 3000 | 2000 | 4750 | 2 | 4 | 2 | 16 | -22 | 9 | 37 | 3500 | 14 | 76410 |
| Evolved 3 | 2103 | 30380 | 48 | 4 | 1 | 12 | 2000 | 2000 | 4750 | 8 | 1 | 1 | 16 | -21 | 13 | 41 | 3500 | 14 | 14905 |

[40] Seven simulation numbers were not assigned in this batch: 1984-1990.

## 10.3 Scenario Habitat-B (HabitatDynamicAbsence)

The settings for this scenario: FFFFFFTFFF. It features habitat reconstruction and dynamic hominins with absence activated. The results for the MatchedIntervalCoverage variable are shown in Figure 50. The top three scores for the Standard (cyan) and for the GA Evolved (orange) are given in Table 28. In total there are 3264 simulations. The maximum score for Standard is 13432, for Evolved 13145 when following the rule to stop after 100 iterations without improvement. This is actually less than the standard set, a difference of minus 2%. If the rule above is augmented to include the condition that the results should actually be better than the Standard set the first individual to comply is 3102, with a score of 13758, an improvement of less than plus 2%. A very substantial improvement of 15.6% is actually found in simulation 3264, with a value of 15534 (included in the table). After all simulations #3 from the Standard set is still in the top three. Note that the usage of a maximum foraging range is not activated in this scenario.




**Figure 50: Simulation results for scenario Habitat-B.**

**Table 28: Best three simulation results and parameter values for scenario Habitat-B.**

| FFFFFTFFF | ScenarioNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 3 | 13432 | 36 | 5 | 6 | 8 | 4000 | 3750 | 3000 | 3 | 2 | 4 | 76 | -29 | 20 | 37 | 2850 | 3 | 30 |
| Standard 2 | 1186 | 12722 | 43 | 4 | 11 | 2 | 3250 | 2500 | 2250 | 6 | 1 | 3 | 96 | -15 | 13 | 37 | 3200 | 13 | 38 |
| Standard 3 | 160 | 11799 | 46 | 15 | 6 | 7 | 2500 | 3750 | 2000 | 2 | 3 | 3 | 126 | -19 | 11 | 40 | 2550 | 2 | 29 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 3264 | 15534 | 39 | 9 | 3 | 8 | 2250 | 4500 | 4000 | 8 | 2 | 2 | 16 | -10 | 15 | 28 | 3200 | 8 | 83 |
| Evolved 2 | 3102 | 13758 | 39 | 9 | 4 | 8 | 2475 | 4500 | 4000 | 8 | 2 | 2 | 16 | -10 | 15 | 28 | 3200 | 7 | 43 |
| Evolved 3 | 3 | 13432 | 36 | 5 | 6 | 8 | 4000 | 3750 | 3000 | 3 | 2 | 4 | 76 | -29 | 20 | 37 | 2850 | 3 | 30 |

## 10.4 Scenario Habitat-C (HabitatDynamicMaxrange)

The settings for this scenario are: FFFTFFFFFF. It features habitat reconstruction and dynamic hominins, with the usage of a maximum foraging range activated. The results for the MatchedIntervalCoverage variable are shown in Figure 51. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 29. In total there are 2008 simulations that are used from this batch. The maximum score for the Standard set is 28532, for the Evolved set it is 30206, an improvement of 5.9%.

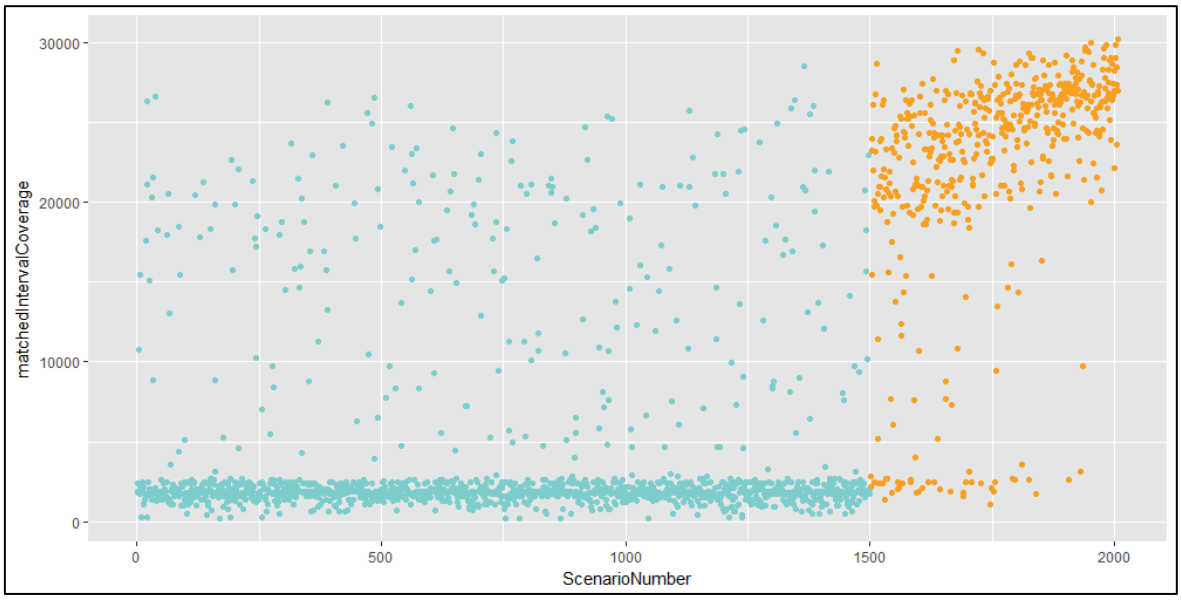



**Figure 51: Simulation results for scenario Habitat-C.**

**Table 29: Best three simulation results and parameter values for scenario Habitat-C.**

| Settings: FFFTFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 28532 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 39 | 26593 | 47 | 5 | 1 | 11 | 2750 | 3250 | 3750 | 5 | 5 | 2 | 56 | -28 | 12 | 23 | 3000 | 10 | 4053 |
| Standard 3 | 486 | 26542 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2008 | 30206 | 50 | 1 | 1 | 12 | 3500 | 3250 | 5000 | 8 | 1 | 1 | 16 | -22 | 9 | 33 | 2750 | 14 | 138160 |
| Evolved 2 | 1952 | 30001 | 43 | 3 | 1 | 2 | 2000 | 2475 | 3250 | 7 | 4 | 2 | 14 | -25 | 9 | 37 | 3150 | 15 | 30712 |
| Evolved 3 | 2002 | 29877 | 49 | 3 | 1 | 8 | 2500 | 4000 | 4675 | 4 | 4 | 2 | 14 | -25 | 9 | 37 | 2750 | 10 | 73985 |

## 10.5 Scenario Habitat-D (HabitatDynamicMaxrangeCoreAbsence)

The settings for this scenario: FFFTFTTFFF. It features habitat reconstruction and dynamic hominins, with maximum foraging range, core areas and absence conditions activated. The results for the MatchedIntervalCoverage variable are shown in Figure 52. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 30. In total there are 2618 simulations that are used from this batch. The maximum score for the Standard set is 18521, for the Evolved set it is 19168, an improvement of 3.5%.

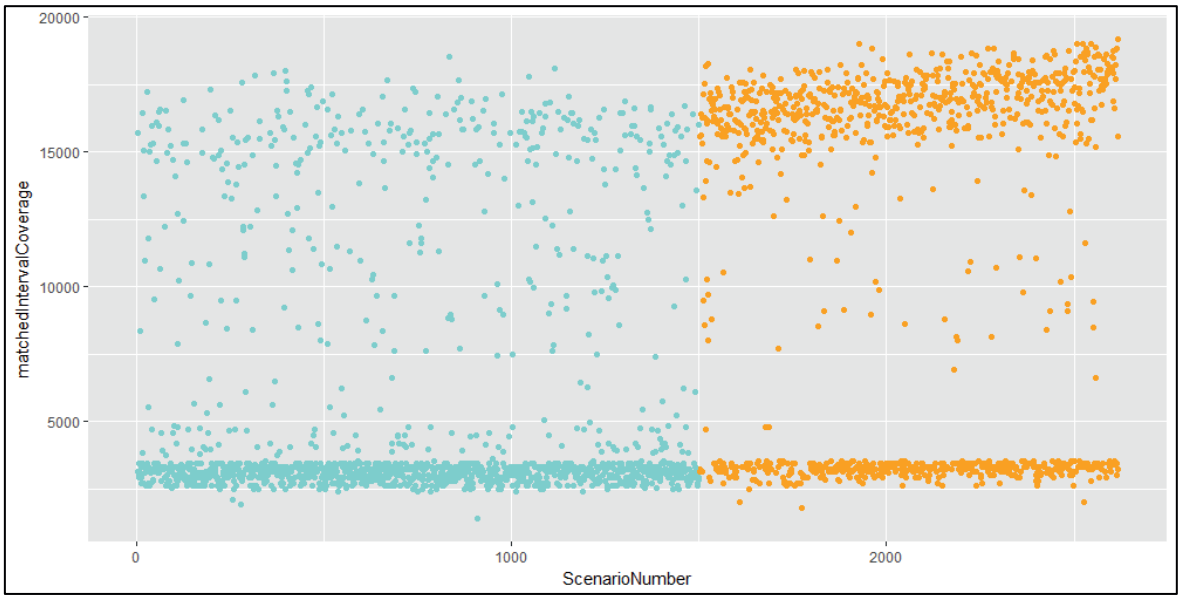



**Figure 52: Simulation results for scenario Habitat-D.**

**Table 30: Best three simulation results and parameter values for scenario Habitat-D.**

| Settings:FFFTFTTFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 832 | 18521 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3000 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |
| Standard 2 | 1115 | 18110 | 38 | 12 | 5 | 11 | 4250 | 4000 | 2750 | 1 | 3 | 5 | 26 | -16 | 18 | 20 | 2950 | 5 | 1 |
| Standard 3 | 395 | 18012 | 38 | 4 | 3 | 4 | 2750 | 5000 | 4000 | 6 | 1 | 4 | 16 | -26 | 14 | 17 | 3300 | 15 | 169 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2618 | 19168 | 1 | 4 | 1 | 1 | 3000 | 2750 | 3267 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 9 |
| Evolved 2 | 2508 | 19024 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3267 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |
| Evolved 3 | 2523 | 19024 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3267 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |

## 10.6 Scenario Habitat-E (HabitatDynamicCoastalMaxrangeCore)

The settings for this scenario are FFTTFTFFFF. It features habitat reconstruction and dynamic hominins, with coastal resources, a maximum foraging range and core areas activated. The results for the MatchedIntervalCoverage variable are shown in Figure 53. The top three scores for the Standard (cyan) and for the GA Evolved (orange) parameter value sets are given in Table 31. In total there are 2420 simulations that are used from this batch. The maximum score for the Standard set is 26370, for the Evolved set it is 29153, an improvement of 10.6%.

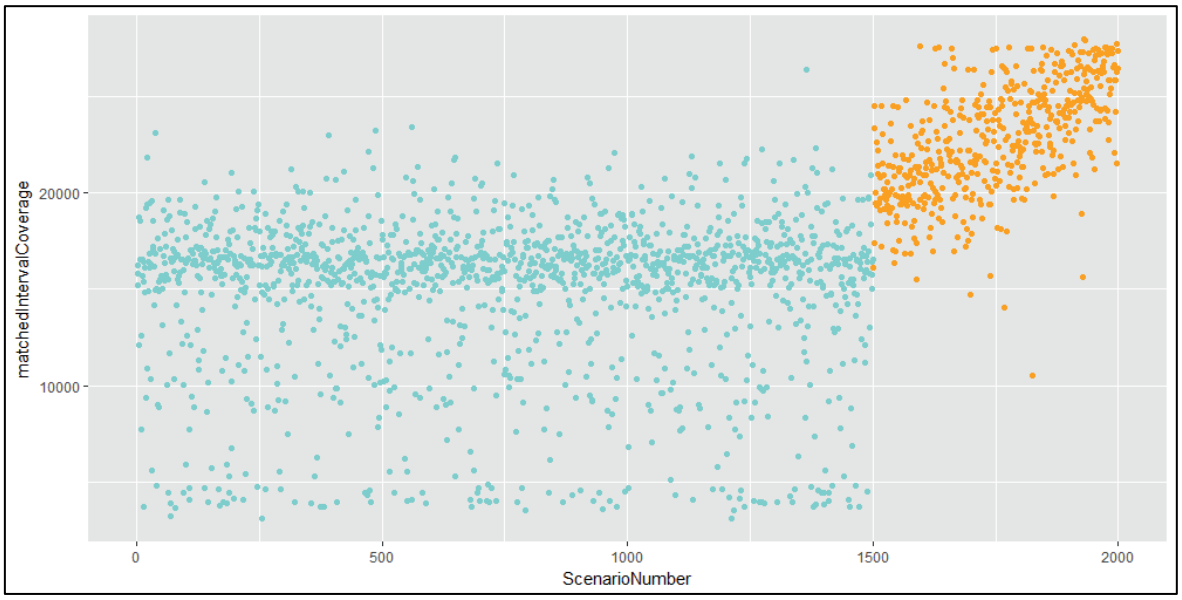



**Figure 53: Simulation results for scenario Habitat-E.**

**Table 31: Best three simulation results and parameter values for scenario Habitat-E.**

| Settings: FFTTFTFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 26370 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 560 | 23422 | 27 | 3 | 1 | 15 | 2500 | 4000 | 3500 | 8 | 4 | 2 | 16 | -16 | 14 | 25 | 3100 | 9 | 301 |
| Standard 3 | 486 | 23216 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2420 | 29153 | 50 | 1 | 1 | 11 | 2500 | 3250 | 3250 | 2 | 4 | 1 | 14 | -29 | 8 | 33 | 3100 | 15 | 207291 |
| Evolved 2 | 2388 | 29053 | 50 | 1 | 1 | 11 | 2500 | 3250 | 3250 | 1 | 4 | 1 | 16 | -23 | 8 | 33 | 3410 | 15 | 207291 |
| Evolved 3 | 2322 | 29050 | 50 | 1 | 1 | 11 | 2500 | 3250 | 3250 | 1 | 4 | 1 | 16 | -26 | 8 | 33 | 3410 | 15 | 207291 |

## 10.7 Scenario Habitat-F (HabitatDynamicCoastalMaxrangeCoreAbsence)

The settings for this scenario: FFTTFTTFFF. It features habitat reconstruction and dynamic hominins, with coastal resources, a maximum foraging range, core areas and forces adherence to the absence condition. The results for the MatchedIntervalCoverage variable are shown in Figure 54. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 32. In total there are 2240 simulations that are used from this batch. The maximum score for the Standard set is 17400, for the Evolved set it is 17486, with almost no improvement of 0.5%.

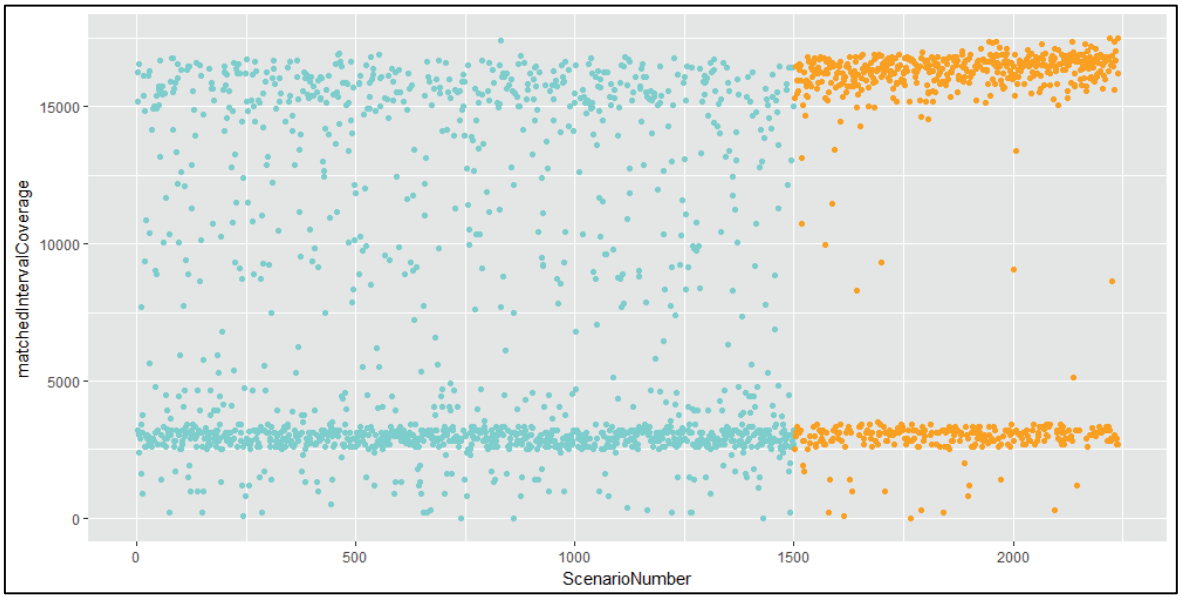



**Figure 54: Simulation results for scenario Habitat-F.**

**Table 32: Best three simulation results and parameter values for scenario Habitat-F.**

| Settings: FFTTFTTFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 832 | 17400 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3000 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |
| Standard 2 | 461 | 16919 | 9 | 2 | 1 | 4 | 4250 | 4500 | 2000 | 8 | 5 | 5 | 16 | -30 | 12 | 29 | 3350 | 9 | 18 |
| Standard 3 | 554 | 16915 | 30 | 6 | 10 | 14 | 2500 | 3000 | 3000 | 9 | 2 | 1 | 96 | -18 | 18 | 23 | 3500 | 9 | 0 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2240 | 17486 | 1 | 6 | 1 | 1 | 3000 | 2750 | 3000 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |
| Evolved 2 | 2220 | 17477 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3000 | 8 | 0 | 1 | 6 | -20 | 11 | 22 | 3200 | 6 | 8 |
| Evolved 3 | 832 | 17400 | 1 | 5 | 1 | 1 | 3000 | 2750 | 3000 | 7 | 1 | 1 | 6 | -20 | 10 | 25 | 3200 | 6 | 8 |

## 10.8 Scenario Habitat-G (HabitatStatic)

The settings for this scenario are FTFFFFFFFF. It features habitat reconstruction and static hominins. The results for the MatchedIntervalCoverage variable are shown in Figure 56. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 34. In total there are 2043 simulations that are used from this batch. The maximum score for the Standard set is 30395, for the Evolved set it is 31016, an improvement of 2%. The use of a maximum foraging range is deactivated and all values for maximum foraging range are random.

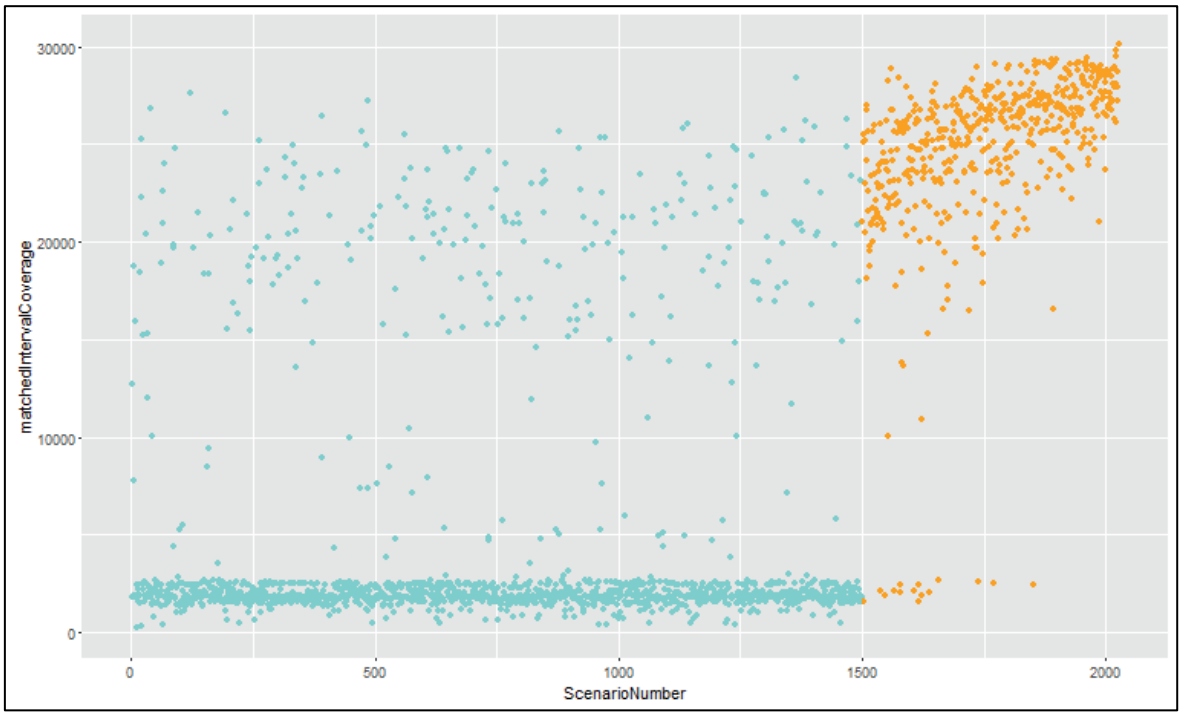



**Figure 55: Simulation results for scenario Habitat-G.**

**Table 33: Best three simulation results and parameter values for scenario Habitat-G.**

| Settings: FTFFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 333 | 30395 | 38 | 6 | 1 | 13 | 4500 | 4250 | 4250 | 1 | 1 | 1 | 66 | -25 | 13 | 27 | 2500 | 3 | 917 |
| Standard 2 | 486 | 30365 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| Standard 3 | 120 | 30340 | 43 | 1 | 1 | 12 | 3750 | 3250 | 5000 | 9 | 1 | 4 | 96 | -29 | 8 | 32 | 2850 | 6 | 62186 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2043 | 31016 | 49 | 1 | 1 | 12 | 3750 | 2000 | 4250 | 5 | 2 | 0 | 46 | -23 | 13 | 32 | 3245 | 7 | 34246 |
| Evolved 2 | 2021 | 30961 | 43 | 1 | 1 | 12 | 2200 | 4000 | 4675 | 5 | 2 | 1 | 16 | -28 | 8 | 36 | 3300 | 4 | 75658 |
| Evolved 3 | 2031 | 30961 | 43 | 1 | 1 | 12 | 2200 | 4000 | 4675 | 5 | 2 | 1 | 16 | -31 | 8 | 36 | 3300 | 4 | 75658 |

## 10.9 Scenario Habitat-H (HabitatStaticCrosswaterCore)

The settings for this scenario are FTFFTTFFFF. It features habitat reconstruction and static hominins who are able to cross larger water bodies and with core areas activated. The results for the MatchedIntervalCoverage variable are shown in Figure 56. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 34. In total there are 2192 simulations that are used from this batch. The maximum score for the Standard set is 30351, for the Evolved set it is 31119, an improvement of 2%. The use of a maximum foraging range is deactivated and all values for maximum foraging range are random, clearly illustrated by the model parameter values for Evolved 1 and Evolved 2 where the only difference is the value for the foraging range.

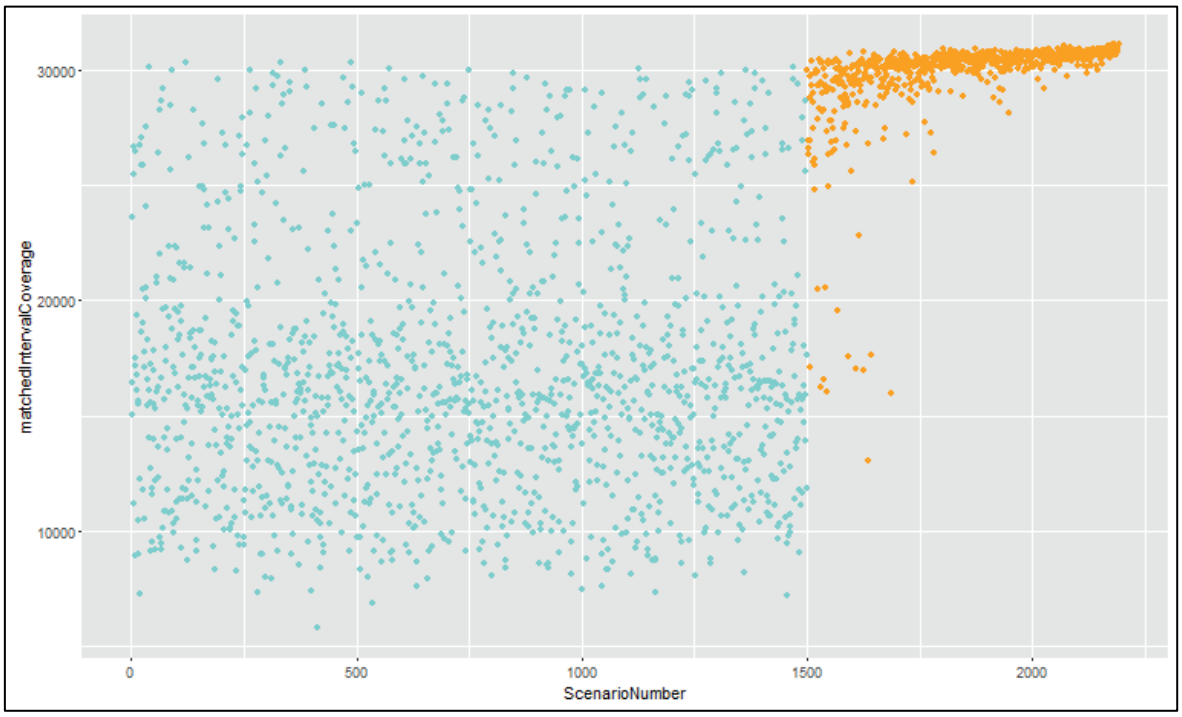



**Figure 56: Simulation results for scenario Habitat-H.**

**Table 34: Best three simulation results and parameter values for scenario Habitat-H.**

| Settings: FTFFTTFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 486 | 30351 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| Standard 2 | 120 | 30346 | 43 | 1 | 1 | 12 | 3750 | 3250 | 5000 | 9 | 1 | 4 | 96 | -29 | 8 | 32 | 2850 | 6 | 62186 |
| Standard 3 | 333 | 30331 | 38 | 6 | 1 | 13 | 4500 | 4250 | 4250 | 1 | 1 | 1 | 66 | -25 | 13 | 27 | 2500 | 3 | 917 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2180 | 31119 | 48 | 1 | 1 | 7 | 4500 | 2250 | 2250 | 1 | 1 | 1 | 56 | -26 | 10 | 40 | 3500 | 15 | 98065 |
| Evolved 2 | 2192 | 31119 | 48 | 1 | 1 | 7 | 4500 | 2250 | 2250 | 1 | 1 | 1 | 56 | -26 | 10 | 40 | 3500 | 13 | 98065 |
| Evolved 3 | 2181 | 31113 | 42 | 1 | 1 | 7 | 2475 | 2025 | 3442 | 1 | 1 | 1 | 66 | -28 | 8 | 28 | 3500 | 13 | 41872 |

## 10.10 Scenario Habitat-I (HabitatStaticMaxrange)

The settings for this scenario: FTFTFFFFFF. It features habitat reconstruction and static hominins with a maximum foraging range activated. The results for the MatchedIntervalCoverage variable are shown in Figure 57. The top three scores for the Standard (cyan) and for the GA Evolved (orange) parameter set are given in Table 35. In total there are 2355 simulations that are used from this batch. The maximum score for the Standard set is 30123, for the Evolved set it is 31142, an improvement of 3.4%.

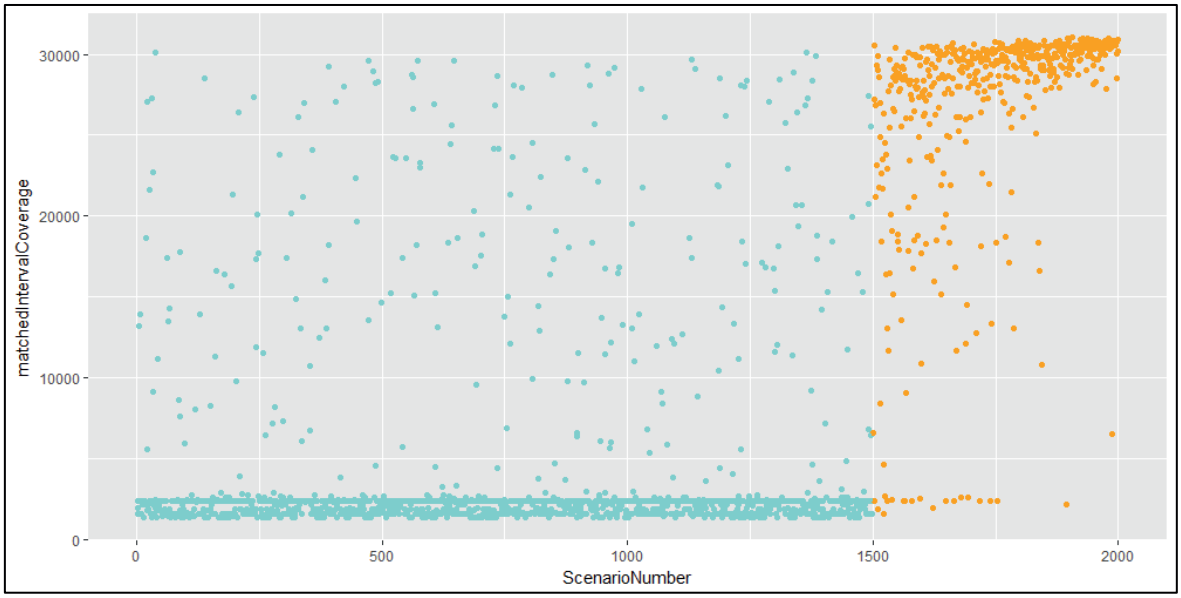



**Figure 57: Simulation results for scenario Habitat-I.**

**Table 35: Best three simulation results and parameter values for scenario Habitat-I.**

| Settings: FTFTFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 39 | 30123 | 47 | 5 | 1 | 11 | 2750 | 3250 | 3750 | 5 | 5 | 2 | 56 | -28 | 12 | 23 | 3000 | 10 | 4053 |
| Standard 2 | 1365 | 30123 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 3 | 1384 | 29927 | 41 | 11 | 1 | 14 | 4250 | 2750 | 4250 | 10 | 5 | 1 | 46 | -29 | 9 | 28 | 2900 | 11 | 1289 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2355 | 31142 | 47 | 1 | 1 | 9 | 3750 | 2750 | 4750 | 2 | 1 | 1 | 56 | -26 | 9 | 35 | 3795 | 15 | 97015 |
| Evolved 2 | 2293 | 31074 | 47 | 4 | 1 | 12 | 2500 | 2700 | 2500 | 6 | 1 | 1 | 50 | -27 | 7 | 35 | 3200 | 13 | 87097 |
| Evolved 3 | 2327 | 31069 | 47 | 1 | 1 | 11 | 2500 | 2750 | 4750 | 4 | 2 | 1 | 56 | -26 | 9 | 35 | 3795 | 15 | 96320 |

## 10.11 Scenario Habitat-J (HabitatStaticCoastalMaxrange)

The settings for this batch: FTTTFFFFFF. It features habitat reconstruction and static hominins, with coastal resources and a maximum foraging range. The results for the MatchedIntervalCoverage variable are shown in Figure 58. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 36. In total there are 2210 simulations that are used from this batch. The maximum score for the Standard set is 29858, for the Evolved set it is 30954, an improvement of 3.7%.

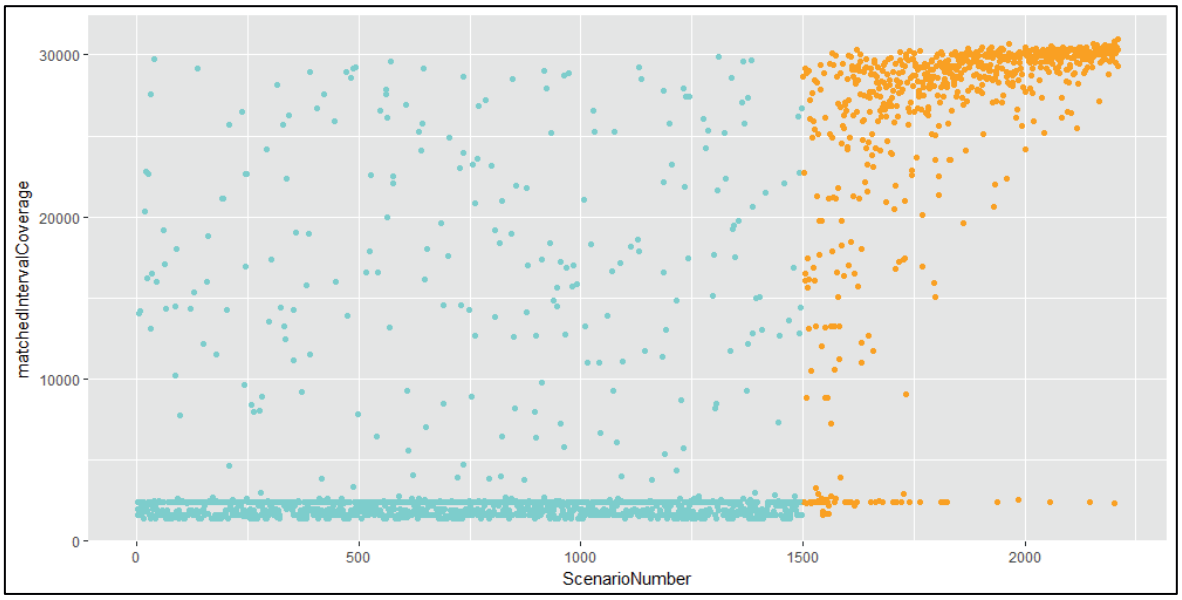



**Figure 58: Simulation results for scenario Habitat-J.**

**Table 36: Best three simulation results and parameter values for scenario Habitat-J.**

| Settings: FTTTFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1309 | 29858 | 30 | 1 | 3 | 6 | 3000 | 4250 | 5000 | 6 | 3 | 1 | 56 | -28 | 15 | 38 | 2700 | 10 | 692 |
| Standard 2 | 39 | 29737 | 47 | 5 | 1 | 11 | 2750 | 3250 | 3750 | 5 | 5 | 2 | 56 | -28 | 12 | 23 | 3000 | 10 | 4053 |
| Standard 3 | 1384 | 29636 | 41 | 11 | 1 | 14 | 4250 | 2750 | 4250 | 10 | 5 | 1 | 46 | -29 | 9 | 28 | 2900 | 11 | 1289 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2210 | 30954 | 35 | 5 | 1 | 9 | 2750 | 2250 | 3000 | 7 | 1 | 1 | 16 | -20 | 8 | 40 | 3100 | 14 | 6643 |
| Evolved 2 | 2201 | 30819 | 35 | 5 | 1 | 9 | 2750 | 2250 | 3000 | 1 | 1 | 1 | 62 | -23 | 8 | 40 | 3100 | 14 | 6643 |
| Evolved 3 | 2153 | 30712 | 31 | 5 | 1 | 9 | 3750 | 3750 | 2000 | 4 | 1 | 1 | 23 | -21 | 7 | 40 | 3300 | 14 | 4434 |

## 10.12 Scenario Energy-A (EnergyDynamic)

The settings for this scenario are coded as: TFFFFFFFFF. This scenario models the landscape using an energy reconstruction and has hominins with dynamic mobility. The results for the MatchedIntervalCoverage variable are shown in Figure 60. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 38. In total there are 2089 simulations that are used from this batch. The maximum score for the Standard set is 28108, for the Evolved set it is 30522, an improvement of 8.6%.

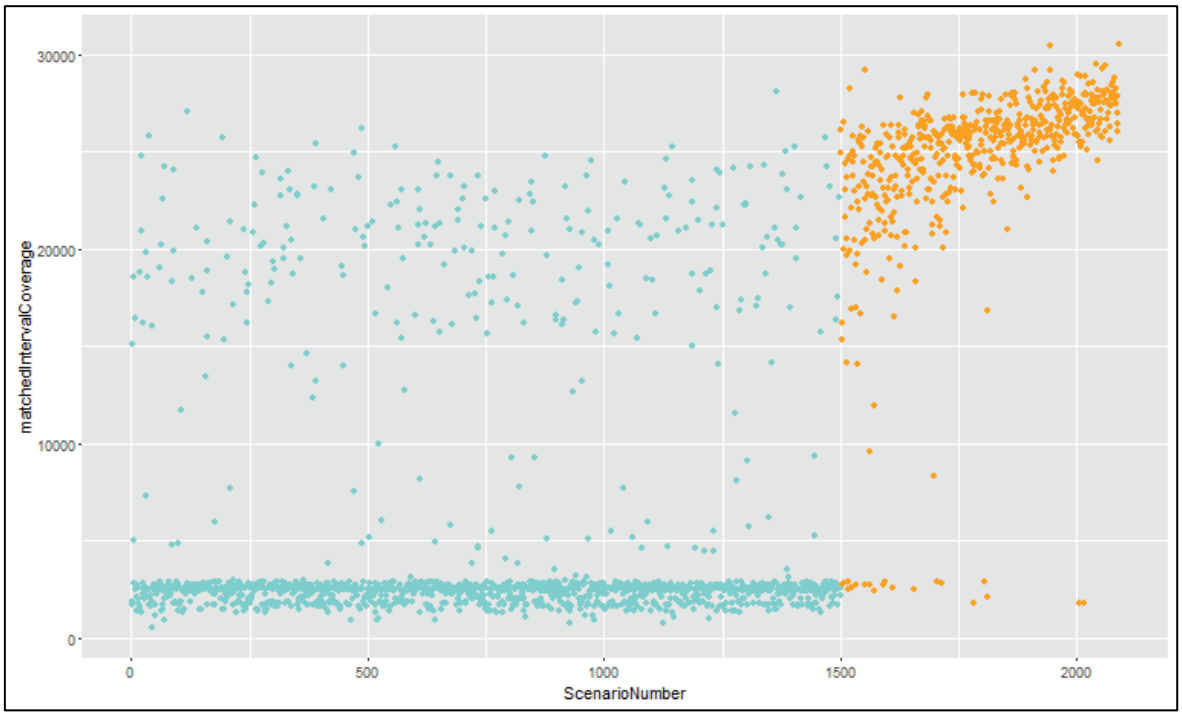



**Figure 59: Simulation results for scenario Energy-A.**

**Table 37: Best three simulation results and parameter values for scenario Energy-A.**

| Settings: TFFFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 28108 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 120 | 27034 | 43 | 1 | 1 | 12 | 3750 | 3250 | 5000 | 9 | 1 | 4 | 96 | -29 | 8 | 32 | 2850 | 6 | 62186 |
| Standard 3 | 486 | 26202 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2089 | 30522 | 43 | 1 | 1 | 9 | 2000 | 2750 | 3250 | 2 | 5 | 2 | 16 | -29 | 9 | 41 | 2700 | 12 | 64869 |
| Evolved 2 | 1945 | 30456 | 43 | 1 | 1 | 12 | 2000 | 3250 | 4500 | 2 | 4 | 4 | 16 | -29 | 9 | 32 | 3300 | 14 | 43815 |
| Evolved 3 | 2040 | 29481 | 44 | 1 | 1 | 12 | 2000 | 2700 | 3250 | 2 | 3 | 2 | 16 | -29 | 11 | 28 | 3190 | 6 | 22280 |

## 10.13 Scenario Energy-B (EnergyDynamicCoastalMaxrange)

The settings for this scenario are coded: TFTTFFFFFF. Simulations in this scenario feature energy level reconstruction and dynamic hominins, with coastal resources and a maximum foraging range. The results for the MatchedIntervalCoverage variable are shown in Figure 60. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 38. In total there are 2278 simulations that are used from this batch. The maximum score for the Standard set is 25249, for the Evolved set it is 29896, an improvement of 18.4%.

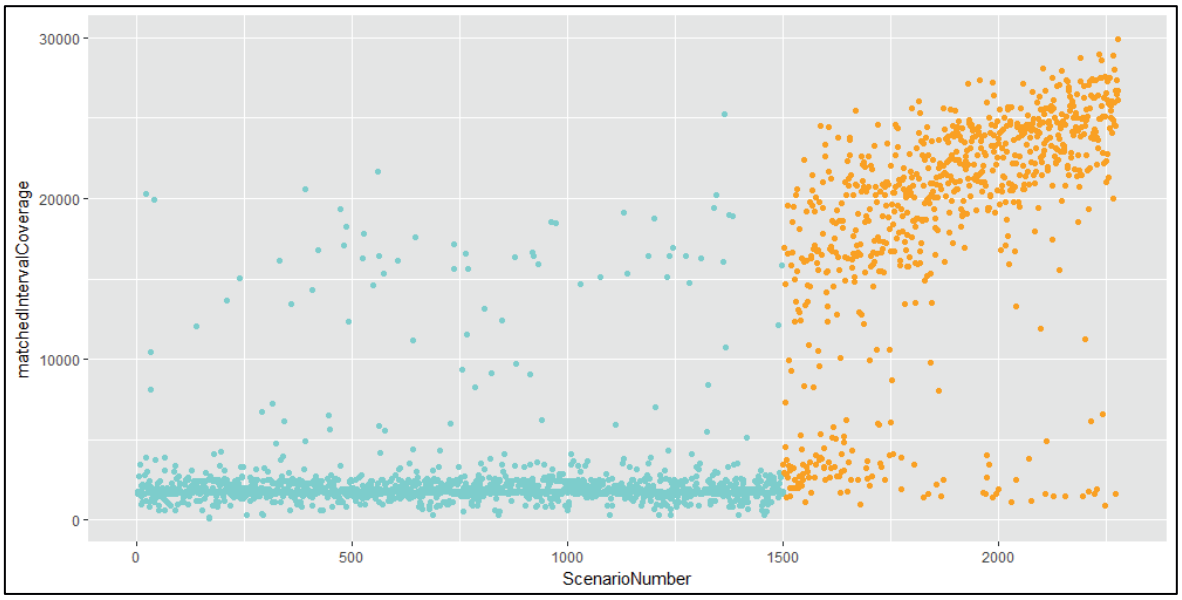



**Figure 60: Simulation results for scenario Energy-B.**

**Table 38: Best three simulation results and parameter values for scenario Energy-B.**

| Settings: TFTTFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 25249 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 560 | 21679 | 27 | 3 | 1 | 15 | 2500 | 4000 | 3500 | 8 | 4 | 2 | 16 | -16 | 14 | 25 | 3100 | 9 | 301 |
| Standard 3 | 390 | 20583 | 47 | 6 | 2 | 2 | 3250 | 4750 | 2000 | 1 | 5 | 5 | 56 | -29 | 9 | 38 | 3150 | 13 | 12842 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2278 | 29896 | 49 | 1 | 1 | 10 | 2000 | 2000 | 2000 | 8 | 1 | 0 | 6 | -26 | 9 | 32 | 3300 | 6 | 113555 |
| Evolved 2 | 2235 | 28946 | 49 | 7 | 1 | 9 | 2000 | 2000 | 2000 | 8 | 1 | 0 | 6 | -23 | 8 | 32 | 3300 | 10 | 26703 |
| Evolved 3 | 2265 | 28915 | 49 | 3 | 1 | 9 | 2000 | 2000 | 2000 | 8 | 1 | 0 | 6 | -23 | 14 | 35 | 3300 | 6 | 14941 |

## 10.14 Scenario Energy-C (EnergyDynamicCoastalMaxrangeAbsence)

The settings for this scenario are coded: TFTTFFTFFF. It features energy level reconstruction and dynamic hominins. Coastal resources and a maximum foraging range are also activated, and the absence condition is imposed. The results for the MatchedIntervalCoverage variable are shown in Figure 61. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 39. In total there are 2118 simulations that are used from this batch. The maximum score for the Standard set is 4766, for the Evolved set it is 6273, an improvement of 31.6%.

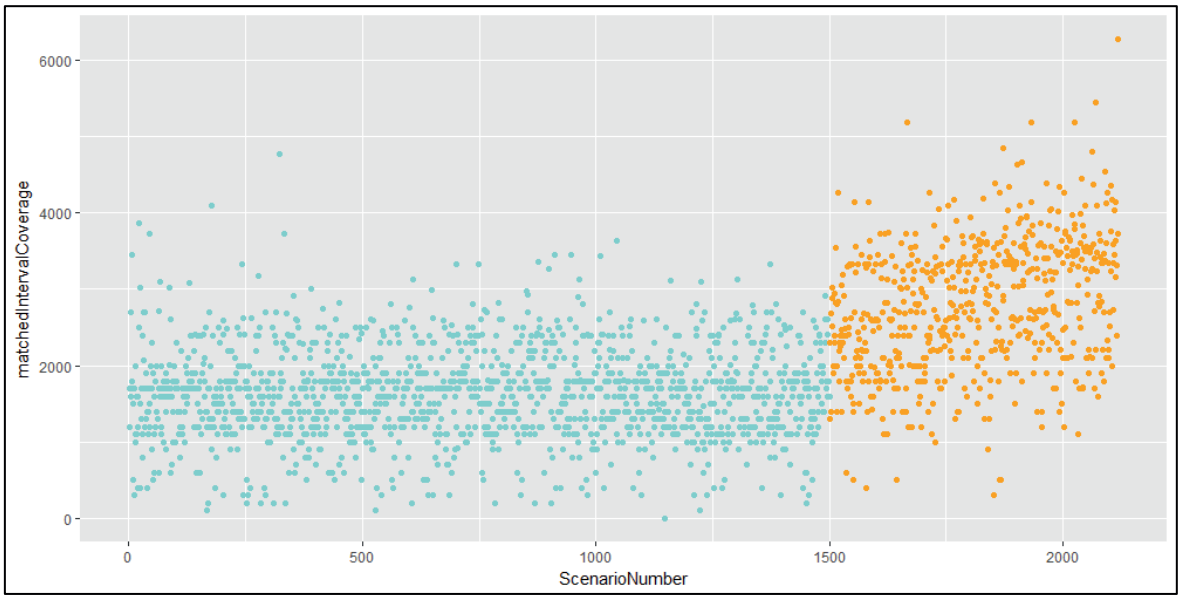



**Figure 61: Simulation results for scenario Energy-C.**

**Table 39: Best three simulation results and parameter values for scenario Energy-C.**

| Settings: TFTTFFTFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 323 | 4766 | 30 | 12 | 1 | 14 | 2500 | 3000 | 3000 | 8 | 2 | 1 | 136 | -27 | 9 | 21 | 3250 | 4 | 71 |
| Standard 2 | 178 | 4091 | 43 | 14 | 3 | 2 | 3250 | 3250 | 3000 | 7 | 5 | 2 | 16 | -8 | 20 | 40 | 2550 | 8 | 39 |
| Standard 3 | 22 | 3861 | 35 | 11 | 4 | 10 | 3750 | 2250 | 2000 | 1 | 3 | 4 | 46 | -26 | 9 | 39 | 3300 | 5 | 115 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2118 | 6273 | 22 | 1 | 1 | 12 | 2750 | 3250 | 4250 | 7 | 1 | 1 | 26 | -11 | 11 | 27 | 3050 | 3 | 427 |
| Evolved 2 | 2070 | 5445 | 22 | 1 | 5 | 14 | 2500 | 3250 | 3000 | 8 | 2 | 1 | 26 | -10 | 9 | 22 | 3250 | 4 | 31 |
| Evolved 3 | 1667 | 5186 | 30 | 12 | 1 | 14 | 2500 | 2700 | 3000 | 8 | 2 | 1 | 136 | -27 | 9 | 21 | 3250 | 4 | 71 |

## 10.15 Scenario Energy-D (EnergyDynamicCoastalMaxrangeCrosswater)

The settings for this scenario are coded: TFTTTFFFFF. These simulations feature energy level reconstruction and dynamic hominins, with coastal resources, a maximum foraging range and the hominins have the ability to cross larger water bodies. The results for the MatchedIntervalCoverage variable are shown in Figure 62. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 40. In total there are 2012 simulations that are used from this batch. The maximum score for the Standard set is 25249, for the Evolved set it is 26112, an improvement of 3.4%. Since only a few more simulations were included in the GA analyse than the more or less arbitrarily minimum of 500, more improvement can be expected with more simulations. And indeed with simulation 2265 a score of 27479 was observed (a substantial improvement of 8.8%, but far beyond the stop criterion of the genetic algorithm and thus not included in this research).

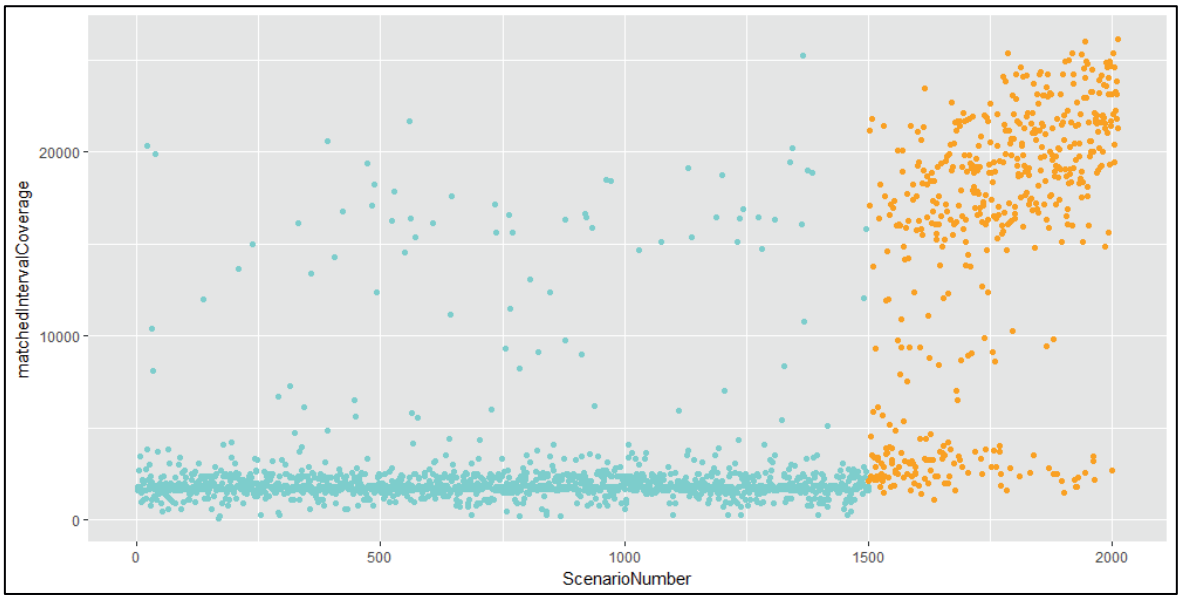



**Figure 62: Simulation results for scenario Energy-D.**

**Table 40: Best three simulation results and parameter values for scenario Energy-D.**

| Settings: TFTTTFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 25249 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 560 | 21679 | 27 | 3 | 1 | 15 | 2500 | 4000 | 3500 | 8 | 4 | 2 | 16 | -16 | 14 | 25 | 3100 | 9 | 301 |
| Standard 3 | 390 | 20583 | 47 | 6 | 2 | 2 | 3250 | 4750 | 2000 | 1 | 5 | 5 | 56 | -29 | 9 | 38 | 3150 | 13 | 12842 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2012 | 26112 | 48 | 1 | 1 | 6 | 2925 | 4500 | 2000 | 2 | 4 | 0 | 16 | -25 | 11 | 21 | 3410 | 13 | 22418 |
| Evolved 2 | 1945 | 25998 | 48 | 1 | 1 | 6 | 2925 | 4500 | 2000 | 2 | 4 | 2 | 16 | -17 | 14 | 38 | 3410 | 13 | 31629 |
| Evolved 3 | 2002 | 25382 | 43 | 3 | 1 | 10 | 2000 | 4500 | 3250 | 2 | 4 | 1 | 16 | -17 | 9 | 37 | 2700 | 12 | 28937 |

## 10.16 Scenario Energy-E (EnergyStaticCoastalMaxrange)

The settings for this scenario are coded: TTTTFFFFFF. It features energy level reconstruction and static hominins with coastal resources and a maximum foraging range. The results for the MatchedIntervalCoverage variable are shown in Figure 63. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 41. In total there are 2176 simulations that are used from this batch. The maximum score for the Standard set is 30035, for the Evolved set it is 30651, an improvement of 2%.

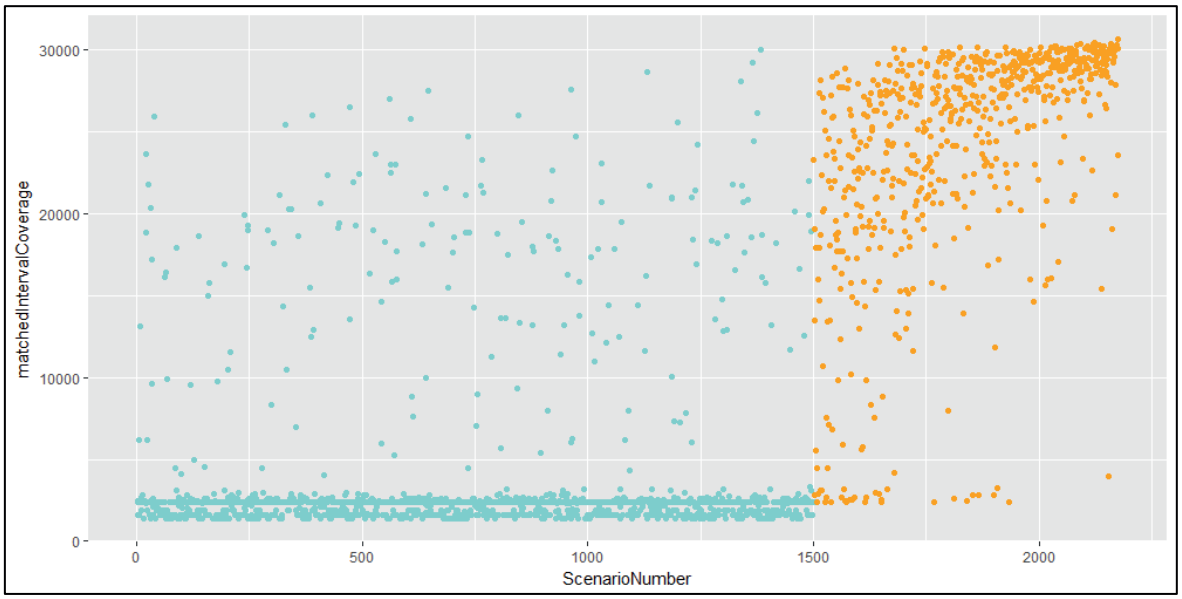



**Figure 63: Simulation results for scenario Energy-E.**

**Table 41: Best three simulation results and parameter values for scenario Energy-E**

| Settings: TTTTFFFFFF | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1384 | 30035 | 41 | 11 | 1 | 14 | 4250 | 2750 | 4250 | 10 | 5 | 1 | 46 | -29 | 9 | 28 | 2900 | 11 | 1289 |
| Standard 2 | 1365 | 29226 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 3 | 1131 | 28670 | 35 | 8 | 2 | 12 | 3000 | 3750 | 4750 | 4 | 1 | 3 | 46 | -26 | 8 | 40 | 3100 | 15 | 1502 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2176 | 30651 | 47 | 1 | 1 | 6 | 2500 | 3500 | 4750 | 2 | 1 | 1 | 16 | -28 | 8 | 37 | 3100 | 15 | 156980 |
| Evolved 2 | 2123 | 30454 | 30 | 2 | 1 | 13 | 4500 | 2475 | 3500 | 10 | 3 | 0 | 76 | -29 | 17 | 21 | 3450 | 15 | 335 |
| Evolved 3 | 2159 | 30360 | 43 | 2 | 1 | 13 | 2000 | 2000 | 3250 | 9 | 1 | 1 | 46 | -17 | 9 | 31 | 3150 | 12 | 29327 |

## 10.17 Additional simulations

This section addresses some of the extra simulation scenarios that were executed.

### 10.17.1 Three simulation scenarios with Random walk activated

This subsection presents three simulation scenarios with Random walk activated. They are used explicitly to explore question seven that investigates movement through the landscape for hominin groups with the effects of random population distribution opposing a directed search for the best resource patches.

**Scenario Energy-B + Randomwalk: TFTTFFFFFT**

This scenario has energy level reconstruction and dynamic hominins, with coastal resources and a maximum foraging range. The results for the MatchedIntervalCoverage variable are shown in Figure 60. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 42. In total there are 2127 simulations that are used from this batch. The maximum score for the Standard set is 26286, for the Evolved set it is 28125, an improvement of 7%.

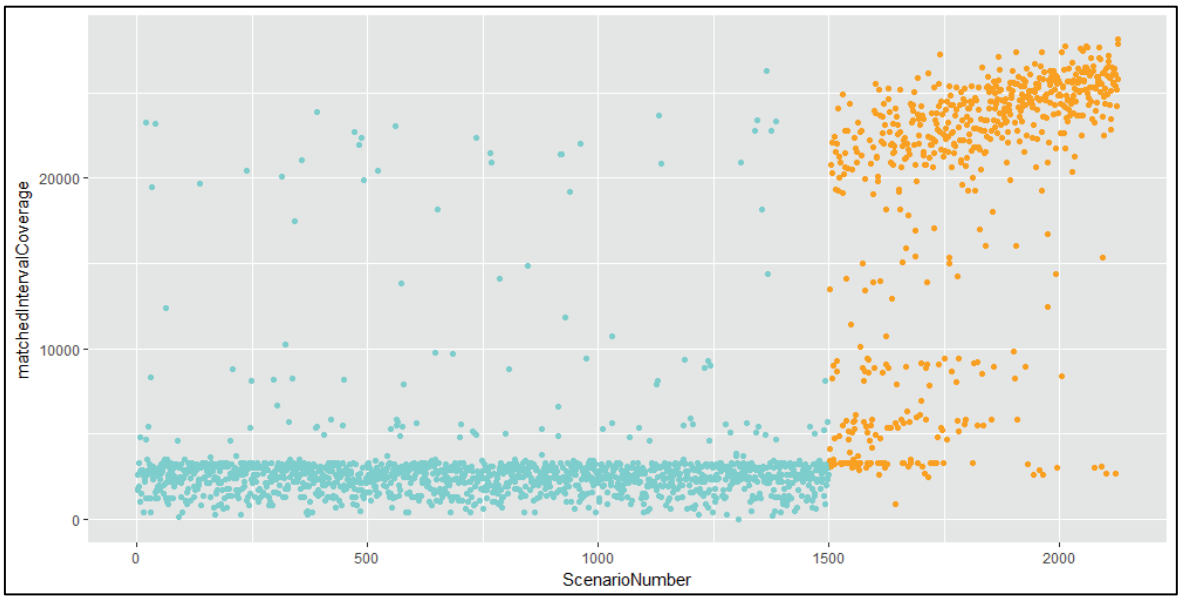



**Figure 64: Simulation results for scenario Energy-BR, with random walk.**

**Table 42: Best three simulation results and parameter values for scenario Energy-BR, (with randomwalk) for Standard and Evolved parameter value sets.**

| Settings: TFTTFFFFT | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 26286 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 390 | 23882 | 47 | 6 | 2 | 2 | 3250 | 4750 | 2000 | 1 | 5 | 5 | 56 | -29 | 9 | 38 | 3150 | 13 | 12842 |
| Standard 3 | 1131 | 23640 | 35 | 8 | 2 | 12 | 3000 | 3750 | 4750 | 4 | 1 | 3 | 46 | -26 | 8 | 40 | 3100 | 15 | 1502 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2127 | 28125 | 30 | 1 | 1 | 12 | 3025 | 5000 | 2250 | 2 | 3 | 1 | 6 | -29 | 9 | 39 | 3150 | 12 | 5991 |
| Evolved 2 | 2126 | 27847 | 43 | 1 | 1 | 12 | 2000 | 3500 | 3250 | 2 | 4 | 1 | 16 | -23 | 9 | 37 | 2700 | 12 | 55429 |
| Evolved 3 | 2057 | 27761 | 50 | 1 | 2 | 12 | 3500 | 3000 | 2000 | 4 | 1 | 1 | 6 | -24 | 11 | 33 | 2700 | 10 | 40880 |

### Scenario Energy-C + Randomwalk: TFTTFFTFFT

It features energy level reconstruction and dynamic hominins. Coastal resources and a maximum foraging range are activated, and the absence condition imposed. The results for the MatchedIntervalCoverage variable are shown in Figure 61. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter set are given in Table 43. In total there are 2079 simulations that are used from this scenario. The maximum score for the Standard set is 3301, for the Evolved set it is 3301, so no improvement is achieved.




**Figure 65: Simulation results for scenario Energy-CR, with random walk.**

**Table 43: Best three simulation results and parameter values for scenario Energy-CR, with randomwalk.**

| Settings: TFTTFFTFT | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1355 | 3301 | 50 | 2 | 13 | 8 | 3750 | 2750 | 2000 | 7 | 2 | 1 | 136 | -30 | 20 | 22 | 3200 | 12 | 28 |
| Standard 2 | 1373 | 3301 | 49 | 10 | 4 | 10 | 3250 | 2000 | 4250 | 10 | 3 | 1 | 116 | -15 | 13 | 27 | 3250 | 2 | 172 |
| Standard 3 | 499 | 3201 | 48 | 14 | 4 | 2 | 4250 | 3500 | 4000 | 2 | 4 | 5 | 66 | -17 | 8 | 31 | 3250 | 6 | 520 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 1355 | 3301 | 50 | 2 | 13 | 8 | 3750 | 2750 | 2000 | 7 | 2 | 1 | 136 | -30 | 20 | 22 | 3200 | 12 | 28 |
| Evolved 2 | 1373 | 3301 | 49 | 10 | 4 | 10 | 3250 | 2000 | 4250 | 10 | 3 | 1 | 116 | -15 | 13 | 27 | 3250 | 2 | 172 |
| Evolved 3 | 1754 | 3301 | 48 | 5 | 5 | 2 | 4250 | 2000 | 2250 | 2 | 5 | 3 | 117 | -27 | 8 | 27 | 3250 | 6 | 2843 |

### Scenario Energy-E +Randomwalk: TTTTFFFFFT

This scenario has energy level reconstruction and static hominins with coastal resources and a maximum foraging range. The results for the MatchedIntervalCoverage variable are shown in Figure 63. The top three scores for the Standard (cyan) and the top three for the GA Evolved (orange) parameter value sets are given in Table 44. In total there are 2037 simulations used from this scenario. The maximum score for the Standard set is 29262, for the Evolved set it is 30439, an improvement of 4%.




**Figure 66: Simulation score for scenario Energy-ER (with random walk).**

**Table 44: Best three simulation results and parameter values for scenario Energy-ER, with randomwalk.**

| Settings: TTTTFFFFT | SimulationNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 486 | 29262 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| Standard 2 | 646 | 29131 | 30 | 11 | 1 | 1 | 3000 | 3750 | 2500 | 6 | 4 | 2 | 66 | -19 | 9 | 40 | 2550 | 15 | 577 |
| Standard 3 | 1384 | 29042 | 41 | 11 | 1 | 14 | 4250 | 2750 | 4250 | 10 | 5 | 1 | 46 | -29 | 9 | 28 | 2900 | 11 | 1289 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 2037 | 30439 | 36 | 4 | 1 | 8 | 4500 | 2000 | 2000 | 7 | 1 | 1 | 56 | -22 | 18 | 40 | 2950 | 12 | 1174 |
| Evolved 2 | 1712 | 30328 | 36 | 4 | 1 | 8 | 4500 | 2000 | 2000 | 7 | 1 | 1 | 56 | -22 | 18 | 40 | 2950 | 14 | 1174 |
| Evolved 3 | 2012 | 30180 | 41 | 4 | 1 | 5 | 3250 | 3500 | 2250 | 7 | 2 | 1 | 76 | -24 | 10 | 28 | 2950 | 15 | 6924 |

### 10.17.2 Duplication of scenario Energy-A for verification purposes

For verification purposes the scenario was rerun, producing the second batch. The first 1500 were still the same Standard simulations, but after that the genetic algorithm will select different individuals in the tournaments and apply different mutations or combinations. The result is summarized in Figure 67 and Table 45.

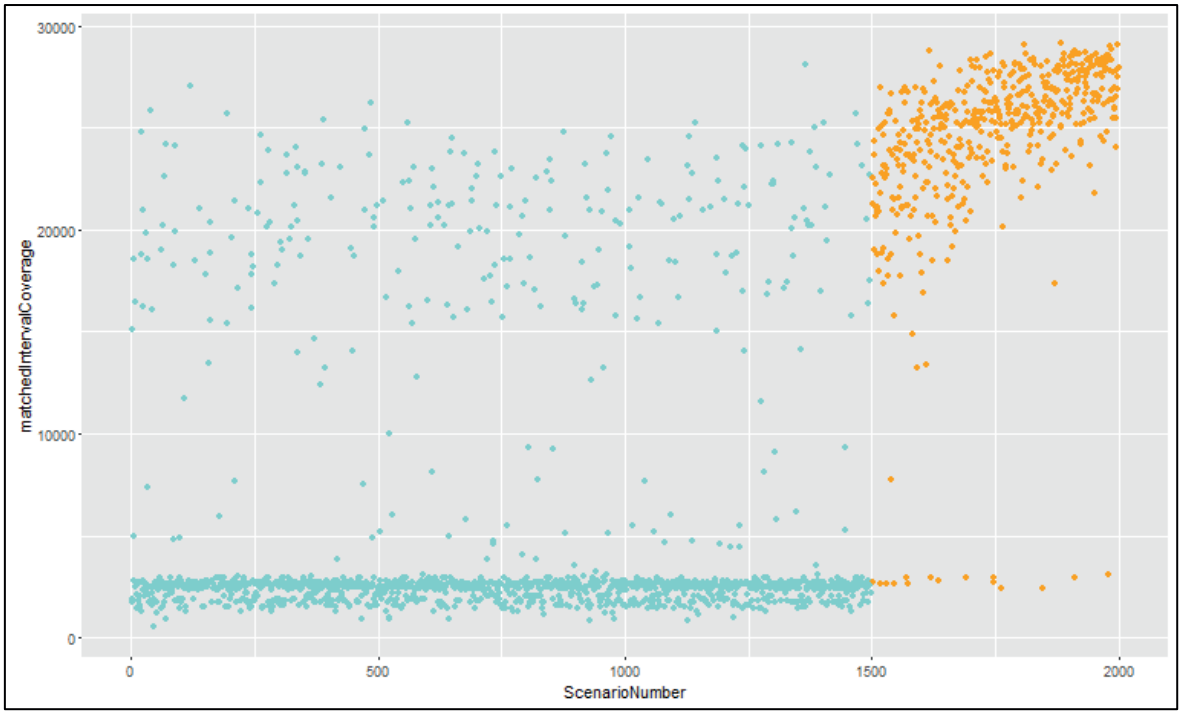



**Figure 67: Simulation results for scenario Energy-A, batch 2.**

**Table 45: Best three simulation results for scenario Energy-A, batch 2.**

| FFFFFFTFFF(2) | ScenarioNumber | matchedIntervalCoverage | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Standard 1 | 1365 | 28108 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| Standard 2 | 120 | 27034 | 43 | 1 | 1 | 12 | 3750 | 3250 | 5000 | 9 | 1 | 4 | 96 | -29 | 8 | 32 | 2850 | 6 | 62186 |
| Standard 3 | 486 | 26202 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| | | | | | | | | | | | | | | | | | | | |
| Evolved 1 | 1884 | 29136 | 43 | 2 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 5 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 39669 |
| Evolved 2 | 1998 | 29067 | 44 | 3 | 1 | 9 | 3250 | 4750 | 2000 | 3 | 5 | 2 | 16 | -18 | 8 | 38 | 2700 | 14 | 52123 |
| Evolved 3 | 1809 | 29030 | 49 | 2 | 1 | 14 | 2500 | 4500 | 3750 | 9 | 3 | 1 | 16 | -19 | 10 | 40 | 2950 | 6 | 78628 |

# 11. ANALYSIS

## 11.1 Introduction

When constructing realistic models, one of the major issues when attempting to describe reality is over-parameterization. Over-parameterized models contain elements that are not constrained by the dataset and where realistic values cannot be determined for all variables. In general, while analysing simulation results one of the implicit aims is to identify those parameters in the model that are significant for the subject being studied (Ligmann-Zielinska *et al.* 2014). This study attempts to construct a realistic but parsimonious Neanderthal model and to find values for the model parameters while exploring hypotheses on Neanderthal behaviour in the landscape. The analysis of the results from all simulations is done in the following sections:

1. First the validity of the approach is tested in Section 11.2 by analysing correlations (a) between input model parameters and (b) between the `matchedIntervalCoverage` output variable and other output variables. There should be limited correlation between model parameters and maximum correlation between output variables.
2. The bulk of this chapter is Section 11.3 which contains a sensitivity analysis of the model parameter values from the different scenarios. This includes a variability assessment, identification of correlations in the Standard and Evolved data sets, relative variable importance calculations, and the identification of monotonic trends in parameter value selection.
3. Section 11.4 presents trends in the simulation output data. This reflects the effect of the genetic algorithm with the evolution of the parameter value set. For each scenario that is explored first the Standard set of simulations is executed. Those model parameter sets that perform best in all scenarios are characterized. This section also presents data on the quality of the archaeological input and a reflection on the different causes of life and death in the simulations.
4. The modelled environment is analysed in detail in Section 11.5. Key locations in the simulation area were monitored with Climate Monitoring Checkpoints throughout the simulation period. Data from these checkpoints is used in the analysis of the environment reconstruction. An important element of the reconstructed environment is the sea level. The effects of fluctuations in this level are analysed by identifying critical points in time and presenting the resulting topography for those points in time.

A set of R-scripts was developed for the automated mathematical analysis of the simulation results and to produce the relevant data. Instead of manual manipulation of data and creation of output representation such scripting is essential to allow proper replication

of the research and reproduction of the results (Marwick 2017). These scripts are used to generate the data, tables and figures in this chapter. The scripts and their output are listed in Appendix 2 and included verbatim in the Supplementary Materials.

## 11.2 Correlating input to input and output to output

The HomininSpace system attempts to estimate parameter values by fitting them to measured archaeological data using an optimization technique based on genetic algorithms. To avoid non-identifiability issues where correlations between model parameters make good estimates for individual parameter values difficult it is important to verify that input parameters are not strongly correlated with each other, and thus can be varied independently.

Here I analyse pair-wise (or one dimensional) correlations in the Standard parameter values set. Since the genetic algorithm modifies model parameters slightly between generations introduced correlations can be expected in the Evolved sets, so I focus on the Standard set only. Strong correlations reveal parameter combinations for which good fitting values would be difficult to obtain. The strongest correlated negative and positive input model parameter to input model parameter correlations, from the Standard set are `Years_Before_Group_Maturity` with `Max_ForagingRange`, correlated at -0,0692 (p = 0,01), and `Subsistence_PreFertileCohort` with `GroupSizeFertile_BeforeMerge`, with a correlation of +0,0601 (p = 0,02).

The correlation values show that even the strongest correlations are not very strong (as expected and required). Both the most negative and most positive correlation are significant with a *p* value higher than or equal to 0,01. All other correlations between model parameters are less strong. Note that these correlations are the same for all scenarios since the values for the input parameters in the Standard set are the same in all scenarios and were randomly generated, independent of other model components.

The output variable `matchedIntervalCoverage` is used to assess the validity of the simulations, with a larger value suggesting a better match with the archaeology. But there are several other output variables included in the output data sets, attempting to characterize the match with the archaeology in a slightly different manner. To ensure a relevant and optimal characterization I show here that these do significantly correlate with the selected variable `matchedIntervalCoverage`. Again the pair-wise correlation from the sensitivity analysis is used to describe the relationships between the different

output variables for the simulated scenarios. Figure 68 presents as a representative sample the results for scenario Habitat-A, with all settings to False. All other scenarios have similar results.

Figure 68 presents the pair wise correlations between all output visit counting variables in a Correlation Matrix, showing the Spearman's correlation coefficients between variables. In such a matrix the same variables are used for the rows as well as for the columns (so variable `matchedIntervalCoverage` is used for the data in the first row and the first column). In this specific format, the matrix presents the correlations in the top-right part with significant correlations indicated by three red dots. In the bottom left part of the matrix variables of the intersecting row-column combination are plotted against each other, illustrating the nature of the correlation.

The weakest correlation between `matchedIntervalCoverage` and the other general visit counting output variables has a significant correlation coefficient of 0.89, which is considered very strong. This is with the `totalCSTVisits` variable that keeps track of all visits to the checkpoint inside or outside an interval. The correlation with the variables that keep track of the times that a group during a visit selects the exact same grid cell as the checkpoint as their home range grid cell are slightly weaker (0.73 for those inside an interval and 0.75 for any visits outside intervals). The correlation with the absence visits is weakest with a significant value of 0.61. All correlations are positive. The strengths of these correlations indicate that the value for `matchedIntervalCoverage` can be used to assess the simulation results without need to perform similar analyses for the other variables.

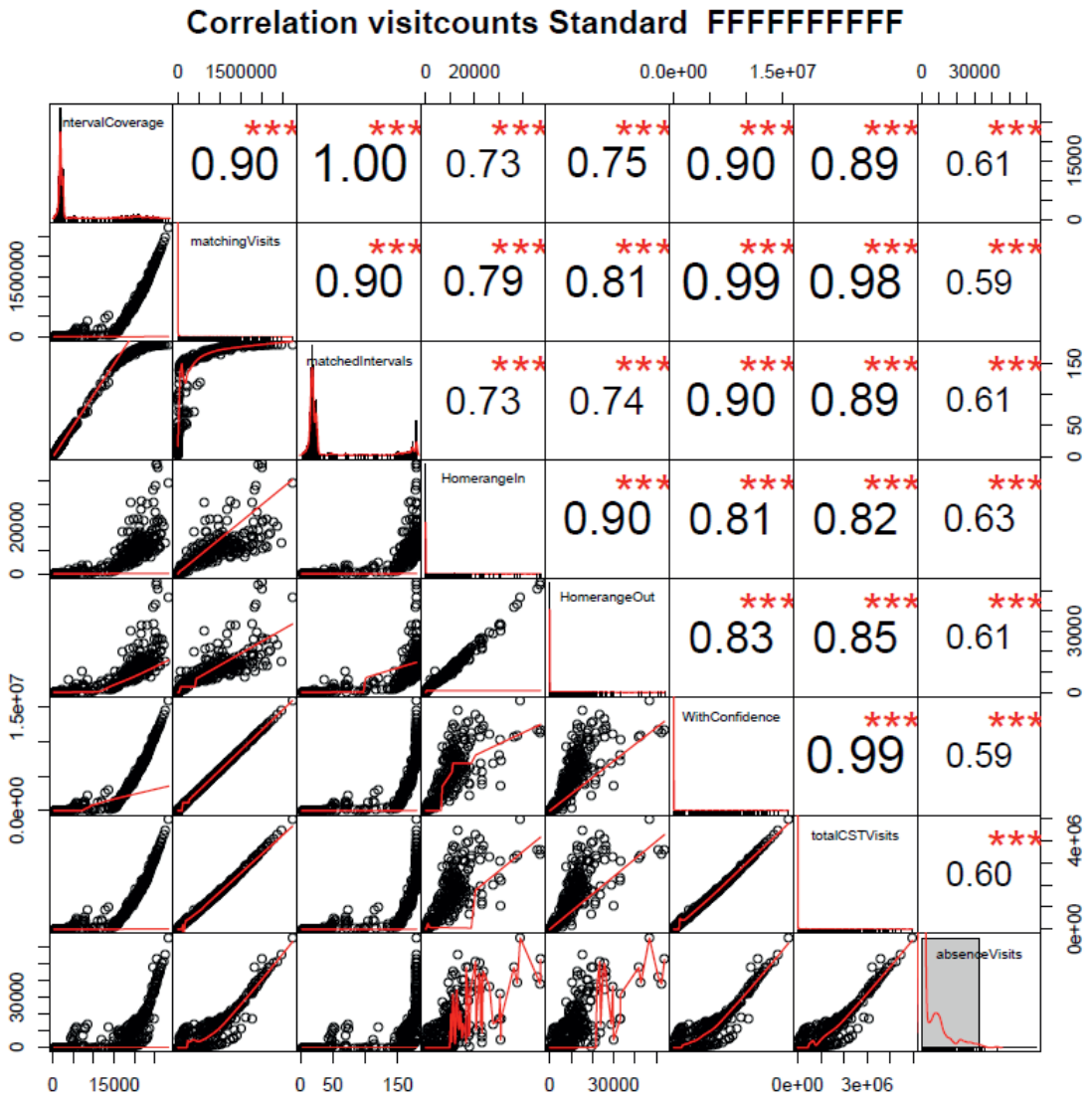



**Figure 68: Correlations between the different visit counting output variables for scenario Habitat-A. Top right part of the matrix contains the correlations. Three red stars indicate a significant correlation. Bottom left part are plotted values.**

## 11.3 Sensitivity analysis in scenario exploration

Sensitivity analysis for simulations is the exploration of the model input parameter values and the associated output results. The aim is to identify those parameters for which small changes cause relevant changes in the output (Ligmann-Zielinska *et al.* 2014). The results can aid in model simplification and improved model accuracy. Here the analysis will contribute to answer specific questions. Note that the use of a genetic algorithm includes a build-in sensitivity analysis when recombining parameter values as the method aims to find those parameters that matter most (Lee *et al.* 2015).

### 11.3.1 Variability in scenarios

Diversity in the simulated models (more specifically, in the model parameter values as optimized by the genetic algorithm) can be graphically explored by displaying the model parameters as visual elements in a larger picture. In HomininSpace there are 16 variables in 18 scenarios with each more than 2000 simulations. Visualizing trends, detecting patterns and identifying outliers in graphs becomes difficult for such large numbers of data items (Guha and Assaf 2018). Graphs for selected variables are easier to interpret, but selection must be guided. Assuming that humans are capable of assimilating data encoded in human faces, complex multivariate data can be presented where the variables each represents the feature of a face (Chernoff 1973). For instance, normalized birthrate values can be used to represent the height of an abstract face (
Figure 69).

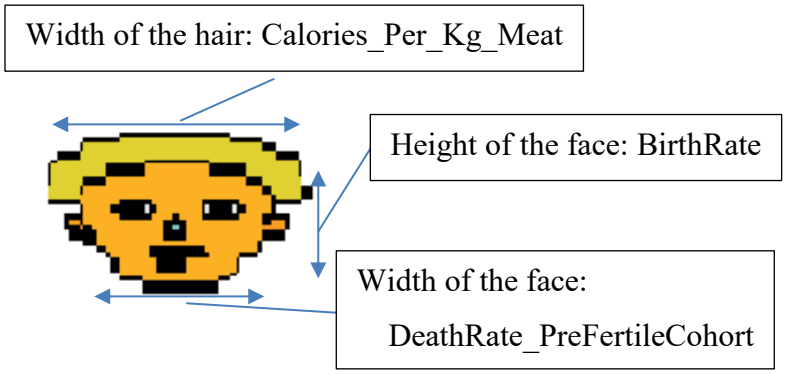



**Figure 69: Explaining three parameter values used to construct Chernoff faces.**

Figure 70 present faces for the top 10 models of all scenarios. The following values are used for the face shape variables:

1-height of face (`BirthRate`);
2-width of face (`DeathRate_PreFertileCohort`);
3-shape of face (`DeathRate_FertileCohort`);
4-height of mouth (`DeathRate_PostFertileCohort`);
5-width of mouth (`Subsistence_PreFertileCohort`);
6-curve of smile (`Subsistence_FertileCohort`);
7-height of eyes (`Subsistence_PostFertileCohort`);
8-width of eyes (`CohortSize_PreFertile`);
9-height of hair (`CohortSize_Fertile`);
10-width of hair (`Calories_Per_Kg_Meat`);
11-styling of hair (`Max_ForagingRange`);
12-height of nose (`Temperature_Tolerance`);
13-width of nose (`Years_Before_Group_Maturity`);

14-width of ears (`GroupSize_BeforeMerge`);
15-height of ears (`GroupSize_BeforeSplit`).

Note that `GroupSizeFertile_BeforeMerge` is not included since that parameter varies little. To add additional discriminative characteristics colors are also coded. For painting elements of a face the colors of are found by averaging of sets of variables: (7,8)-eyes:iris, (1,2,3)-lips, (14,15)-ears, (12,13)-nose, (9,10,11)-hair, (1,2)-face. For the data set a top-10 of best performing models per scenario is selected, ranking based on the `matchedIntervalCoverage` score. An R-script is written to create Figure 70, with the top-10 model parameter values for each scenario ("Analysis draw faces.R", included in the Supplementary Materials, as is the produced output file with the top-10 for each scenario)[42]. Each row contains the data for one scenario, where the first column contains the best performing model parameter value set, the second row the second best performing model, etc. Faces 1 to 10 belong to scenario 1 (Habitat-A); faces 11 – 20 are produced in scenario 2 (Habitat-B), etc.

[42] This uses the *aplpack* package with the *faces* routine to draw the faces.



**Figure 70: Chernoff faces representing model parameter values for all 18 scenarios, illustrating variability within the top-10 best performing models per scenario (line), with the best scoring model numbered 1 (leftmost).**

While most scenarios have more or less variety in their faces, it immediately becomes clear that scenarios 4 and 6 show hardly any variation at all in their top-10. They have evolved very similar solutions (  , with very low birth rates and high death rates for the pre-fertile cohort). These scenarios implement the absence criterion. Also scenario 5 has limited variation upon  with high birth rates and low death rates, almost the opposite of the scenarios 4 and 6. The two outliers in scenario 2 (wide faces, numbers 15 and 16) are solutions characterized by very high death rates in the pre-fertile cohort, accompanied by high birth rates. In scenario 12 (111-120) all faces sport narrow split eyes with only faces numbered 117, 119, and 120 have round eyes. These last three individuals have very high subsistence needs for the post-fertile cohort. Similarly, in scenario 15 some faces have

wide mouths, indicating high subsistence needs for the pre-fertile cohort (143, 145, 146, and to a lesser extent 141 and 148).

The Chernoff faces representing model parameter values illustrate the variability within each scenario (or the lack thereof), and also suggests that certain types of value combinations perform better in one scenario than in another. A detailed interpretation of this variability is however impossible from the images alone and a deeper exploration of the relation between model parameter values and simulation results follows below.

### 11.3.2 Correlations between parameter values and simulation results

A correlation analysis aims to see if two measured variable co vary in a linear relationship and quantifies the strength and direction of this relation. Values range from -1.0 for a negative linear relation via 0 with no linear relation to +1.0 for a positive linearity between the two variables. Care must be taken when interpreting correlations, since a third factor can cause the correlation and therefore correlation is not by definition causation (Aldrich 1995). Correlations that are presented here are Spearman's rho rank correlation coefficients. Missing values are deleted, with mid-ranks used in case of ties (Hollander and Wolfe 1973). P-values are approximated by using t of F distributions, and correlations with P-values smaller than 0.01 are removed because they are considered insignificant and are not shown for clarity (for a general discussion about the choice of the 0.01 value for meaning 'very significant' see Nuzzo (2014)).

Spearman's rho rank correlation is used because this does not require input parameter values to satisfy of any assumptions about the distribution of the data. Spearman does not require a normal distribution as for instance the Pearson correlation method (Fletcher and Lock 2005). The random created values of the Standard set and the recombined values in the Evolved set both violate the Pearson correlation assumptions of variable value distribution. Spearman's rank order correlation measures the strength and direction of the associations between ranked, monotonic variables. The results are calculated per batch and should be interpreted as an index of sensitivity (Grimm and Railsback 2005). Using a more conservative P-value of 0.001 to avoid spatial auto-correlation does not create a different picture (as suggested in Wiegand *et al.* (2004)).

Table 46 presents the correlations for the Standard sets of all batches that were explored. The correlations for the Standard sets must be compared with those for the Evolved parameter sets, with parameter values that have been improved by the GA. The nature of

the values for the randomly generated Standard set and the directionally improved Evolved set are different. Correlations for the complete set would of course average the values for Standard and Evolved sets. The Spearman correlation coefficients for the Evolved set are presented in .

**Table 46: For all Standard sets the significant Spearman correlation coefficients for parameters versus matchedIntervalCoverage. Green are positive correlations, red negative. Darker colors indicate stronger correlations.**

| Scenario (Standard) | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 0.45 | -0.24 | -0.45 | | | 0.10 | 0.07 | | | -0.17 | 0.21 | -0.14 | -0.08 | | -0.13 | | 0.66 |
| Habitat-B | 0.38 | -0.22 | -0.40 | | | 0.12 | 0.08 | | | -0.20 | 0.25 | -0.12 | -0.09 | | -0.12 | | 0.58 |
| Habitat-C | 0.38 | -0.19 | -0.45 | | | | | | | -0.19 | 0.09 | -0.14 | | | -0.09 | 0.22 | 0.59 |
| Habitat-D | -0.20 | 0.12 | 0.19 | | | | | 0.09 | | | -0.25 | -0.09 | | | | -0.43 | -0.24 |
| Habitat-E | 0.33 | -0.15 | -0.42 | | | | | | | -0.18 | 0.14 | -0.31 | | | | 0.39 | 0.51 |
| Habitat-F | -0.35 | 0.19 | 0.35 | | | -0.07 | -0.08 | | | | -0.17 | -0.07 | | | 0.10 | -0.19 | -0.52 |
| Habitat-G | 0.41 | -0.24 | -0.41 | | | 0.09 | 0.08 | | | -0.10 | 0.17 | -0.13 | | | -0.09 | | 0.61 |
| Habitat-H | 0.46 | -0.27 | -0.36 | -0.18 | | | 0.08 | -0.09 | | 0.16 | 0.19 | -0.15 | | | -0.10 | | 0.71 |
| Habitat-I | 0.26 | -0.13 | -0.33 | | | | | | | -0.12 | | | | | | 0.60 | 0.41 |
| Habitat-J | 0.27 | -0.12 | -0.34 | | | | | | | -0.12 | | | | | | 0.60 | 0.41 |
| Energy-A | 0.45 | -0.25 | -0.48 | | 0.08 | 0.13 | 0.10 | | | -0.17 | 0.17 | -0.13 | -0.08 | | -0.14 | | 0.68 |
| Energy-B | 0.25 | -0.12 | -0.36 | | 0.07 | 0.07 | | | | -0.15 | | -0.11 | | 0.08 | | 0.31 | 0.45 |
| Energy-BR | 0.25 | -0.11 | -0.33 | | | | | | | -0.16 | | -0.21 | | | | 0.54 | 0.38 |
| Energy-C | 0.13 | | -0.17 | | 0.07 | 0.10 | | | | -0.10 | | -0.07 | | | -0.07 | | 0.23 |
| Energy-CR | 0.07 | | -0.14 | | | | | | | -0.10 | | -0.10 | | | | 0.32 | 0.11 |
| Energy-D | 0.26 | -0.12 | -0.36 | | 0.07 | 0.07 | | | | -0.16 | | -0.11 | | 0.08 | | 0.32 | 0.45 |
| Energy-E | 0.19 | -0.07 | -0.30 | | | | | -0.09 | | -0.14 | -0.08 | | | | | 0.71 | 0.33 |
| Energy-ER | 0.19 | -0.09 | -0.31 | | | | | -0.09 | | -0.14 | -0.08 | | | | | 0.72 | 0.34 |

When analysing the correlation results for the Standard sets, the following general patterns stand out:

- Overall there is a strong positive correlation between birth-rate and simulation results, and a corresponding negative correlation with the death rate of the fertile cohort. There is also a noticeable negative correlation overall with the death rate of the pre-fertile cohort;
- There is an overall but not large negative correlation with group size fertile before split, and a similar but smaller correlation with the temperature tolerance. The group size before split is more positive;
- The maximum foraging range is generally strong and positively correlated with the simulation results;
- The viability index (a value calculating the effectiveness of the demographic model alone) has a very strong and positive correlation with the simulation results.

These are general trends where the actual values for the correlations differ per batch. There are two noticeable exceptions, and those are batches Habitat-D and Habitat-G. Values for

correlations for both batches are opposite of the described trends above. Both batches explore a combination of a limited foraging range, the production from core areas and a limiting absence restriction. Since most checkpoints are in the south of the simulation area, where also the core areas are located, most matches with the archaeology can here be obtained from Neanderthals that do not disperse quickly to the north, where the absence check (access to the UK) is implemented. Lower birth rates and higher death rates will actually accomplish that (since new groups will be produced from core areas anyhow), as will smaller foraging ranges (since the foraging range also defines how far a dispersal event can reach per time step). Note that it is the combination that restricts the parameter values, since Habitat-B (absence) and Habitat-C (maximum foraging range) do not exhibit this divergence from the general pattern.

A positive correlation for simulation results with birth rates, and negative correlations with death rates for the pre-fertile and even stronger for the fertile cohort suggests that more Neanderthals is better. This is a general phenomenon that is easily understood: the more groups there are in space and time, the better the overall chances for an improved match with the archaeology. There is one limit to an omni-present Neanderthal, and that is a limiting carrying capacity. There is only so much food in the environment, which also fluctuates with a changing climate and is consumed by present Neanderthals. However, many correlations can be better understood with the tendency to fill the area with Neanderthal groups in mind. The non-correlation with the death rate for post-fertile Neanderthals can be understood in this context: they do no longer contribute to more Neanderthals, and a higher death rate would not influence the number of new Neanderthals.

Also, groups that live longer can help with a better match with the archaeology. For instance a larger foraging range will allow a group to retrieve resources from a larger area and this will keep Neanderthals alive, even when circumstance deteriorate. Those batches that are not correlated with the maximum foraging do not have the setting Maximum Foraging range activated (batches Habitat-A, Habitat-B, Habitat-E, and Habitat-H; note that Energy-C with an activated Maximum range also does not correlate this value to the simulation results). A lower value for the temperature tolerance means better survival chances when average temperatures drop, and this would explain the negative correlation with this parameter. Since only few checkpoints are in areas known to become cold, the correlation is not very strong.

An overall negative correlation with the parameter group size fertile before merge means that smaller values lead to better results. A lower value means that groups will more quickly attempt to merge with other groups if there are fewer individuals in the fertile cohort, increasing survival changes for the group members. A more positive correlation of the GroupSize_BeforeSplit parameter means that it requires a larger group before the group is forced to split into two groups. Smaller groups are more vulnerable, especially in the areas where the climate deteriorates more or faster (generally, the northern sections of the simulation area).

Therefore it is very surprising that there is overall no negative correlation with the subsistence parameters. One expects that lower energy needs would allow for more Neanderthals. For instance this can explain the effect of the (small) negative correlation with the Calories_Per_Kg_Meat. Also surprising is the absence of any correlation with the CohortSize_Fertile, since a larger value for this parameter allows a female Neanderthal to conceive more Neanderthal babies. One would also expect a positive correlation with the DeathRate_PostFertileCohort, since post-fertile Neanderthals only consume resources and are not expected to contribute to survival of the group as such. Years_Before_Group_Maturity and GroupSize_BeforeMerge are also parameters that are not correlated with the simulations results at all.

**Table 47: The Spearman correlation coefficients of the Evolved sets for parameters versus matchedIntervalCoverage. Green are positive, red are negative correlations.**

| Scenario (Evolved) | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 0.25 | -0.35 | -0.49 | 0.16 | -0.19 | | | | | -0.35 | -0.67 | -0.12 | -0.19 | 0.41 | | 0.21 | 0.72 |
| Habitat-B | 0.21 | -0.20 | -0.24 | | | | | | | | | | | | | | 0.59 |
| Habitat-C | 0.38 | -0.16 | -0.55 | 0.14 | -0.21 | | 0.23 | | | | -0.41 | -0.22 | -0.31 | 0.12 | | 0.12 | 0.76 |
| Habitat-D | -0.13 | | | | | -0.09 | | | | -0.08 | -0.17 | | -0.09 | | 0.11 | | |
| Habitat-E | 0.70 | -0.67 | -0.56 | 0.20 | -0.10 | 0.21 | -0.11 | | 0.24 | -0.55 | -0.75 | -0.32 | -0.56 | | 0.23 | | 0.87 |
| Habitat-F | -0.24 | | | | | -0.10 | -0.14 | | | -0.16 | -0.12 | | | | 0.10 | | -0.27 |
| Habitat-G | 0.33 | -0.28 | -0.61 | 0.17 | -0.13 | 0.13 | 0.16 | | -0.14 | -0.37 | | -0.24 | -0.36 | | | | 0.74 |
| Habitat-H | 0.33 | -0.22 | -0.62 | 0.17 | | -0.35 | | -0.32 | -0.49 | -0.56 | -0.20 | | -0.21 | 0.22 | | | 0.70 |
| Habitat-I | 0.24 | -0.30 | -0.64 | 0.22 | | -0.40 | 0.34 | | -0.36 | -0.36 | -0.27 | -0.41 | -0.58 | 0.17 | 0.21 | 0.40 | 0.80 |
| Habitat-J | | -0.19 | -0.57 | | | | | | -0.18 | -0.34 | -0.23 | -0.26 | -0.44 | 0.12 | 0.15 | 0.26 | 0.70 |
| Energy-A | 0.20 | -0.56 | -0.62 | | -0.33 | | | | | | -0.34 | -0.30 | -0.40 | 0.16 | -0.21 | 0.21 | 0.83 |
| Energy-B | 0.26 | -0.26 | -0.37 | | -0.22 | -0.15 | -0.11 | 0.13 | -0.10 | -0.46 | -0.54 | -0.20 | -0.32 | 0.09 | 0.18 | -0.17 | 0.65 |
| Energy-BR | 0.29 | -0.20 | -0.37 | | -0.15 | | | -0.15 | | -0.25 | -0.56 | -0.21 | -0.32 | 0.23 | -0.11 | 0.12 | 0.66 |
| Energy-C | 0.25 | -0.13 | -0.26 | 0.23 | | -0.12 | | | | -0.43 | | | -0.12 | | | -0.40 | 0.42 |
| Energy-CR | | | | | | | | | | | | -0.11 | | | | | |
| Energy-D | 0.25 | -0.37 | -0.47 | | -0.15 | | | | | | -0.40 | | -0.27 | | | | 0.72 |
| Energy-E | 0.12 | -0.18 | -0.42 | 0.13 | -0.32 | | | | -0.23 | -0.42 | -0.39 | -0.20 | -0.35 | 0.19 | 0.17 | 0.32 | 0.61 |
| Energy-ER | | -0.26 | -0.42 | 0.16 | | | | 0.11 | -0.32 | -0.49 | -0.38 | -0.35 | -0.16 | | | 0.27 | 0.58 |

When comparing the correlations for the Evolved parameter values with those in the Standard batches one thing is obvious: there are more correlations in the Evolved batches, and most correlations are larger. That means that the values for the parameters as generated by the genetic algorithm are more strongly associated with the simulation results than the randomly generated values from the Standard set. This is to be expected since the aim of the genetic algorithm is to positively influence the simulation results by changing the parameter values, so more correlations will be found, also for parameters that are not correlated at all in the Standard set. Since the genetic algorithm changes mostly a few parameter values, others might seem correlated simply due to the fact that they remain untouched for several generations. Very surprizing is the generally positive correlation of GroupSize_BeforeSplit in the Standard set is overall changed to a very negative correlation in the Evolved sets. Also remarkable is the evolved negative correlation with CohortSize_Prefertile.

### 11.3.3 Relative Variable Importance (RVI) - what makes a Neanderthal tick?

The relative importance of variables attempts to rank parameters by building a predictive model from the simulation results, suggesting which subset of parameters is most

important for creating a good match with the archaeology (Johnson 2000). This can be done by creating a linear regression model and calculating the relative contribution (importance) of individual variables. The method thus assumes that a linear model exists to explain the output. This study uses the *remaimpo* package in R for calculating such a *lm* model. After this the parameters are ranked according to importance scaled to 100 using the *calc.relimp* method. The dependant variable is `matchedIntervalCoverage`. For the relative variable importance, all values below 5% are not shown. The RVI is calculated for the Standard set of all scenarios (Table 48) and the Evolved set (Table 49).

**Table 48: Relative variable importance assuming a linear model for all Standard sets.**

| Scenario | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 0.40 | 0.14 | 0.36 | | | | | | | | | | | | | |
| Habitat-B | 0.32 | 0.09 | 0.26 | | | | | | | 0.14 | | | | | | |
| Habitat-C | 0.34 | 0.11 | 0.39 | | | | | | | | | | | | | 0.07 |
| Habitat-D | 0.17 | 0.08 | 0.17 | | | | | | | 0.27 | | | | | | 0.26 |
| Habitat-E | 0.09 | | 0.21 | | | | | | 0.07 | | 0.13 | | | | | 0.44 |
| Habitat-F | 0.38 | 0.11 | 0.30 | | | | | | | 0.10 | | | | | | |
| Habitat-G | 0.40 | 0.14 | 0.36 | | | | | | | | | | | | | |
| Habitat-H | 0.38 | 0.14 | 0.28 | | | | | | | 0.06 | | | | | | |
| Habitat-I | 0.30 | 0.10 | 0.34 | | | | | | | | | | | | | 0.18 |
| Habitat-J | 0.31 | 0.10 | 0.35 | | | | | | | | | | | | | 0.16 |
| Energy-A | 0.41 | 0.14 | 0.35 | | | | | | | | | | | | | |
| Energy-B | 0.23 | 0.07 | 0.37 | | | | | | | | | | | | | 0.22 |
| Energy-BR | 0.17 | | 0.25 | | | | | | | | 0.11 | | | | | 0.32 |
| Energy-C | 0.21 | | 0.38 | | 0.06 | 0.09 | | | | 0.10 | | | | | 0.06 | |
| Energy-CR | | | 0.11 | | | | | | | 0.05 | | 0.06 | | | | 0.67 |
| Energy-D | 0.23 | 0.07 | 0.37 | | | | | | | | | | | | | 0.22 |
| Energy-E | 0.26 | 0.07 | 0.33 | | | | | | | | | | | | | 0.26 |
| Energy-ER | 0.25 | 0.08 | 0.32 | | | | | | | | | | | | | 0.26 |

**Table 49: Relative Variable Importance for the Evolved sets from all scenarios.**

| Scenario | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 0.10 | 0.13 | 0.40 | | | | | | 0.09 | 0.06 | | 0.05 | 0.06 | | | |
| Habitat-B | 0.26 | 0.13 | | | 0.12 | | | 0.16 | | 0.12 | | | | | | |
| Habitat-C | 0.17 | 0.08 | 0.31 | | | | | | | | 0.07 | 0.07 | 0.09 | | | |
| Habitat-D | 0.24 | | 0.24 | | | | | | 0.06 | 0.13 | | | | 0.08 | 0.07 | |
| Habitat-E | 0.15 | 0.16 | 0.14 | | | | | | 0.12 | 0.20 | | 0.10 | | | | |
| Habitat-F | 0.26 | 0.12 | 0.18 | 0.06 | | | 0.07 | | | 0.12 | | | | | 0.07 | |
| Habitat-G | 0.11 | 0.19 | 0.37 | | | | | | | | | 0.09 | | | | |
| Habitat-H | 0.13 | 0.07 | 0.57 | | | | | | 0.09 | | | | | | | |
| Habitat-I | | | 0.32 | | | | | | | 0.08 | | 0.10 | | | 0.25 | |
| Habitat-J | 0.11 | 0.11 | 0.33 | | | | | | 0.07 | | | 0.12 | | | 0.12 | |
| Energy-A | 0.21 | 0.19 | 0.35 | | | | | | | | 0.05 | 0.07 | 0.05 | | | |
| Energy-B | 0.06 | 0.07 | 0.19 | | | | | | 0.15 | 0.19 | | 0.08 | | | | |
| Energy-BR | 0.09 | 0.06 | 0.17 | | | | | | | 0.25 | 0.13 | 0.08 | | | 0.07 | |
| Energy-C | 0.10 | 0.05 | 0.17 | 0.07 | | | | | 0.29 | | | | | | 0.21 | |
| Energy-CR | | | 0.11 | 0.05 | | 0.25 | | 0.06 | | 0.06 | 0.17 | | | | 0.14 | |
| Energy-D | 0.09 | 0.15 | 0.29 | | | | | | | 0.20 | | 0.07 | | | 0.10 | |
| Energy-E | 0.08 | | 0.22 | | | | | | 0.11 | 0.09 | | 0.06 | | | 0.22 | |
| Energy-ER | 0.09 | 0.09 | 0.23 | | | | | | 0.15 | 0.11 | 0.09 | 0.06 | | | 0.12 | |

### 11.3.4 Monotonic trends in parameter values

A variable in a series of observations that consistently increases or decreases through time displays a monotonic upward (or downward) trend. Such a trend does not have to be linear but the presence of these trends suggests a purposeful change. To detect monotonic trends in time series again the Spearman's rho test is useful (D'Abrera and Lehmann 1975; Sneyers 1991). The method does not explain the causes of any trend. Here the test is used to statistically identify changes that the genetic algorithm has induced.

Therefore input parameter values are correlated against the simulation number that ranks the simulations through time in the scenarios. A (strong) correlation indicates that the parameter changes into one direction as the experiment progresses. Such changes could be induced by the genetic algorithm attempting to optimize the simulation results. Table 50 lists for the Evolved series of all scenarios the results for the Spearman's rho correlations of parameters and simulation numbers. Values are only calculated for the Evolved data sets since there should be by definition no direction in the Standard parameter value sets. Note that more simulations in a scenario (Habitat-I) will result in more and stronger patterns (compare to Habitat-A).

**Table 50: Correlations between model input parameters and scenario number, for the Evolved datasets. Positive correlations are green colored and darker colors indicate stronger correlations.**

| Scenario (Evolved) | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 0.17 | -0.26 | -0.31 | 0.12 | | | | | | -0.23 | -0.52 | -0.13 | -0.13 | 0.34 | | 0.17 | 0.52 |
| Habitat-B | | -0.15 | | | | | | | -0.13 | | | | 0.12 | | | | 0.15 |
| Habitat-C | 0.25 | | -0.37 | | | | 0.23 | | | | -0.31 | -0.16 | -0.19 | | | | 0.46 |
| Habitat-D | -0.12 | | -0.09 | | | | | 0.09 | -0.15 | -0.27 | -0.19 | -0.09 | -0.11 | | | 0.16 | 0.14 |
| Habitat-E | 0.62 | -0.59 | -0.46 | 0.18 | | 0.21 | -0.09 | | 0.23 | -0.48 | -0.65 | -0.28 | -0.51 | | 0.17 | | 0.76 |
| Habitat-F | -0.15 | -0.11 | -0.17 | | | -0.13 | | | | -0.16 | | | | | 0.13 | | |
| Habitat-G | 0.20 | -0.15 | -0.47 | 0.16 | | 0.13 | | | | -0.23 | | -0.18 | -0.26 | | | | 0.52 |
| Habitat-H | 0.27 | -0.13 | -0.48 | 0.12 | | -0.29 | | -0.27 | -0.39 | -0.45 | -0.15 | | -0.17 | 0.16 | | 0.10 | 0.53 |
| Habitat-I | 0.18 | -0.26 | -0.54 | 0.19 | 0.11 | -0.34 | 0.31 | | -0.33 | -0.29 | -0.22 | -0.37 | -0.50 | 0.13 | 0.18 | 0.30 | 0.67 |
| Habitat-J | | -0.12 | -0.45 | | 0.11 | | | | -0.14 | -0.27 | -0.18 | -0.21 | -0.38 | | 0.16 | 0.23 | 0.51 |
| Energy-A | 0.14 | -0.40 | -0.41 | | -0.29 | | | | | | -0.26 | -0.26 | -0.30 | 0.14 | -0.19 | 0.21 | 0.61 |
| Energy-B | 0.19 | -0.21 | -0.30 | | -0.20 | -0.14 | | 0.13 | | -0.39 | -0.48 | -0.18 | -0.27 | | 0.12 | -0.17 | 0.53 |
| Energy-BR | 0.18 | -0.14 | -0.24 | | -0.18 | | | -0.14 | | -0.15 | -0.48 | -0.15 | -0.23 | 0.19 | | | 0.48 |
| Energy-C | 0.16 | | -0.25 | 0.19 | | -0.17 | | | | -0.31 | | | | | | -0.27 | 0.31 |
| Energy-CR | | | | | | | -0.16 | | | | | -0.13 | | | | | |
| Energy-D | 0.14 | -0.27 | -0.36 | | | | | | | | -0.39 | | -0.22 | | | | 0.48 |
| Energy-E | | -0.14 | -0.39 | 0.13 | -0.27 | | | | -0.18 | -0.34 | -0.29 | -0.18 | -0.28 | 0.13 | 0.12 | 0.18 | 0.49 |
| Energy-ER | | -0.17 | -0.30 | 0.16 | | | | | -0.23 | -0.27 | -0.24 | -0.22 | | | | 0.22 | 0.41 |

### 11.3.5 Creating impossible Neanderthals and other cheats

In the Standard batch of simulations the values for the parameters are randomly selected from a given range. In the Evolved batch, the GA has taken good performing parameter value sets and mutated these or combined them with others in order to create even more successful parameter sets. The mutations are done to one randomly selected parameter value, and involve an increase or decrease of the original value with ten percent. This occasionally produces parameter values that fall outside of the initial range. For most parameters this was allowed because avoidance of user bias is one of the underlying aims of this study. When extreme values are produced it is interesting why the system evolves the given parameter in the chosen direction. Sometimes however evolved values were clearly impossible. The following parameters were occasionally manipulated beyond the initial boundaries:

- **Birth-rate**: birth-rates evolved to values larger than 50%, or below 1%;
- **Death rates** for all cohorts evolved to values smaller than 1%;
- **Subsistence** needs for all cohorts were reduced to 0 or less;
- The value for **years before group maturity** became occasionally less than 1 year;

- The values for **group size before merge**, **group size fertile cohort before merge** and **group size before split** sometimes became less than 1;
- **Cohort sizes** for fertile and pre-fertile cohort were manipulated to less than 1;
- **Maximum foraging ranges** fell to values below 1 or became larger than 20.

These changes are followed by the genetic algorithm to avoid certain imposed limits (where only a certain range is offered, for example the birth and death rates) or to change functionality (for instance the group maturity variable). The evolved directionality can be surprizing. It is important to realize that the algorithm just manipulates values and then runs simulations to see if these changes are successful. The algorithm does not know the *meaning* of parameters. Therefore it is natural for the algorithm to change death rates to zero, since Neanderthals that never die will be very present everywhere and thus match the archaeology very well. Also birth rates above 50, ignoring any physical limits for Neanderthal women, are for the system very successful parameter values, especially combined with low energy needs.

To avoid clearly impossible Neanderthals, the system was subsequently restricted in the allowed mutation directions for certain parameters: birth rates, death rates, and subsistence needs are now not allowed below one (1), and a maximum foraging range cannot impose a limit of less than one (1) or more than 20. Cohort sizes are kept at a minimum size of one (1). The other mutations were not restricted and the resulting evolved values violating pre-conditions are presented in Table 51. Note that the empty columns here remain to illustrate the effects of imposed restrictions. The strongest violations are scenario Energy-B, with 79 out of a total of 2,278 simulations having a value of 0 or lower for GroupSizeFertile_BeforeMerge (effectively disabling this variable and thus allowing all groups to always merge; but many more scenarios violate this criterium), and Habitat-E with 48 simulations with a Birthrate larger than 50% (out of a total of 2,420).

**Table 51: Number of model parameter value sets per scenario with parameter values manipulated beyond the initial limits. Darker colors identify larger table values.**

| Scenario | BirthRate<=0 | BirthRate>50 | DeathRate_FertileCohort <=0 | DeathRate_FertileCohort <=0 | DeathRate_PostFertileCohort <=0 | Subsistence_PreFertileCohort <=0 | Subsistence_FertileCohort <=0 | Subsistence_PostFertileCohort <=0 | Years_Before_Goup_Maturity<=0 | GroupSize_BeforeMerge <=0 | GroupSizeFertile_BeforeMerge<=0 | GroupSize_BeforeSplit <=0 | CohortSize_PreFertile<=0 | CohortSize_Fertile <=0 | Max_ForagingRange <=0 | Max_ForagingRange>20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | | 11 | | | | | | | | 5 | 6 | | | | | |
| Habitat-B | | 3 | | | | | | | 2 | | | | | | | |
| Habitat-C | | 25 | | | | | | | 1 | | 1 | | | | | |
| Habitat-D | | | | | | | | | 1 | 7 | 10 | | | | | |
| Habitat-E | | 48 | | | | | | | 3 | 4 | 16 | | | | | |
| Habitat-F | | 1 | | | | | | | | 9 | 6 | | | | | |
| Habitat-G | | 1 | | | | | | | | 1 | 8 | | | | | |
| Habitat-H | | 8 | | | | | | | 14 | 17 | 36 | | | | | |
| Habitat-I | | 5 | | | | | | | 2 | 3 | 8 | | | | | |
| Habitat-J | | 6 | | | | | | | | 4 | 14 | | | | | |
| Energy-A | | 9 | | | | | | | 2 | 14 | 1 | | | | | |
| Energy-B | | 12 | | | | | | | 5 | 1 | 79 | | | | | |
| Energy-BR | | 11 | | | | | | | 1 | | 2 | | | | | |
| Energy-C | | 10 | | | | | | | 10 | 2 | 20 | | | | | |
| Energy-CR | | 2 | | | | | | | | 10 | 3 | | | | | |
| Energy-D | | | | | | | | | 1 | | 16 | | | | | |
| Energy-E | | | | | | | | | | 6 | 21 | | | | | |
| Energy-ER | | 1 | | | | | | | | 6 | 14 | | | | | |

## 11.4 Trends in the simulation output

A general observation of the simulation results could suggest that the model behaves in a linear fashion, where more Neanderthals mean a better match with the archaeological data. More or less as expected since presence is a success criterion and will only be limited by resources and the presence of other hominins. But correlations from the previous section show that this relation is not straightforward. This section therefore focusses on relevant patterns in the simulation output, in order to understand the simulation process in more detail.

### 11.4.1 The simulation score

The `MatchedIntervalCoverage` variable is a crucial element in the HominінSpace simulation system, because simulation results are assessed on this value alone. The implemented optimization method, genetic algorithms, requires a single number to optimize upon, and therefore is it essential to understand how this value is constructed. Each checkpoint reports at the end of a simulation several output variables, including a

local [checkpoint name]_matchedIntervalCoverage number. This value indicates how well the simulated visits to the checkpoint match with the archaeological attested hominin presence (subsection 7.5.1). All local values are all included in the simulation output file and summed together into this `MatchedIntervalCoverage` variable[43]. Here I describe patterns in the simulation score throughout the simulated scenarios.

Figure 71 presents graphs with `MatchedIntervalCoverage` scores, sorted from low to high (here *not* sorted on simulation number). Overall, the score patterns for all scenarios are similar: up to a certain value the simulations do not result in a very high score. Then there is a (small) range presenting a tipping point: suddenly very high values are scored. Then the graphs plateau with a small but noticeable increase at the end of the spectrum. These figures illustrate three important aspects of the score: (1) until the tipping point which lies generally around 1500 the scores are low and most often from the Standard set (cyan coloured dots); (2) after the tipping point the score increases (sometimes steeply) and then levels again until you reach the small tip at the end which is constructed of very high scores with almost exclusively evolved results; (3) for all scenarios the exact score pattern of `MatchedIntervalCoverage` is different and depends on the model parameters values and settings.

[43] All relevant simulations from all scenarios are copied into the file "AllData.csv" which is included in the Supplementary Materials.

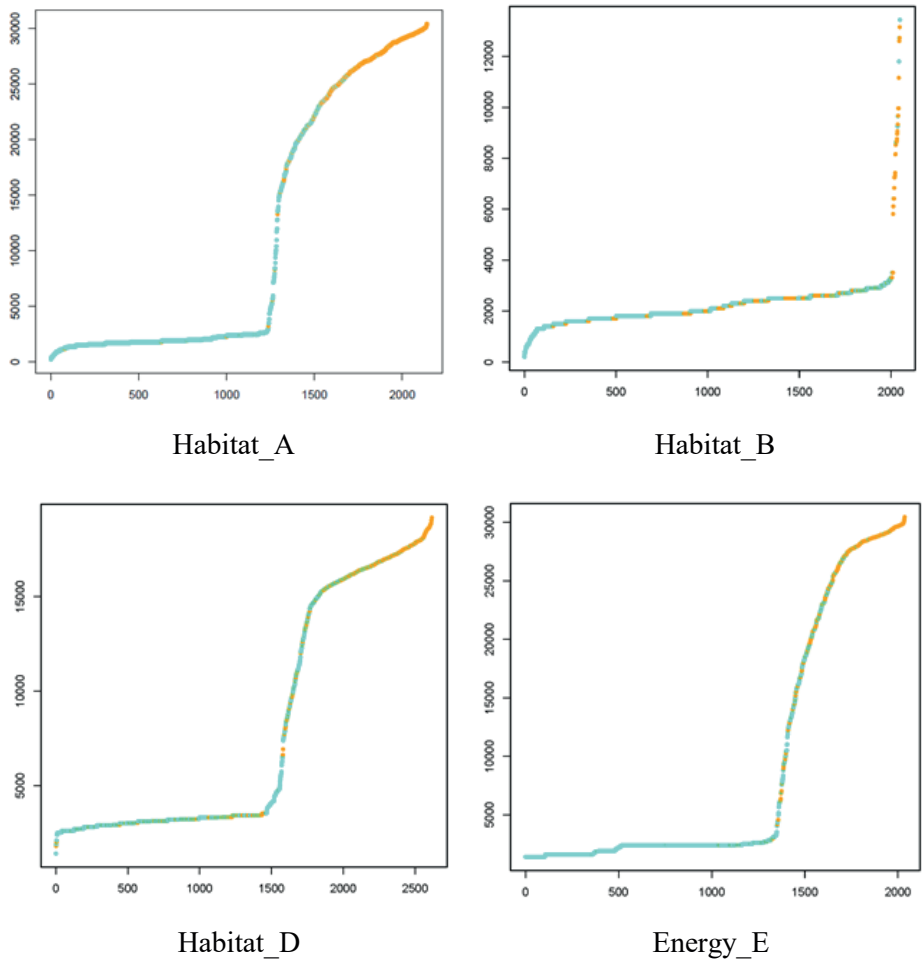



**Figure 71: Overview of sorted `MatchedIntervalCoverage` values for different scenarios. Simulations are sorted on score. Scores from the Standard set are colored cyan.**

By adding the scores for all simulations for all scenarios together and sorting on resulting sums it is possible to rank the scores for all checkpoints[44]. Values range from 663,545 for Saint-Amand-les-Eaux to a maximum of 29,701,374 for Pech de l-Aze II . This is all scores added together. Note that in individual scenarios the contributions of individual checkpoints differ widely. Excluding the monitoring checkpoints and the checkpoints with absence intervals, the checkpoints that have the lowest `_matchedIntervalCoverage` scores in total are: Saint-Amand-les-Eaux (663,545), Ault (814,400), Savy (814,400), Ormesson (867,377), and Beauvais 1 (1,004,270). These low scoring checkpoints all have one or two short intervals only, all positioned temporarily at or even overlapping with the end of the simulation period (some of these intervals will never be completely matched, since the simulations stop at 50 ka).

[44] The resulting excel sheet is included in the Supplementary Materials as “Summed data only.xlsx”.

The best scoring checkpoints are Grotte de Coudoulous II (10,762,343), Abri Bourgeois-Delaunay (12,030,833), Chez-Pinaud Jonzac (13,464,009), Coudoulous I(14,982,272), and Pech de l-Aze II (29,701,374). The site Pech de l-Aze II has by far the most intervals which helps to explain its very high score (each matched interval counts). Both Coudoulous I and Grotte de Coudoulous II have many very accurate (and thus very short and easily filled) intervals that add bonus to the score when completely matched (and are also positioned in the same grid cell).

### 11.4.2 Characterizing the best models from the Standard set

For the discussion in this subsection I refer to the Top-10 results from the Standard model sets, for all scenarios. These results were retrieved from the simulation output files and are stored for easy reference in the Excel data file named “Results – Top10 Standard sets all scenarios.xlsx”, included in the Supplementary Materials. The models in the Standard set have randomly generated parameter values and are executed in each scenario. By chance, some of these models are well suited for simulation in HomininSpace in certain scenarios while unfit for others. See  for good performing models and the associated parameter values.

Roughly there are two successful model groups: those with a very low `ViabilityIndex` value (below 190, with some exceptions) which are very successful in simulations with the absence criterion activated, and those with a viability index of 190 and up, scoring high in most other scenarios. Since there are only four scenarios with absence activated the successful model parameter sets for these scenarios are not represented in .

A low viability index can be due to a low birthrate (e.g. 832, 876, 461, 80, 280), high mortality figures (3, 157, 542, 912) or combinations thereof (3, 83, 533), supplemented by additional measures that can include more positive values for temperature tolerance (178, 857, 1008; sometimes combined with very high values for group size before split (754)), and small foraging ranges (43, 323, 1009). A high viability index is realized with a high birth rate (390, 486, 1355, 1373), low mortality rates (43, 486, 1309), especially for the fertile cohort (120, 1365, 1384), more negative temperature tolerances (39, 1309, 1384), and large foraging ranges (395, 646, 1131).

Models that are successful in multiple scenarios (more than six times) with high values for the viability index are 39, 390, 472, 486, 1131, 1365, 1373, and 1384. 1365 scored by far

the best, but 486 has high places in the top-10 of many scenarios. A high viability index gives a good chance on a high simulation score but is no guarantee. For instance, number 120 has in an index of 62186 (maximum in this list), and finds itself only 4 times in any top 10. Model 1365 (13 times top 10, with 8 x place 1) has an index of 28722, but 1131 (9 times in the top 10) has an index of only 1502. Model 560 has an index of 301 (7 times top 10). Note that almost all small scale foraging societies show the positive growth rates associated with larger viability index figures (Gurven and Kaplan 2007, 347)[45].

**Table 52: Presenting the best performing models from the Standard set. The first column gives the number of times the simulation number appears in the top-10 scores over all scenarios.**

| Times in Top 10 | SimulationNumber | 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 | ViabilityIndex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13 | 1365 | 43 | 3 | 1 | 12 | 2000 | 5000 | 3250 | 2 | 4 | 2 | 16 | -17 | 9 | 37 | 2700 | 12 | 28722 |
| 12 | 39 | 47 | 5 | 1 | 11 | 2750 | 3250 | 3750 | 5 | 5 | 2 | 56 | -28 | 12 | 23 | 3000 | 10 | 4053 |
| 9 | 390 | 47 | 6 | 2 | 2 | 3250 | 4750 | 2000 | 1 | 5 | 5 | 56 | -29 | 9 | 38 | 3150 | 13 | 12842 |
| 9 | 1131 | 35 | 8 | 2 | 12 | 3000 | 3750 | 4750 | 4 | 1 | 3 | 46 | -26 | 8 | 40 | 3100 | 15 | 1502 |
| 9 | 1384 | 41 | 11 | 1 | 14 | 4250 | 2750 | 4250 | 10 | 5 | 1 | 46 | -29 | 9 | 28 | 2900 | 11 | 1289 |
| 8 | 486 | 49 | 2 | 2 | 12 | 2750 | 4750 | 4250 | 5 | 2 | 1 | 56 | -23 | 8 | 25 | 2550 | 10 | 41287 |
| 7 | 560 | 27 | 3 | 1 | 15 | 2500 | 4000 | 3500 | 8 | 4 | 2 | 16 | -16 | 14 | 25 | 3100 | 9 | 301 |
| 6 | 472 | 33 | 4 | 2 | 4 | 2500 | 3000 | 4250 | 8 | 1 | 3 | 56 | -17 | 9 | 35 | 3200 | 12 | 1978 |
| 5 | 1339 | 35 | 4 | 2 | 3 | 3750 | 3500 | 3000 | 8 | 1 | 1 | 16 | -23 | 12 | 17 | 2750 | 14 | 280 |

### 11.4.3 Causes of life and death

For each simulation administrative data is stored on the total cumulative number of hominin individuals and groups through time. The variables for these simulation output results are named `created_hominins` and `created_groups`. Within HomininSpace there are two ways in which new hominins can be introduced into a simulation: via the hominin reproduction mechanism and via creation inside core areas (that is, produced by so-called factories if these are activated). Where the size and composition of reproduced groups depend on the parent group (half the size of the parent cohorts), core areas produce new groups (one per time step, and only when conditions are favourable) with a standard group size and structure: 25 individuals of which 8 are in the pre-fertile age, 11 fertile and 6 are post-fertile individuals[46].

[45] By definition, since prolonged negative growth rate leads to extinction.
[46] These numbers are numeric constants in the source code and can be changed at will. For this research however they were not changed between simulations.

The number of created hominins are stored in the variables `Offspring`[47] for reproduced hominin groups and in the variables `Iberia` and `Italy` for groups created in the two predefined core areas (if these are activated). The number of created groups and hominins help in interpretation of the simulation processes and results. Figure 72 suggests that the number of created hominins is probably very important for obtaining higher simulation scores (later simulations all have high created hominins numbers). For scenario Habitat-A the relation between the simulation score and the number of created hominins is visualized in Figure 73. Also the number of created groups are important, but extreme values are not always necessary for high simulation scores.

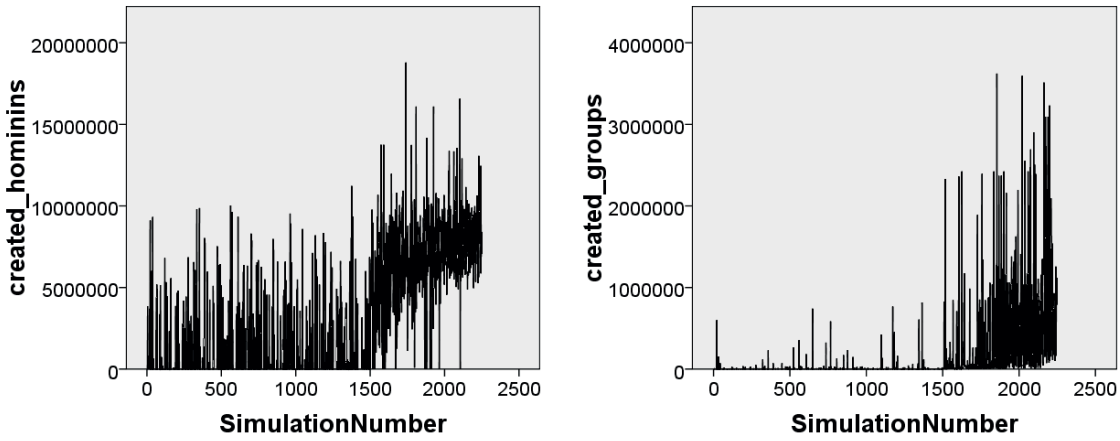



**Figure 72: Overview of created_hominins and created_groups versus SimulationNumber from scenario Habitat-A. The first 1500 simulations are the Standard set.**

[47] Note that the Offspring variable does not include starting populations and produced groups thereof.

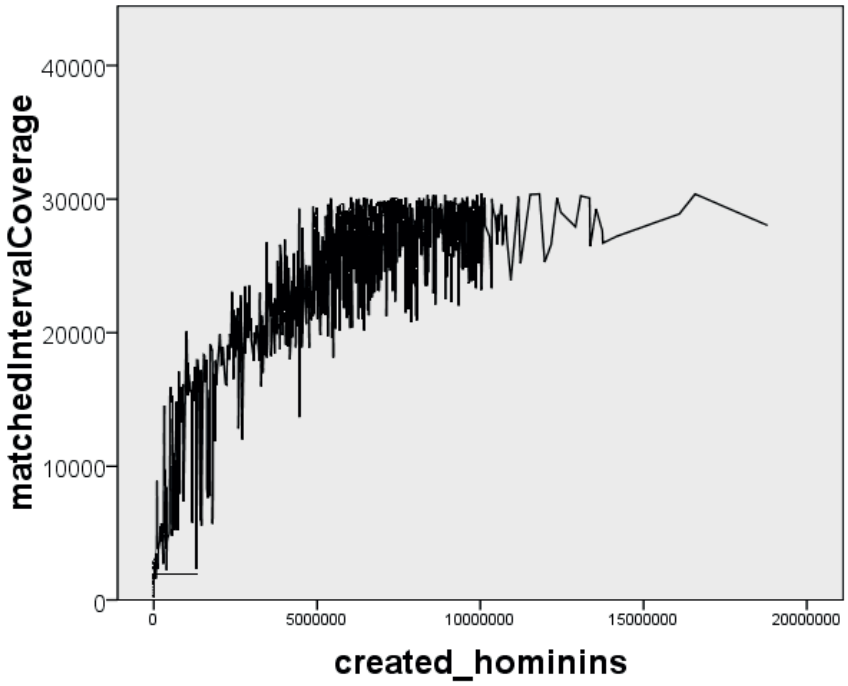



**Figure 73: Relation between the number of created hominins per simulation and the simulation score (scenario Habitat-A).**

Hominins can die in many different ways. In the model underlying the HominínSpace simulation system the mortality rates for all cohorts are applied to all hominin groups every year, reducing the population as a whole. In addition to this natural attrition that varies due to the individual circumstances of each group, a group can succumb as a whole to cold, flooding, hunger (insufficient resources), and to being too small to sustain itself. In all these cases the whole group ceases to exist and disappears. The numbers of hominins that perish in these manners are administrated and totals are reported in the simulation output at the end of each simulation. They also cease to exist in the simulation because they cross the border of the simulation area (if this is possible). Groups that disappear because they merge into other groups are counted separately. Table 53 lists the maxima for all these variables per scenario, to illustrate the differences between scenarios. Note that death by flooding can only happen to static hominins, save very unfortunate exceptions where groups have no choice than to move onto or remain in flooded grid cells.

**Table 53: Maxima for the numbers of created hominins and groups, and maxima for the number of different possible deaths in any simulation. Color coding is per column and darker colors indicate higher scores within the column.**

| Scenario | Max CreatedHominins | Max CreatedGroups | Max CreatedOffspring | Max Iberia | Max Italy | Max DeathsCold | Max DeathsFlooding | Max DeathsHunger | Max DeathsMerging | Max DeathsTooSmall |
|---|---|---|---|---|---|---|---|---|---|---|
| Habitat-A | 18795432 | 3622260 | 3621991 | | | 165967 | | 8857 | 3620674 | 251099 |
| Habitat-B | 1138269 | 93108 | 93087 | | | 834 | | 10 | 92964 | 104 |
| Habitat-C | 13721186 | 8090115 | 8089607 | | | 233517 | | 10448 | 7977343 | 269545 |
| Habitat-D | 8904608 | 1972062 | 1830538 | 151187 | 194166 | 377484 | | 5963 | 1511144 | 1379726 |
| Habitat-E | 19827306 | 5470156 | 5469788 | 158762 | 185920 | 254238 | 1 | 55062 | 5384471 | 1405852 |
| Habitat-F | 8899456 | 1588281 | 1514281 | 158762 | 185920 | 254238 | | 6429 | 1395871 | 1405852 |
| Habitat-G | 13539102 | 2946697 | 2946546 | | | 78030 | | 11097 | 2918412 | 119754 |
| Habitat-H | 14481376 | 4604894 | 4522427 | 150057 | 184333 | 359509 | 17 | 7556 | 4590155 | 1497503 |
| Habitat-I | 15777251 | 6095733 | 6095362 | | | 52375 | 18 | 5748 | 5984937 | 148395 |
| Habitat-J | 20138638 | 3286710 | 3286511 | | | 69212 | 311 | 5256 | 3278679 | 207564 |
| Energy-A | 5801491 | 2314134 | 2313957 | | | 53017 | | 6523 | 2312831 | 60427 |
| Energy-B | 12826469 | 6211677 | 6211276 | | | 163483 | 1 | 33627 | 6163162 | 656422 |
| Energy-BR | 6960929 | 3954676 | 3954421 | | | 48730 | 65 | 11059 | 3931901 | 88748 |
| Energy-C | 1364805 | 9245 | 9232 | | | 24 | | 22 | 9197 | 732 |
| Energy-CR | 2791 | 404 | 242 | | | 28 | | 14 | 186 | 56 |
| Energy-D | 7612793 | 1718926 | 1718811 | | | 137895 | | 7010 | 1712553 | 93514 |
| Energy-E | 10221491 | 1746704 | 1746589 | | | 25829 | 91 | 3381 | 1739833 | 83280 |
| Energy-ER | 9091567 | 2163739 | 2163621 | | | 54276 | 90 | 6098 | 2106022 | 105571 |

The output variables detailing mortality are named: `deathsCrossing`, `deathsCrossingIberia`, `deathsCold`, `deathsFlooding`, `deathsHunger`, `deathsMerging`, and `deathsTooSmall`. These numbers are included in the simulation results to aide in interpreting the simulation process and results. The geographical distribution of the deaths as they occurred in the simulated environment can be visualized in maps (subsection 8.7.2). They are also used to illustrate the batch simulation process and the effect of the evolution implemented in the Genetic Algorithm. An example of this is presented in Figure 74, which illustrates that in scenario Habitat-E the Evolved hominins perish in colder circumstances than the Standard hominins. In this scenario there is a relatively strong negative correlation between `Temperature_Tolerance` and simulation number, resulting in hominins that can sustain much colder temperatures.

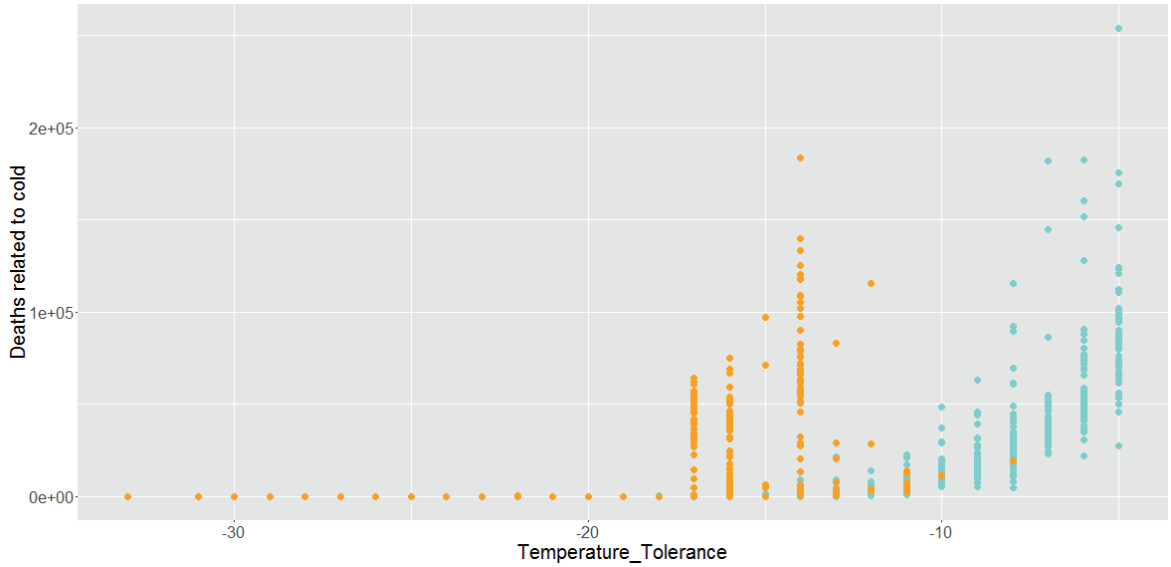



**Figure 74: Depicted the numbers of groups that perished due to cold temperatures versus temperature tolerance. In orange evolved hominins, in cyan results from the Standard set (scenario Habitat-E).**

The lives of individual hominins are not tracked in the HominínSpace system, but groups can survive for many years, outlasting the individual hominins they are composed of. The lifetime or age of the hominin groups is variable and depends on group composition, model parameters and local circumstances. Sometimes groups can survive up to 50,000 years or more. The average group age per simulation is depicted for scenario Energy_E to illustrate the possible variation within the resulting models in Figure 75. Note that for some very successful simulations the average age is not necessarily very high.

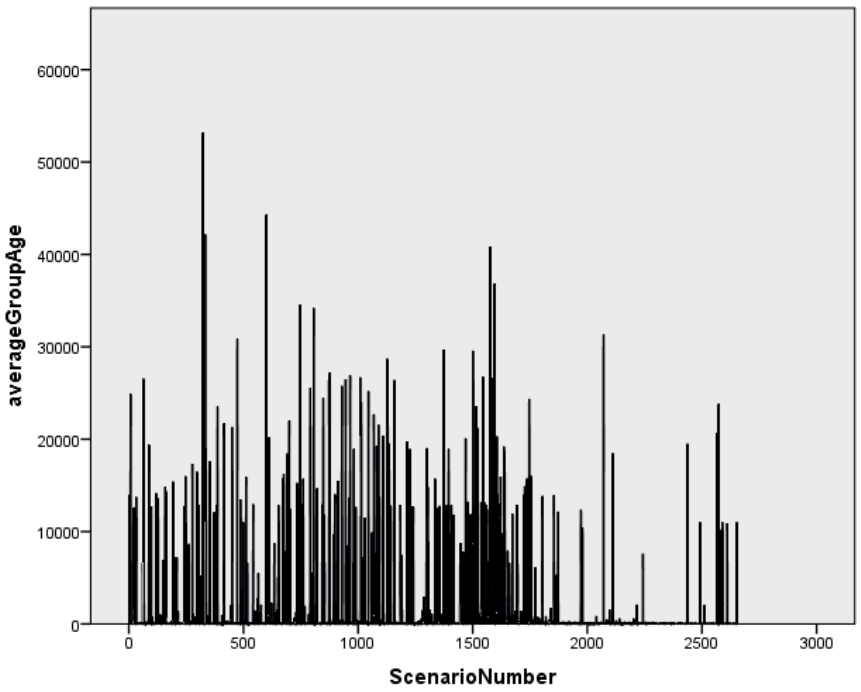



**Figure 75: Average group age for scenario Energy_B.**

### 11.4.4 Influence of data quality on the simulation score

Each interval has been assigned a quality indication. This quality value can be used in assessing the total simulation score with all checkpoint intervals versus the score excluding the bad quality intervals. To facilitate this comparison two results `matchingVisits` and `matchingVisitsWithConfidence` are included in the simulation results file. The variable `matchingVisits` counts the number of years that has visits inside an interval. The `matchingVisitsWithConfidence` stores the interval visit results with the confidence level as a factor included in the count (see subsection 6.3.1). To force a strong influence of the confidence level, a visit of confidence level one (1) is counted once, for level two (2) a visit adds four to the count and for level three (3) nine points are added to the score. Thus `matchingVisitsWithConfidence` always scores at least the same as `matchingVisits`, but depending on the confidence levels of the visited intervals can score much higher. In total there are 29 intervals that were classified as unreliable with level one, 206 with level two and 236 with level three (n=471).

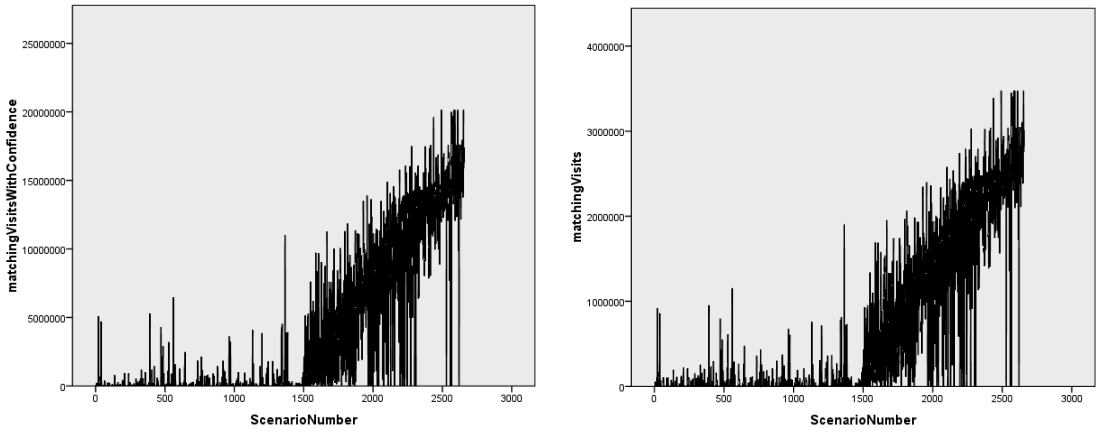



**Figure 76: Output of the simulation variables `matchingVisitsWithConfidence` (left) and `matchingVisits` (both from Scenario Energy-B).**

Both graphs have an almost identical shape, with the absolute values being different and the values for `matchingVisitsWithConfidence` consistently almost a factor 5 higher. This illustrates clearly that the influence of the unreliable dates is constant throughout the simulations. This suggests that the presence of these few unreliable dates has a minimum influence on the resulting output.

The correlation between `matchingVisits` and `matchingVisitsWithConfidence` is for all scenarios nearly perfect, with a significant correlation value of +1.00, the maximum score possible. Correlations between the simulation score (`MatchedIntervalCoverage`) and both variables are also

similar, both significant and with a value around +0.96, indicating a near perfect correlation (understandable since the simulation score is constructed using primarily the matching visits results). Since the simulation score is based on matching visits it includes matches for intervals that are classed as unreliable, but overall this inclusion should therefore not influence the results.

Intervals are constructed using the one standard deviation value for lower and upper bounds. This presents a 68% confidence interval around the mean. Thus a less reliable date has a larger standard deviation which results in a larger interval. To ensure that a larger interval does not produce a higher score, the reported score per interval is divided by the length of the interval, resulting in a coverage percentage, not an absolute number. All percentages added together and multiplied by 100 to avoid rounding form the `MatchedIntervalCoverage` score.

## 11.5 Analysing the modelled environment

The environment in HominінSpace uses the reconstructed topography and the reconstructed climate parameters temperature and precipitation. A fluctuating sea level defines available landmasses where the climate determines how much energy there is available in the landscape. The energy level per grid cell is calculated directly by extrapolation or via habitat reconstruction. The reconstruction is analysed here visually with the user interface of the application and by using the data from the climate recording checkpoints. Identified are larger patterns and local deviations.

Table 2 identifies local and global climatic maxima and minima in the data supplied by Bintanja and van de Wal (2008). Those points in time that are part of the simulation period are used to illustrate the topography and energy reconstructions. All dates are in ka, with temperature offsets between brackets in degrees Celsius. The minima are at: 112 (-9.70°), 88 (-12.07°), 78 (-11.85°), and 61 (-13.31°), and maxima can be found around 124 (2.19°), 120 (0.57°), 97 (-6.53°) and 83 (-6.51°). With the starting point at 131 ka (-5.6°) and simulation termination at 50 ka (-12.0°) this encompasses in total ten points in time to capture the reconstructions. These moments are listed in Table 54 ordered by time step, with the given temperature given as offset to the current global mean temperature.

**Table 54: Relevant points in time (climate extremes) for screen capturing.**

| *Description* | *Time step* | *Date (years ago)* | *Temperature (Celsius)* |
|---|---|---|---|
| *Start of the simulation period* | 1 | 131,000 | -5.6 |
| An absolute maximum around 124 ka | 6,200 | 124,800 | 2.19 |
| A local maximum around 120 ka | 9,600 | 121,400 | 0.57 |
| A local minimum around 112 ka | 18,200 | 112,800 | -9.70 |
| A local maximum around 97 ka | 33,800 | 97,200 | -6.53 |
| A local minimum around 88 ka | 42,300 | 88,700 | -12.07 |
| A local maximum around 83 ka | 47,200 | 83,800 | -6.51 |
| A local minimum around 78 ka | 52,900 | 78,100 | -11.85 |
| A local minimum around 61 ka | 69,500 | 61,500 | -13.31 |
| *End of the simulation period* | 81,000 | 50,000 | -12.0 |

To assess changes in energy level values in the simulation the climate recording checkpoints are used. To interpret the observed dynamics it is important to realize that both energy computational methods use values for precipitation and temperature that are calculated by interpolation between and extrapolation from current day and reconstructed LGM values. Table 55 presents the recording checkpoints with their LGM and current day values. Note that for climate monitor 1 the LGM and current day values for precipitation are the same, so there is minimal variance throughout the simulation period.

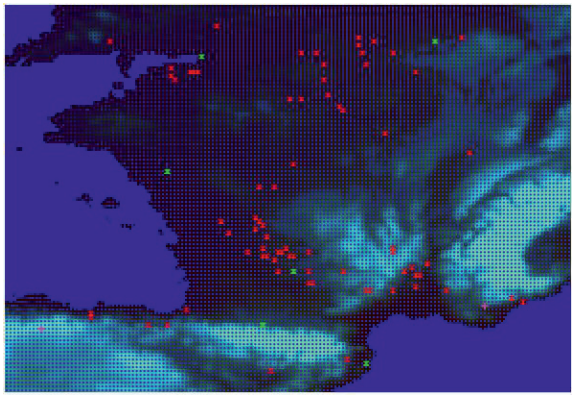
**Table 55: The locations of the climate recording checkpoints and their modelled LGM and current day precipitation (P) and temperature (T) values.**

| # | *Name* | *Lat* | *Long* | *LGM T* | *LGM P* | *Cur T* | *Cur P* |
|---|---|---|---|---|---|---|---|
| **1** | Climate Monitor 1 Atlantic | 47.0 | -2.0 | 4.4 | 828.5 | 11.9 | 828.5 |
| **2** | Climate Monitor 2 Occitanie | 44.4 | 1.3 | 7.1 | 934.5 | 12.2 | 814.6 |
| **3** | Climate Monitor 3 Mediterranean | 42.0 | 3.2 | 11.8 | 655.0 | 15.9 | 583.6 |
| **4** | Climate Monitor 4 Pyrenees | 43.0 | 0.5 | 5.8 | 1143.9 | 10.4 | 892.2 |
| **5** | Climate Monitor 5 Scladina | 50.5 | 5.0 | -5.1 | 870.0 | 9.4 | 865.6 |
| **6** | Climate Monitor 6 Channel | 50.0 | -1.1 | -4.0 | 823.8 | 11.4 | 800.5 |

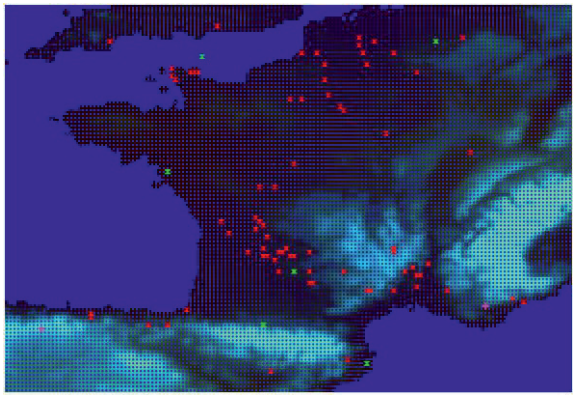
### 11.5.1 Topography

The only factor in HominínSpace influencing the topography is the changing sea level. Large areas of land currently under water in the Channel between the European mainland and the British Islands, and at the Atlantic coastal side of France become dry land when the sea level drops due to colder temperatures. The effects are less pronounced for the Iberian Peninsula and minimal in the Mediterranean areas. Screenshots for the indicated moments in time are given in Figure 77, with date and temperature offset. These figures illustrate that the available land mass varies considerably through time, with large areas becoming

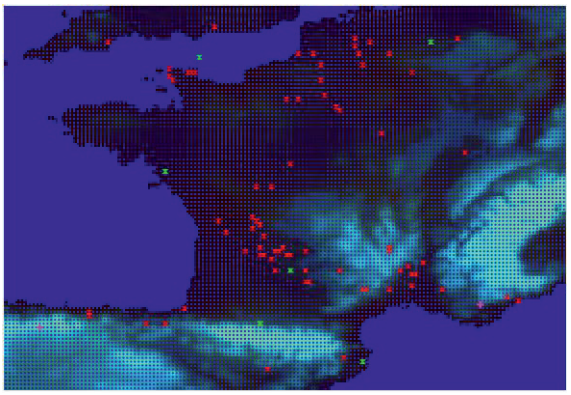
submerged when water level rises due to warmer global climates and melting of glaciers, and subsequently becoming dry land when the temperatures and sea level drop. Coloring of grid cells is according to the relative height to sea level.



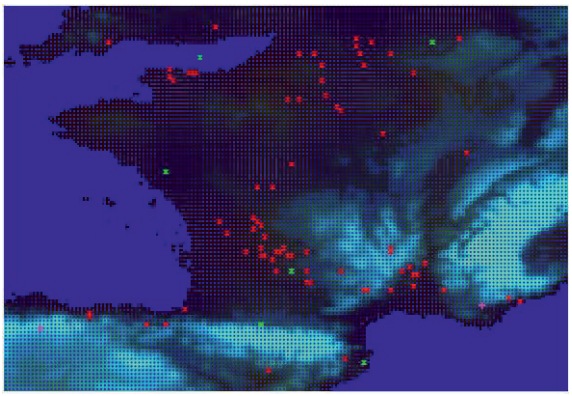
131,000 years ago (-5.6°)



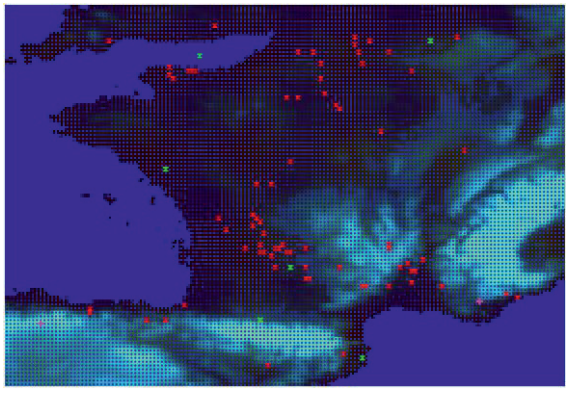
124,800 years ago (2.19°)



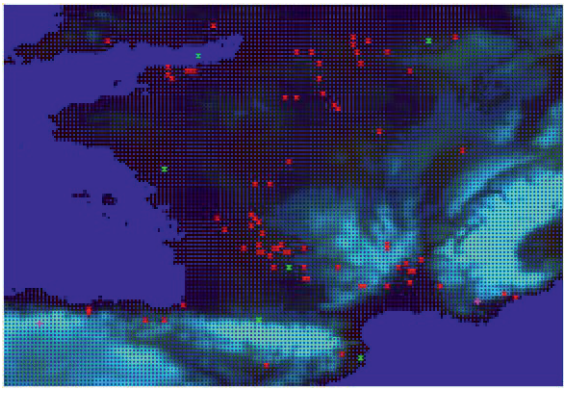
121,400 years ago (0.57°)



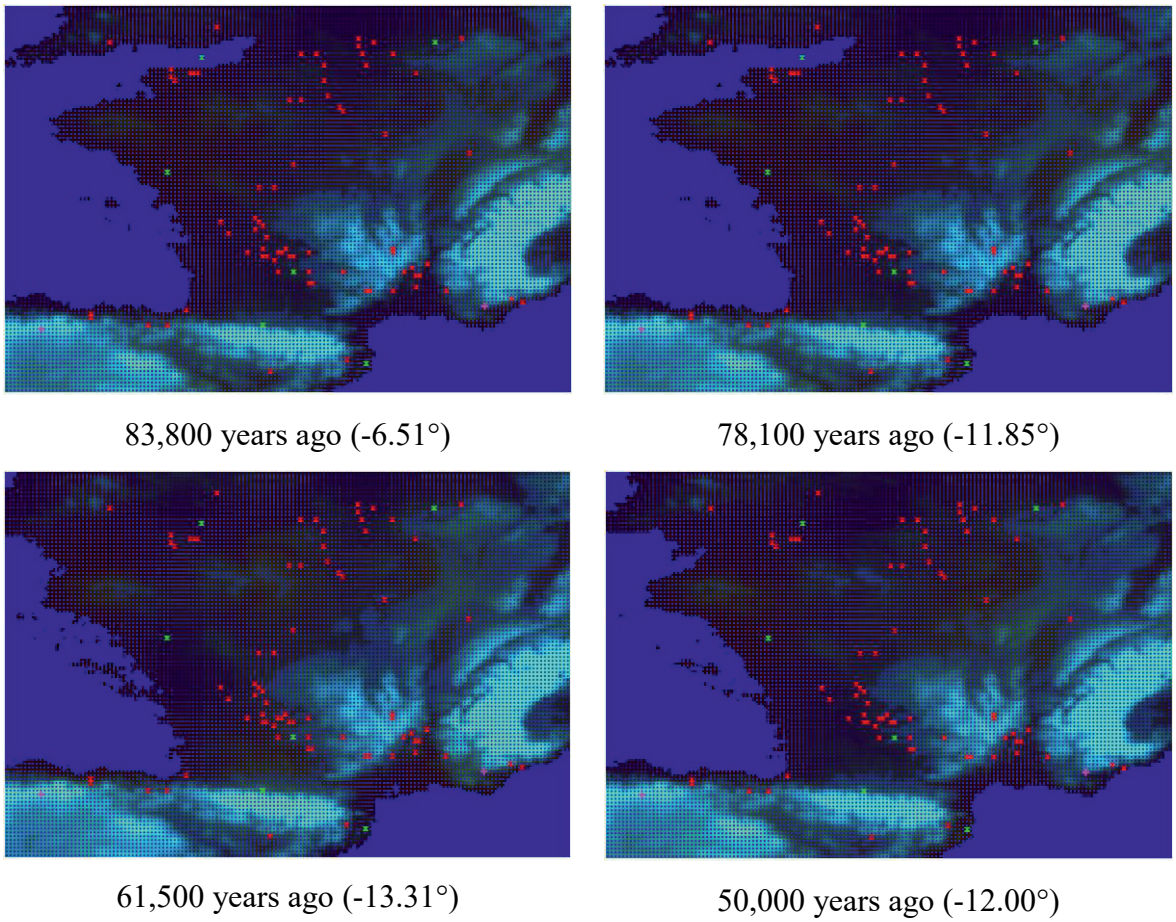
112,800 years ago (-9.7°)



97,200 years ago (-6.53°)



88,700 years ago (-12.07°)



83,800 years ago (-6.51°) 78,100 years ago (-11.85°)

61,500 years ago (-13.31°) 50,000 years ago (-12.00°)

**Figure 77: Reconstruction of the topography during the simulation period for selected moments in time. Lighter blue indicate higher elevation.**

Climate recording checkpoints monitor, among other environmental data, the topography of the grid cell they are located in. One checkpoint in particular is positioned to monitor the changing sea level, and that is “Climate Monitor 6 Channel”, located at 50.0 latitude and -1.1 longitude (Figure 78). Climate recording checkpoints are depicted as small green crosses in Figure 77, and checkpoint 6 is the top left green cross. Today it would be located right in the middle of the Channel but at the start of the simulation period it finds itself on land very close to shore where at the end of the simulation it lies deep inland. Right after the start of the simulation the grid cell where this checkpoint is located becomes submerged, resurfacing around 110,500 years ago. Then drowning again shortly after that (106,101 years ago), only to become dry land around 74,700 years ago until the end of the simulation. The recorded topography type is visualized in Figure 78. Note that whenever the grid cell is dry land located directly next to a sea grid cell it becomes a beach type grid cell (topography type 2 in Figure 78).

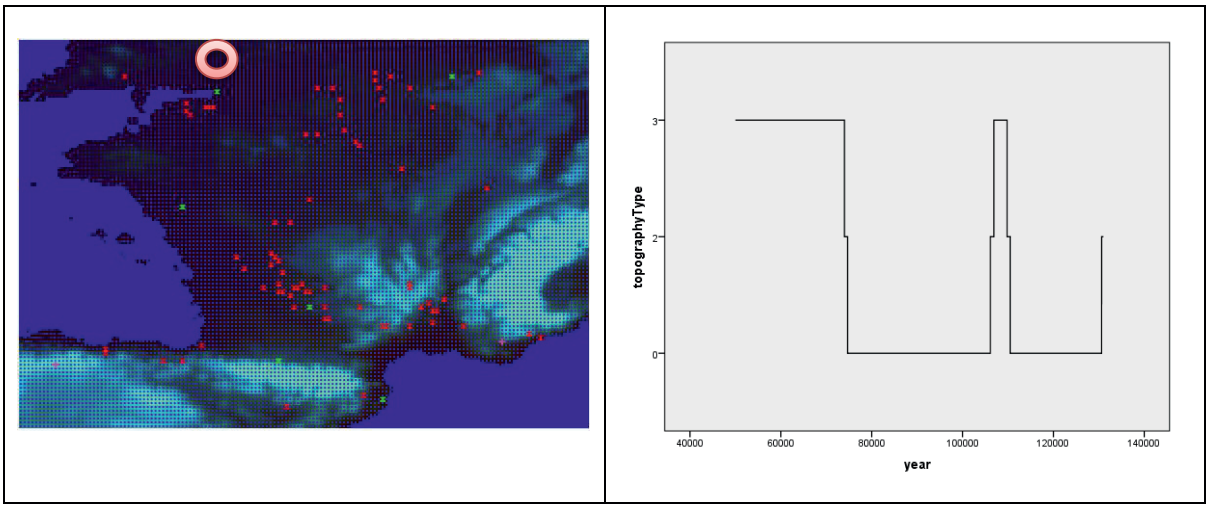


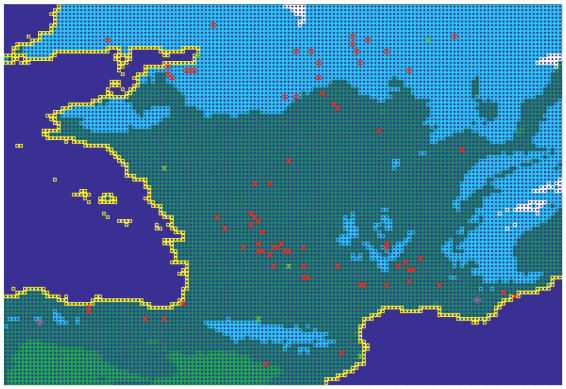

**Figure 78: Location of Climate Monitor 6 Channel (left) and the topographical output for the simulation period. Topography type 0 in the graph is water, 2 is beach and 3 is plain.**

The results show that a land bridge exists to the UK at the location of this checkpoint for a limited period of time during the simulation. At other locations (to the north) access to the British Islands would have been longer possible. The British Islands were only really separated from the European main land during a relatively short period in the simulations. This period of island status in HomininSpace starts around time step 4,220 and ends at 15,815 (that is between 126.8 ka and 115 ka), spanning roughly 11,000 years. These results are similar to for instance the reconstruction by Ashton and Lewis (2002), Figure 3. These authors identify a similar time frame with direct access to Britain, but prevent hominins to cross the dry Channel area in most of MIS 3 with climatic factors. When they include a progressive subsidence of the North Sea Basin floor, access from the main land was limited to a very small window in time around 125 ka (Ashton *et al.* 2011).

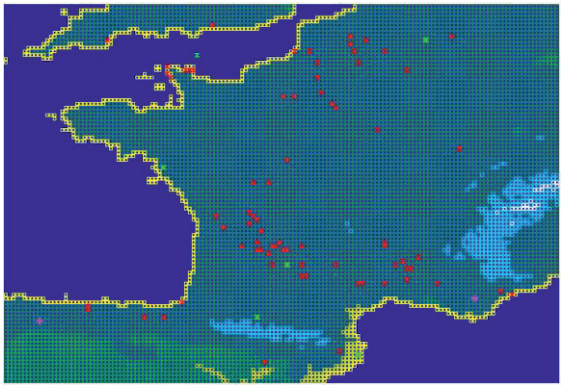
Note that height in the topography does not have an explicit effect on movement of hominins. In other words, topography does not prevent nor promote hominin groups to certain areas in the landscape. However, the reconstructed climate parameters for especially elevated grid cells result in limited availability of resources in for instance the wider Alpine and Pyrenean regions (see the next subsections). Thus these areas are less attractive for hominins which results in lower presence densities during simulations. Mountainous areas specifically do not prevent access by modelled Neanderthal hominins as Neanderthal presence is attested in occasional archaeological finds from higher altitudes (Domingo *et al.* 2017).

### 11.5.2 Energy via habitat reconstruction

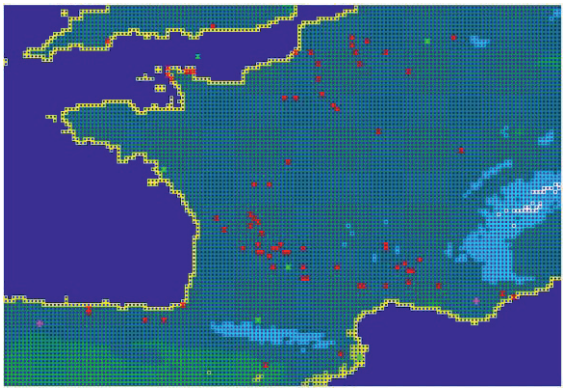
To visualize changes in the reconstructed habitat Figure 79 presents screenshots from identified global and local climate minima and maxima within the simulation period. Each habitat type is defined by precipitation and temperature parameter values, and has associated energy levels (subsection 4.4.3). In the following figures white coloured grid cells are tundra, light blue is boreal forest, dark green is evergreen forest and lighter green is grassland (see Table 22). The yellowish coloured patches in the Mediterranean area in warmer periods are very productive woodlands. The reconstructed habitats are independent of the chosen scenario settings and cannot be manipulated by the foraging hominins.



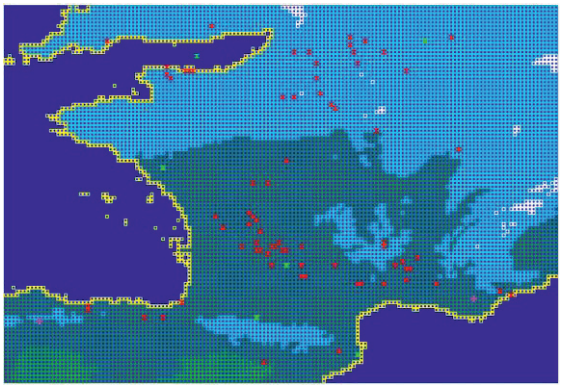
131,000 years ago (-5.6°)



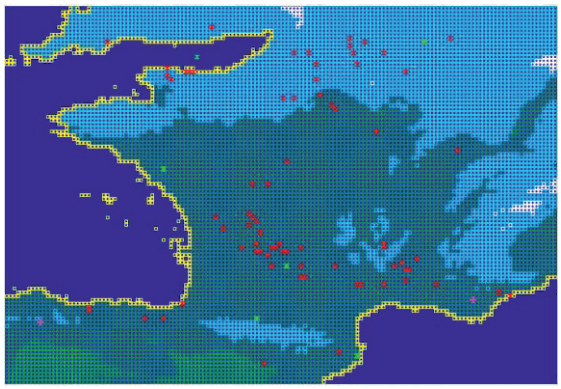
124,800 years ago (2.19°)



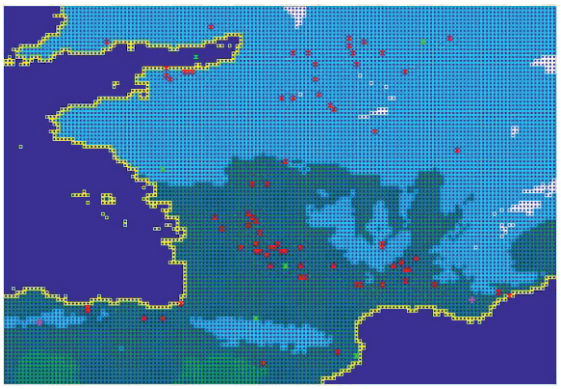
121,400 years ago (0.57°)



112,800 years ago (-9.7°)



97,200 years ago (-6.53°)



88,700 years ago (-12.07°)

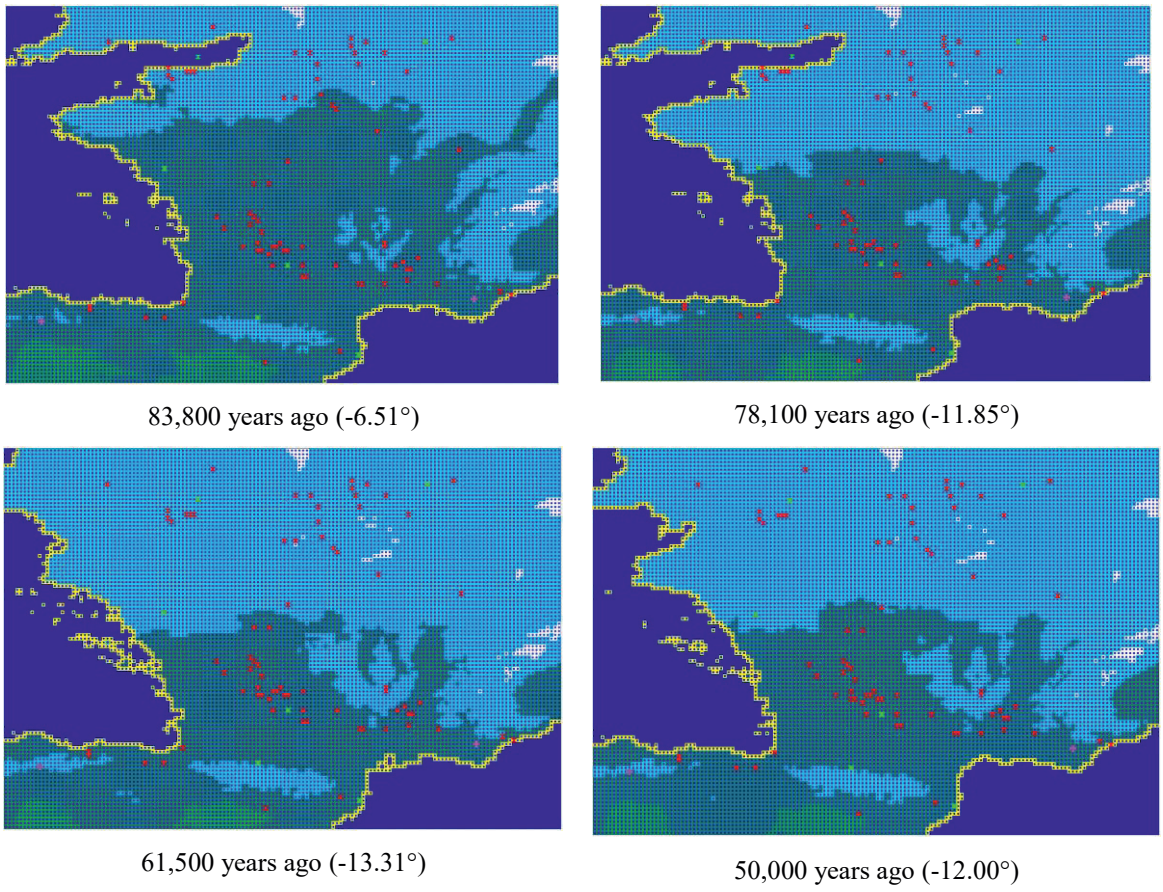



**Figure 79: Habitat reconstruction for the simulation period.**

Six climate recording checkpoints are installed when a simulation is started. Two of them were positioned specifically to monitor changing habitats: “Climate Monitor 1 Atlantic” and “Climate Monitor 3 Mediterranean”. See Figure 80 for the resulting changes through time. Near the Atlantic coast the habitat oscillates between what is referred to as cool forest and boreal forest. Near the Mediterranean Sea there is a period of (very productive) grass land, where during the warmer middle of MIS 5e there even exist a habitat that is referred to as warm grassland, or a habitat comparable to the current day African Savannah (yellowish grid cells in Figure 79).

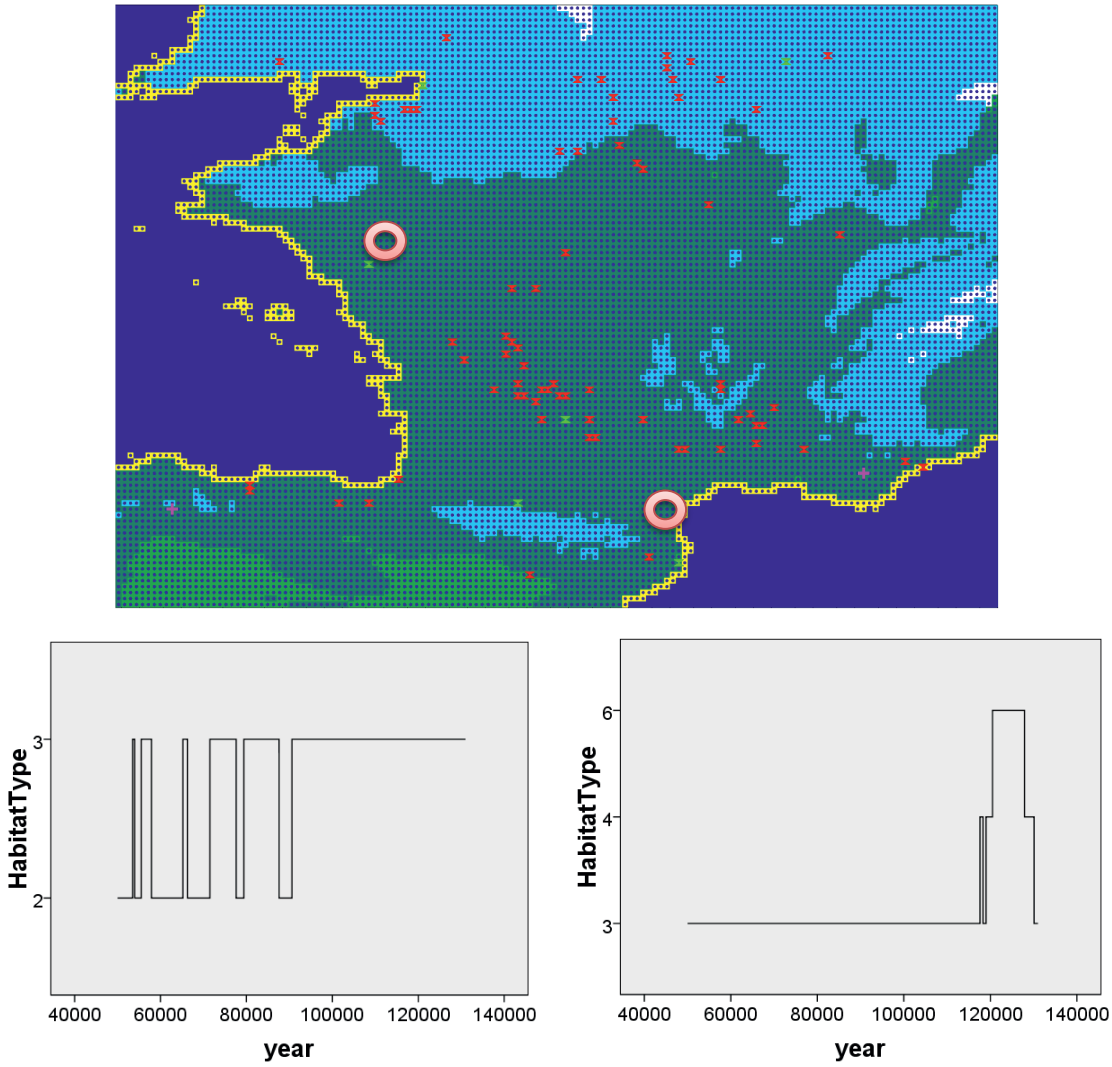



**Figure 80: Geographical positions and Habitat types for the checkpoints Climate Monitor 1 Atlantic (bottom left) and Climate Monitor 3 Mediterranean (right). Habitat type 2 is boreal forest, 3 is cool forest, 4 is cool grass, and 6 is warm grass.**

Climate Monitor 5 is positioned at the Scladina archaeological site, to monitor changes in this specific area. It is a site for which extensive climate reconstruction efforts have been made (López-García *et al.* 2017), some of which can be compared against the reconstructions in HomininSpace for the simulation period. Figure 81 plots the resulting habitat reconstruction in HomininSpace together with the reconstruction from survey C4 (López-García *et al.* 2017, 629). In HomininSpace throughout the simulation period a forested environment is reconstructed, characterizes for most of the time as boreal forest, and only during a short period before 120 ka this was replaced by cool forest. The reconstructed C4 survey also contains mostly woodland, with a varying percentage of open environment, never surpassing 50% of the non-water landscape. In HomininSpace the Scladina area the climatic parameters never throughout the simulation period allow reconstruction of extensive grass lands.

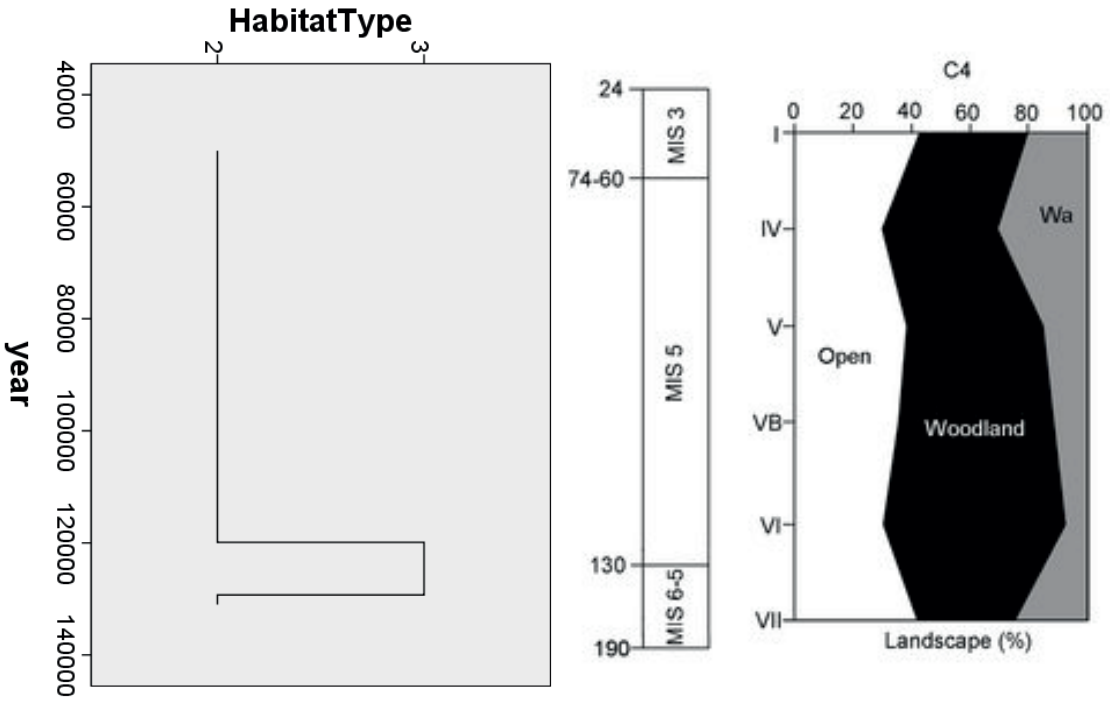
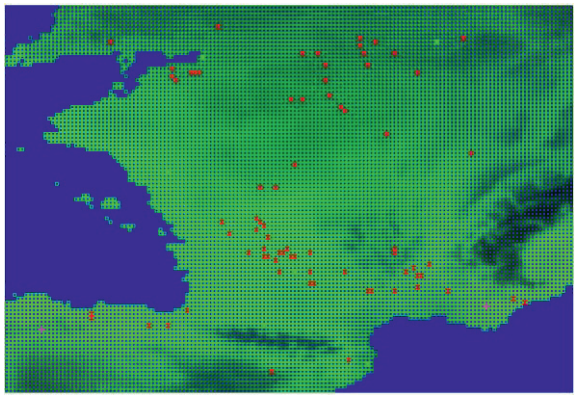


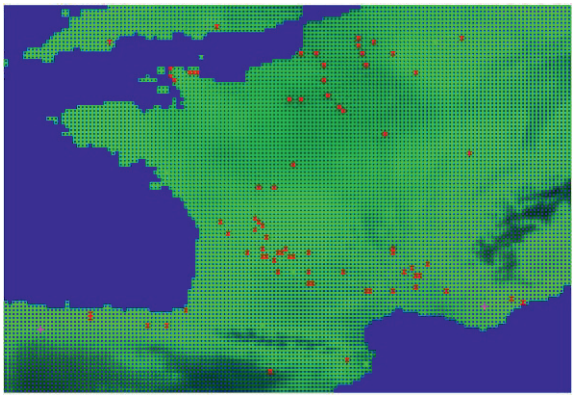

**Figure 81: Reconstructing habitat types for the Scladina area. Results from HomininSpace (left, with HabitatType 2 = booreal forest, and 3 = cool forest) and López-García *et al.* (2017).**

### 11.5.3 Energy via direct extrapolation

Energy in the landscape can also be recreated by using temperature and precipitation values to extrapolate from known environmental data. Figure 82 illustrates the continuous energy landscape reconstruction in HomininSpace (subsection 4.4.2), for selected moments in time. Per grid cell the resulting energy levels are calculated. In this figure darker colours indicate less energy in the landscape.



131,000 years ago (-5.6°)



124,800 years ago (2.19°)

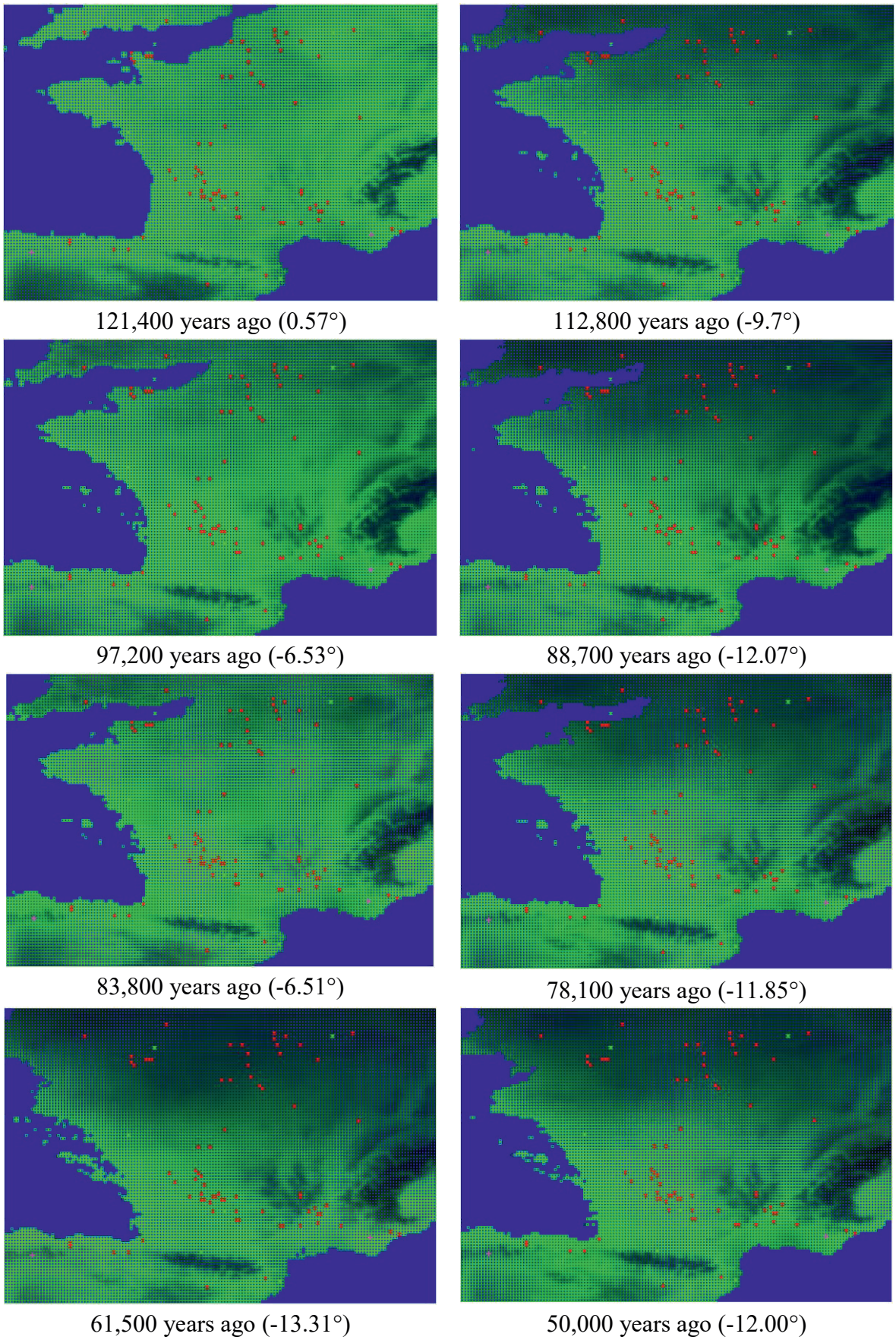



**Figure 82: Reconstruction of the energy levels in the landscape for the complete simulation period. Darker colors indicate less energy.**

To further quantify the energy levels in the landscape, Figure 83 presents the energy levels for all climate recording checkpoints. Note that recording checkpoint 6 (located in the Channel and colored red in the figure) is submerged during longer periods and has energy levels of zero during those time frames. For beach type grid cells the energy levels are effectively doubled which explains the non-continuities for that same checkpoint which after and before submergence becomes beach and then turns into regular land area when the water retreats further or becomes flooded when the sea level rises further.

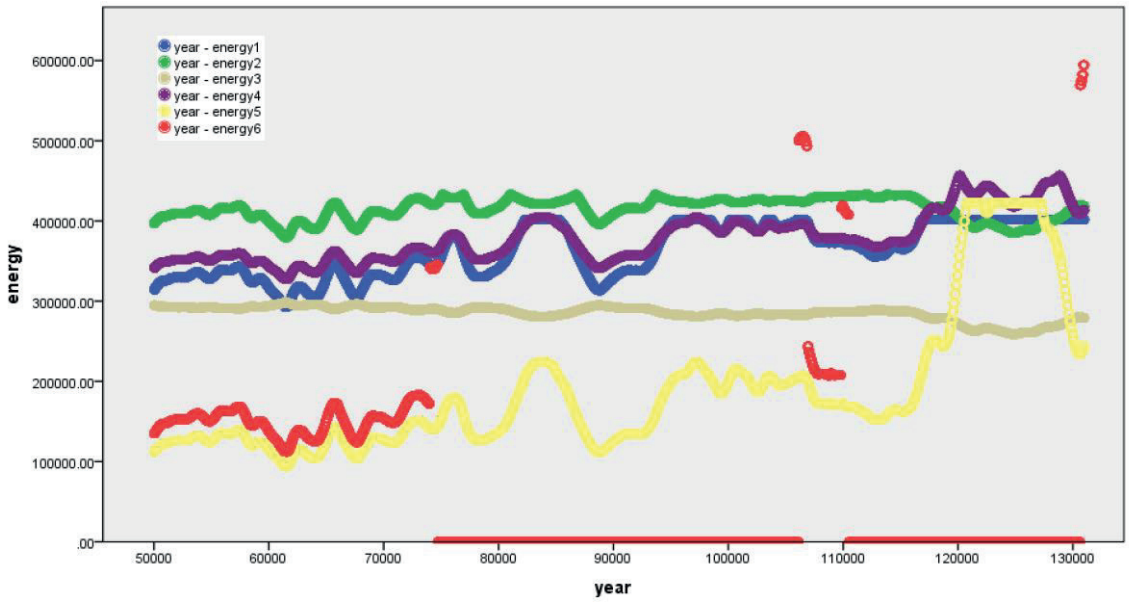



**Figure 83: Energy levels per year recorded by the Climate Recording Checkpoints for the whole simulation period. Blue line is for Checkpoint 1, Red for 6.**

Both checkpoints 5 (Scladina) and 6 (Channel) are located in the top (or northern) area of the map. This area suffers most from deteriorating of the environment due to colder climates. Both of them offer substantially less energy than the others that are further south. The very noticeable peak (can be inferred from the beach production in 6) in the beginning of the simulation is due to the significant improvement during MIS 5e. The effects are very pronounces for northern sites due to the large temperature differences between LGM and current day conditions. Furthermore it is quite interesting to find that the conditions in the Mediterranean area hardly change at all, but does not produce the most energy. Best conditions can be found in Occitanie (2) and near the Pyrenees (4), where the latter is most benign and outperforming 2 during MIS 5e conditions. Note that the production for Occitanie is also very constant through time.

These patterns are as can be expected. Energy levels are directly derived from reconstructed precipitation and temperature values (see Figure 84 and Figure 85). Interesting is the effect for the Pyrenees area (purple colored in the figures), where a

relatively high precipitation level results in an unexpected high energy production throughout the simulation period making this a very attractive place for hominin groups.

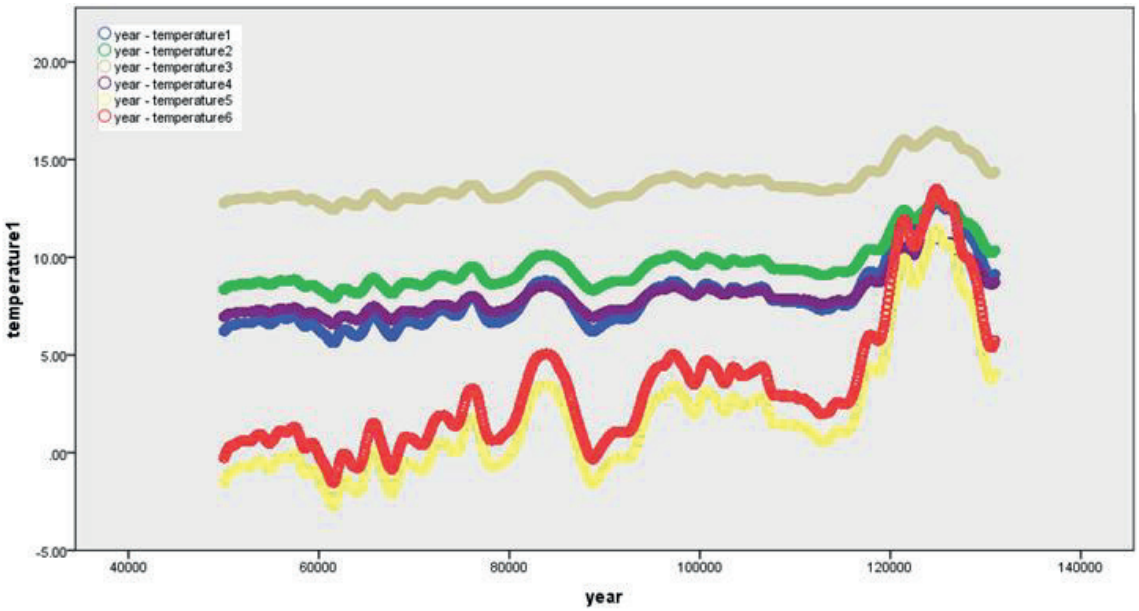



**Figure 84: Recorded values for reconstructed temperatures per Climate Recording Checkpoint.**

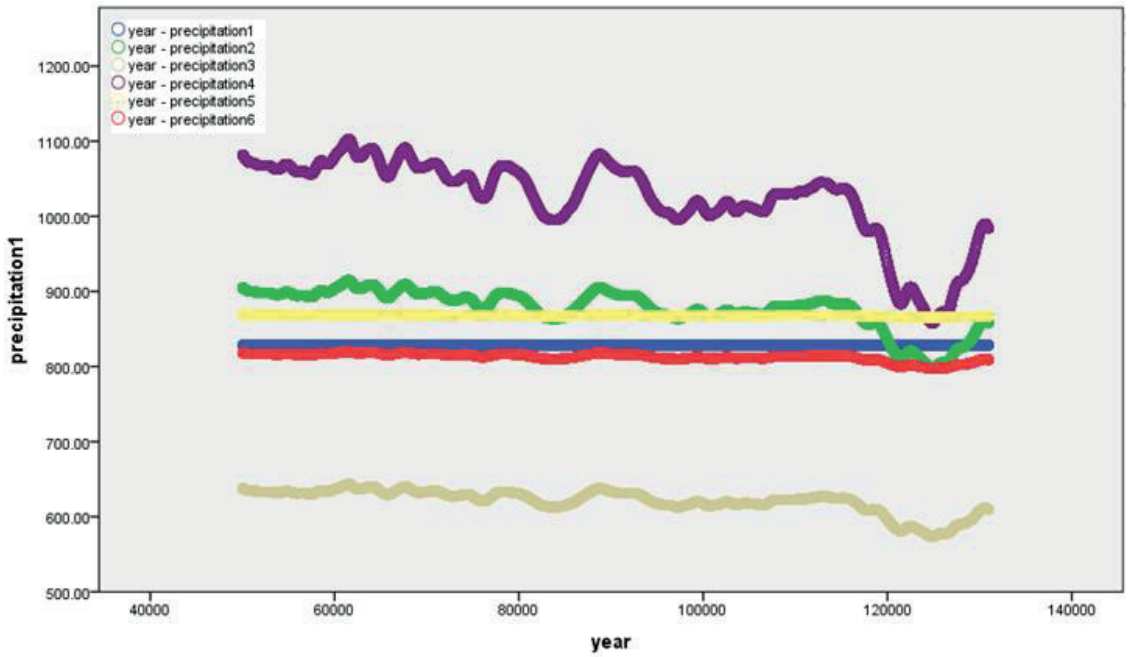



**Figure 85: Recorded values for reconstructed precipitation values per Climate Recording Checkpoint.**

Note that the modelled environment is the same in all simulations for all scenarios. The simulated hominins have no influence on the reconstructed sea level, temperature, precipitation, or primary production values. They can however consume some or all of the calculated secondary biomass, which influences availability of such biomass for the next year.

# 12. DISCUSSION

## 12.1 Introduction

Our interpretation of the past is always coloured: we find what we expect to see in the past because we know what to expect in the past (Kelly 1995). This makes agent-based modelling of people and behaviours in the past especially challenging since expectations might steer model design decisions and might bias implementation of the agents and their behaviour, and steer interpretation of modelling efforts. One of the most important results and the main research objective of the process of developing HomininSpace is therefore the insight gained into the requirements and structural elements for building an unbiased simulation model of past hominins (Section 12.2).

Modelling is meant to assign weight to different plausible parameterized scenarios of the past. It is particularly useful for comparing alternative explanations when trying to understand events and patterns in the past. But it is more than likely that Neanderthal behaviour operated differently than its simulated counterpart. And the resolution of the archaeological data is rather limited for the chosen Palaeolithic period of interest. Still this study assumes that, even though there are only relatively few data points, behaviours can be explored and inferences can be made about Neanderthals using data on their presence in the past. Scenarios were designed to answer specific questions illustrating the usability of the developed HomininSpace tool. The results for the simulations from the different scenarios were presented in Chapter 1 and analysed in Chapter 1. Section 12.3 discusses these outcomes per question. That section will also discuss general aspects of parameter evolution observed in all scenarios.

## 12.2 System development

In this study a simulation system has been developed that allows researchers to explore hypotheses and answer questions about hominins and their behaviour in reconstructed landscapes of the deep past. When programming a simulation system the main goals according to Axelrod (1997) are verification (that is, internal validity), usability and extendibility. In the development of HomininSpace much attention has been paid to verification (and validation) of the system and the underlying model. These include peer review of system and documentation, code review, replication and reproduction of (elements of) the system (Crooks *et al.* 2015).

To allow proper replication and reproduction of modelling results all underlying assumptions must be explicitly stated. Debate is only possible when these are unambiguously communicated. When a model is implemented and turned into a working simulation, many additional decisions are made. For instance, the design can state that groups move through the landscape, but only in the implementation phase the programmer must decide if a grid based environment is selected which limits the movement options. Such issues are hidden in the source code of an implemented model. Note that also the design methodology implemented in the Overview, Design concepts and Details protocol (ODD) does not guarantee that such details are communicated.

The peer review process has shown that to allow proper replication access to the source code is required. Analysis of the source code can provide insight in the *underlying* assumptions of the modeller (Will 2009). Note that these assumptions can also be unconsciously made, unintended of even unwanted, and code review and model replication are instruments to identify these. An example of such a (hidden) assumption in the HomininSpace system: when there is no reconstructed sea level available for any given year, the next (more recent) value is used (so no linear interpolation between data points).

The structure of the model underlying HomininSpace is grounded on ethnographic observations, but parameter values are not taken from the ethnographic record but derived by fitting simulation results against a selected material culture represented by presence data. Those models in which simulated presence compares well with archaeological presence most likely represent possible past behaviours that could have created the archaeological record. Variation between individual Neanderthals is explicitly not coded in the reproduction process, but is applied to the parameter values by the Genetic Algorithm at intra-simulation level. That means that there is evolution between simulations, but not within a simulation. This is a simplification of the real world inspired by the relative constant character of the Neanderthal archaeological record through time. It avoids biased parameter value assignment and allowing quantification in the correlation of results with parameter values.

Some model elements were easy to implement, others were surprisingly complex. Here I use the time to design, implement, test and describe a feature to indicate how hard it was to add some example functionalities:

- Two alternative ways to calculate secondary biomass: each one week;
- Coastal resources, including extra entry in the setting file: two days;

- Mobility type of dynamic versus static hominins (very few coding lines but structural change of the model): two months;
- Implementation of the foraging behaviour is one of the most important elements of the model. It took three months to implement and test all functionality. Including a maximum foraging range as a parameter, and using such ranges according to setting values then required minimal effort: estimated at two days;
- It is not trivial to determine if moving from one grid cell to the other requires crossing of larger water systems. After some recursive efforts an iterative implementation was chosen: total implementation time about one month;
- Implementation of the Death Penalty for absence violations is rather straight forward: the addition of a type for intervals, the inclusion of absent interval counts in the output, and the filtering in the genetic algorithm. Implementation and testing of this functionality took about four days;
- To enable random moving hominin groups: one hour.

### 12.2.1 Characterizing Neanderthal presence using radiometrically dated archaeology

In HomininSpace radiometric dates of presence of Neanderthals in western Europe are instrumental in determining a fitness value of simulation results. Dates are used in an absolute sense with one standard deviation which means that the actual presence in the past has occurred with a probability of 66% within a time frame defined by one standard deviation plus and minus the radiometrically calculated date. That also means that one third of the measured data points are wrong, with archaeology producing hominin presence actually occurring outside the given interval.

Only presences attested by radiometric dates are included in the database. Positive relatively or culturally dated sites would have provided many more checkpoints if included in the database. Especially for areas that now contain few data points (see for instance Discamps *et al.* (2011) with more data on north of France). It is acknowledged that a database as used in HomininSpace can contain only a selection of sites, possibly limiting the resolution of the inferences. It is also acknowledged that the selection procedure to obtain the dates is heavily biased, for many reasons. Figure 87 visually illustrates the bias in the interval data. Especially the northern part of the simulation area has little presence data.

The dates in HomininSpace are used to validate presence events, not demographic population change for which these would be unfit (Attenbrow and Hiscock 2015). An omni-present hominin species would provide the best match with the archaeology. Modelled hominins that can quickly move through the landscape, can spread thin but hold on in adverse circumstances, can recover from setbacks or simply flood the area are good performing individuals. When absence is introduced this forces the modelled hominins into

a different direction. Fast reproducing but very short living species (that do not reach the absence areas) are the result. The actual location and amount of resources is less important since there are many ways to overcome local shortages (move, spread, avoid).

One of the elements in the database is the confidence level that is attached to individual data points. The effect of this confidence level is included in the output of each simulation, in the form of the `matchingVisitsWithConfidence` variable. The correlation between `matchingVisits`, `matchingVisitsWithConfidence`, and also `totalCSTVisits` are consistently nearly perfect for all scenarios (subsection 11.4.4). An example is shown in Figure 86. Note that in the lower left section of the depicted matrix a plot with a straight line from bottom left to top right indicates a perfect correlation between the variables in the corresponding row and column. Such a correlation between the mentioned variables indicates that including the confidence level in the score has no noticeable influence on the observed patterns.

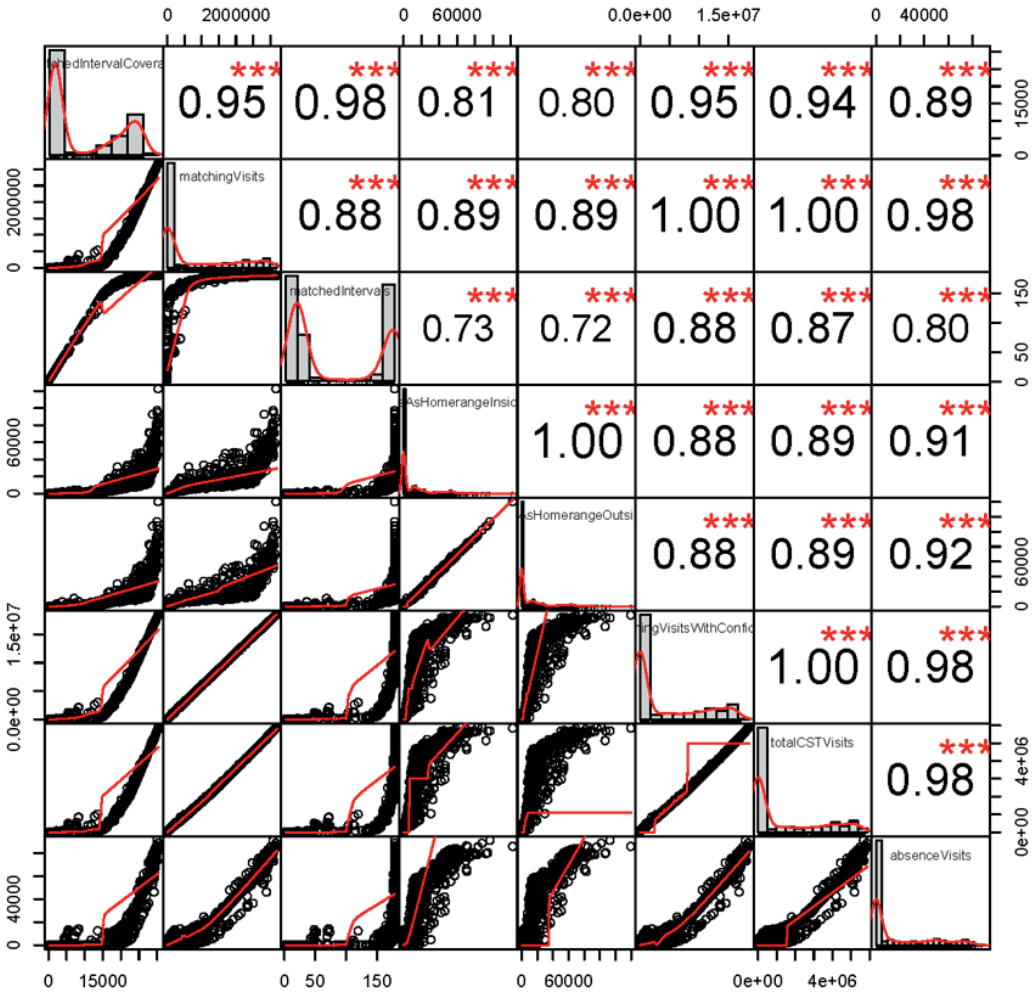



**Figure 86: Correlations between all visit counting output variables for scenario Habitat-A. For instance, the correlation between matchedIntervalCoverage (simulation score) and matchedIntervals (count of all matched intervals) is 0.98, and very significant (three red stars, as have all correlations).**

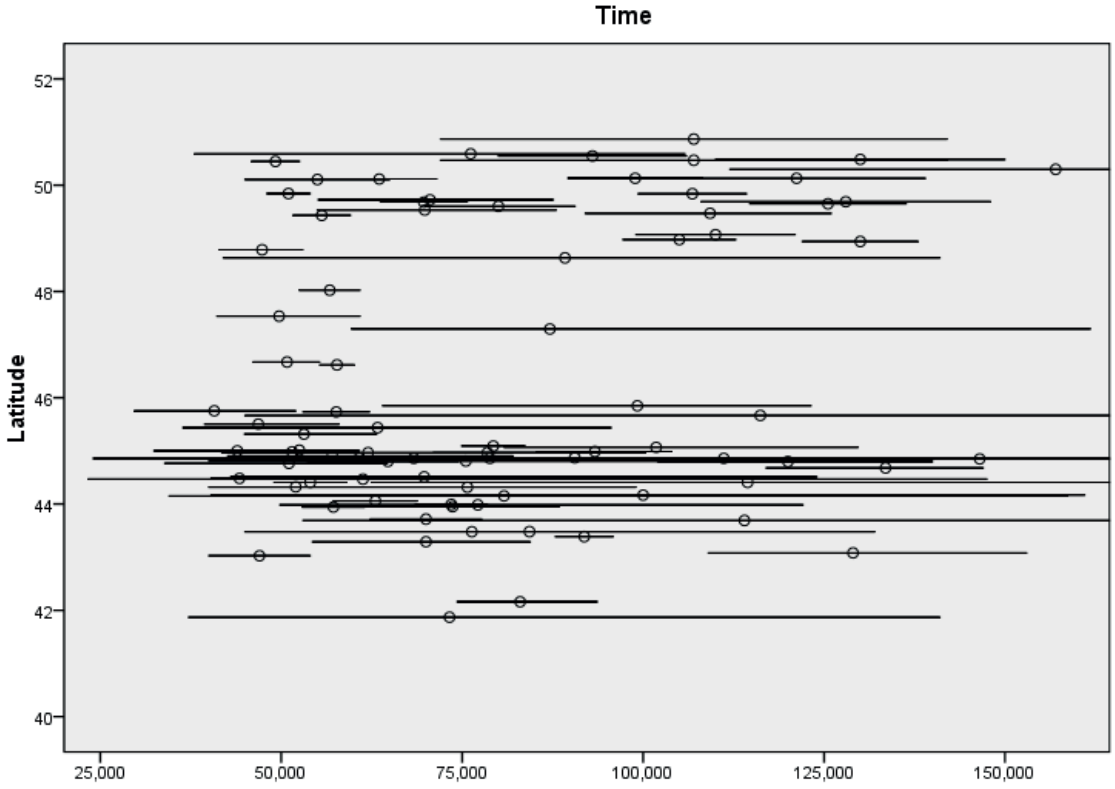



**Figure 87: Illustrating the bias in the checkpoint dataset: overview of the used intervals in the simulations. X-axis gives Time, with each interval defined by starting point and end point in years ago. Vertical axis is the latitude of the site.**

Presence data is a good medium to allow characterization of hominin models. Although different types of hominins fit the data equally well this provides alternative explanations and illustrates the possibilities of the flexible hominin model. The resolution and distribution of the available data is very limited, but for the geographic and temporal scale of the underlying research this is sufficient. A more temporal restricted (say around 50 ka) or geographically reduced area (say between 44 and 46 degrees) would provide a much higher modelling resolution. But most model variants with a positive viability index are omni present in smaller scale areas and time frames which would leave less discriminating power in the model. The large empty topographical and temporal spaces are actually beneficial, even if not used as absence criteria themselves.

### 12.2.2 Genetic Algorithms – do we need them?

Is it useful to apply a genetic algorithm in modelling research? Does it find Neanderthals that would not have been found using statistics alone? A common modelling approach would be to run simulations with ‘sensible’ parameter values and do a sensitivity analysis in order to: (1) find the best values for these parameters, and (2) reduce the parameter set of the model. Analysis of the results from executing the Standard sets in the scenarios in

this study (subsection 11.3.2) identified the importance of the demographic variables, especially birth rate and death rates for the pre-fertile and fertile sections. But after using the GA for generating at least 500 new individuals exploring the parameter space and improving all scores, different additional correlations can be identified over all scenario results.

Standing out are new positive correlations of the results with the death rate for the post-fertile section and the cohort size of the fertile section, and also the overall positive correlation of the calories per kilo meat (which showed a negative tendency in the Standard set). Strong negative correlations became visible for cohort size pre-fertile and especially with group size before split. When taking the statistical suggestions from the sensitivity analysis of the Standard set one could easily dismiss the variables for the death rate for the post-fertile section, and for the cohort sizes for fertile and pre-fertile sections. These are variables that appear to have become important in some Evolved sets. Also the slightly negative correlation of the calories per kilo meat and slightly positive correlations of group size before split and subsistence of the pre-fertile cohort were (sometimes strongly) reversed. With this improved knowledge parameter value selection would be different.

The choice of the stop criterion is an important element in the implementation of GAs. Since by definition it cannot be known if the absolute maximum score has been obtained it is necessary to define a stop criterion. However, any subsequent simulation might still improve the simulation results. The stop criterion in HomininSpace is the absence of any improvement within the last 100 generated models. Improvement was obtained however in some of those batches that did execute more simulations, sometimes after more than 300 new generations (see Habitat-B, or especially the extended results for Energy-D).

In the genetic algorithm in HomininSpace, new generations and their parents stay included in the population from which new parents are selected. A consequence of this is that some families tend to dominate the results, occasionally with the top 10 being all close member of the same lineage (with only variation in non-relevant parameters). Since a stop criterion has to be implemented, and one of the results of this study is the identification of the multiple ways the parameters values can be used to implement successful hominins, the stop criterion of 100 new generations without improvement was maintained. And thus analysis focuses on those parameter values obtained before the stop criterion is reached. It is acknowledged that defining a good stop criterion is important in the design of a study that uses GA methodologies.

In the GA methodology mutation and combination techniques are used to create new individuals. To illustrate that both techniques work well and both contribute to the end result Figure 88 colours the evolved models. Those models that result from a mutation in one of the parameters are coloured in green, and where two parameter sets are combined in red. Although combination (red) creates most of the worst scoring individuals and also the best one, green dots are scattered throughout the simulation score space, often contributing via local maxima. This is as expected since mutation will only slightly modify an successful model expecting only minor changes in the new simulation score, where combination can break a successful formula or create new successful lineages on a different basis.

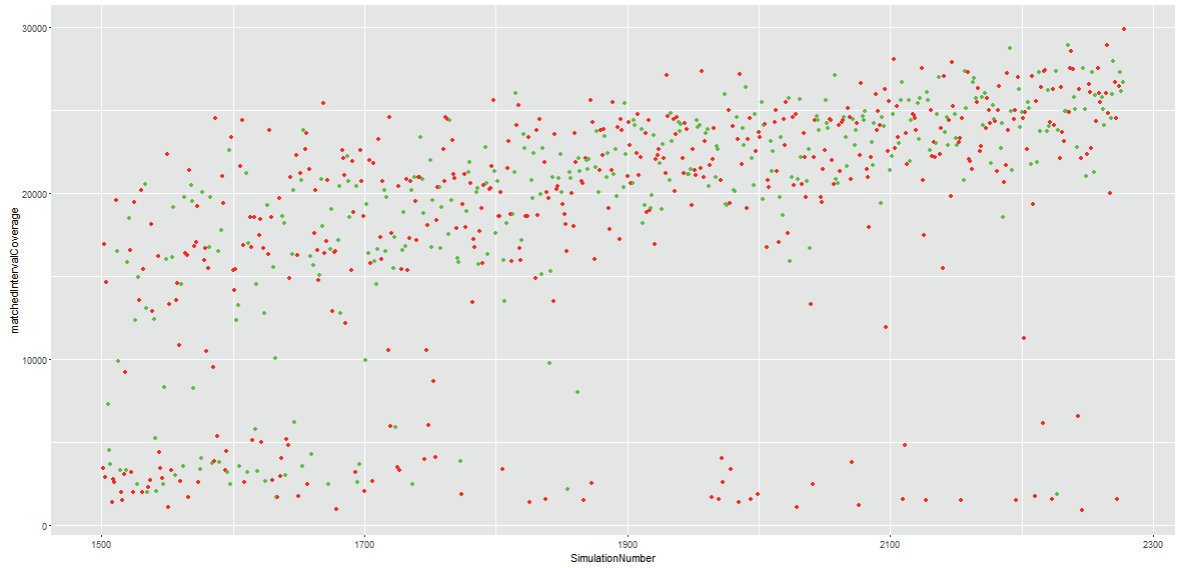



**Figure 88: Mapping models resulting from mutation (in green) and combination (in red). Horizontal axis the simulation number (from 1500 onwards), the Y-axis presents the simulation score. Data from the Evolved set in scenario Energy-B.**

The usage of GA in HomininSpace allows the system to explore the parameter space with better results than would be possible by manual selection of parameter values (biased) or exhaustive search (too much time) of value spaces. It also performs better compared to using the statistical suggestions for parameter value selection and model construction. As such, better hominins were created, i.e. those that match the archaeology more. The flexibility of the GA enables the system to find good performing individuals, sometimes even by modification outside given value boundaries or effectively changing the function of a parameter. These aspects of the GA are very important when deciding what to include or exclude in a model, and how to explore the parameter value space.

### 12.2.3 Realistic environment reconstruction

One of the foundations of the simulations in this study is the environment reconstruction. This creates an energy distribution through space and time, using reconstructed values for temperature and precipitation for each time step during the simulation period. What influences the realism of the reconstruction? Two different methods were implemented to compute the available resources for the hominin agents, both of them using interpolated values from Worldclim.org. This data is known to result in overpredictions of suitable habitats (Maria and Udo 2017).

Climate reconstruction in general is based on a sequence of proxy data for temperature and precipitation with absolute dates associated with periods of certain climates. It can and has been shown that dates for reconstructed global climate events can deviate between locations by 5,000 years or more (Sier *et al.* 2011). For instance, the start of the Eemian (MIS 5e) is about 6000 years later in the north of Europe when compared to the 'same' event in the south. This is mainly due to the fact that the Eemian is defined as a certain pollen sequence in the record showing amelioration of the climate with specific vegetation patterns. These patterns occur much later in northern Europe due to the fact that vegetation reacts relatively slowly to improvements of the climate. Thus global variations in temperature and precipitation will have different effects on local environments. HomininSpace does not accommodate such delays and implements changes immediately. The modelled resource availability will therefore, especially for vegetation reconstructions, be temporally off and less useful to validate simulation results against radiometrically obtained absolute dates.

Flora reacts quickly to climate deterioration. At Quaternary time scales ($10^1$-$10^5$ years) vegetation dynamics are mainly forced by climate change (Huntley and Webb III 1989; Webb III 1986). When the climate deteriorates sufficiently, arboreal taxa will swiftly retreat to refugia often in the south. When the climate is not very hostile, local refugia like those located north of the Alps can preserve healthy populations. More extreme climates or longer durations of cold spells will drive trees further to the south. Most tree species take a few hundred years to recolonize former distribution areas after minor cold periods while thousands of years can be needed for recovering from the more severe glacial periods (Müller *et al.* 2003). Although the composition of biomes in the past differed significantly from those of today, characteristic taxa for certain biomes survived as elements within other biomes. For instance, temperate deciduous trees that persisted throughout the LGM remained scattered through cool mixed forests in Eastern Europe (Prentice *et al.* 2000).

These other biomes could function as source for typical taxa in fluctuating climates, making a direct link between refugia and repopulation areas less essential. Within HomininSpace, reconstructed biome types will appear anywhere where climate conditions would allow them, ignoring any topographically imposed challenges.

The comparison of the modelled habitat reconstruction in the Scladina area with the reconstruction from López-García *et al.* (2017) illustrates the limitations of the reconstruction method in HomininSpace, but also of the limited information that is available. The reconstruction in HomininSpace uses two (extreme) data points to reconstruct temperatures and precipitation levels. For Scladina (Table 55) these are temperatures of -5.1 and +9.4 degrees, and 870 and 865 mm for the associated precipitation levels. López-García *et al.* (2017) use six data points and find a more restricted temperature range between -2 and +6, and a much wider precipitation range of 700 up to 1200mm. They calculate the minima and maxima for both climate parameters for the middle of MIS 3 (with no later data available) and MIS 5c (not 5e, as in HomininSpace, probably since earlier data points are lacking).

López-García *et al.* (2017) use the available proxy data (small mammal assemblages) to create a local approximation of the $^{18}O$ isotope curve for the site Caverne Marie-Jeanne. Such a curve provides a much better locally tailored interpolation mechanism between the given climate data points, most likely resulting in a high resolution reconstruction of past environments. HomininSpace uses only two data points, and can vary the precipitation for this area only between the values of 865 and 870. More climate points allow a more detailed local reconstruction. For the Scladina area thus a highly detailed picture can be constructed for the given period. The question remains however how local and thus how useful this picture is for a realistic reconstruction encompassing a wider and more varied area. Given the available data the chosen reconstruction methods in HomininSpace provide the most realistic reconstruction possible for the given period of time.

### 12.2.4 Probability in HomininSpace

One of the fundamental issues with a stochastic process is the uncertainty about the quality of the result. Only if the stochastic nature of the parameters in the process is sufficiently known, statements about the results can include quantification of any uncertainty. In HomininSpace the effect of random events has been reduced considerably for individual simulations. Specifically, probability is not implemented as a decisive factor in mortality to avoid major impact from random events (cf. Boyd *et al.* (2011)). Randomness is present in

way finding when selecting between two destinations that have exactly the same amount of resources. The occurrences and thus the effect of these specific random events are minimal.

Furthermore, resource extraction is implemented in such a manner that the random order of group activation cannot directly cause extinction (by preventing that one group consumes all resources before another group can take its turn). The implemented extraction can however influence the simulation score to some extent since a checkpoint that is located within a foraging range is not necessarily visited each time step. The resource contents of grid cells in the foraging range are consumed one after the other until resource requirements are met. Remaining grid cells are not visited this time step, and therefore checkpoints if they happen to be skipped in resource extraction are not visited. Such time steps do not count towards the simulation score of that interval. This is why no interval is visited for 100%, and explains mostly why two simulations with the same parameter set but different seeds for the random number generator result in (slightly) different scores.

The effect of random resource extraction is partly compensated for by the score calculations where visiting an interval already accounts for half the possible points for that interval. Visits further depend on foraging range, resource needs (group size), available resources (climate, previous consumption), and the presence of other hominin groups. Since the overall effects of randomness are limited it is not needed to obtain a so-called probability distribution for each unique parameter value combination, which would require multiple reruns for each combination (Crema 2018). These limited effects allow for a highly deterministic simulation in which the effects of parameter value choice can be calculated with great certainty. This assumed determinism does not reflect the real world where random events can have a major impact.

However, probability does play a major role in the genetic algorithm where it is considered a fundamental property of the method itself and essential for its functioning, responsible for allowing the algorithm to escape attraction of local optima. It is difficult to define stop criteria assuring near-maximum results. Especially in the chosen implementation where offspring does not replace parents. When new individuals are added to the population this increases the risk of single family dominance of the tournaments, but allows variation of offspring. The stop criterion was experimentally defined and presents a compromise between the number of simulations and the increase in score. This is illustrated by the results for scenario Energy-D where far beyond the stop criterion a very good scoring

individual is found. However, adjusting the stop criterion to include this individual would fail to stop most other scenarios requiring many more simulations for which resources are lacking and would increase the danger of single family dominance that would obscure other potentially successful individuals.

It is almost impossible to exclude probability effects in realistic models. When introducing genetic algorithms into a system randomness is included as well. Therefore running many simulation remain essential, if not in order to quantify uncertainty then to explore the variability of vast parameter spaces.

## 12.3 Answering questions with HominінSpace

The scenarios use settings that activate or deactivate model elements aimed to answer the following questions:

1. How does Energy reconstruction of the environment compare to Habitat reconstruction when matching the archaeology with modelled hominins?
2. What is the effect of adding coastal resources to the model?
3. How does the addition of ebb and flow dispersal with the available resources influence the match with the archaeology?
4. What is an optimal maximum foraging range or would an unlimited range be better when matching archaeology?
5. What happens if we assume that Neanderthals were able to cross large open water systems?
6. Absence data indicates locations where Neanderthals were NOT present for some period of time. What influence has adding such data when comparing simulation results with only archaeology attested presence?
7. How can population core areas that produce new Neanderthal groups under certain conditions be implemented?

And further three neutral model questions. The first two were only used in model development and not included in the experiments. Only the effect of number setting 10 that implements random movement of groups is further explored due to its relevance in the literature:

8. How does the reconstructed topography influence the dispersal characteristics of Neanderthal groups?
9. How does the reconstructed energy distribution in the simulation area influence the dispersal characteristics of Neanderthal groups?
10. What are the results if the Neanderthals are implemented with random movement instead of going for the maximum amount of available resources?

These questions and combinations thereof were explored in the scenarios that contain the simulations that were executed in this research. The next subsections discuss the results of

those explorations per question, but first I will focus on the evolution of parameter values. When referring to simulations results in this discussion, the maximum obtained score (best match with the archaeology) for a scenario is given between brackets (see Table 25 for a summary of all results).

### 12.3.1 Parameter value evolution in all scenarios

In most simulations the actual number of Neanderthals that are in the area at a given time is low. Being thin on the ground is made possible by several parameters that include large foraging ranges, dynamic moving groups, and large movement distances. The analysis by Castellano *et al.* (2014) of Neandertal DNA show that Neandertal populations might indeed have been small and that they lived separated from each other.

The `ViabilityIndex` is a value associated with demographic aspects of the hominin model, and represents the potential growth factor of the population. Even though the `ViabilityIndex` is not a real parameter in the sense that it can be manipulated directly, it is included in the correlation statistics to assess the importance of the demographic submodel. The correlation of this submodel with the `matchedIntervalCoverage` simulation score is in all scenarios higher than for all other model parameters (see Table 46 and ), with notable exceptions for those scenarios where absence data plays a role. This suggests that population growth potential is a very strong indicator for a successful model. There are models however with a very high `ViabilityIndex` that still yield low simulation scores, and models with a low growth factor that score very well.

Quantifying the parameters that are used to model hominins in the past is one of the key elements of HominínSpace. Strong correlation between simulation results and input parameter values is suggestive of a (linear) relationship, and the relative variable importance statistic can indicate which parameters are most important in this relationship. The genetic algorithm however does not use such information, but selects and mutates parameter values from already successful value sets attempting to obtain even better performing values. This is a fundamentally different approach implemented in this research and compared against possible statistical inferences.

Generalizing over all simulations, the sizes of the fertile cohorts are rather large and the models feature short childhoods with small pre-fertile cohort sizes. The death rate for the post-fertile cohort shows a mostly positive correlation with the scenario number. There are some positive correlations for the subsistence for the post-fertile cohort with increasing

scenario numbers and also the values in the top 10 for the subsistence needs of the post-fertile cohort compare generally higher to the pre-fertile and fertile cohorts. This suggests that the post-fertile cohort in the underlying model is less important than the other cohorts. This contrasts the so-called grand-mother hypothesis that attaches additional benefits for younger generations to the presence of post-fertile individuals. And contra-intuitive energy related variables appear also non-significant for scoring against the archaeology.

When inspecting the non-significant parameters for the same set of solutions, it becomes clear that values for these vary more than for the significant parameters (Scherjon 2016). This makes sense since they influence the final result less and are therefore less likely to be subjected to evolution. The maximum possible score in the simulations that were executed in this research (with the given configuration of Neanderthal sites and radiometrically determined intervals) is a total value of 39,200. The highest score obtained from all simulations is 31,142 in the Habitat-I scenario. This result is closely followed however by scores in Habitat-H, Habitat-J, Energy-E and other scenarios. From the simulation results it appears that demographic variables (high birth rate, low death rates) are important since these facilitate quick recovery when climate conditions improve.

The Neanderthal model as implemented is a very flexible model, able to adapt to most circumstances by varying parameter value combinations. As an example I discuss the top-3 of scenario Energy-E (EnergyStaticCoastalMaxrange) Standard set: 1384, 1365 and 1131. Each implements a different strategy to get good results. 1384 has a low ViabilityIndex due to very high pre-fertile and post-fertile mortality rates, but a very small fertile cohort size combined with reduced energy needs for the fertile cohort, and a very good tolerance for cold. 1365 features the opposite, a very high ViabilityIndex due to high birth and low pre-fertile death rates. The fertile cohort is very large but requires much energy, and the groups are not very tolerant for cold. 1131 then has a relatively low birth rate but large fertile cohort size, high cold tolerance, large foraging range, and high energy yield from meat. Each of these models yield similar simulation results and do well in most scenarios, illustrating the flexibility of the model in implementing different types of models and the effectiveness of the Genetic Algorithm in identifying and promoting promising models.

Finally quite interesting is the comparison of the results in the duplication effort of Energy-A with the first scenario run (subsection 10.17.2). Simulation scores were not very different (the first attempt scored 30522 in 2089 simulations, the second reached 29136 in 1998 simulations) but the evolved models were not exactly the same. Most obvious

differences evolved in the duplication are a higher death rate for the pre-fertile cohort, much higher subsistence needs for the fertile cohort, and a temperature tolerance that allowed less colder circumstances (-19 versus -29). This illustrates that more solutions exist within the parameter space that can result in high simulation scores, with a process of selection and mutation resembling genetic drift creating specific individuals. The result matches anthropological observations that different hunter-gatherer cultures behave differently in similar landscapes, and that even groups from one culture can display behavioural differences in the same landscape (Vandermeer 2006; Grøn 2018; Groß *et al.* 2018). Note that quite a few of the parameters and settings model aspects that are at least culturally affected if not controlled.

### 12.3.2 Q1 - How does the Energy reconstruction of the environment compare to Habitat reconstruction?

Habitat reconstruction will create distinct areas with uniform energy levels where Energy reconstruction will create an energy distribution along a continuous scale. It is expected that creating areas with a much higher energy level than surrounding areas will result in containment of hominin groups, where areas with much lower values will hinder movement (compare desert environments). If the preference of Neanderthal hominins for a certain habitat is reflected in the used archaeological data there should be marked differences in the simulation results for both approaches.

There are two set of scenarios that can be compared to address this question: Habitat-A (30436) versus Energy-A (30522), and Habitat-J (30954) versus Energy-E (30651). Pairwise these scores are virtually the same, and both the Spearman correlation coefficients as well as the relative variable importance for all parameters are very similar (Table 46, Table 48). When comparing the Evolved individuals, or model parameter sets, for Habitat-A and Energy-A there are marked differences (Table 47): birthrates for Habitat-A are higher, years before group maturity are very low for Energy-A, and Temperature tolerance is much lower in Energy-A than in Habitat-A. Interestingly, for Habitat-A and Energy-A the top three best scoring individuals in the Standard set are exactly the same ones, suggesting a similar scoring scheme.

When comparing Habitat-J with Energy-E, there are also some noticeable differences: for Habitat-J the birth rates are relatively low and death rates for the pre-fertile cohort very high, suggesting that it is possible to score well in the right habitats with such limitations. Note that GroupSizeFertile_BeforeMerge is zero in Energy-E, simulation number two.

This means that this condition is never satisfied, effectively disabling this parameter. This individual while in the top three has also a very low value for the viability index and a very high value for the CohortSize_PreFertile.

The manner in which the energy in a landscape is distributed clearly influences the different Neanderthals that are produced by the system when attempting to evolve better matches with the archaeology. However, the spread through the landscape of these different hominin types does not produce significantly different scores. From these results it is impossible to say which energy landscape reconstruction gives better matches with the real world since both perform equally well. It furthermore appears that Neanderthals do not show preference for certain habitats, at least when using energy distribution alone to create distinct areas.

### 12.3.3 Q2 - What is the effect of adding coastal resources to the model?

Coastal resources were used in the following scenarios: Habitat-E, Habitat-F, Habitat-J, Energy-B through Energy-E, and Energy-BR, Energy-CR and Energy-ER. Results from Habitat-J can directly be compared to Habitat-I. Simulation maxima are virtually the same (Table 25), and correlations for the Standard set are very similar. Correlations for the Evolved sets deviate, and quite interestingly, as one of few scenarios Habitat-J has no correlation at all of birth rate versus simulation results. When comparing the top-3 of evolved individuals those from Habitat-J indeed have relatively low birth rates (35), high mortality rates for pre-fertile and post-fertile cohorts, and small group sizes before split. This all results in a relatively low viability index, but that does not prevent a high simulation score. In other scenarios similar individuals appear, for instance Evolved 2 from scenario Energy-E (very low birth rate, high mortality post-fertile, and a viability index of 335). Such high scores with very low viability indices are only present in the top-3 of scenarios with coastal resources.

Coastal resources will keep hominins close to the coast (see Figure 89, or even more pronounced for the dynamic hominins from scenario-B in Figure 90), where there is consequently more competition. But there are few coastal Neanderthal sites. Submerged coastlines have maybe hidden any direct evidence, but there is indirect evidence in the form of transported shells, the use of sea animals in some inland sites, etc. Since we have few checkpoints along the coast, Neanderthals once living there do not significantly contribute to the simulation score. Coastal resources can function as a focal area for hominin group dispersal, since coastal resources generally give more kcal than their

surroundings. However, when sea levels drops more land is exposed. Here are no checkpoints and when the coastal region is wide enough (see Figure 24) those checkpoints that are now located more inland are no longer reached by foraging efforts, since resources are obtained mostly from the coastal areas. This limits the potential advantages when success is measured using the simulation score.

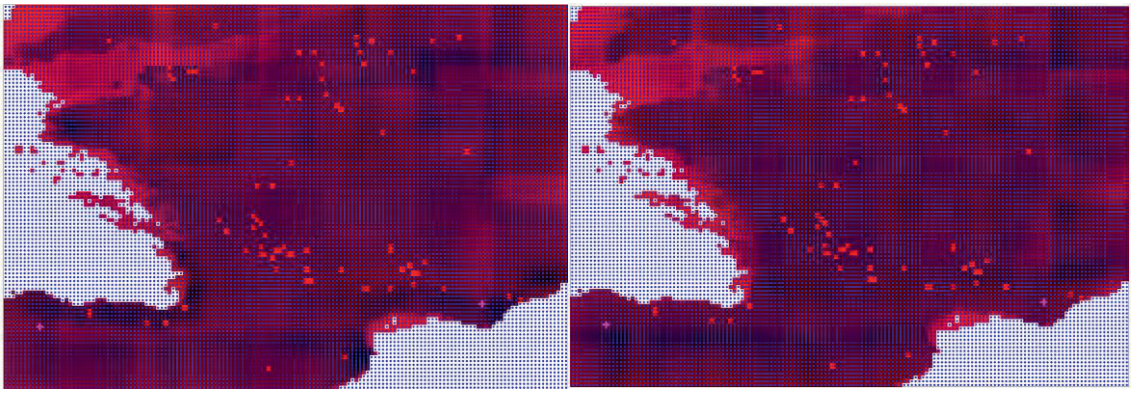


**Figure 89: Illustrating the effects of using coastal resources. Left simulation 1384 with coastal resources turned on (scenario Habitat-J). On the right the same simulation with coastal resources disabled (Habitat-I).**

The score for the simulations depicted in Figure 89 are respectively 29,268 with coastal resources, and 29,869 without. This is a simulation with a birth percentage of 41 and very high mortality rates for pre-fertile (11) and post-fertile (14) cohorts. Checkpoints that score better with coastal resources are Abri des Pecheurs, Covalejos Cave, El Castillo, Fermanville-La Mondree, Grossoeuvre, Grotte du Figuier, Grotte du Lazaret, Havrincourt 1, Jupiter, La Rochette, Le Moustier, Le Prisse, Ormesson, Payre, Saint-Amand-les-Eaux, and Walou Cave. In the figure it can be observed that coastal areas are visited more often, especially near the Mediterranean and the Iberian Atlantic coastlines (darker colors). However, even though the map without coastal resources is overall darker coloured, for most sites the scores for both simulations are rather similar. This means that in both simulations during most intervals hominins were present to create the archaeology, even though percentages of these intervals indicate more presence events during scenario Habitat-I.

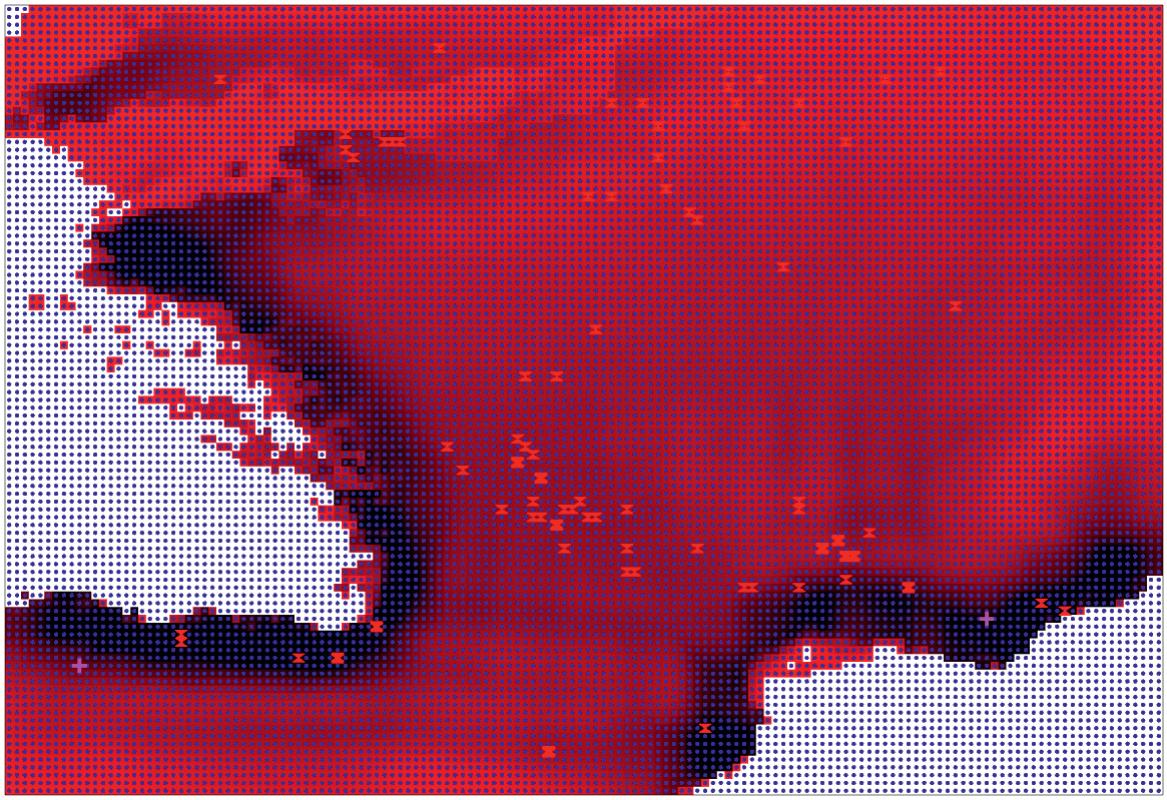


**Figure 90: The attractiveness of coastal resources. Simulation 1365 from Scenario Energy-B but with dynamic hominins.**

### 12.3.4 Q3 - How does the addition of ebb and flow dispersal with the available resources influences the match with the archaeology?

Research suggests that anatomically modern humans at least in some cases follow a strategy where populations expand and contract using refugia (Williams *et al.* 2013). For example, the general human population density in Australia declined during the LGM, with some areas abandoned and not re-occupied until much later. Williams *et al.* (2013) suggest that specific areas, especially those watered by melt water, acted as (cryptic) refugia. These populations are described as highly mobile and relocations have been identified suggesting that groups of people move entirely into more attractive and preferred habitats. Such behaviour is implemented in HomininSpace and referred to as dynamic mobility, with groups continuously moving into areas that contain the most resources. This is opposed to static mobility, where the modelled groups after selecting an area for foraging cannot move to another area (not even when the environment deteriorates and not enough resources remain).

Static dispersal is used in scenarios Habitat-G through Habitat-J and in Energy-E and Energy-ER. Thus all other scenarios implement a dynamic strategy. The results in Habitat-G (30121) can directly be compared against Habitat-A (30436), Habitat-I (31142) against

Habitat-C (30206), and Energy-E (30651) can be compared against the results from Energy-B (29896). Energy-ER (30439) accordingly compares against Energy-BR (28125). In all these comparisons the only difference is the mobility character, static versus dynamic. There is no consistent difference, with Habitat-I having a higher score than Habitat-C, but the others scoring less. Static hominins do not score significantly higher (nor lower) than dynamic hominins what suggests that the archaeological record cannot be used to infer preference for a specific mobility strategy from those implemented in HomininSpace.

### 12.3.5 Q4 - What is an optimal maximum foraging range?

A maximum foraging range is used in scenarios Habitat-C through Habitat-F, Habitat-I and Habitat-J, Energy-B through Energy-E, Energy-BR, Energy-CR, and Energy-ER. In order to qualify the effect of a maximum foraging range, it is possible to compare Habitat-C (30206) with Habitat-A (30436) (for dynamic hominins), and Habitat-I (31142) with Habitat-G (30121) (implementing static hominins). These scenarios all have similar maximum scores, as have Energy-B and Energy-E, as well as Energy-ER, and although a bit lower even Energy-BR. Note that scenarios without a maximum foraging range have no limit imposed on the actual foraging ranges, which means these can theoretically encompass the whole simulation area. Note that the unit for the foraging range is number of grid cells.

In those scenarios where the absence condition is imposed the maximum foraging range is not or negatively correlated with the simulation score. In these contexts a smaller foraging range is better since the ultimate aim here is actually to make it difficult for hominin groups to survive (subsection 12.3.7). And indeed, in all other scenarios where the foraging range is used it shows a strong positive correlation with the simulation score. That means that a higher maximum foraging range generally results in a higher simulation score. But during a simulation the actual foraging ranges that groups exploit are not necessarily equal to the maximum range. Circumstances that include group size, resources, and competition often permit much smaller ranges (see Figure 91).

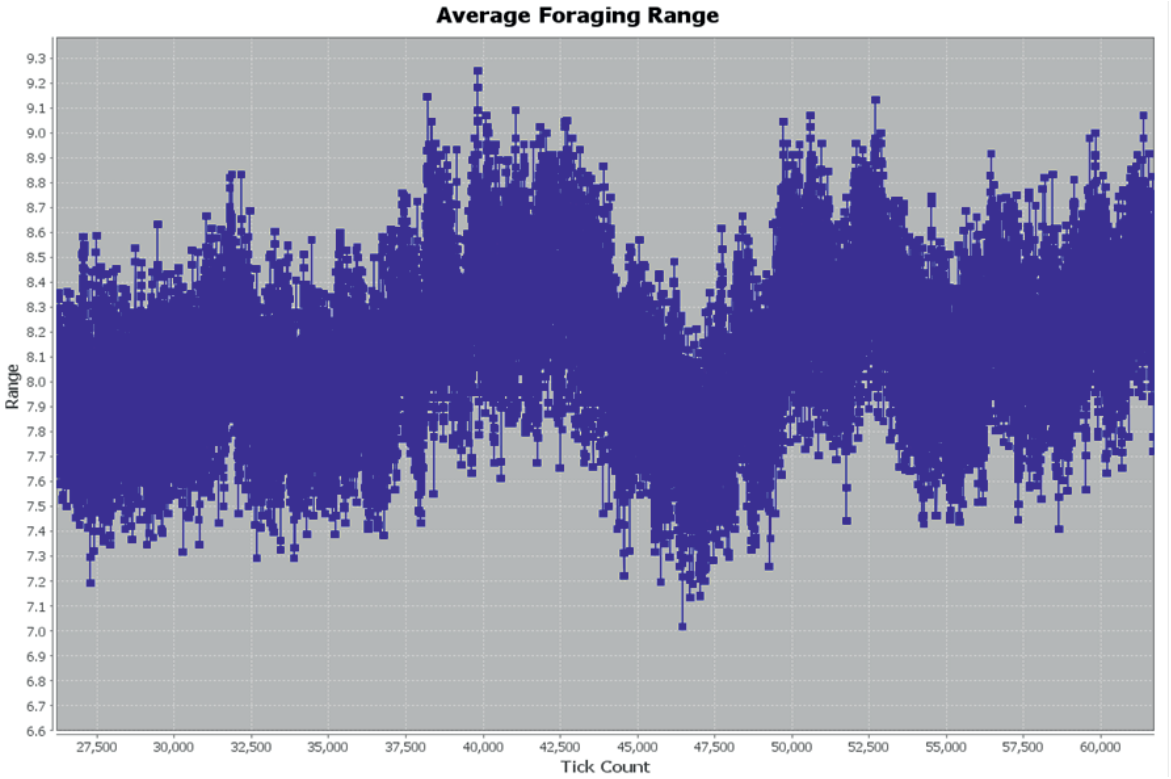



**Figure 91: Overview of the average foraging range for part of the simulation period. Largest value around 9.3, minimum value around 7.0, standard deviation around 2. Figure created in HominінSpace, by calculating the average foraging range of all present groups for each time step. Simulation 1384, scenario Habitat-I, the maximum value for the foraging range is 11.**

The best maximum foraging range depends on the settings and other parameter values and varies per scenario. That is clearly illustrated by the fact that for each scenario with the maximum foraging range activated the top-10 scores vary in the resulting values for this variable. Or, a large value does not guarantee a high score. In some scenarios (Habitat-C, Habitat-E, Habitat-I) values hover around 15 for a maximum score, in others (Energy-D and especially Energy-B) values can be much lower for good simulation scores. Figure 92 illustrates as an example that with a maximum foraging range of 10 grid cells (100 km) good simulation scores (> 30,000) can already and consistently be obtained in scenario Habitat-C (number 3 in the evolved set has actually a maximum foraging range of 10). Note that the relative low values for 6 and 17 are incidental, where models that have this value for other reasons do not score well and therefore produce no offspring.

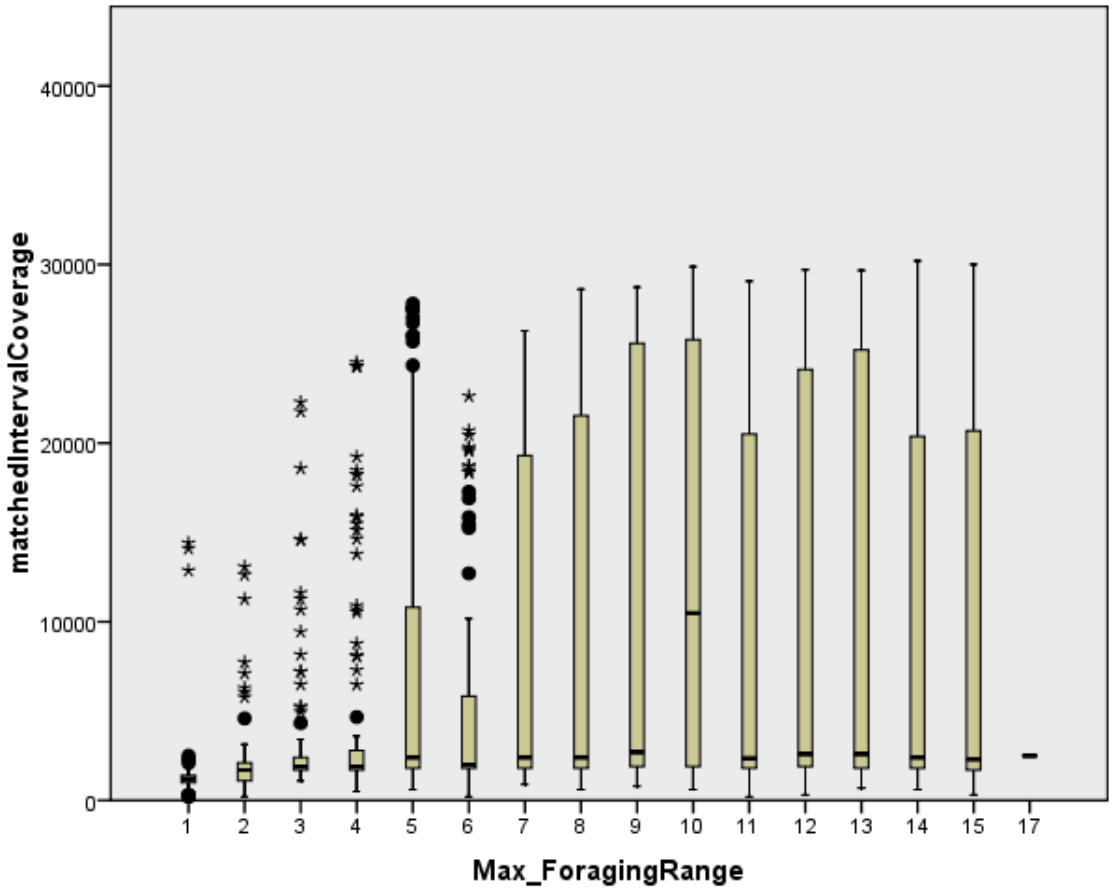



**Figure 92: Simulation score versus maximum foraging range. For low values of the foraging range good scores are difficult to obtain. Increasing the maximum beyond 9 does not influence the score much. Values are from scenario Habitat-C.**

From these results we could say that it is unlikely that Neanderthals would have a very limited foraging range. But for a good match with the archaeology an unlimited foraging range is not necessary. Likely values for any maximum would be over 90 km.

### 12.3.6 Q5 - What happens if we assume that Neanderthals were able to cross large open water systems?

Crossing open water is enabled in scenarios Habitat-H and Energy-D. In Habitat-H next to the ability to cross open water also a core area producing hominins was included. The score was very high (31119) but difficult to compare with other scenarios. Energy-D (26112) can be compared to Energy-B (29896) and scores actually less than that scenario without crossing.

Illustrating the effect of crossing larger water bodies is Figure 93, a density map from scenario Energy-D for simulation number 1365. This can be compared against results without that ability, in this case Energy-B (see Figure 90). The scores are again very comparable, and there are few differences visible on these maps. The most striking

differences are the darker colours for and thus more visits to some of the islands off the Atlantic coast of France (marked by a circle in Figure 93).

Note that the exclusion of the ability to cross larger water bodies does not cause the model to match the absence data any better, suggesting that flooded land bridges are not keeping hominin groups out of England.

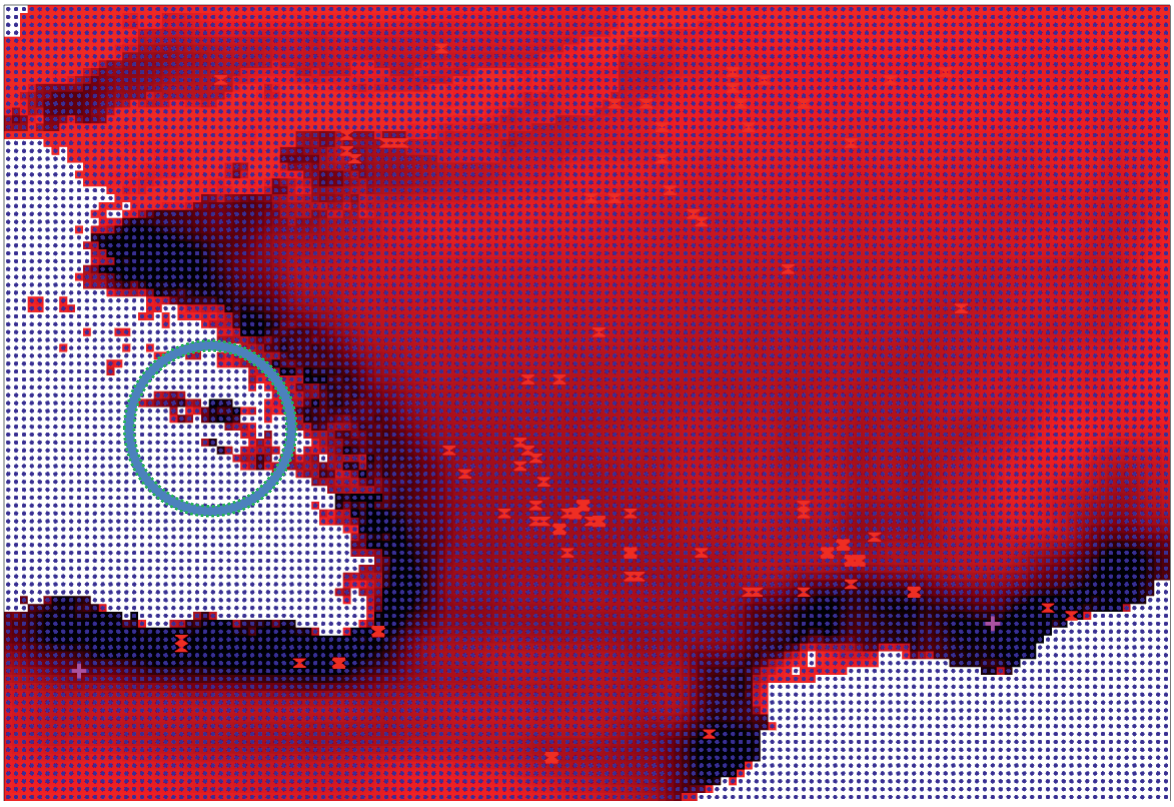


**Figure 93: Illustrating the ability to cross larger waterbodies to some of the islands off the coast of France. Produced in Scenario Energy-D with dynamic hominins, simulation 1365.**

### 12.3.7 Q6 - What influence has adding absence data to the model?

The absence condition is activated in scenarios Habitat-B, Habitat-D, Habitat-F, and Energy-C. It is implemented as a death penalty, meaning that if the condition is violated the simulation is aborted and not included in the results and thus cannot be used to generate offspring. This is the setting that has the most effect upon the model and the simulation results. When the absence condition is activated, the score is invariably lowered, sometimes to a large extent. The maximum score with absence activated is 19186 (Habitat-D). The only absence intervals are defined for areas in England. This translates into the drive to stay out of England and still score as many points as possible. Note that points cannot be obtained from England itself. Hominins cannot go there, but that absence does

not cost points. The loss of points comes from other areas that cannot be reached when England is unreachable for the model configuration.

The successful individuals in the scenarios that have the absence condition imposed all have an extremely low value for the viability index, for instance by implementing low birth rates and high mortality figures. This can be understood from the drive to not reach England. These groups will live to reproduce and produce archaeology, but not long enough to reach England, coming from productive areas in the south (core areas or the generally high productivity areas near the southern coast lines and mountain ranges). This also explains some of the temperature tolerances (only to -10 for the two best performing models in scenario Habitat-B, for example) since an inability to withstand colder temperatures helps in not reaching England. Similarly, small foraging ranges will reduce the capability to obtain sufficient resources in the north, where generally the environment is less productive (for instance the consistent value of only 6 for the top 3 best evolved individuals in scenario Habitat-D, compared to values of 10-14 for scenario Habitat-C).

The most extreme evolution can be observed in models that combine absence conditions with core areas (that constantly produce new groups). The core areas are located in the south, and to obtain (high) scores the new hominin groups should *not* reach England. One way to achieve this is with very low birth rates (possible because new groups are produced by the core areas anyway), combined with high death rates or other mechanisms to prevent living in the north. The birth rates of the best evolved models in scenario Habitat-D all have a birth rate of 1 (!), as have those from Habitat-F. All these models have small foraging ranges, higher values for years before group maturity, very small values for group size before merge is allowed, and high death rates for the pre-fertile cohort. All these key figures help to *not* reach England.

Using presence data only creates a perverse drive to have as many Neanderthals everywhere, all the time. But introducing absence data can instil an even more perverse force on the evolution of better matching hominins. If such a force creates hominins that are clearly impossible there is a chance that the chosen absence conditions do not reflect reality.

### 12.3.8 Q7 – What is the effect of using population core areas?

Population core areas were implemented in scenarios Habitat-D, Habitat-E, Habitat-F, and Habitat-H. A core area is represented by a hominin factory in a single grid cell, which

produces up to one new hominin group per time step if conditions are favourable. All core areas produce similar numbers of hominins (see Table 53). On average they add almost two hominins per time step (or one group per 12 years) to the simulation. Habitat-H (31119) can be compared to Habitat-G (30121), which only lacks the setting for allowing the crossing of water bodies, or to Habitat-I (31142) with lacks crossing water and has a maximum foraging range. The simulation scores are very similar for these scenarios, with or without core areas.

The number of deaths is administrated per simulation, and differentiated per type (Table 53). All scenarios with core areas have similar large numbers of hominins that are eliminated because their group size falls below the sustainable minimum size (each almost 1.5 million deaths). No other scenario setting results in such mortality rates for this variable by far. Other mortality numbers are also high. For death due to cold the four scenarios with core areas form the top four. For deaths due to hunger Habitat-E has the maximum number. Other figures do not stand out. Only Habitat-H has static hominins, and here deaths due to flooding is not exceptional, and for all core are scenarios the deaths due to merging are no different than in other scenarios. Also the created hominin count is not remarkable.

Habitat-D and Habitat-F are scenarios where the absence criterion is forced upon the models. And although the simulation scores are not very high, the core areas function as hominin pump that allows the population to maintain a very low birth rate while still scoring points. These are scenarios where the value for birth rate can be one, and combined with high mortality rates for the pre-fertile cohort this ensures proper extinction before England can be reached.

Habitat-E is a scenario that became dominated by all very similar individuals that have an extreme birth rate of 50%, which would no doubt have been even higher if this would have been allowed. A maximum score of 29153 (not exceptionally high) is only reached in evolved simulation number 2420, suggesting that improvement was slow, steady and difficult to achieve. This is typical for single dominance, where one type of successful individuals dominates the tournaments. That this is the case is well illustrated by the all very similar Chernoff faces for this scenario in Figure 70.

In scenario H the settings and parameter values in the Standard set already allow very high simulation scores, difficult to improve but resulting in one of the highest scores overall. Individuals here have high birth rates, low mortality rates, early maturation but hesitant to

merging, short childhoods and prolonged fertile periods, and very large foraging ranges. As in other scenarios with cores added, but very visible in this scenario is the fact that these extra hominins boost the simulation score for all models in the Standard set, see Figure 56.

Adding core areas to an environment does not fundamentally change the model. The additional hominins are considerable in numbers but do not improve the simulation score to great extent. The constant trickle of new hominins does allow extreme values for certain parameters (like the birthrate of 1 for absence models). However, it appears that core areas as they are implemented in HomininSpace are not a necessary asset.

### 12.3.9 Q10 - What are the results if the Neanderthals are implemented with random movement?

Hominins that implement random movement move through the landscape irrespective of available resources. This opposes the directed movement where groups always move towards the area with the most available resources. Random moving hominins were implemented in three scenarios: Energy-BR, Energy-CR, and Energy-ER. Each of these scenarios adds random movement to another scenario against which the results can be compared. Energy-BR (28125) against Energy-B (29896), Energy-CR (3301) against Energy-C (6273), and Energy-ER (30439) can be compared with Energy-E (30651). The hominins in Energy-B implement dynamic movement, Energy-E has static hominins and Energy-C imposes the absence criterion.

Random movement results consistently in slightly lower simulation scores, but scores for Energy-BR and Energy-ER are not significantly different from their non-random counterparts. An exception is Energy-C that implements the absence criterion and with random movement the simulation score, which was already very low, is even halved. This is expected since random moving hominins easily move into absence areas accidently, eliminating any hominin model that has groups moving north (Energy-C implements dynamic movement). This leaves only models that inherently produce very low scores.

Energy-E implements static movement which means that groups select areas in which they will remain. Thus the effect of random movement is expected to be little. When random movement is activated the resulting best scoring individuals (the top three) are rather uniform, with relatively low birth rates and a viability index in the order of several thousands. Interestingly, there is not much difference with the score for the dynamic

moving hominins in Energy-B. However, when random walk is activated that scenario has consistently fewer created hominins and lower mortality numbers (20-60% of non-random Energy B). In short, scenario Energy-BR reaches the same score but with fewer hominins. Constantly moving randomly through the landscape will not generate the most optimal energy yield from the environment, resulting in lower population totals. But apparently still enough to produce high simulation scores.

# 13. CONCLUSIONS

A model is no more than a rough simplification describing real world aspects and events. The purpose of a simulation model is to allow exploration of questions and ideas in order to identify the mechanisms that could provide a possible explanation of the observed phenomena. The outputs of a simulation derive from the interactions of the modelled parts, through a process referred to as emergence. To obtain a most parsimonious model and the most optimal fit with the data, model parameters can be manipulated to focus the comparison results. More focus increases trust in the proposed model elements and selected values.

HomininSpace was designed to contribute to the discussion of existing ideas about mobility of past hominins. In the process of implementation many parameters were identified that play a role in a model that takes subsistence requirements and translates those into motivations to move through a landscape. This final section of the thesis will start in a bottom-up fashion with suggestions for improvement of those parameters, through a more elaborate simulation structure and finally a wider application of this simulation environment.

Underlying the HomininSpace simulation system is a minimal model that describes population structure, subsistence, demographics and social interaction. Sixteen model parameters can be manipulated to create new model instances with different characteristics. Parameter values are modified with a Genetic Algorithm, an Artificial Intelligence method aiming to find good performing models through optimization techniques inspired by evolution. In a case study focussing on Middle Pleistocene hominins in western Europe, several questions that remain unresolved in the current debate on Neanderthals were explored. These include questions on the usage of coastal resources, the usability and maximal extent of a foraging range, the ability to cross large open water systems, the presence of population core areas, the distribution of energy in the environment, the use of absence data, and the character of the mobility when groups of Neanderthals explore the landscape.

In HomininSpace, a fitness value is computed based on comparison with archaeology to allow ranking of alternative models and thus exploration of hypotheses. To summarize some quick results: coastal resources make life easier for past hominins, especially for female individuals; a large foraging range allows hominin groups to cope with setbacks;

there are no better simulation scores when hominins are able to cross water bodies; population core areas facilitate certain model configurations but are by no means essential; energy in the landscape can be described using habitat types or with a continuous energy distribution but for Neanderthals this made no difference; the introduction and enforcement of absence data points in England steers the system towards unrealistic model configurations illustrating the difficulty the model has with staying out of England; and finally the mobility character of past hominins, static or dynamic, made no difference in the match with the archaeology at all.

To understand the simulation results, it is important to realize that the fitness value is constructed by matching *presence* data in the archaeological record. These data points are spread through the simulation area, both in space and time. The theoretical best match would be attained by a hominin that is present everywhere all the time (!). Such a hominin cannot be sustained by the limited amount of resources produced by the environment. So the system, implemented in a GA, searches for sub-optimal solutions. One of these is a family of solutions that represents a hominin type which is not, as such, recognized in the literature. This is a hominin that constantly travels through the landscape in very small groups (5-15 individuals), hardly interacts with other groups, and with a very high birth-rate that is close to the physical limit a female modern human body would be able to sustain (see for instance the results from Energy-B).

Because unexpected solutions that perform well are constructed it can be concluded that the implemented Genetic Algorithm works. It improves upon results from randomly constructed (unbiased) parameter values and falsifies previous research (Scherjon 2015a) that suggested that the evolved Dynamic Neanderthals have an equal or better fit than all evolved or manually constructed Static ones. The search for optimal solutions is generic, systematic and produces better results than informed manual selection of parameter values (Scherjon 2016). The character of the parameter value development for the set of most optimal solutions confirms the statistical analysis on the significance of certain parameters on the fitness value. This information can be used in future parameter reduction efforts on the model, but care must be taken: some of these parameters which are unimportant in some scenarios are important in others and vice versa (for examples see the discussion about the Temperature_Tolerance and mortality rates for the post-fertile cohort in subsection 12.3.1).

## 13.1 Developing HominінSpace – a tool for modelling hominin mobility

One of the aims when building the HominінSpace simulation system was the development of a tool that would be able to simulate modelled hominins in a realistic, reconstructed environment without relying on model parameter value choice by the user. Instead, the system should construct values by comparing simulation results with archaeological data and using an optimization algorithm to modify model parameter value sets in order to produce an optimal fit when used in simulations. The building of such a system is challenging, and finding a balance between what is possible, the level of realism (and the amount of details) that is to be implemented, and what can be realised given the computational capabilities and other resources requires an open mind at every step and careful consideration of alternatives and a focus on the research aims.

One of the most important lessons learned from this project is that the communication of the model is of utmost importance throughout the development process. An implemented model not only consists of model parameters and formulas, but also the structure in which those are embedded (the source code that implements the model and the runtime environment where the model is simulated). When implementing a model numerous assumptions remain hidden in the implemented classes and transmitted data between objects, in functions, in data stores, in input files, etc. Those assumptions form an integral element of a model, and must be clearly communicated. The use of an ODD is instrumental in this since the ultimate aim of an ODD is not the documentation of a model, but to provide all information needed to recreate the simulations. This means that all underlying assumptions in a simulation system must be stated in the ODD. The model developed in this study is extensive and a re-implementation is not easily done, and the review of the ODD and the reimplementation of parts of the model within the framework of this study can only address parts of the model, not the complete system.

Reproduction of simulation results has been shown to be an essential part of modelling efforts. Using community developed best practices and standards can improve reproducibility of any research (Freedman *et al.* 2015). An ODD and even a re-implementation cannot prove that a model is correct, but by making sure that in the design and in the implementation all effort is given towards proper building of the model, one can make plausible that a correct model has been created. Further analysis must make sure that the model is verified and validated. To allow proper replication by other researchers

complete replication files should be publicly available (Dafoe 2014; Marwick 2017). Therefore the model of HomininSpace, including source code, data files and documentation is made available through open access at OpenABM.

Code review is the systematic examination of computer source code by third party researchers and has been shown to be very effective in eliminating errors, finding inconsistencies and improving the quality of the reviewed work early in the development cycle (Ackerman *et al.* 1989; Fagan 1976; Kemerer and Paulk 2009). Peer reviews may also address readability, consistency in coding style and purpose, potential performance issues, simplicity of programming efforts (Whitner and Balci 1989) or evolvability defects (interesting for projects with expected long life cycles, see Mäntylä and Lassenius (2009)).

The model and source code of HomininSpace has been reviewed by Iza Romanowska, a computer programmer and modeller with experience in Palaeolithic archaeology (Romanowska *et al.* 2017). In the formal code inspection model reviewers must consider the effects of proposed changes, discuss these with the developers and follow a checklist when reviewing the code (Fagan 1976). In modern, light weight reviewing procedures such strict review criteria are released (Rigby and Storey 2011) with the risk of neglecting the overall design features (Bacchelli and Bird 2013). In the current study, by positioning the code review in a replication effort, such dangers are averted even though the review itself was informally organised. Code inspection is also used to detect underlying assumptions and instrumental in completely understanding the implemented system.

### 13.1.1 Usability of the tool

Since a model is always a simplification of reality, modelling efforts are never an end product. Instead modelling results will be integrated into larger research projects, to be used as additional argumentation material and hypotheses exploration. A good Neanderthal model depends on the cooperation of scholars capable of writing software that implements an explicit model and executes the simulations and other scholars with a profound knowledge of and access to Neanderthal materials and data defining a useful model. An ABM is a very useful complex systems modelling tool that can link individual behaviours to population level patterns that archaeologists seek to understand. It is a modelling tool that is able to simulate change through time and capable of modelling the interactions between individuals and the environment in a decentralised, bottom-up manner. Modelled Neanderthals implement parameters to explore a range of conditions that cannot be

observed first hand and do not form analogies with today but are used as blocks for systematic theory building.

Modelling appears to hold a reputation of objectivity but in general it is ultimately the designer of the system who decides what to include in a model, and which values are assigned to the model parameters. To avoid such non-neutral input HomininSpace has been designed to find values for model parameters in an unbiased and objective manner, using Genetic Algorithms. A Genetic Algorithm (GA) has been implemented to systematically explore the parameter space that results from the chosen range of parameter values in the underlying model. Simulation results are compared against an archaeological record of actual presence data, resulting in a quantitative fitness value per simulation.

It is shown that the implemented GA is capable of finding optimal fitting parameter value combinations that result in a higher fitness value than informed manually selected parameter values. It performs a directed systematic search of the parameter value space. With GA researchers' expectations and biases are overcome and a better understanding of the parameter impact on modelling results is possible. As such, a genetic algorithm or a similar systematic technique improves the results of a large scale (in simulation time, number of model parameters and/or parameter value spaces) modelling system.

Genetic Algorithm techniques applied in HomininSpace are, unfortunately, computationally expensive, since the calculation of the fitness value is the actual simulation run with the evolved parameter set. Therefore the following elements must be considered carefully when constructing a GA for an archaeological ABM: the total computational costs, stochasticity in the genetic algorithm, choice of GA operators, the stop criterion and the chosen fitness function. As with all stochastic modelling, obtaining a good or the best model is never guaranteed. And note that due to the nature of the GA improvement process the best scores tend to represent individuals from a few successful lineages only.

When applied to the research question on Neanderthal mobility alternatives, it must be concluded that the scores for Dynamic and Static strategies are very comparable. The fitness values for improved individuals are within the same order of magnitude, and there is no statistically significant difference between Static and Dynamic hominins (contrasting previous research). However, it is noteworthy that the best matching simulations were always those which, albeit by a narrow (and insignificant) margin, have hominins that are implementing a Dynamic mobility strategy.

A simulation is scored by matching the relevant archaeological data. In this research each radiometrically obtained data point is used as an interval around the calculated point in time, min and max delimited by the standard deviation. Dating results from the same layer are combined. A total of 83 checkpoints is created, illustrating the deficiency in accurate chronological information for the simulation area and period to match against. A match is a simulated presence at an absolute point in time, where one point is added to the simulation score for the matched presence itself. Up to one more point can be gained depending on how many more visits occur during the interval (as a percentage with a maximum possible score of 100%). Visits outside an interval do not count, neither as bonus points nor as minus points.

The scoring mechanism does justice to the fact that having at least one visit is most important, since we cannot be sure how much archaeology any visit produces (often one visit could theoretically have created the site in question). And there is logic in the assumption that more visits mean more archaeology to match. Visits outside the intervals cannot be interpreted as violations, since it is not sure that there are no visits if no archaeology is found (visits could have produced no archaeology to test against, the archaeology might not have been retrieved yet, more precise dating results are not available, or visits could have been made to the next cave).

Critically reviewing the model and simulation one might conclude that there is a lack of heterogeneity and that only energy requirements rigorously steer Neanderthals. Such critiques and others match some of the things said previously about equation-based models (Brughmans and Poblome 2017) and ABM has gained popularity just because it can be used to address such issues. Agents in HomininSpace are homogeneous in one simulation but vary between simulations. They are similar in character but vary in size throughout any simulation. And parameters related to the metabolic budget are amongst the least important ones in all simulations. With HomininSpace a virtual laboratory is created where hominins evolve and where the best adapted individuals are allowed to create offspring that maybe even better matches the archaeology.

### 13.1.2 Who were those Neanderthals?

Quite telling is the fact that the HomininSpace system is able to find good performing Neanderthals in most circumstances, suggestive of the flexibility in the implemented model. The characteristics of these Neanderthals differ per scenario, aiding in interpretation and assessment of the implemented hypotheses and answering the questions

that this study aims to explore. There appears to be no difference between the continuous energy reconstruction and a delimited habitat area reconstruction. Modelled Neanderthal hominins were irrespective of the manner of energy distribution capable of using the environment to their benefit. And the comparison between dynamic and static hominins failed to identify a best strategy. Both implementations allowed the hominins to cope with the challenging environment.

Introduction of random movement made no difference in some cases (for instance in combination with static hominins) but could also result in very ineffective models (see below). Adding coastal resources to the model seems to make no difference at first glance (same simulation scores) but when model parameters are inspected these additional resources allow a more relaxed demographic sub-model with more realistic values with less physical stress for female Neanderthals (e.g. birth rates below 40%). Adding the ability to cross open water does not influence the simulation scores, but allowed hominins to reach England earlier.

Absence data states that hominins were *not* in an area for a given time period. It is very difficult to define proper temporary and geographical boundaries for absence data, and as such, only a suggested absence of England is included in the different scenarios. Striking is that the absence condition leads the system into producing relatively "strange" hominins, often stretching the boundaries of realism and the model. With the flexibility to adjust the model into many different configurations that generally score well, the model appears to have difficulty staying out of England. Introduction of model element like random movement only makes matters worse. And where limiting the foraging range does not seem to hinder foraging hominins in other scenarios (in the sense that other model parameters can be adjusted to mitigate the limiting effects), in combination with the absence condition it is actively used by the model to restrict hominin presence to the north. Especially when a population pump in the form of the core areas is activated ensuring constant supply of hominins the system can create groups that have no chance of survival for any prolonged period of time.

Very short lived and fast reproducing Neanderthals score well with the archaeological data while being unable to reach England. But in general the scenarios that have absence conditions imposed score relatively low, with Scenario Energy-CR as an ultimate example with a maximum score below 4000 points. This apparent struggle and the sometimes unexpected turns that the system takes indicate that the model cannot easily match the

absence of archaeological data for this period and area. This could indicate that there were other mechanisms that kept Neanderthals out of England, or that there is archaeology still to be found or properly dated (Gibbard 1988).

In general, absence data is considered vital to avoid the perverse drive of producing as much offspring as possible in order to create an optimal match with the presence in the real world. But that is now replaced by a similar perverse system of sufficient reproduction to score but with model parameters that ensure *non* viability. Note that any other absence locations in the simulation area would have resulted in similar effects but even lower scores, since these would lack the potential absence causing mechanisms of flooded land bridges or latitudes that sport very cold climates warding off hominins. It can be concluded that absence data is a potentially very useful asset to create more realistic simulation models, but that those models must include very explicit and potentially effective mechanisms to cause that absence in order to avoid even more perverse models.

The basic hominin model as implemented is generally very capable of adjusting to different circumstances while still producing good simulation results. The HomininSpace simulation system will automatically substantiate which model under what hypothesis best represents in all likelihood past hominin behaviour in the landscape and will thus lend some empirical credibility to the effects of the included interpretive hypotheses. It will also fill in model parameter values that reflect the effect of environment, other hominins and settings in exploring possible behaviours in the past. In this manner answers can be given for specific questions, after translating the questions into hypotheses that were implemented as elements of the model.

The system prefers certain parameter value combinations for answering each question, but some trends can be identified. High birth and low mortality rates for the (pre) fertile cohorts were evident. Very clear is also that the system does not care for hominins of post-fertile conditions. Mortality rates here are generally very high, especially when compared to fertile cohort numbers. These numbers are sometimes accompanied by increased subsistence needs for elderly, often even surpassing the needs of the fertile cohort, made possible by the limited number of surviving individuals.

The subsistence requirements vary widely but whatever value is chosen it can be compensated by an unlimited or very large foraging range. The same is true for the amount of calories extracted from a kilogram of edible meat, including proteins and fat. Even when resources are scarce large foraging ranges enable hominins to gather sufficient foods

allowing survival albeit in very low numbers. These very large ranges ward of competition but also paradoxically reduce score potential since checkpoints that are located within foraging range are not necessarily visited every year. Randomly moving through the landscape appears not necessarily detrimental for survival or for a good simulation score (but can be for population densities).

Tolerance for colder temperatures is not essential for getting high scores. Normally, values where hominins tolerate around -20°C as the coldest temperature of the year do well. Combined with the absence condition the value can be surprisingly high (around -10°C), contributing to early death in the north. Finally, the variables that steer group dynamics vary widely. Especially the years before group maturity and the group size before split can take values from a wide range, varying within and between scenarios. Group size and group size of the fertile cohort before merge also vary, although from a much smaller range. Occasionally the functionality these variables implement is turned off completely by the system assigning zero to them.

The power of this model to overcome different circumstances, to recover quickly after setbacks and the ability to cover large distances enables the modelled hominins to generate substantial matches with the given archaeology. That is independent of the usage of coastal resources, the ability to cross larger water bodies, the apparent non-importance of subsistence quantifications in some combinations, etc. On the other hand, it must be pointed out that there probably was not a set template for a Neanderthal individual. There was no standardized Neanderthal. True realism is more divers than a set of parameters can capture.

## 13.2 Implications of this study

In the literature there are constant attempts to improve demographic parameter values in Neanderthal models. The energy needs for age groups are refined and many chosen values (especially for pre- and post-fertile age groups) are supported in the literature. Implementation of the basic concept of refugia is modelled ever more realistically. The contributions of this study however suggest that variation within a basic hominin model that includes some essential demographic parameters is sufficiently capable of exploiting limited resources. Allowing parameter value variation to optimally adjust to specific circumstances creates good matches with the real world data.

Essential for a good modelling performance is the implementation of a flexible way of assigning model parameter values, allowing the system to create that combination of parameter values that produces the most optimal match with the archaeological data. In other words, the flexibility observed in the anthropological literature regarding the variation in modern humans would allow earlier hominins to be very adaptive and capable of creating the archaeological data sets as we know today. I see no reason to assume that such flexibility is lacking from, let say, the basic model of the Neanderthal species.

This study also illustrates the importance of well-chosen variables to include or exclude from the model and simulated environment. More important, it is made very clear that biased value assignment to model parameters is unnecessarily limiting the power of the model, compared to flexible value selection between given boundaries. Even a limited, very basic hominin model is then capable of matching archaeological presence data optimally, excluding the need for implementing unwarranted ideas about either environment or behaviour.

Variation within the basic hominin model allows individuals to very effectively exploit the environment negating the effects of explicit handicaps like a lacking ability to exploit coastal resources, the inability to cross larger water bodies or the behavioural limitations from having core areas. When that flexibility is removed and biased parameter values are selected the most optimal models are avoided leading to circular reasoning and self-fulfilling prophecies.

This research acknowledges what Kelly (1995) has shown in his well-wrought book '*The foraging spectrum*': that modern day as well as historical hunter-gatherers can be extremely different from each other and therefore most likely also differ from prehistoric foragers as well. There is a wide variety in almost all aspects of the hunter-gatherer life (White 2013). It is thus not possible to speak of *the* hunter-gatherer. Nor of *the* Neanderthal since there is no reason to assume less variety in pre-modern humans.

The implementation of the genetic algorithm as an essential part of the HominінSpace modelling system is warranted by the need for an effective parameter value space searching mechanism. This is a contribution to the study of agent-based modelling methodology and simulation techniques itself.

## 13.3 Avenues for further research

The output of the HomininSpace tool is varied and diverse. This thesis has only explored some of the simulation results, mainly aimed at answering the underlying research questions. But the emerging patterns can offer insight into many characteristics of hominin behaviour and the reconstructed environment. Graphs of the matches through time with bars for each period with the predicted amount of hits and the registered hits could better identify matching archaeology. Analysis of the matches within each checkpoint would aid in this aspect as well.

Some social aspects of hominin behaviour have not been included in the model to limit variability. But some are not difficult to implement. For instance, territoriality in many difference variants can be implemented. The different origins of the hominin group (from the different refugia for instance) could play a role here, both in sharing resources as well as social interaction (exchange of individuals, cultural transmission of ideas/technology, merging of groups). A genetic profile for each individual or group could help to illustrate and analyse the effects of such social behaviour. Extending HomininSpace with social aspects would enhance the usefulness as a tool to test hypotheses.

Where the main conclusion regarding the behavioural mobility of the past hominins in the Middle Pleistocene is that they could have included at least partly a static aspect of mobility in the way they utilize the landscape around them, a natural step further would be to investigate further the nature of that mobility. Questions regarding territoriality, group identification, cultural diversity or genetic profile distribution can be answered with only small adaptations to the existing system. For instance allowing other groups to move through a foraging range can be based on cultural type, relative group size, etc. A genetic footprint can be attached to groups, and density maps based on these profiles can be produced. Implementing some social interactions, which could include avoidance of other groups based on intrinsic group characteristics, would add significantly to the utility value of HomininSpace.

The multi-dimensional variable space created by implementing the many but essential model parameters within HomininSpace (as well as for the modelled hominins, the modelled environment and the topography) has hardly been explored, due to limited resources (time and computing power). A few promising directions have already been identified with the variation in values for cold resistance and simulations with mixed types of hominins. It would be revealing to analyse what kind of combination between static and

dynamic groups produces the best matches with the archaeology. A systematic exploration of the different variables would allow the exact mapping of the influence of each parameter on the behaviour and thus the archaeology of the modelled hominins. Such an endeavour is possible using the automated parameterization feature of the Repast modelling system, possibly assisted by an automatic collection of results. This is also very useful for regression tests when further developing the software, ensuring that new functionality does not compromise earlier results. Based on the promising results of the research presented in this thesis I would highly recommend that such an exploration is undertaken.

The use of HominínSpace allows users a thorough understanding of effect of a wide range of parameters on hominin behaviour. Not only does the tool illustrate the effects of these parameters, but the ability to manipulate the variables and see the actual consequences of chosen values will animate otherwise purely theoretical concepts. The simulation period is easily adjusted, making it possible to focus on specific periods in the past. With a little more effort the simulation area can be changed since model data is basically available for any chosen locality. The model is sufficiently generic that with minor adjustments it can be applied in modelling any top predator. The possibility of actually implementing old and new theories will allow students of any level to answer questions and test hypotheses on a wide range of concepts and will aid to conceptualize ideas from the literature.

### 13.3.1 Recommendations when reconstructing past environments

Environmental reconstructions of the past are today created in more and more detail. This study illustrates that such reconstructions can be very helpful, and that more detail is needed to allow a more thorough analysis of past behaviour of hominins. Many options for improvement have been identified. Adding additional information to improve habitat reconstruction is a recommendable option. From actual observations it appears that soil data holds information that improves predictive models (Wisiol 1984, 474; Prentice *et al.* 1992). Accordingly, including soil information in the simulation and environment reconstruction is one way of creating a more reliable habitat reconstruction. Information about the modern geographic distribution of soils is available in the Zobler Soil Database, to be downloaded from the Oak Ridge National Laboratory Distributed Active Archive Center Oak Ridge Tennessee, USA (Zobler 1986). The Zobler soil dataset consists of latitude and longitude coordinates (1° x1° grid) for all the identified soil types under the FAO Soil Classification System (26 soil types).

This study is an attempt to use temperature and precipitation as the proximal predictors on resource gradients (as described by Guisan and Zimmermann 2000). Other resources besides energy might steer hominin movement through the landscape. One possibility is lithic availability (Fernandes *et al.* 2008). Flint, the main and preferred raw material for making tools, is mined from carbonate rocks. Flint bearing rock is distributed throughout the world (Ford and Williams 2007) and can be used to attract hominins. A digital version of the world map of carbonate rocks is available from http://web.env.auckland.ac.nz/our_research/karst/, accessed 17 September 2013.

From the simulations in this research energy, although essential in itself, appears to be less differentiating as expected. Other resources can be implemented as attractors inducing movement through the landscape and could include lithics like flint or minerals like ochre and manganese oxides. Research further suggests that specifically presence of many herbivores maintained the Mammoth steppe, and extended this habitat into a wider diversity of climates (from Spain to Siberia) than those in which it would occur naturally (Zimov *et al.* 2012). As such it is more insensitive to climate variation and thus climatic parameters alone are insufficient to predict the prevalence of this steppe. The specific conditions that allowed the Mammoth steppe to exist no longer prevail, with the possible exception at the Pleistocene Park, an artificially created and maintained area aimed to represent that steppe environment (Zimov 2005). HomininSpace models the environment using climate parameters only and as such cannot include the Mammoth Steppe as reconstructed habitat, where this is suggested to be very productive in edible resources.

### 13.3.2 Suggestions for simulating hominins

Palaeolithic archaeology is a fascinating research discipline. New methods and techniques are implemented rapidly and can gain wide acceptance (for instance the palaeoproteomic analysis described in Welker *et al.* 2016). This study has shown that agent-based modelling can contribute to understanding parts of the deep past. And although not a new method in itself, I argue that simulation can and should be used more often to explore the unknown elements of our origin, if only because it forces the researcher to explicitly state his or her ideas about the past in a formal and testable manner.

Elements of HomininSpace can be improved upon. Most questions about Neanderthal behaviour in the landscape were only touched upon in this work and would deserve a dedicated study with the single aim to further our knowledge focussing on a single question. In such a study the results from this research can be used as a basic data set, a

starting point from which the modelled aspects can be tailored into a more extensive explicit model. Characterizing proper absence is very important (e.g. Bertran *et al.* 2013). Defining high resolution presence intervals, surrounded by assumed absence might provide a good selection mechanism. There is more data available about areas in eastern European. Core areas are now implemented by producing new hominin groups according to local circumstances, and they could vary their output and implement a variable absorption protocol (now in the form of open or closed borders, irrespective of climate, population density or other modelled attributes).

In HomininSpace, hominins are assumed to extract their main energy from mammalian resources. It is not unlikely that much if not all food consumed consisted of plant materials. The problem is how to quantify this over the landscape. Data exists on secondary biomass in a variety of habitats, but what is needed when Neanderthals are assumed vegetarians would be a quantification of edible materials for the different reconstructed habitats or energy levels.

Influencing the habitat they live in is something modern humans are very good at. Fire use in the landscape is a tool that would have allowed Neanderthals to do likewise (Scherjon *et al.* 2015). Increasing resources in a number of ways is one of the known side effects of repeated burning. We have seen that adding coastal resources relaxes pressure on some model parameters. Creating additional resources inland, appearing where Neanderthals are might have similar or additional effects which would be interesting to explore.

The capability to run the development environment on Windows, Linux and on high performance computers ensures availability of the tool on many platforms and in almost any computing environment. The intuitive source structure invites students to implement new functionality. A limited foraging range or social interaction with other groups as examples have been shown easy to implement. The use of Java as a development environment allows a large degree of freedom in the use of additional libraries and existing software. The performance of the tool allows fast development tracts in assessing effects for certain traits or in the implementation of new functionality. With the execution of multiple time steps per second, depending on chosen computer hardware and environment size, the effects of new functionality or variation in parameter values are immediately visible.

I hope that by implementing an explicit model of Neanderthals and via showing the usefulness of simulations to address questions the potential power of computational

modelling will inspire other researchers to apply these methods in their own research. And that this research illustrates that the inherent flexibility in a basic hominin model which is sufficient to avoid imposing unwarranted handicaps and self-fulfilling prophecies.

## 13.4 Last words

With the implementation of HomininSpace a new tool has become available to answer questions and test hypotheses on behaviour of hominins in large, realistic landscapes. The implementation of Neanderthal mobility in roughly the area currently known as France illustrates the complexity of answering such questions, the possibilities and limitations of the available data, and the feasibility of implementing and simulating widely spread but thinly distributed societies over a prolonged period of time. I hope that the experience gained from this development as put down in this thesis can offer guidance to other researchers pursuing their research objectives with modelling and simulation. And that with HomininSpace a new and powerful tool to explain archaeological patterns and to answer new and old questions is added to the toolset of archaeologists and paleoanthropologists.

# REFERENCES

The software and data used in this research is credited according to the guidelines given at GitHub, the data storage and versioning service[48].

### Paleoclimate data sets

- NOAA, National Climatic Data Center: Paleoclimatology data: http://www.ncdc.noaa.gov/paleo/model.html (accessed 29 May 2013, verified 17 September 2017).
- WorldClim – Global Climate Data – Past conditions: http://www.worldclim.org/past (verified 17 September 2017).
- PMIP – Paleoclimate Modelling Intercomparison Project Phase II: http://pmip2.lsce.ipsl.fr/ (verified 17 September 2017).
- CCSM – Community Climate System Model, a Global Climate Model (Wikipedia entry): http://en.wikipedia.org/wiki/Community_Climate_System_Model (verified 17 September 2017).
- MIROC3.2 (Model for Interdisciplinary Research on Climate) – model information: http://www-pcmdi.llnl.gov/ipcc/model_documentation/MIROC3.2_medres.htm (verified 17 September 2017).

### Topographical data sets

- VTP – Undersea Terrain Elevation (Bathymetry) – model description: http://vterrain.org/Elevation/Bathy/ (verified 17 September 2017).
- World Map of Carbonate Rock Outcrops v3.0, http://web.env.auckland.ac.nz/our_research/karst/, accessed 17 September 2013. Data is available as shape files (verified 17 September 2017).

### Software tools used in this study

- SAM – Spatial Analysis in Macroecology, v4.0: http://www.ecoevol.ufg.br/sam/ (verified 17 September 2017).
- Repast Simphony, version 2.2: https://repast.github.io/download.html (accessed 17 November 2016, verified 17 September 2017).
- IBM SPSS Statistics - http://www-01.ibm.com/software/analytics/spss/downloads/ (verified 17 September 2017).

### Literature

Abar, S., G.K. Theodoropoulos, P. Lemarinier and G.M.P. O'Hare, 2017. Agent Based Modelling and Simulation tools: A review of the state-of-art software. *Computer Science Review* 24 (Supplement C): 13-33.

Ackerman, A.F., L.S. Buchwald and F.H. Lewski, 1989. Software inspections: an effective verification process. *IEEE software* 6 (3): 31-36.

Adler, D.S., G. Barz, A. Belfer-Cohen and O. Bar-Yosef, 2006. Ahead of the Game: Middle and Upper Palaeolithic Hunting Behaviors in the Southern Caucasus. *Current Anthropology* 47: 89-118.

Agustí, J. and X. Rubio-Campillo, 2017. Were Neanderthals responsible for their own extinction? *Quaternary International* 431: 232-237.

[48] http://mr-c.github.io/shouldacite/index.html, accessed 21-04-2016.

Aiello, L.C. and P. Wheeler, 2003. Neanderthal Thermoregulation and the Glacial Climate. In *Neanderthals and Modern Humans in the European Landscape during the Last Glaciation: Archaeological Results of the Stage 3 Project*, edited by T. H. Van Andel and W. Davies. Cambridge: McDonald Institute for Archaeological Research Monograph Series, 147-166.

Akazawa, T., Y. Nishiaki and K. Aoki, 2013. *Dynamics of Learning in Neanderthals and Modern Humans Volume 1: Cultural Perspectives*: Springer.

Aldrich, J., 1995. Correlations genuine and spurious in Pearson and Yule. *Statistical science*: 364-376.

Aleksandra, B.D., M.E. Jovan and A.D. Rakic, 1997. Genetic algorithms for continuous optimization problems - a concept of parameter-space size adjustment. *Journal of Physics A: Mathematical and General* 30 (22): 7849.

Alexander, R.M., 2002. Energetics and Optimization of Human Walking and Running: The 2000 Raymond Pearl Memorial Lecture. *American Journal of Human Biology* 14: 641-648.

Allaby, M., 2005. A dictionary of Ecology (3rd ed). edited by M. Allaby. Oxford: Oxford University Press.

Allen, J.R.M., U. Brandt, A. Brauer, H.-W. Hubberten, B. Huntley, J. Keller, M. Kraml, A. Mackensen, J. Mingram, J.F.W. Negendank, N.R. Nowaczyk, H. Oberhansli, W.A. Watts, S. Wulf and B. Zolitschka, 1999. Rapid environmental changes in southern Europe during the last glacial period. *Nature* 400 (6746): 740-743.

Amante, C. and B.W. Eakins, 2009. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. *NOAA Technical Memorandum NESDIS NGDC-24*.

Ammerman, A.J. and L.L. Cavalli-Sforza, 1984. *The Neolithic transition and the genetics of populations in Europe*: Princeton University Press Princeton.

Anderson, R.P., 2003. Real vs. Artefactual Absences in Species Distributions: Tests for Oryzomys albigularis (Rodentia: Muridae) in Venezuela. *Journal of Biogeography* 30 (4): 591-605.

Angus, S. and B. Hassani-Mahmooei, 2015. "Anarchy" Reigns: A Quantitative Analysis of Agent-Based Modelling Publication Practices in JASSS, 2001-2012. *Journal of Artificial Societies and Social Simulation* 18 (4): 16.

Antoine, P., J.P. Coutard, P. Gibbard, B. Hallegouet, J.P. Lautridou and J.C. Ozouf, 2003. The Pleistocene rivers of the English channel region. *Journal of Quaternary Science* 18 (3-4): 227-243.

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

Anwar, N., K. MacDonald, W. Roebroeks and A. Verpoorte, 2007. The Evolution of the Human Niche: Integrating Models with the Fossil Record. In *Guts and Brains. An Integrative Approach to the Hominin Record*, edited by W. Roebroeks. Leiden: Leiden University Press, 235-269.

Appenzeller, T., 2013. Neanderthal culture: Old masters. *Nature* 497: 302.

Archer, W. and D.R. Braun, 2013. Investigating the Signature of Aquatic Resource Use within Pleistocene Hominin Dietary Adaptations. *PLoS ONE* 8 (8): e69899.

Arenas, M., O. François, M. Currat, N. Ray and L. Excoffier, 2013. Influence of admixture and paleolithic range contractions on current European diversity gradients. *Molecular biology and evolution* 30 (1): 57-61.

Ashton, N., 2002. Absence of Humans in Britain during the last interglacial (oxygen isotope stage 5e). In *Le Dernier Interglaciaire et les occupations humaines du Paléolithique moyen*, edited by A. Tuffreau and W. Roebroeks. Lille: CERP/Universite des Sciences et Technologies de Lille, 93-103.

Ashton, N., C.R.E. Harris and S.G. Lewis, 2018. Frontiers and route-ways from Europe: the Early Middle Palaeolithic of Britain. *Journal of Quaternary Science*: n/a-n/a.

Ashton, N. and S.G. Lewis, 2002. Deserted Britain: declining populations in the British late Middle Pleistocene. *Antiquity* 76: 388-396.

Ashton, N., S.G. Lewis and R. Hosfield, 2011. Mapping the Human Record: Population Change in Britain During the Early Palaeolithic. In *The Ancient Human Occupation of Britain*, edited by N. M. Ashton, S. G. Lewis and C. B. Stringer. Amsterdam: Elsevier, 39-51.

Ashton, N. and B. Scott, 2016. The British Middle Palaeolithic. *Quaternary International* 411 (Part A): 62-76.

Attenbrow, V. and P. Hiscock, 2015. Dates and demography: are radiometric dates a robust proxy for long-term prehistoric demographic change? *Archaeology in Oceania* 50 (S1): 30-36.

Augustine, D. and S. McNaughton, 2006. Interactive Effects of Ungulate Herbivores, Soil Fertility, and Variable Rainfall on Ecosystem Processes in a Semi-arid Savanna. *Ecosystems* 9 (8): 1242-1256.

Axelrod, R., 1997. Advancing the Art of Simulation in the Social Sciences. In *Simulating Social Phenomena*, edited by R. Conte, R. Hegselmann and P. Terna. Berlin, Heidelberg: Springer Berlin Heidelberg, 21-40.

Babbage, C., 1864. *Passages from the life of a philosopher*. London: Longman, Green, Longman, Roberts, & Green.

Bacchelli, A. and C. Bird, 2013. Expectations, outcomes, and challenges of modern code review. Proceedings of the 2013 international conference on software engineering.

Back, T., F. Hoffmeister and H.-P. Schwefel, 1991. A survey of evolution strategies. Proceedings of the fourth international conference on genetic algorithms.

Bailey, H.P., 1960. A method of determining the warmth and temperateness of climate. *Geografiska Annaler* 42 (1): 1-16.

Baker, R.R., 1978. *The evolutionary ecology of animal migration*. Vol. 1012: Hodder and Stoughton London.

Bankes, S.C., 2002. Agent-based modeling: A revolution? *Proceedings of the National Academy of Sciences of the United States of America* 99 (Suppl 3): 7199-7200.

Banks, W.E., F. d'Errico, A.T. Peterson, M. Kageyama, A. Sima and M.-F. Sánchez-Goñi, 2008. Neanderthal Extinction by Competitive Exclusion. *PLoS ONE* 3 (12): e3972.

Banks, W.E., F. d'Errico, A.T. Peterson, M. Vanhaeren, M. Kageyama, P. Sepulchre, G. Ramstein, A. Jost and D. Lunt, 2008. Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. *Journal of Archaeological Science* 35 (2): 481-491.

Barceló, J.A., F. del Castillo, L. Mameli and E. Moreno, 2008. The Computer Simulation of Social Dynamics and Historical Evolution. *Unpublished Paper presented at the II Jornadas sobre Simulación Social y Análisis de Sociedades Artificiales Barcelona, 20-21 noviembre 2008*.

Barnes, K., 2010. Georgia's fossils on view. *Nature* 463: 162.

Barrows, T.T., J.O. Stone, L.K. Fifield and R.G. Cresswell, 2002. The timing of the Last Glacial Maximum in Australia. *Quaternary Science Reviews* 21 (1–3): 159-173.

Bates, S.J., J. Sienz and D.S. Langley, 2003. Formulation of the Audze–Eglais Uniform Latin Hypercube design of experiments. *Advances in Engineering Software* 34 (8): 493-506.

Belovsky, G.E., 1988. An optimal foraging-based model of hunter-gatherer population dynamics. *Journal of Anthropological Archaeology* 7 (4): 329-372.

Benito, B.M., J.-C. Svenning, T. Kellberg-Nielsen, F. Riede, G. Gil-Romera, T. Mailund, P.C. Kjaergaard and B.S. Sandel, 2016. The ecological niche and distribution of Neanderthals during the Last Interglacial. *Journal of Biogeography* 44 (1): 51-61.

Benke, A.C. and A.D. Huryn, 2006. Secondary production of macroinvertebrates. In *Methods in stream ecology 2nd ed*, edited by F. R. Hauer and G. A. Lamberti. Burlington: Academic Press, 691-710.

Bentley, R.A., 2006. Strontium Isotopes from the Earth to the Archaeological Skeleton: A Review. *Journal of Archaeological Method and Theory* 13 (3): 135-187.

Bernbeck, R., 1997. *Theorien in der Archäologie*: Francke Tübingen.

Bertran, P., L. Sitzia, W.E. Banks, M.D. Bateman, P.-Y. Demars, M. Hernandez, M. Lenoir, N. Mercier and F. Prodeo, 2013. The Landes de Gascogne (southwest France): periglacial desert and cultural frontier during the Palaeolithic. *Journal of Archaeological Science* 40 (5): 2274-2285.

Biesele, M. and S. Barclay, 2001. Ju/'hoan women's tracking knowledge and its contribution to their husbands' hunting success. *African Study Monographs* 26: 67-84.

Binford, L.R., 1980. Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. *American Antiquity* 45 (1): 4-20.

Binford, L.R., 1982. The Archaeology of Place. *Journal of Anthropological Archaeology* 1: 5-31.

Binford, L.R., 1991. When the going gets tough, the tough gets going: Nunamiut local groups, camping patterns and economic organization. In *Ethnoarchaeological approaches to mobile campsites: hunter-gatherer and pastoralist case studies*, edited by C. S. Gamble and W. A. Boismier. Ann Arbor: International monographs in prehistory, Ethnoarchaeological series 1, 25-138.

Binford, L.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.

Bintanja, R., R.S.W. van de Wal and J. Oerlemans, 2005. Modelled atmospheric temperatures and global sea levels over the past million years. *Nature* 437 (7055): 125-128.

Bird, D.W., R.B. Bird and B.F. Codding, 2009. In pursuit of mobile prey: Martu hunting strategies and archaeofaunal interpretation. *American Antiquity* 74: 3-29.

Bird, D.W. and R. Bliege Bird, 2005. Martu children´s foraging strategies in the Western Desert, Australia. In *Hunter gatherer childhoods*, edited by B. S. Hewlett and M. E. Lamb. New Brunswick, New Jersey: Transaction Publishers, 129-46.

Birdsell, J.B., 1968. Some predictions for the pleistocene based on equilibrium systems among recent hunter-gatherers. In *Man the Hunter*, edited by R. B. Lee and I. DeVore. Chicago: Aldine, 229-240.

Blackwell, B., M. Bisson, A. Skinner and P. Beelitz, 2007. ESR dating bovid teeth from the Neandertal layer at La Ferrassie, France. GSA Denver Annual Meeting.

Bocherens, H., G. Baryshnikov and W. Van Neer, 2013. Were bears or lions involved in salmon accumulation in the Middle Palaeolithic of the Caucasus? An isotopic investigation in Kudaro 3 cave. *Quaternary International*.

Bocherens, H., D.G. Drucker, D. Billiou, M. Patou-Mathis and B. Vandermeersch, 2005. Isotopic evidence for diet and subsistence pattern of the Saint-Césaire I Neanderthal: review and use of a multi-source mixing model. *Journal of Human Evolution* 49: 71-87.

Bocquet-Appel, J.-P. and A. Tuffreau, 2009. Technological Responses of Neanderthals to Macroclimatic Variations (240,000–40,000 BP). *Human Biology* 81 (2-3): 287-307.

Bogin, B., 1999. *Patterns of human growth*. Vol. 23. Cambridge: Cambridge University Press.

Boismier, W.A., C. Gamble and F. Coward, 2012. *Neanderthals among mammoths: excavations at Lynford Quarry, Norfolk UK*: English Heritage Monographs.

Boivin, N., D.Q. Fuller, R. Dennell, R. Allaby and M.D. Petraglia, 2013. Human dispersal across diverse environments of Asia during the Upper Pleistocene. *Quaternary International* 300: 32-47.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America* 99 (Suppl 3): 7280-7287.

Bond, W., F. Woodward and G. Midgley, 2005. The global distribution of ecosystems in a world without fire. *New Phytologist* 165 (2): 525-538.

Borrill, P.L. and L. Tesfatsion, 2010. 11 Agent-based modeling: the right mathematics for the social sciences? In *The Elgar Companion to Recent Economic Methodology*, 228.

Bosch, M.D., M.A. Mannino, A.L. Prendergast, F.P. Wesselingh, T.C. O'Connell and J.-J. Hublin, 2017. Year-round shellfish exploitation in the Levant and implications for Upper Palaeolithic hunter-gatherer subsistence. *Journal of Archaeological Science: Reports*.

Boyd, R., P.J. Richerson and J. Henrich, 2011. Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology* 65 (3): 431-444.

Bradtmöller, M., A. Pastoors, B. Weninger and G.-C. Weniger, 2012. The repeated replacement model – Rapid climate change and population dynamics in Late Pleistocene Europe. *Quaternary International* 247: 38-49.

Brantingham, P.J., 2003. A neutral model of stone raw material procurement. *American Antiquity* 68 (3): 487-509.

Braun, D.R., J.W.K. Harris, N.E. Levin, J.T. McCoy, A.I.R. Herries, M.K. Bamford, L.C. Bishop, B.G. Richmond and M. Kibunjia, 2010. Early hominin diet included diverse terrestrial and aquatic animals 1.95 Ma in East Turkana, Kenya. *Proceedings of the National Academy of Sciences* 107 (22): 10002-10007.

Broadfield, D.C., R.L. Holloway, K. Mowbray, A. Silvers, M.S. Yuan and S. Márquez, 2001. Endocast of Sambungmacan 3 (Sm 3): a new *Homo erectus* from Indonesia. *The Anatomical Record* 262 (4): 369-379.

Broennimann, O., M.C. Fitzpatrick, P.B. Pearman, B. Petitpierre, L. Pellissier, N.G. Yoccoz, W. Thuiller, M.-J. Fortin, C. Randin, N.E. Zimmermann, C.H. Graham and A. Guisan, 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* 21 (4): 481-497.

Broodbank, C., 2006. The origins and early development of Mediterranean maritime activity. *Journal of Mediterranean Archaeology* 19 (2).

Broodbank, C., 2014. So... What? Does the Paradigm Currently Want to Budge So Much? *Journal of Mediterranean Archaeology* 27 (2): 267-272.

Brooks, R., 2001. The relationship between matter and life. *Nature* 409 (6818): 409.

Brughmans, T. and J. Poblome, 2017. The case for computational modelling of the Roman economy: a reply to Van Oyen. *Antiquity* 91 (359): 1364-1366.

Burnham, K.P. and D.R. Anderson, 2003. *Model selection and multimodel inference: a practical information-theoretic approach*: Springer Science & Business Media.

Bynoe, R., J.K. Dix and F. Sturt, 2016. Of mammoths and other monsters: historic approaches to the submerged Palaeolithic. *Antiquity* 90 (352): 857-875.

Calvez, B. and G. Hutzler, 2006. Automatic Tuning of Agent-Based Models Using Genetic Algorithms. In *Multi-Agent-Based Simulation VI: International Workshop, MABS 2005, Utrecht, The Netherlands, July 25, 2005, Revised and Invited Papers*, edited by J. S. Sichman and L. Antunes. Berlin, Heidelberg: Springer Berlin Heidelberg, 41-57.

Carrión, J.S., C. Finlayson, S. Fernández, G. Finlayson, E. Allué, J.A. López-Sáez, P. López-García, G. Gil-Romera, G. Bailey and P. González-Sampériz, 2008. A coastal reservoir of biodiversity for Upper Pleistocene human populations: palaeoecological investigations in Gorham's Cave (Gibraltar) in the context of the Iberian Peninsula. *Quaternary Science Reviews* 27 (23): 2118-2135.

Castellano, S., G. Parra, F.A. Sánchez-Quinto, F. Racimo, M. Kuhlwilm, M. Kircher, S. Sawyer, Q. Fu, A. Heinze, B. Nickel, J. Dabney, M. Siebauer, L. White, H.A. Burbano, G. Renaud, U. Stenzel, C. Lalueza-Fox, M. de la Rasilla, A. Rosas, P. Rudan, D. Brajković, Ž. Kucan, I. Gušic, M.V. Shunkov, A.P. Derevianko, B. Viola, M. Meyer, J. Kelso, A.M. Andrés and S. Pääbo, 2014. Patterns of coding variation in the complete exomes of three Neandertals. *Proceedings of the National Academy of Sciences*, 111(18), 6666-6671.

Castle, C.J. and A.T. Crooks, 2006. Principles and concepts of agent-based modelling for developing geospatial simulations. *Working paper 110*, Centre for Advanced Spatial Analysis, University College London.

Cegielski, W.H. and J.D. Rogers, 2016. Rethinking the role of Agent-Based Modeling in archaeology. *Journal of Anthropological Archaeology* 41: 283-298.

Chase, P.G., 1989. How different was Middle Palaeolithic subsistence? A zooarchaeological perspective on the Middle to Upper Palaeolithic transition. In *The Human Revolution: Behavioral and Biological Perspectives on the Origins of Modern Humans*, edited by P. Mellars and C. Stringer. New Jersey: Princeton University Press, 321-338.

Chernoff, H., 1973. The use of faces to represent points in k-dimensional space graphically. *Journal of the American statistical Association* 68 (342): 361-368.

Christiansen, J.H. and M.R. Altaweel, 2005. Understanding ancient societies: A new approach using agent-based holistic modeling. *Structure and Dynamics* 1 (2).

Churchill, S.E., 2005. Calorie count reveals Neandertals out-ate hardiest modern hunters. *Science* 307: 840.

Churchill, S.E., 2006. Bioenergetic Perspectives on Neanderthal Thermoregulatory and Activity Budgets. In *Neanderthals Revisited: New Approaches and Perspectives*, edited by K. Harvati and T. Harrison. Dordrecht: Springer, 113-134.

Churchill, S.E., 2009. Energetic competition between Neandertals and anatomically modern humans. *PaleoAnthropology* 96: 116.

Clarke, D., 1973. Archaeology: the loss of innocence. *Antiquity* 47 (185): 6.

Cliquet, D., N. Mercier, H. Valladas, L. Froget, D. Michel, B. van Vliet-Lanoë and G. Vilgrain, 2003. Apport de la thermoluminescence sur silex chauffés à la chronologie de sites paléolithiques de Normandie : nouvelles données et interprétations. *Quaternaire* 14 (1): 51 - 64.

Close, A.E., 2009. The Middle-Upper Palaeolithic hiatus of insular north Africa. In *The Mediterranean from 50 000 to 25 000 BP: Turning points and new directions*, edited by M. a. C. S. Camps. Oxford, UK: Oxbow Books, 35-50.

Codding, B.F. and T.L. Jones, 2013. Environmental productivity predicts migration, demographic, and linguistic patterns in prehistoric California. *Proceedings of the National Academy of Sciences* 110 (36): 14569-14573.

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

Coello Coello, C.A., 2002. Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. *Computer Methods in Applied Mechanics and Engineering* 191 (11–12): 1245-1287.

Cohen, K. and P. Gibbard, 2011. Global chronostratigraphical correlation table for the last 2.7 million years: Submission on Quaternary Stratigraphy (International Commission on Stratigraphy). *Cambridge, England: Electronic document available at http://quaternary.stratigraphy.org/charts/, verified 9 April 2019.*

Cohen, K.M., K. MacDonald, J.C.A. Joordens, W. Roebroeks and P.L. Gibbard, 2012. The earliest occupation of north-west Europe: a coastal perspective. *Quaternary International* 271: 70-83.

Cole, L.E.S., S.A. Bhagwat and K.J. Willis, 2014. Recovery and resilience of tropical forests after disturbance. *Nature Communications* 5: 3906.

Collier, N. and M. North, 2010. Repast Java getting started.

Collins, W.D., C.M. Bitz, M.L. Blackmon, G.B. Bonan, C.S. Bretherton, J.A. Carton, P. Chang, S.C. Doney, J.J. Hack, T.B. Henderson, J.T. Kiehl, W.G. Large, D.S. McKenna, B.D. Santer and R.D. Smith, 2006. The Community Climate System Model version 3 (CCSM3). *Journal of Climate* 19 (11): 2122-2143.

Consortium, P.k., 2013. Continental-scale temperature variability during the past two millennia. *Nature Geoscience* 6 (5): 339-346.

Cortés-Sánchez, M., A. Morales-Muñiz, M.D. Simón-Vallejo, M.C. Lozano-Francisco, J.L. Vera-Peláez, C. Finlayson, J. Rodríguez-Vidal, A. Delgado-Huertas, F.J. Jiménez-Espejo, F. Martínez-Ruiz, M.A. Martínez-Aguirre, A.J. Pascual-Granged, M.M. Bergadà-Zapata, J.F. Gibaja-Bao, J.A. Riquelme-Cantal, J.A. López-Sáez, M. Rodrigo-Gámiz, S. Sakai, S. Sugisaki, G. Finlayson, D.A. Fa and N.F. Bicho, 2011. Earliest Known Use of Marine Resources by Neanderthals. *PLoS ONE* 6 (9): e24026.

Crema, E.R., 2018. Statistical inference and archaeological simulations. *SAA Archaeological Record*: 20-23.

Crofoot, M.C. and I.C. Gilby, 2012. Cheating monkeys undermine group strength in enemy territory. *Proceedings of the National Academy of Sciences* 109 (2): 501-505.

Crooks, A., A. Croitoru, X. Lu, S. Wise, J.M. Irvine and A. Stefanidis, 2015. Walk this way: Improving pedestrian agent-based models through scene activity analysis. *ISPRS International Journal of Geo-Information* 4 (3): 1627-1656.

Crooks, A.T., 2007. The Repast Simulation/Modelling System for geospatial Simulation. In *Agent-Based Models for Spatial Systems in Social Sciences & Economic Science with Heterogeneous Interacting Agents (ABM–S4–ESHIA)* Agelonde, La Londe les Maures, France.

Cucart-Mora, C., S. Lozano and J. Fernández-López de Pablo, 2018. Bio-cultural interactions and demography during the Middle to Upper Palaeolithic transition in Iberia: An agent-based modelling approach. *Journal of Archaeological Science* 89: 14-24.

D'Abrera, H. and E. Lehmann, 1975. *Nonparametrics: statistical methods based on ranks*: Holden-day.

D'Errico, F., W.E. Banks, M. Vanhaeren, V. Laroulandie and M. Langlais, 2011. PACEA geo-referenced radiocarbon database. *PaleoAnthropology* 2011: 1-12.

Dafoe, A., 2014. Science deserves better: the imperative to share complete replication files. *PS: Political Science & Politics* 47 (1): 60-66.

Dalén, L., L. Orlando, B. Shapiro, M. Brandström-Durling, R. Quam, M.T.P. Gilbert, J.C.D. Fernández-Lomana, E. Willerslev, J.L. Arsuaga and A. Götherström, 2012. Partial genetic turnover in neandertals: continuity in the east and population replacement in the west. *Molecular biology and evolution* 29 (8): 1893-1897.

Damuth, J., 1991. Of size and abundance. *Nature* 351 (6324): 268-269.

Daura, J., M. Sanz, N. García, E. Allué, M. Vaquero, E. Fierro, J.S. Carrión, J.M. López-García, H.A. Blain, A. Sánchez-Marco, C. Valls, R.M. Albert, J.J. Fornós, R. Julià, J.M. Fullola and J. Zilhão, 2013. Terrasses de la Riera dels Canyars (Gavà, Barcelona): the landscape of Heinrich Stadial 4 north of the "Ebro frontier" and implications for modern human dispersal into Iberia. *Quaternary Science Reviews* 60: 26-48.

Davidson, I. and W. Noble, 1992. Why the first colonisation of the Australian region is the earliest evidence of modern human behaviour. *Archaeology in Oceania* 27(3): 135-142.

Davies, B., S.J. Holdaway and P.C. Fanning, 2016. Modelling the palimpsest: an exploratory agent-based model of surface archaeological deposit formation in a fluvial arid Australian landscape. *The Holocene* 26 (3): 450-463.

Davison, K., P. Dolukhanov, G.R. Sarson and A. Shukurov, 2006. The role of waterways in the spread of the Neolithic. *Journal of Archaeological Science* 33 (5): 641-652.

De Vynck, J.C., R. Anderson, C. Atwater, R.M. Cowling, E.C. Fisher, C.W. Marean, R.S. Walker, and K. Hill, 2016. Return rates from intertidal foraging from Blombos Cave to Pinnacle Point: Understanding early human economies. *Journal of Human Evolution* 92: 101-115.

Del Grosso, S., W. Parton, T. Stohlgren, D. Zheng, D. Bachelet, S. Prince, K. Hibbard and R. Olson, 2008. Global potential net primary production predicted from vegetation class, precipitation, and temperature. *Ecology* 89 (8): 2117-2126.

Dennell, R., 2003. Dispersal and colonisation, long and short chronologies: how continuous is the Early Pleistocene record for hominids outside East Africa? *Journal of Human Evolution* 45 (6): 421-440.

Dennell, R. and W. Roebroeks, 2005. An Asian perspective on early human dispersal from Africa. *Nature* 438 (7071): 1099-104.

Dennell, R.W., J. Louys, H.J. O'Regan and D.M. Wilkinson, 2014. The origins and persistence of Homo floresiensis on Flores: biogeographical and ecological perspectives. *Quaternary Science Reviews* 96: 98-107.

Dennell, R.W., M. Martinón-Torres and J.M. Bermúdez de Castro, 2011. Hominin variability, climatic instability and population demography in Middle Pleistocene Europe. *Quaternary Science Reviews* 30 (11-12): 1511-1524.

Derricourt, R., 2005. Getting "Out of Africa": Sea Crossings, Land Crossings and Culture in the Hominin Migrations. *Journal of World Prehistory* 19: 119-132.

Dershowitz, N. and Y. Gurevich, 2008. A natural axiomatization of computability and proof of Church's Thesis. *The Bulletin of Symbolic Logic*: 299-350.

Diem, K. and C. Lentner, 1962. Documenta Geigy: Scientific tables. edited by K. Diem and C. Lentner: Geigy Pharmaceuticals.

Discamps, E. and J.-P. Faivre, 2017. Substantial biases affecting Combe-Grenal faunal record cast doubts on previous models of Neanderthal subsistence and environmental context. *Journal of Archaeological Science* 81: 128-132.

Discamps, E., J. Jaubert and F. Bachellerie, 2011. Human choices and environmental constraints: deciphering the variability of large game procurement: from Mousterian to Aurignacian times (MIS 5-3) in southwestern France. *Quaternary Science Reviews* 30 (19-20): 2755-2775.

Dobson, F.S., 1982. Competition for mates and predominant juvenile male dispersal in mammals. *Animal behaviour* 30 (4): 1183-1192.

Domingo, R., J.L. Peña-Monné, T. de Torres, J.E. Ortiz and P. Utrilla, 2017. Neanderthal highlanders: Las Callejuelas (Monteagudo del Castillo, Teruel, Spain), a high-altitude site occupied during MIS 5. *Quaternary International* 435: 129-143.

Drogoul, A. and J. Ferber, 1994. Multi-agent simulation as a tool for modeling societies: Application to social differentiation in ant colonies. In *Artificial Social Systems*: Springer, 2-23.

Du Toit, J.T. and D.H.M. Cumming, 1999. Functional significance of ungulate diversity in African savannas and the ecological implications of the spread of pastoralism. *Biodiversity & Conservation* 8 (12): 1643-1661.

Dunbar, R.I.M., 1993. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* 16: 681-735.

Dusseldorp, G.L., 2009. *A View to a Kill: Investigating Middle Palaeolithic subsistence using an Optimal Foraging perspective.* Published Manuscript, Human Origins, Leiden University, Leiden.

Dynesius, M. and R. Jansson, 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences* 97 (16): 9115-9120.

Edmonds, B. and J.J. Bryson, 2004. The insufficiency of formal design methods the necessity of an experimental approach-for the understanding and control of complex mas. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-* Volume 2.

Eisenberg, J. and G. McKay, 1974. Comparison of ungulate adaptations in the New World and Old World tropical forests with special reference to Ceylon and the rainforests of Central America. *The behavior of ungulates and its relation to management (V. Geist and F. Walther, eds.). International Union for Conservation of Nature, Morges, Switzerland*: 585-602.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberon, S. Williams, M.S. Wisz and N.E. Zimmermann, 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29 (2): 129-151.

Ellis, E.C. and N. Ramankutty, 2008. Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment* 6 (8): 439-447.

Epstein, J.M., 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. STU - Student edition ed: Princeton University Press.

Eriksson, A., L. Betti, A.D. Friend, S.J. Lycett, J.S. Singarayer, N. von Cramon-Taubadel, P.J. Valdes, F. Balloux and A. Manica, 2012. Late Pleistocene climate change and the global expansion of anatomically modern humans. *Proceedings of the National Academy of Sciences* 109 (40): 16089-16094.

Fa, D.A., J.C. Finlayson, G. Finlayson, F. Giles-Pacheco, J. Rodríguez-Vidal and J.M. Gutiérrez-López, 2016. Marine mollusc exploitation as evidenced by the Gorham's Cave (Gibraltar) excavations 1998–2005: The Middle–Upper Palaeolithic transition. *Quaternary International* 407 (Part B): 16-28.

Fagan, M.E., 1976. Design and Code Inspections to Reduce Errors in Program Development. *IBM Systems Journal* 15 (3): 182-211.

Fairbanks, R.G., 1989. A 17, 000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* 342 (6250): 637-642.

Fan, Y., H. Li and G. Miguez-Macho, 2013. Global Patterns of Groundwater Table Depth. *Science* 339 (6122): 940-943.

Féblot-Augustins, J., 1999. Raw material transport patterns and settlement systems in the European Lower and Middle Palaeolithic: continuity, change and variability. In *The Middle Palaeolithic Occupation of Europe*, edited by W. Roebroeks and C. S. Gamble. Leiden: University of Leiden, 193-214.

Féblot-Augustins, J., 2009. Revisiting European Upper Paleolithic raw material transfers: the demise of the cultural ecological paradigm? In *Lithic Materials and Paleolithic Societies*, edited by B. Adams and B. S. Blades: Blackwell Publishing, 25-46.

Feddema, J.J., K.W. Oleson, G.B. Bonan, L.O. Mearns, L.E. Buja, G.A. Meehl and W.M. Washington, 2005. The importance of land-cover change in simulating future climates. *Science* 310 (5754): 1674-1678.

Fernandes, P., J.-P. Raynal and M.-H. Moncel, 2008. Middle Palaeolithic raw material gathering territories and human mobility in the southern Massif Central, France: first results from a petro-archaeological study on flint. *Journal of Archaeological Science* 35 (8): 2357-2370.

Ferràndez-Cañadell, C., F. Ribot and L. Gibert, 2014. New fossil teeth of Theropithecus oswaldi (Cercopithecoidea) from the Early Pleistocene at Cueva Victoria (SE Spain). *Journal of Human Evolution* 74: 55-66.

Fick, S.E. and R.J. Hijmans, 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37 (12): 4302-4315.

Finlayson, C., 2009. *The humans who went extinct: why Neanderthals died out and we survived*: Oxford University Press.

Finlayson, C. and F. Giles Pacheco, 2000. The southern Iberian Peninsula in the late Pleistocene: geography, ecology, and human occupation. *Neanderthals on the edge*: 139-154.

Finlayson, C., F. Giles Pacheco, J. Rodríguez-Vidal, D.A. Fa, J.M. Gutierrez López, A. Santiago Pérez, G. Finlayson, E. Allue, J. Baena Preysler, I. Cáceres, J.S. Carrión, Y. Fernández Jalvo, C.P. Gleed-Owen, F.J. Jimenez Espejo, P. López, J. Antonio López Sáez, J. Antonio Riquelme Cantal, A. Sánchez Marco, F. Giles Guzman, K. Brown, N. Fuentes, C.A. Valarino, A. Villalpando, C.B. Stringer, F. Martinez Ruiz and T. Sakamoto, 2006. Late survival of Neanderthals at the southernmost extreme of Europe. *Nature* 443: 850-853.

Fiore, I., M. Gala, M. Romandini, E. Cocca, A. Tagliacozzo and M. Peresani, 2016. From feathers to food: Reconstructing the complete exploitation of avifaunal resources by Neanderthals at Fumane cave, unit A9. *Quaternary International* 421 (9): 134-153.

Fisher, R.A., 1937. The wave of advance of advantageous genes. *Annals of Eugenics* 7 (4): 355-369.

Fletcher, M. and G.R. Lock, 2005. *Digging Numbers: Elementary Statistics for Archaeologists*. Second ed. Vol. 33, *Oxford University School of Archaeology Monograph*: Oxford University Committee for Archaeology.

Flores, J., 1997. A mathematical model for Neanderthal extinction. *arXiv preprint physics/9712024*.

Flores, J., 1998. A mathematical model for Neanderthal extinction. *Journal of theoretical Biology* 191 (3): 295-298.

Flores, J., 2011. Diffusion coefficient of modern humans outcompeting Neanderthals. *Journal of theoretical biology* 280 (1): 189-190.

Ford, D.C. and P. Williams, 2007. *Karst hydrogeology and geomorphology*: John Wiley & Sons.

Forrest, S., 1993. Genetic algorithms: principles of natural selection applied to computation. *Science* 261 (5123): 872.

Forseth, I.H., 2012. Terrestrial Biomes. *Nature Education Knowledge*.

Forster, M.R., 2000. Key Concepts in Model Selection: Performance and Generalizability. *Journal of Mathematical Psychology* 44 (1): 205-231.

Frakes, S., 2003. NGS Toolkit, Part 4: General Geodetic Computations. *Professional Surveyor Magazine* 23 (9).

Freedman, L.P., I.M. Cockburn and T.S. Simcoe, 2015. The Economics of Reproducibility in Preclinical Research. *PLOS Biology* 13 (6): e1002165.

Freeman, J. and J.M. Anderies, 2015. The socioecology of hunter–gatherer territory size. *Journal of Anthropological Archaeology* 39: 110-123.

Froehle, A.W. and S.E. Churchill, 2009. Energetic competition between Neandertals and Anatomically Modern Humans. *PaleoAnthropology* 2009: 96-116.

Gamble, C., 1993. *Timewalkers:The Prehistory of Global Colonization*. Cambridge: Harvard University Press.

Gamble, C.S. and J. Steele, 1999. Hominid ranging patterns and dietary strategies. In *Hominid evolution. Lifestyles and survival strategies*, edited by H. Ullrich. Berlin: Edition Archaea, 396-409.

Gaona, P., P. Ferreras and M. Delibes, 1998. Dynamics and viability of a metapopulation of the endangered Iberian lynx (Lynx pardinus). *Ecological monographs* 68 (3): 349-370.

Gibbard, P.L., 1988. The history of the great northwest European rivers during the past three million years. *Philosophical Transactions of the Royal Society of London* Series 318: 559-602.

Gilbert, N., 2008. *Agent-based models*, *Quantitative Applications in the Social Sciences*. Thousand Oaks, California (USA): Sage Publications.

Gilbert, N. and P. Ahrweiler, 2009. The Epistemologies of Social Simulation Research. In *Epistemological Aspects of Computer Simulation in the Social Sciences: Second International Workshop, EPOS 2006, Brescia, Italy, October 5-6, 2006, Revised Selected and Invited Papers*, edited by F. Squazzoni. Berlin, Heidelberg: Springer Berlin Heidelberg, 12-28.

Gilbert, N. and S. Bankes, 2002. Platforms and methods for agent-based modeling. *Proceedings of the National Academy of Sciences of the United States of America* 99 (Suppl 3): 7197-7198.

Gilmour, M., A. Currant, R. Jacobi and C. Stringer, 2007. Recent TIMS dating results from British Late Pleistocene vertebrate faunal localities: context and interpretation. *Journal of Quaternary Science* 22 (8): 793-800.

Gilpin, W., M.W. Feldman and K. Aoki, 2016. An ecocultural model predicts Neanderthal extinction through competition with modern humans. *Proceedings of the National Academy of Sciences*.

Goldberg, D.E., 1989. *Genetic algorithms in search, optimization, and machine learning*. Vol. 1989: Addison-wesley Publishing Company, inc.

Gosling, J., D.C. Holmes and K. Arnold, 2005. *The Java programming language*: Addison Wesley Professional.

Gosling, J. and H. McGilton, 1996. The java language environment, a white paper. *Sun Developer Network White Paper.* .

Goval, E., 2008. *Définitions, analyses et caractérisations des territoires des Néandertaliens au Weichsélien ancien en France septentrionale (Approches technologiques et spatiales, élargissement au Nord-Ouest de l'Europe)*, Préhistoire, Université des sciences et technologies de Lille, Lille.

Grant, K., E. Rohling, M. Bar-Matthews, A. Ayalon, M. Medina-Elizalde, C.B. Ramsey, C. Satow and A. Roberts, 2012. Rapid coupling between ice volume and polar temperature over the past 150,000 [thinsp] years. *Nature* 491 (7426): 744-747.

Green, R.E., J. Krause, A.W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M.H.-Y. Fritz, N.F. Hansen, E.Y. Durand, A.-S. Malaspinas, J.D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H.A. Burbano, J.M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E.S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Ž. Kucan, I. Gušic, V.B. Doronichev, L.V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R.W. Schmitz, P.L.F. Johnson, E.E. Eichler, D. Falush, E. Birney, J.C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich and S. Pääbo, 2010. A Draft Sequence of the Neandertal Genome. *Science* 328 (5979): 710-722.

Grimm, V., 1994. Mathematical models and understanding in ecology. *Ecological modelling* 75: 641-651.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S.K. Heinz and G. Huse, 2006. A standard protocol for describing individual-based and agent-based models. *Ecological modelling* 198 (1): 115-126.

Grimm, V., K. Frank, F. Jeltsch, R. Brandl, J. Uchmański and C. Wissel, 1996. Pattern-oriented modelling in population ecology. *Science of the Total Environment* 183 (1-2): 151-166.

Grimm, V. and S.F. Railsback, 2005. Individual-based Modeling and Ecology:(Princeton Series in Theoretical and Computational Biology).

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.

Grøn, O., 2018. Some problems with modelling the positions of prehistoric hunter-gatherer settlements on the basis of landscape topography. *Journal of Archaeological Science: Reports* 20: 192-199.

Groß, D., A. Zander, A. Boethius, S. Dreibrodt, O. Grøn, A. Hansson, C. Jessen, S. Koivisto, L. Larsson, H. Lübke and B. Nilsson, 2018. People, lakes and seashores: Studies from the Baltic Sea basin and adjacent areas in the early and Mid-Holocene. *Quaternary Science Reviews* 185: 27-40.

Groucutt, H.S., M.D. Petraglia, G. Bailey, E.M. Scerri, A. Parton, L. Clark-Balzan, R.P. Jennings, L. Lewis, J. Blinkhorn and N.A. Drake, 2015. Rethinking the dispersal of Homo sapiens out of Africa. *Evolutionary Anthropology: Issues, News, and Reviews* 24 (4): 149-164.

Guha, A.H. and G.J. Assaf, 2018. Representing Collected Road Condition Data with Chernoff Faces for Evaluation of Pavement Conditions. Civil Infrastructures Confronting Severe Weathers and Climate Changes Conference.

Guibert, P., F. Bechtel, L. Bourguignon, M. Brenet, I. Couchoud, A. Delagnes, F. Delpech, L. Detrain, M. Duttine, M. Folgado, J. Jaubert, C. Lahaye, M. Lenoir, B. Maureille, J.-P. Texier, A. Turq, E. Vieillevigne and G. Villeneuve, 2008. Une base de données pour la chronologie du Paléolithique moyen dans le Sud-Ouest de la France. In *Les sociétés du Paléolithique dans un Grand Sud-ouest de la France: nouveaux gisements, nouveaux résultats, nouvelles méthodes*, edited by J. Jaubert, J.-G. Bordes and I. Ortega: CNRS, 19-40.

Guiot, J., A. Pons, J.L. de Beaulieu and M. Reille, 1989. A 140,000-year continental climate reconstruction from two European pollen records. *Nature* 338: 309-313.

Guisan, A. and N.E. Zimmermann, 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135 (2): 147-186.

Gunz, P., S. Neubauer, B. Maureille and J.-J. Hublin, 2010. Brain development after birth differs between Neanderthals and modern humans. *Current Biology* 20 (21): R921-R922.

Gupta, S., J.S. Collier, A. Palmer-Felgate and G. Potter, 2007. Catastrophic flooding origin of shelf valley systems in the English Channel. *Nature* 448 (7151): 342-5.

Gurven, M. and H. Kaplan, 2007. Hunter-gatherer longevity: cross-cultural perspectives. *Population and development review* 33: 321-365.

Guthrie, R.D., 1982. Mammals of the mammoth steppe as palaeoenvironmental indicators. In *Palaeoecology of Beringia*, edited by D. M. Hopkins, J. V. Matthews, C. E. Schweger and S. B. Young. New York: Academic Press, 307-26.

Guthrie, R.D., 1990. *Frozen Fauna of the Mammoth Steppe: the story of Blue Babe*. Chicago: Chicago University Press.

Hajdinjak, M., Q. Fu, A. Hübner, M. Petr, F. Mafessoni, S. Grote, P. Skoglund, V. Narasimham, H. Rougier, I. Crevecoeur, P. Semal, M. Soressi, S. Talamo, J.-J. Hublin, I. Gušić, Ž. Kućan, P. Rudan, L.V. Golovanova, V.B. Doronichev, C. Posth, J. Krause, P. Korlević, S. Nagel, B. Nickel, M. Slatkin, N. Patterson, D. Reich, K. Prüfer, M. Meyer, S. Pääbo and J. Kelso, 2018. Reconstructing the genetic history of late Neanderthals. *Nature* 555: 652.

Hamilton, K., 2007. Analyses of IPCC Climate Model Data at IPRC. *IPRC Climate.* Newsletter. Internation Pacific Research Center.

Hamilton, M.J., B.T. Milne, R.S. Walker and J.H. Brown, 2007. Nonlinear scaling of space use in human hunter-gatherers. *Proceedings of the National Academy of Sciences* 104 (11): 4765-4769.

Hansson, L., L. Fahrig and G. Merriam, 1995. *Mosaic landscapes and ecological processes*. The Netherlands: Springer Science & Business Media.

Hardy, K., J. Brand-Miller, K.D. Brown, M.G. Thomas and L. Copeland, 2015. The Importance of Dietary Carbohydrate in Human Evolution. *The Quarterly Review of Biology* 90 (3): 251-268.

Harrison, S.P. and I.C. Prentice, 2003. Climate and $CO_2$ controls on global vegetation distribution at the last glacial maximum: analysis based on palaeovegetation data, biome modelling and palaeoclimate simulations. *Global Change Biology* 9 (7): 983-1004.

Hartmann, S., 1996. The World as a Process. In *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, edited by R. Hegselmann, U. Mueller and K. Troitzsch: Springer Netherlands, 77-100.

Hasumi, H. and S. Emori, 2004. *K-1 coupled GCM (MIROC) description*. Center for Climate System Research, University of Tokyo, Tokyo.

Hawe, G.I., G. Coates, D.T. Wilson and R.S. Crouch, 2012. Agent-based simulation for large-scale emergency response: A survey of usage and implementation. *ACM Computing Surveys (CSUR)* 45 (1): 8.

Hawkins, B.A., R. Field, H.V. Cornell, D.J. Currie, J.-F. Guégan, D.M. Kaufman, J.T. Kerr, G.G. Mittelbach, T. Oberdorff, E.M. O'Brien, E.E. Porter and J.R.G. Turner, 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84 (12): 3105-3117.

Haxeltine, A. and I.C. Prentice, 1996. BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. *Global Biogeochemical Cycles* 10 (4): 693-709.

Hays, J.D., J. Imbrie and N.J. Shackleton, 1976. Variations in the Earth's Orbit: Pacemaker of the Ice Ages. *Science* 194 (4270): 1121-1132.

Henry, A.G., A.S. Brooks and D.R. Piperno, 2011. Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proceedings of the National Academy of Sciences* 108 (2): 486-491.

Herrera, F., M. Lozano and J.L. Verdegay, 1998. Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis. *Artificial Intelligence Review* 12 (4): 265-319.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247-276.

Hewlett, B.S., 1991. Demography and childcare in preindustrial societies. *Journal of Anthropological Research* 47: 1-47.

Heyes, P. and K. MacDonald, 2015. Neandertal energetics: Uncertainty in body mass estimation limits comparisons with Homo sapiens. *Journal of Human Evolution* 85: 193-197.

Higham, T., T. Compton, C. Stringer, R. Jacobi, B. Shapiro, E. Trinkaus, B. Chandler, F. Groning, C. Collins, S. Hillson, P. O'Higgins, C. FitzGerald and M. Fagan, 2011. The earliest evidence for anatomically modern humans in northwestern Europe. *Nature* 479(7374), 521.

Higham, T., K. Douka, R. Wood, C.B. Ramsey, F. Brock, L. Basell, M. Camps, A. Arrizabalaga, J. Baena, C. Barroso-Ruiz, C. Bergman, C. Boitard, P. Boscato, M. Caparros, N.J. Conard, C. Draily, A. Froment, B. Galvan, P. Gambassini, A. Garcia-Moreno, S. Grimaldi, P. Haesaerts, B. Holt, M.-J. Iriarte-Chiapusso, A. Jelinek, J.F. Jorda Pardo, J.-M. Maillo-Fernandez, A. Marom, J. Maroto, M. Menendez, L. Metz, E. Morin, A. Moroni, F. Negrino, E. Panagopoulou, M. Peresani, S. Pirson, M. de la Rasilla, J. Riel-Salvatore, A. Ronchitelli, D. Santamaria, P. Semal, L. Slimak, J. Soler, N. Soler, A. Villaluenga, R. Pinhasi and R. Jacobi, 2014. The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* 512 (7514): 306-309.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25 (15): 1965-1978.

Hill, K., A.M. Hurtado and R.S. Walker, 2007. High adult mortality among Hiwi hunter-gatherers: Implications for human evolution. *Journal of Human Evolution* 52: 443-454.

Hill, R.A. and R.I.M. Dunbar, 2003. Social Network Size in Humans. *Human Nature* 14(1): 53-72.

Hobbs, N.T., 1996. Modification of Ecosystems by Ungulates. *The Journal of Wildlife Management* 60 (4): 695-713.

Hodgkins, J., C.W. Marean, A. Turq, D. Sandgathe, S.J.P. McPherron and H. Dibble, 2016. Climate-mediated shifts in Neandertal subsistence behaviors at Pech de l'Azé IV and Roc de Marsal (Dordogne Valley, France). *Journal of Human Evolution* 96: 1-18.

Hofer, H. and M.L. East, 1993. The commuting system of Serengeti spotted hyaenas: how a predator copes with migratory prey. II. Intrusion pressure and commuters' space use. *Animal Behaviour* 46 (3): 559-574.

Hoffecker, J.F. and N. Cleghorn, 2000. Mousterian hunting patterns in the northwestern Caucasus and the ecology of the Neanderthals. *International Journal of Osteoarchaeology* 10 (5): 368-378.

Hoffmann, D.L., C.D. Standish, M. García-Diez, P.B. Pettitt, J.A. Milton, J. Zilhão, J.J. Alcolea-González, P. Cantalejo-Duarte, H. Collado, R. de Balbín, M. Lorblanchet, J. Ramos-Muñoz, G.-C. Weniger and A.W.G. Pike, 2018. U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science* 359 (6378): 912-915.

Holland, D. and N. Quinn, 1987. *Cultural models in language and thought*: Cambridge University Press.

Holland, J.H., 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*: The University of Michigan Press, Ann Arbor, MI.

Hollander, M. and D.A. Wolfe, 1973. *Nonparametric statistical methods*, *Wiley Series in Probability and Statistics*. New York: John Wiley & Sons, inc.

Holliday, T.W. and A.B. Falsetti, 1995. Lower limb length of European early modern humans in relation to mobility and climate. *Journal of Human Evolution* 29 (2): 141-153.

Holmes, K.M., 2007. Using Pliocene palaeoclimatic data to postulate dispersal pathways of early hominins. *Palaeogeography, Palaeoclimatology, Palaeoecology* 248: 96-108.

Hoogakker, B.A., R.S. Smith, J.S. Singarayer, R. Marchant, I.C. Prentice, J.R. Allen, R.S. Anderson, S.A. Bhagwat, H. Behling and O. Borisova, 2016. Terrestrial biosphere changes over the last 120 kyr. *Climate of the Past* 12 (1): 51-73.

Horan, R.D., E. Bulte and J.F. Shogren, 2005. How trade saved humanity from biological exclusion: an economic theory of Neanderthal extinction. *Journal of Economic Behavior & Organization* 58 (1): 1-29.

Hublin, J.-J., 2000. Modern-Nonmodern Hominid Interactions: A Mediterranean Perspective. In *The Geography of Neandertals and Modern Humans in Europe and the Greater Mediterranean*, edited by O. Bar-Yosef and D. Pilbeam. Cambridge: Peabody Museum of Archaeology and Ethnology, Harvard University, 157-182.

Hublin, J.-J., 2009. The Origin of Neandertals. *Proceedings of the National Academy of Sciences* 106 (38): 16022-16027.

Hublin, J.-J., 2015. The modern human colonization of western Eurasia: when and where? *Quaternary Science Reviews* 118: 194-210.

Hublin, J.-J., 2017. The last Neanderthal. *Proceedings of the National Academy of Sciences*.

Hublin, J.-J., A. Ben-Ncer, S.E. Bailey, S.E. Freidline, S. Neubauer, M.M. Skinner, I. Bergmann, A. Le Cabec, S. Benazzi, K. Harvati and P. Gunz, 2017. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature* 546: 289.

Hublin, J.-J. and W. Roebroeks, 2009. Ebb and flow or regional extinctions? On the character of Neandertal occupation of northern environments. *Comptes Rendus Palevol* 8 (5): 503-509.

Hublin, J.-J., S. Talamo, M. Julien, F. David, N. Connet, P. Bodu, B. Vandermeersch and M.P. Richards, 2012. Radiocarbon dates from the Grotte du Renne and Saint-Césaire support a Neandertal origin for the Châtelperronian. *Proceedings of the National Academy of Sciences* 109: 18743–18748.

Huntley, B. and T. Webb III, 1989. Migration: species' response to climatic variations caused by changes in the earth's orbit. *Journal of Biogeography*: 5-19.

IBM Corp., 2015. IBM SPSS Statistics for Windows. *Version 23.0.0.3* 64-bit.

Iman, R.L. and W. Conover, 1980. Small sample sensitivity analysis techniques for computer models. with an application to risk assessment. *Communications in statistics-theory and methods* 9 (17): 1749-1842.

Imbrie, J., J.D. Hays, D.G. Martinson, A. McIntyre, A.C. Mix, J.J. Morley, N.G. Pisias, W.L. Prell and N.J. Shackleton, 1984. The orbital theory of Pleistocene climate: Support from a revised chronology of the marine delta $^{18}$O record. In *Milankovitch and climate: Understanding the response to astronomical forcing*, edited by A. I. Berget et al., D. Reidel Publishing company.

Isaac, G.L., 1968. Traces of Pleistocene Hunters: An East African Example. In *Man the Hunter*, edited by R. B. Lee and I. DeVore: Aldine Publishing Company, Chicago, 253-361.

Isarin, R.F. and S.J. Bohncke, 1999. Mean July temperatures during the Younger Dryas in northwestern and central Europe as inferred from climate indicator plant species. *Quaternary Research* 51 (2): 158-173.

Ives, A.R. and S.R. Carpenter, 2007. Stability and Diversity of Ecosystems. *Science* 317 (5834): 58.

Janis, C.M., J. Damuth and J.M. Theodor, 2000. Miocene ungulates and terrestrial primary productivity: Where have all the browsers gone? *Proceedings of the National Academy of Sciences* 97 (14): 7899-7904.

Janssen, M.A., 2009. Understanding artificial anasazi. *Journal of Artificial Societies and Social Simulation* 12 (4): 13.

Janssen, M.A., L.N.i. 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.

Jennings, R., C. Finlayson, D. Fa and G. Finlayson, 2011. Southern Iberia as a refuge for the last Neanderthal populations. *Journal of Biogeography* 38 (10): 1873-1885.

Jiménez-Valverde, A. and J.M. Lobo, 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica* 31 (3): 361-369.

Johnson, J.W., 2000. A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression. *Multivariate Behavioral Research* 35 (1): 1-19.

Johnston, A.S.A., M.E. Hodson, P. Thorbek, T. Alvarez and R.M. Sibly, 2014. An energy budget agent-based model of earthworm populations and its application to study the effects of pesticides. *Ecological Modelling* 280: 5-17.

Joordens, J.C.A., F. d'Errico, F.P. Wesselingh, S. Munro, J. de Vos, J. Wallinga, C. Ankjaergaard, T. Reimann, J.R. Wijbrans, K.F. Kuiper, H.J. Mucher, H. Coqueugniot, V. Prie, I. Joosten, B. van Os, A.S. Schulp, M. Panuel, V. van der Haas, W. Lustenhouwer, J.J.G. Reijmer and W. Roebroeks, 2015. Homo erectus at Trinil on Java used shells for tool production and engraving. *Nature* 518 (7538): 228-231.

Joussaume, S. and K.E. Taylor, 1995. Status of the Paleoclimate Modeling Intercomparison Project (PMIP). Proceedings of the first international AMIP scientific conference. WCRP Report.

Kageyama, M., P. Braconnot, L. Bopp, V. Mariotti, T. Roy, M.-N. Woillez, A. Caubel, M.-A. Foujols, E. Guilyardi, M. Khodri, J. Lloyd, F. Lombard and O. Marti, 2013. Mid-Holocene and last glacial maximum climate simulations with the IPSL model: part II: model-data comparisons. *Climate Dynamics* 40 (9-10): 2469-2495.

Kaplan, H.S., K. Hill, J. Lancaster and A.M. Hurtado, 2000. A theory of human life history evolution: diet, intelligence and longevity. *Evolutionary Anthropology* 9 (4): 156-185.

Kay, D. and M. van Harmelen, 2012. Analytics for the whole institution; balancing strategy and tactics. *CETIS Analytics Series* 1 (2): 1067598-1.

Keeley, L.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.

Kelly, R.L., 1995. *The Foraging Spectrum. Diversity in Hunter-Gatherer Lifeways*. Washington, D.C.: Smithsonian Institution Press.

Kelly, R.L., 2003. Colonization of new land by hunter-gatherers. Expectations and implications based on ethnographic data. In *Colonization of Unfamiliar Landscapes: The archaeology of adaptation.*, edited by M. Rockman and J. Steele. London/New York: Routledge, 44-58.

Kemerer, C.F. and M.C. Paulk, 2009. The impact of design and code reviews on software quality: An empirical study based on psp data. *IEEE transactions on software engineering* 35 (4): 534-550.

Klein, R.G., 2008. Out of Africa and the Evolution of Human Behavior. *Evolutionary Anthropology* 17: 267-281.

Klein, R.G., 2009a. Darwin and the recent African origin of modern humans. *Proceedings of the National Academy of Sciences* 106 (38): 16007-16009.

Klein, R.G., 2009b. *The Human Career: Human Biological and Cultural Evolution*. Third Edition : University of Chicago Press, Chicago.

Klein, R.G. and T.E. Steele, 2013. Archaeological shellfish size and later human evolution in Africa *PNAS* 110 (27): 10910-10915.

Kohler, T.A., 2012. Complex systems and archaeology. *Archaeological theory today*: 93-123.

Kohler, T.A. and S.E. Van der Leeuw, 2007. Introduction. Historical Socionatural Systems and Models. In *The model-based archaeology of socionatural systems*, edited by T. A. Kohler and S. E. Van der Leeuw. Santa Fe, New Mexico: School for Advanced Research Press, 1-12.

Kolodny, O. and M.W. Feldman, 2017. A parsimonious neutral model suggests Neanderthal replacement was determined by migration and random species drift. *Nature Communications* 8 (1): 1040.

Kondo, Y., K. Sano, T. Omori, A. Abe-Ouchi, W.-L. Chan, S. Kadowaki, M. Naganuma, R. O'ishi, T. Oguchi, Y. Nishiaki and M. Yoneda, 2018. Ecological Niche and Least-Cost Path Analyses to Estimate Optimal Migration Routes of Initial Upper Palaeolithic Populations to Eurasia. In *The Middle and Upper Paleolithic Archeology of the Levant and Beyond*, edited by Y. Nishiaki and T. Akazawa. Singapore: Springer Singapore, 199-212.

Konigsberg, L.W. and S.R. Frankenberg, 2013. Demography. In *Research Methods in Human Skeletal Biology*, edited by E. A. diGangi and M. K. Moore. New York City (USA): ElsevierAcademic Press, 293-322.

Kowarik, K., H. Reschreiter and G.P.X. Wurzer, 2012. Modelling Prehistoric Mining. *IFAC Proceedings Volumes* 45 (2): 17-29.

Kremer, M., 1993. Population growth and technological change: one million BC to 1990. *The Quarterly Journal of Economics* 108 (3): 681-716.

Krohne, D., 2001. *General ecology*, 2nd edn. Pacific Grove, CA: Brooks/Cole.

Kuhn, S.L., D.A. Raichlen and A.E. Clark, 2016. What moves us? How mobility and movement are at the center of human evolution. *Evolutionary Anthropology: Issues, News, and Reviews* 25 (3): 86-97.

Kuhn, S.L. and M.C. Stiner, 2006. What's a Mother to Do? The Division of Labor among Neandertals and Modern Humans in Eurasia. *Current Anthropology* 47 (6): 953-980.

Kyriacou, K., J.E. Parkington, A.D. Marais and D.R. Braun, 2014. Nutrition, modernity and the archaeological record: Coastal resources and nutrition among Middle Stone Age hunter-gatherers on the western Cape coast of South Africa. *Journal of Human Evolution* 77 (Supplement C): 64-73.

Lake, M.W., 2000. MAGICAL computer simulation of Mesolithic foraging. *Dynamics in Human and Primate Societies: agent-based modelling of social and spatial processes*: 107-143.

Lake, M.W., 2014. Trends in Archaeological Simulation. *Journal of Archaeological Method and Theory* 21 (2): 258-287.

Lalueza-Fox, C., A. Rosas, A. Estalrrich, E. Gigli, P.F. Campos, A. García-Tabernero, S. García-Vargas, F. Sánchez-Quinto, O. Ramírez, S. Civit, M. Bastir, R. Huguet, D. Santamaría, M.T.P. Gilbert, E. Willerslev and M. de la Rasilla, 2011. Genetic evidence for patrilocal mating behavior among Neandertal groups. *Proceedings of the National Academy of Sciences* 108 (1): 250-253.

Lambeck, K. and J. Chappell, 2001. Sea level change through the last glacial cycle. *Science* 292: 679-685.

Law, A.M. and M.G. McComas, 1991. Secrets of successful simulation studies. *Proceedings of the 1991 Winter Conference on Simulation* (23rd).

Lawing, A.M. and P.D. Polly, 2011. Pleistocene Climate, Phylogeny, and Climate Envelope Models: An Integrative Approach to Better Understand Species' Response to Climate Change. *PLoS ONE* 6 (12): e28554.

Layton, R. and S. O'Hara, 2010. Human Social Evolution: A Comparison of Hunter-gatherer and Chimpanzee Social Organization. Proceedings of the British Academy.

Lee, J.-S., T. Filatova, A. Ligmann-Zielinska, B. Hassani-Mahmooei, F. Stonedahl, I. Lorscheid, A. Voinov, J.G. Polhill, Z. Sun and D.C. Parker, 2015. The Complexities of Agent-Based Modeling Output Analysis. *Journal of Artificial Societies and Social Simulation* 18 (4).

Lee, R.B., 1968. What hunters do for a living, or, how to make out on scarce resources. In *Man the Hunter*, edited by R. B. Lee and I. DeVore. New York: Aldine Publishing Company, 30-48.

Lee, R.B. and I. DeVore, 1968. Preface. In *Man the Hunter*, edited by R. B. Lee and I. DeVore. New York: Aldine Publishing Company, vii-ix.

Lehmann, J., A.H. Korstjens and R.I.M. Dunbar, 2007. Fission–fusion social systems as a strategy for coping with ecological constraints: a primate case. *Evolutionary Ecology* 21 (5): 613-634.

Leppard, T.P., 2015. Passive Dispersal versus Strategic Dispersal in Island Colonization by Hominins. *Current Anthropology* 56 (4): 7.

Leppard, T.P. and C. Runnels, 2017. Maritime hominin dispersals in the Pleistocene: advancing the debate. *antiquity* 91 (356): 510-519.

Levins, R., 1966. The strategy of model building in population biology. *American Scientist* 54 (4): 421-431.

Lewis, S.G., N.M. Ashton and R. Jacobi, 2011. Testing Human Presence During the Last Interglacial (MIS 5e): A Review of the British Evidence. In *The Ancient Human Occupation of Britain*, edited by N. M. Ashton, S. G. Lewis and C. B. Stringer. Amsterdam: Elsevier, 125-164.

Li, J., Y. Cui, J. Liu, W. Shi and Y. Qin, 2013. Estimation and analysis of net primary productivity by integrating MODIS remote sensing data with a light use efficiency model. *Ecological Modelling* 252 (0): 3-10.

Lieberman, D.E., 2002. Neandertal and Early Modern Human Mobility Patterns. In *Neandertals and Modern Humans in Western Asia*, edited by T. Akazawa, K. Aoki and O. Bar-Yosef: Springer US, 263-275.

Liefvendahl, M. and R. Stocki, 2006. A study on algorithms for optimization of Latin hypercubes. *Journal of Statistical Planning and Inference* 136 (9): 3231-3247.

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

Ligmann-Zielinska, A., D.B. Kramer, K. Spence Cheruvelil and P.A. Soranno, 2014. Using Uncertainty and Sensitivity Analyses in Socioecological Agent-Based Models to Improve Their Analytical Performance and Policy Relevance. *PLoS ONE* 9 (10): e109779.

Linge, N. and A. Sutton, 2015. *30 Years of Mobile Phones in the UK*: Amberley Publishing Limited.

Lisiecki, L.E. and M.E. Raymo, 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic d18O records. *Paleoceanography* 20: PA1003.

Locht, J.-L., É. Goval and P. Antoine, 2010. Reconstructing Middle Palaeolithic hominid behaviour during OIS 5 in northern France. In *Settlement Dynamics of the Middle Paleolithic and Middle Stone Age: Volume III*, edited by N. J. Conard and A. Delagnes: Kerns Verlag, Tübingen, 329-356.

López-García, J.M., H.-A. Blain, I. Lozano-Fernández, E. Luzi and A. Folie, 2017. Environmental and climatic reconstruction of MIS 3 in northwestern Europe using the small-mammal assemblage from Caverne Marie-Jeanne (Hastière-Lavaux, Belgium). *Palaeogeography, Palaeoclimatology, Palaeoecology* 485: 622-631.

Lowe, J.J. and M.J. Walker, 2014. *Reconstructing quaternary environments*: Routledge.

Lyman, R.L., 1979. Available meat from faunal remains: A consideration of techniques. *American Antiquity*: 536-546.

Macal, C.M. and M.J. North, 2005. Tutorial on agent-based modeling and simulation. Proceedings of the 37th conference on Winter simulation.

Macal, C.M. and M.J. North, 2009. Agent-based modeling and simulation. In *Winter Simulation Conference*. Austin, Texas: Winter Simulation Conference.

MacDonald, K., 2007. Cross-cultural Comparison of Learning in Human Hunting. Implications for Life History Evolution. *Human Nature* 18 (4): 386-402.

MacDonald, K., M. Martinón-Torres, R.W. Dennell and J.M. Bermúdez de Castro, 2012. Discontinuity in the record for hominin occupation in south-western Europe: Implications for occupation of the middle latitudes of Europe. *Quaternary International* 271: 84-97.

MacDonald, K., W. Roebroeks and A. Verpoorte, 2009. An energetics perspective on the Neandertal record. In *The evolution of hominin diets: Integrating approaches to the study of Palaeolithic subsistence.*, edited by J.-J. Hublin and M. P. Richards. Leipzig: Springer, 211-220.

MacDonell, G., 1995. Environmental resources and range size: a study of modern and ancients hunter gathers. *Nebraska Anthropologist* 12.

Malleson, N., 2006. An agent-based model of burglary in Leeds, School of Geography, Uniersity of Leeds, Leeds.

Mäntylä, M.V. and C. Lassenius, 2009. What Types of Defects Are Really Discovered in Code Reviews? *IEEE Transactions on Software Engineering* 35 (3): 430-448.

Marc, P., 2001. Digital Natives, Digital Immigrants Part 2: Do They Really Think Differently? *On the Horizon* 9 (6): 1-6.

Marean, C.W. and S.Y. Kim, 1998. Mousterian large-mammal remains from Kobeh Cave: Behavioural implications for Neanderthals and early modern humans. *Current Anthropology* 39 (Suppl. 3): S79-S113.

Maria, B. and S. Udo, 2017. Why input matters: Selection of climate data sets for modelling the potential distribution of a treeline species in the Himalayan region. *Ecological Modelling* 359: 92-102.

Marín, J., P. Saladié, A. Rodríguez-Hidalgo and E. Carbonell, 2017. Neanderthal hunting strategies inferred from mortality profiles within the Abric Romaní sequence. *PLoS ONE* 12 (11): e0186970.

Maris, M.N.A., B. de Boer and J. Oerlemans, 2012. A climate model intercomparison for the Antarctic region: present and past. *Climate of the Past* 8: 803-814.

Marks, J., 2012. My ancestors, myself. Fossil genomics is opening new windows to the past. But the view through them isn't as clear as we like to think. . In *Aeon 12*.

Martelle Hayter, H., 2011. Hunter-gatherers and the Ethnographic Analogy: Theoretical Perspectives. *Totem: The University of Western Ontario Journal of Anthropology* 1 (1): 8.

Marwick, B., 2017. Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory* 24 (2): 424-450.

Mateos, A., I. Goikoetxea, W.R. Leonard, J.Á. Martín-González, G. Rodríguez-Gómez and J. Rodríguez, 2014. Neandertal growth: What are the costs? *Journal of Human Evolution* 77: 167-178.

Mathur, P., 2007. Agent-based Modeling of Urban Phenomena in GIS. In *Urban Spatial Analytics, Capstone Project*: University of Pennsylvania.

Mayr, E., 1950. Taxonomic categories in fossil homnids. *Cold Spring Harbor Symposium on Quantitative Biology* 15: 108-118.

McGlade, J., 2014. Simulation as narrative: contingency, dialogics, and the modeling conundrum. *Journal of Archaeological Method and Theory* 21 (2): 288-305.

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.

Medina-Elizalde, M., 2013. A global compilation of coral sea-level benchmarks: Implications and new challenges. *Earth and Planetary Science Letters* 362 (0): 310-318.

Melchionna, M., M. Di Febbraro, F. Carotenuto, L. Rook, A. Mondanaro, S. Castiglione, C. Serio, V.A. Vero, G. Tesone, M. Piccolo, J.A.F. Diniz-Filho and P. Raia, 2018. Fragmentation of Neanderthals' pre-extinction distribution by climate change. *Palaeogeography, Palaeoclimatology, Palaeoecology* 496: 146-154.

Melis, A.P. and M. Tomasello, 2013. Chimpanzees' (Pan troglodytes) strategic helping in a collaborative task. *Biology Letters* 9 (2).

Mellars, P.A., 1992. Archaeology and the population-dispersal hypothesis of modern human origins in Europe. *Philosophical Transactions of the Royal Society of London* 337: 225-234.

Meyer, M., J.-L. Arsuaga, C. de Filippo, S. Nagel, A. Aximu-Petri, B. Nickel, I. Martínez, A. Gracia, J.M.B. de Castro, E. Carbonell, B. Viola, J. Kelso, K. Prüfer and S. Pääbo, 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* 531: 504.

Michalewicz, Z., 1992. *Genetic algorithms + data structures = evolution programs*: Springer-Verlag.

Miller, B.L. and D.E. Goldberg, 1995. Genetic algorithms, tournament selection, and the effects of noise. *Complex systems* 9 (3): 193-212.

Miller, J.H. and S.E. Page, 2007a. *Complex Adaptive Systems: An introduction to computational models of social life.*: Princeton, NJ: Princeton University Press.

Miller, J.H. and S.E. Page, 2007b. Social Science in Between, from Complex Adaptive Systems: *An Introduction to Computational Models of Social Life*. Introductory Chapters.

Mithen, S. and M. Reed, 2002. Stepping out: a computer simulation of hominid dispersal from Africa. *Journal of Human Evolution* 43 (4): 433-462.

Monserud, R.A. and R. Leemans, 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling* 62 (4): 275-293.

Montuire, S. and E. Desclaux, 1997. Palaeoecological analysis of mammalian faunas and environmental evolution in the South of France during the Pleistocene. *Boreas* 26 (4): 355-365.

Moore, G., 1965. Moore's law. *Electronics Magazine* 38 (8): 114.

Mosbrugger, V., T. Utescher and D.L. Dilcher, 2005. Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences of the United States of America* 102 (42): 14964-14969.

Moum, J.N., A. Perlin, J.D. Nash and M.J. McPhaden, 2013. Seasonal sea surface cooling in the equatorial Pacific cold tongue controlled by ocean mixing. *Nature* 500 (7460): 64-67.

Mulder, V., V. Heri and T. Wickham, 2000. Traditional honey and wax collection with Apis dorsata in the upper Kapuas Lake region, West Kalimantan. In *Borneo Research Bulletin* edited by C. Sather: Borneo Research Council, 246-260.

Müller, B., S. Balbi, C.M. Buchmann, L. de Sousa, G. Dressler, J. Groeneveld, C.J. Klassert, Q.B. 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. Weise, 2014. Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling & Software* 55 (0): 156-163.

Müller, U.C., J.ö. Pross and E. Bibus, 2003. Vegetation response to rapid climate change in central europe during the past 140,000 yr based on evidence from the Füramoos pollen record. *Quaternary Research* 59 (2): 235-245.

Murdock, G.P., 1967. *Ethnographic Atlas: a Summary*. Pittsburgh: University of Pittsburgh Press.

Muttoni, G., G. Scardia and D.V. Kent, 2018. Early hominins in Europe: The Galerian migration hypothesis. *Quaternary Science Reviews* 180: 1-29.

Nabel, J.E., N. Zurbriggen and H. Lischke, 2013. Interannual climate variability and population density thresholds can have a substantial impact on simulated tree species' migration. *Ecological Modelling* 257: 88-100.

Nakazawa, Y. and C.J. Bae, 2017. Quaternary paleoenvironmental variation and its impact on initial human dispersals into the Japanese Archipelago. *Palaeogeography, Palaeoclimatology, Palaeoecology*.

Niazi, M.A. and A. Temkin, 2017. Why teach modeling & simulation in schools? *Complex Adaptive Systems Modeling* 5 (1): 7.

Nichols, J.D., 2001. Using models in the conduct of science and management of natural resources. In *Modeling in natural resource management: development, interpretation, and application*, edited by T. M. Shenk and A. B. Franklin. Washington, DC: Island Press.

Nikolai, C. and G. Madey, 2009. Tools of the trade: A survey of various agent based modeling platforms. *Journal of Artificial Societies and Social Simulation* 12 (2): 2.

North, M.J., N.T. Collier, J. Ozik, E.R. Tatara, C.M. Macal, M. Bragen and P. Sydelko, 2013. Complex adaptive systems modeling with Repast Simphony. *Complex Adaptive Systems Modeling* 1 (1): 1-26.

North, M.J., N.T. Collier and J.R. Vos, 2006. Experiences creating three implementations of the repast agent modeling toolkit. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 16 (1): 1-25.

North, M.J., T.R. Howe, N.T. Collier and J.R. Vos, 2005. The repast simphony runtime system. Agent 2005 Conference on Generative Social Processes, Models, and Mechanisms, at Argonne National Laboratory, Argonne, IL USA.

Nuzzo, R., 2014. Scientific method: statistical errors. *Nature News* 506 (7487): 150.

O'Connell, J.F., 2006. How did modern humans displace Neanderthals? Insights from hunter-gatherer ethnography and archaeology. *When Neanderthals and modern humans met*: 43-64.

O'Connor, S., J. Louys, S. Kealy and S.C.S. Carro, 2017. Hominin Dispersal and Settlement East of Huxley's Line: The Role of Sea Level Changes, Island Size, and Subsistence Behavior. *Current Anthropology* 58(S17), S567-S582.

Oesterheld, M., O.E. Sala and S.J. McNaughton, 1992. Effect of animal husbandry on herbivore-carrying capacity at a regional scale. *Nature* 356 (6366): 234-236.

Olfert, O., R.M. Weiss and D. Kriticos, 2011. Application of General Circulation Models to Assess the Potential Impact of Climate Change on Potential Distribution and Relative Abundance of Melanoplus sanguinipes (Fabricius) (Orthoptera: Acrididae) in North America. *Psyche: A Journal of Entomology* 2011: 9.

Otto-Bliesner, B.L., E.C. Brady, G. Clauzet, R. Tomas, S. Levis and Z. Kothavala, 2006. Last Glacial Maximum and Holocene Climate in CCSM3. *Journal of Climate* 19 (11): 2526-2544.

Otto-Bliesner, B.L., C.D. Hewitt, T.M. Marchitto, E. Brady, A. Abe-Ouchi, M. Crucifix, S. Murakami and S.L. Weber, 2007. Last Glacial Maximum ocean thermohaline circulation: PMIP2 model intercomparisons and data constraints. *Geophysical Research Letters* 34 (12): L12706.

Pais, C.M., M.G. Colazo, M. Fernandez, S. Bulatovich and H. Fernandez, 2017. Dengue Agent-Based Model in South American Temperate Zone. In *Bioinformatics and Biomedical Engineering: 5th International Work-Conference, IWBBIO 2017, Granada, Spain, April 26–28, 2017*, edited by I. Rojas and F. Ortuño. Granada, Spain: Springer International Publishing.

Pearson, R.G., W. Thuiller, M.B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T.P. Dawson and D.C. Lees, 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography* 33 (10): 1704-1711.

Peeters, H. and J.-W. Romeijn, 2016. Epistemic considerations about uncertainty and model selection in computational archaeology: a case study on exploratory modeling. In *Uncertainty and Sensitivity Analysis in Archaeological Computational Modeling*: Springer, 37-58.

Peng, R.D., 2011. Reproducible Research in Computational Science. *Science* 334 (6060): 1226.

Pennington, 2001. Hunter-gatherer demography. In *Hunter-gatherers: an interdisciplinary perspective*, edited by C. Panter-Brick, R. H. Layton and P. Rowley-Conwy. Cambridge: Cambridge University Press.

Pereda, M., D. Zurro, J.I. Santos, I. Briz i Godino, M. Álvarez, J. Caro and J.M. Galán, 2017. Emergence and Evolution of Cooperation Under Resource Pressure. *Scientific Reports* 7: 45574.

Pfeifer, B., K. Kugler, M.M. Tejada, C. Baumgartner, M. Seger, M. Osl, M. Netzer, M. Handler, A. Dander and M. Wurz, 2008. A cellular automaton framework for infectious disease spread simulation. *The open medical informatics journal* 2: 70.

Phillips, B.L., G.P. Brown, J.K. Webb and R. Shine, 2006. Invasion and the evolution of speed in toads. *Nature* 439 (7078): 803-803.

Phillips, S.J., R.P. Anderson and R.E. Schapire, 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190 (3–4): 231-259.

Pitt, W.C., P.W. Box and F.F. Knowlton, 2003. An individual-based model of canid populations: modelling territoriality and social structure. *Ecological Modelling* 166 (1): 109-121.

Polack, F.A., P.S. Andrews, T. Ghetiu, M. Read, S. Stepney, J. Timmis and A.T. Sampson, 2010. *Reflections on the simulation of complex systems for science*. 15th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS).

Power, R., D. Salazar-García, D. A, M. Rubini, K. Harvati, W. M.J, H. J.J and H. A, 2018. Dental calculus indicates widespread plant use within the stable Neanderthal dietary niche. *Journal of Human Evolution,* 119, 27-41.

Premo, L., J. Murphy, J. Scholnick, B. Gabler and J. Beaver, 2005. Making a case for agent-based modeling. *Society for Archaeological Sciences Bulletin* 28 (3): 11-13.

Prentice, I.C., W. Cramer, S.P. Harrison, R. Leemans, R.A. Monserud and A.M. Solomon, 1992. Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. *Journal of biogeography*: 117-134.

Prentice, I.C., D. Jolly and B. participants, 2000. Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biogeography* 27 (3): 507-519.

Pryor, A., A. Pullen, D. Beresford-Jones, J. Svoboda and C. Gamble, 2016. Reflections on Gravettian firewood procurement near the Pavlov Hills, Czech Republic. *Journal of Anthropological Archaeology* 43: 1-12.

Pulliam, H.R., 1988. Sources, sinks, and population regulation. *The American Naturalist* 132 (5): 652-661.

Pulliam, H.R., 1996. Sources and sinks: empirical evidence and population consequences. In *Population dynamics in ecological space and time*, edited by O. E. J. Rhodes, R. K. Chester and M. H. Smith. Chicago: University of Chicago Press, Illinois, USA, 45-69.

R Core Team, 2017. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria.

Radloff, F.G.T. and J.T. Du Toit, 2004. Large predators and their prey in a southern African savanna: a predator's size determines its prey size range. *Journal of Animal Ecology* 73 (3): 410-423.

Railsback, S.F., S.L. Lytinen and S.K. Jackson, 2006. Agent-based simulation platforms: Review and development recommendations. *Simulation* 82 (9): 609-623.

Rangel, T.F.L., J.A.F. Diniz-Filho and L.M. Bini, 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global ecology and biogeography* 15 (4): 321-327.

Rangel, T.F.L., J.A.F. Diniz Filho and L.M. Bini, 2010. SAM: a comprehensive application for spatial analysis in macroecology. *Ecography* 33 (1): 46-50.

Rasmussen, S.O., M. Bigler, S.P. Blockley, T. Blunier, S.L. Buchardt, H.B. Clausen, I. Cvijanovic, D. Dahl-Jensen, S.J. Johnsen, H. Fischer, V. Gkinis, M. Guillevic, W.Z. Hoek, J.J. Lowe, J.B. Pedro, T. Popp, I.K. Seierstad, J.P. Steffensen, A.M. Svensson, P. Vallelonga, B.M. Vinther, M.J.C. Walker, J.J. Wheatley and M. Winstrup, 2014. A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy. *Quaternary Science Reviews* 106: 14-28.

Revay, P. and C. Cioffi-Revilla, 2018. Survey of evolutionary computation methods in social agent-based modeling studies. *Journal of Computational Social Science* 1 (1): 115-146.

Reyes-Centeno, H., S. Ghirotto, F. Détroit, D. Grimaud-Hervé, G. Barbujani and K. Harvati, 2014. Genomic and cranial phenotype data support multiple modern human dispersals from Africa and a southern route into Asia. *Proceedings of the National Academy of Sciences* 111 (20): 7248-7253.

Richards, M., K. Harvati, V. Grimes, C. Smith, T. Smith, J.-J. Hublin, P. Karkanas and E. Panagopoulou, 2008. Strontium isotope evidence of Neanderthal mobility at the site of Lakonis, Greece using laser-ablation PIMMS. *Journal of Archaeological Science* 35 (5): 1251-1256.

Richards, M.P., P.B. Pettitt, E. Trinkaus, F.H. Smith, M. Paunović and I. Karavanić, 2000. Neanderthal diet at Vindija and Neanderthal predation: The evidence from stable isotopes. *Proceedings of the National Academy of Sciences* 97 (13): 7663-7666.

Richards, M.P. and E. Trinkaus, 2009. Isotopic evidence for the diets of European Neanderthals and early modern humans. *Proceedings of the National Academy of Sciences* 106 (38): 16034-16039.

Richter, D., H. Dibble, P. Goldberg, S.P. McPherron, L. Niven, D. Sandgathe, S. Talamo and A. Turq, 2013. The late Middle Palaeolithic in Southwest France: New TL dates for the sequence of Pech de l'Azé IV. *Quaternary International* 294: 160-167.

Riede, F. and J.B. Pedersen, 2018. Late Glacial Human Dispersals in Northern Europe and Disequilibrium Dynamics. *Human Ecology*.

Rigby, P.C. and M.-A. Storey, 2011. Understanding broadcast based peer review on open source software projects. *Proceedings of the 33rd International Conference on Software Engineering*. IEEE.

Roebroeks, W., 2006. The human colonisation of Europe: where are we? *Journal of Quaternary Science* 21 (5): 425-435.

Roebroeks, W., 2010. *The Neandertal Experiment*. Amsterdam: Stichting Nederlands Museum voor Anthropologie en Praehistorie.

Roebroeks, W., N.J. Conard and T. van Kolfschoten, 1992. Dense forests, cold steppes and the Palaeolithic settlement of Northern Europe. *Current Anthropology* 33 (5): 551-586.

Roebroeks, W., J.J. Hublin and K. MacDonald, eds, 2011. *Continuities and Discontinuities in Neandertal Presence: A Closer look at Northwestern Europe*. Edited by N. M. Ashton, S. G. Lewis and C. B. Stringer. The Ancient Human Occupation Of Britain ed. Vol. 14, *Developments in Quaternary Science*. Amsterdam: Elsevier.

Roebroeks, W. and M. Soressi, 2016. Neandertals revised. *Proceedings of the National Academy of Sciences* 113 (23): 6372-6379.

Rogers, J.D. and W.H. Cegielski, 2017. Opinion: Building a better past with the help of agent-based modeling. *Proceedings of the National Academy of Sciences* 114 (49): 12841-12844.

Romanowska, I., C. Gamble, S. Bullock and F. Sturt, 2017. Dispersal and the Movius line: testing the effect of dispersal on population density through simulation. *Quaternary International* 431: 53-63.

Roussel, M., M. Soressi and J.-J. Hublin, 2016. The Châtelperronian conundrum: blade and bladelet lithic technologies from Quinçay, France. *Journal of Human Evolution* 95: 13-32.

RStudio: Integrated development environment for R 1.0.153, Boston, MA.

Ruebens, K., S.J.P. McPherron and J.-J. Hublin, 2015. On the local Mousterian origin of the Châtelperronian: Integrating typo-technological, chronostratigraphic and contextual data. *Journal of Human Evolution* 86: 55-91.

Rufà, A., R. Blasco, F. Rivals and J. Rosell, 2015. Who eats whom? Taphonomic analysis of the avian record from the Middle Paleolithic site of Teixoneres Cave (Moià, Barcelona, Spain). *Quaternary International*, 421, 103-115.

Salazar-García, D.C., R.C. Power, A. Sanchis Serra, V. Villaverde, M.J. Walker and A.G. Henry, 2013. Neanderthal diets in central and southeastern Mediterranean Iberia. *Quaternary International,* 318, 3-18.

Sanchez Goñi, M.F. and S.P. Harrison, 2010. Millennial-scale climate variability and vegetation changes during the Last Glacial: Concepts and terminology. *Quaternary Science Reviews* 29 (21-22): 2823-2827.

Sanchez Goñi, M.F., A. Landais, W.J. Fletcher, F. Naughton, S. Desprat and J. Duprat, 2008. Contrasting impacts of Dansgaard–Oeschger events over a western European latitudinal transect modulated by orbital parameters. *Quaternary Science Reviews* 27 (11): 1136-1151.

Sandve, G.K., A. Nekrutenko, J. Taylor and E. Hovig, 2013. Ten Simple Rules for Reproducible Computational Research. *PLOS Computational Biology* 9 (10): e1003285.

Schelling, T.C., 1978. *Micromotives and Macrobehavior,* Fels lectures on public policy analysis. New York: WW Norton.

Scherjon, F., 2011. SteppingIn – Modelling and Simulating Modern Humans moving into Palaeolithic Europe, Faculty of archaeology, University/College of Higher Education: University of Leiden, Leiden, The Netherlands.

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.

Scherjon, F., 2015a. *HomininSpace, modelling and simulating hominins moving through time and real geographic space* (thesis), Human Origins, Faculty of Archaeology, Leiden University, Leiden, The Netherlands.

Scherjon, F., 2015b. Large scale population modelling in the deep past. Issues and concerns. HomininSpace, a case study in points. In *Computer Applications and Quantitative Methods in Archaeology (CAA)*. Siena, Spain.

Scherjon, F., 2015c. Neanderthals on the move. Or not. Poster presentation. In *European Society for the Study of Human Evolution (ESHE)*. London, UK.

Scherjon, F., 2016. Evolving hominins in HomininSpace—Genetic algorithms and the search for the 'perfect'Neanderthal. *Proceedings Computer Applications & Quantitative Methods in Archaeology*, at Oslo.

Scherjon, F., C. Bakels, K. MacDonald and W. Roebroeks, 2015. Burning the Land: An Ethnographic Study of Off-Site Fire Use by Current and Historically Documented Foragers and Implications for the Interpretation of Past Fire Practices in the Landscape. *Current Anthropology* 56 (3): 299-326.

Schiffer, M.B., 1985. Is There a "Pompeii Premise" in Archaeology? *Journal of Anthropological Research* 41: 18-41.

Schmitt, T. and Z. Varga, 2012. Extra-Mediterranean refugia: The rule and not the exception? *Frontiers in zoology* 9 (1): 1-12.

Schmitz, H., 2018. Schnek: A C++ library for the development of parallel simulation codes on regular grids. *Computer Physics Communications*.

Scott, B. and N.M. Ashton, 2011. The Early Middle Palaeolithic: The European Context. In *The Ancient Human Occupation of Britain*, edited by N. M. Ashton, S. G. Lewis and C. B. Stringer. Amsterdam: Elsevier, 91-112.

Shaman, J., S. Solomon, R.R. Colwell and C.B. Field, 2013. Fostering advances in interdisciplinary climate science. *Proceedings of the National Academy of Sciences* 110 (Supplement 1): 3653-3656.

Shaw, C.N. and J.T. Stock, 2013. Extreme mobility in the Late Pleistocene? Comparing limb biomechanics among fossil< i> Homo</i>, varsity athletes and Holocene foragers. *Journal of human evolution*.

Shea, John J., 1998. Neandertal and Early Modern Human Behavioral Variability A Regional-Scale Approach to Lithic Evidence for Hunting in the Levantine Mousterian. *Current Anthropology* 39 (s1): S45-S78.

Shen, C.-C., T. Lee, K.-K. Liu, H.-H. Hsu, R.L. Edwards, C.-H. Wang, M.-Y. Lee, Y.-G. Chen, H.-J. Lee and H.-T. Sun, 2005. An evaluation of quantitative reconstruction of past precipitation records using coral skeletal Sr/Ca and δ18O data. *Earth and Planetary Science Letters* 237 (3–4): 370-386.

Short, R., 1987. The biological basis for the contraceptive effects of breast feeding. *International Journal of Gynecology & Obstetrics* 25: 207-217.

Shultz, S., C. Opie and Q.D. Atkinson, 2011. Stepwise evolution of stable sociality in primates. *Nature* 479 (7372): 219-222.

Sibly, R.M., V. Grimm, B.T. Martin, A.S. Johnston, K. Kułakowska, C.J. Topping, P. Calow, J. Nabe-Nielsen, P. Thorbek and D.L. DeAngelis, 2013. Representing the acquisition and use of energy by individuals in agent-based models of animal populations. *Methods in Ecology and Evolution* 4 (2): 151-161.

Sier, M.J., J.M. Parés, P. Antoine, J.-L. Locht, M.J. Dekkers, N. Limondin-Lozouet and W. Roebroeks, 2015. Evidence for the Blake Event recorded at the Eemian archaeological site of Caours, France. *Quaternary International* 357: 149-157.

Sier, M.J., W. Roebroeks, C.C. Bakels, M.J. Dekkers, E. Brühl, D. De Loecker, S. Gaudzinski-Windheuser, N. Hesse, A. Jagich, L. Kindler, W.J. Kuijper, T. Laurat, H.J. Mücher, K.E.H. Penkman, D. Richter and D.J.J. van Hinsbergen, 2011. Direct terrestrial-marine correlation demonstrates surprisingly late onset of the last interglacial in central Europe. *Quaternary Research* 75 (1): 213-218.

Silva, M., M. Brimacombe and J.A. Downing, 2001. Effects of body mass, climate, geography, and census area on population density of terrestrial mammals. *Global Ecology and Biogeography* 10 (5): 469-485.

Skinner, M., 1997. Dental wear in immature Late Pleistocene European hominines. *Journal of archaeological science* 24 (8): 677-700.

Slimak, L. and Y. Giraud, 2007. Circulations sur plusieurs centaines de kilomètres durant le Paléolithique moyen. Contribution à la connaissance des sociétés néandertaliennes. *Comptes Rendus Palevol* 6 (5): 359-368.

Smith, T.M., P. Tafforeau, D.J. Reid, R. Grün, S. Eggins, M. Boutakiout and J.-J. Hublin, 2007. Earliest evidence of modern human life history in North African early *Homo sapiens*. *Proceedings of the National Academy of Sciences* 104 (15): 6128-6133.

Smith, T.M., P. Tafforeau, D.J. Reid, J. Pouech, V. Lazzari, J.P. Zermeno, D. Guatelli-Steinberg, A.J. Olejniczak, A. Hoffman, J. Radovcic, M. Makaremi, M. Toussaint, C. Stringer and J.-J. Hublin, 2010. Dental evidence for ontogenetic differences between modern humans and Neanderthals. *Proceedings of the National Academy of Sciences,* 107(49), 20923-20928.

Smith, W.H.F. and D.T. Sandwell, 1997. Global Sea Floor Topography from Satellite Altimetry and Ship Depth Soundings. *Science* 277 (5334): 1956-1962.

Sneyers, R., 1991. *On the statistical analysis of series of observations*. Technical Note - World Meteorological Organization 143.

Snodgrass, J.J. and W.R. Leonard, 2009. Neandertal energetics revisited: Insights into population dynamics and life history evolution. *PaleoAnthropology*: 220-237.

Soltis, J., R. Boyd and P.J. Richerson, 1995. Can Group-Functional Behaviors Evolve by Cultural Group Selection?: An Empirical Test. *Current Anthropology* 36 (3).

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

Sørensen, B., 2011. Demography and the extinction of European Neanderthals. *Journal of Anthropological Archaeology* 30: 17-29.

Sorensen, M.V. and W.R. Leonard, 2001. Neandertal energetics and foraging efficiency. *Journal of Human Evolution* 40 (6): 483-95.

Soressi, M., 2016. Archaeology: Neanderthals built underground. *Nature* 534 (7605): 43-44.

Soressi, M. and M.A. Hays, 2003. Manufacture, Transport, and use of Mousterian Biface. In *Multiple Approaches to the Study of Bifacial Technologies*, edited by M. Soressi and H. L. Dibble. Philadelphia: University of Pennsylvania Press.

Speth, J.D., 2010. *The Paleoanthropology and Archaeology of Big-game Hunting: Protein, Fat, Or Politics?*, *Interdisciplinary Contributions to Archaeology*. New York, NY: Springer Science+Business Media, LLC.

Speth, J.D., 2013. Thoughts about hunting: Some things we know and some things we don't know. *Quaternary International* 297: 176-185.

Steegmann, T.A., F.J. Cerny and T.W. Holliday, 2002. Neandertal cold adaptation: Physiological and energetic factors. *American Journal of Human Biology* 14 (5): 566-583.

Steele, J., J. Adams and T. Sluckin, 1998. Modelling Paleoindian Dispersals. *World Archaeology* 30 (2): 286-305.

Steinmetz, R., W. Chutipong, N. Seuaturien, E. Chirngsaard and M. Khaengkhetkarn, 2010. Population recovery patterns of Southeast Asian ungulates after poaching. *Biological Conservation* 143 (1): 42-51.

Stewart, F.L. and P.W. Stahl, 1977. Cautionary note on edible meat poundage figures. *American Antiquity*: 267-270.

Stewart, J.R., 2007. Neanderthal extinction as part of the faunal change in Europe during Oxygen Isotope Stage 3. *Acta Zoologica Cracoviensia-Series A: Vertebrata* 50 (1-2): 93-124.

Stewart, J.R. and A.M. Lister, 2001. Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution* 16(11): 608-613.

Stewart, J.R., A.M. Lister, I. Barnes and L. Dalén, 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences* 277 (1682): 661-671.

Stiegelmeyer, S.M. and M.C. Giddings, 2013. Agent-based modeling of competence phenotype switching in Bacillus subtilis. *Theoretical Biology and Medical Modelling* 10 (1): 23.

Strasser, T.F., E. Panagopoulou, C.N. Runnels, P.M. Murray, N. Thompson, P. Karkanas, F.W. McCoy and K.W. Wegmann, 2010. Stone Age seafaring in the Mediterranean: evidence from the Plakias region for Lower Palaeolithic and Mesolithic habitation of Crete. *Hesperia*: 145-190.

Straus, L.G., 2001. Africa and Iberia in the Pleistocene. *Quaternary International* 75 (1): 91-102.

Stringer, C., 2003. Human evolution: Out of Ethiopia. *Nature* 423 (6941): 692-695.

Stringer, C., 2016. The origin and evolution of Homo sapiens. *Phil. Trans. R. Soc. B* 371 (1698): 20150237.

Surovell, T.A., 2003. Simulating Coastal Migration in New World Colonization. *Current Anthropology* 44 (4): 580-591.

Surovell, T.A. and P.J. Brantingham, 2007. A note on the use of temporal frequency distributions in studies of prehistoric demography. *Journal of Archaeological Science* 34 (11): 1868-1877.

Surovell, T.A., J. Byrd Finley, G.M. Smith, P.J. Brantingham and R. Kelly, 2009. Correcting temporal frequency distributions for taphonomic bias. *Journal of Archaeological Science* 36 (8): 1715-1724.

Syswerda, G., 1989. Uniform Crossover in Genetic Algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms*: Morgan Kaufmann Publishers Inc.

Tafelmaier, Y., A. Pastoors and G.-C. Weniger, 2017. Cultural contact over the Strait of Gibraltar during the Middle Palaeolithic? Evaluating the visibility of cultural exchange. *Munibe Antropologia-Arkeologia* (68).

Tatara, E., M. North, T. Howe, N. Collier and J. Vos, 2006. An indroduction to repast simphony modeling using a simple predator-prey example. Proceedings of the Agent 2006 Conference on Social Agents: Results and Prospects.

Teyssandier, N., 2008. Revolution or evolution: the emergence of the Upper Paleolithic in Europe. *World Archaeology* 40 (4): 493-519.

Thackeray, J.F., 2013. Human evolution from an African perspective. With reference to Charles Darwin palaeoclimate and the problem of defining a species. In *35ste Kroon-voordracht.* . Stichting Nederlands museum voor Anthropologie en Praehistorie.

Thomas, C.D., A. Cameron, R.E. Green, M. Bakkenes, L.J. Beaumont, Y.C. Collingham, B.F. Erasmus, M.F. De Siqueira, A. Grainger and L. Hannah, 2004. Extinction risk from climate change. *Nature* 427 (6970): 145-148.

Timm, I.J., F. Lorig, E. Holzchen and C. Hertler, 2014. Multi-scale agent-based simulation of long-term dispersal processes. Social Simulation Conference.

Tobias, R. and C. Hofmann, 2004. Evaluation of free Java-libraries for social-scientific agent based simulation. *Journal of Artificial Societies and Social Simulation* 7 (1).

Trinkaus, E., 1986. The Neandertals and Modern Human Origins. *Annual Review of Anthropology* 15: 193-218.

Trinkaus, E., 2011. Late Pleistocene adult mortality patterns and modern human establishment. *Proceedings of the National Academy of Sciences* 108 (4): 1267-1271.

Tsanova, T., 2006. Les débuts du Paléolithique supérieur dans l'Est des Balkans. Réflexion à partir de l'étude taphonomique et techno-économique des ensembles lithiques des sites de Bacho Kiro (couche 11), Temnata (couches VI et 4) et Kozarnika (niveau VII), Sciences du Vivant, Géoscience, Sciences de l'environnement, Université Bordeaux 1.

Turing, A.M., 1948. *Intelligent Machinery, report for the National Physical Laboratory*: The Turing Archive for the History of Computing (www.AlanTuring.net/intelligent_machinery, verified 9 April 2019).

Turing, A.M., 1954. Solvable and unsolvable problems. *Science News-ens. fr* 39.

Usman, F., K. Murakami, A. Dwi Wicaksono and E. Setiawan, 2017. Application of Agent-Based Model Simulation for Tsunami Evacuation in Pacitan, Indonesia. *MATEC Web Conf.* 97: 01064.

Valeggia, C. and P.T. Ellison, 2009. Interactions between metabolic and reproductive functions in the resumption of postpartum fecundity. *American Journal of Human Biology* 21 (4): 559-566.

van Andel, T.H., 2002. The Climate and Landscape of the Middle Part of the Weichselian Glaciation in Europe: The Stage 3 Project. *Quaternary Research* 57: 2-8.

van Andel, T.H. and W. Davies, eds, 2003. *Neanderthals and modern humans in the European landscape during the last glaciation: archaeological results of the Stage 3 Project*. Cambridge: The McDonald Institute for Archaeological Research.

van Gijssel, K., 2006. A continent-wide framework for local and regional stratigraphies: application of genetic sequence and event stratigraphy to the Middle Pleistocene terrestrial succession of Northwest and Central Europe. Ph.D. thesis, Faculty of Archaeology, Leiden University, Leiden.

Vandermeer, J., 2006. Oscillating Populations and Biodiversity Maintenance. *BioScience* 56 (12): 967-975.

Vaquero, M., J. Maroto, A. Arrizabalga, J. Baena, E. Baquedano, E. Carrión, J.F. Jordà, M. Martinón, M. Menéndez, R. Montes and J. Rosell, 2006. The Neandertal-Modern Human Meeting in Iberia: A Critical View of the Cultural, Geographical and Chronological Data. In *When Neanderthals and Modern Humans Met*, edited by N. J. Conard. Tübingen: Kerns Verlag, 419-439.

Verpoorte, A., 2006. Neanderthal energetics and spatial behaviour. *Before Farming* 2006/3: 1-6.

Verpoorte, A., 2009. Limiting factors on early modern human dispersals: The human biogeography of late Pleniglacial Europe. *Quaternary International* 201: 77-85.

Vieillevigne, E., L. Bourguignon, I. Ortega and P. Guibert, 2008. Analyse croisée des données chronologiques et des industries lithiques dans le grand sud-ouest de la France (OIS 10 à 3) *PALEO* 20.

Villa, P. and W. Roebroeks, 2014. Neandertal Demise: An Archaeological Analysis of the Modern Human Superiority Complex. *PLoS ONE* 9: e96424.

Waelbroeck, C., L. Labeyrie, E. Michel, J.C. Duplessy, J.F. McManus, K. Lambeck, E. Balbon and M. Labracherie, 2002. Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records. *Quaternary Science Reviews* 21 (1–3): 295-305.

Wagenmakers, E.-J., 2003. How many parameters does it take to fit an elephant? *Journal of Mathematical Psychology* 47 (5): 580-586.

Wall-Scheffler, C.M. and M.J. Myers, 2013. Reproductive costs for everyone: How female loads impact human mobility strategies. *Journal of Human Evolution* 64 (5): 448-456.

Waltari, E., R.J. Hijmans, A.T. Peterson, Á.S. Nyári, S.L. Perkins and R.P. Guralnick, 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* 2 (7).

Weaver, T.D., H. Coqueugniot, L.V. Golovanova, V.B. Doronichev, B. Maureille and J.-J. Hublin, 2016. Neonatal postcrania from Mezmaiskaya, Russia, and Le Moustier, France, and the development of Neandertal body form. *Proceedings of the National Academy of Sciences*.

Weaver, T.D. and J.-J. Hublin, 2009. Neandertal birth canal shape and the evolution of human childbirth. *Proceedings of the National Academy of Sciences* 106 (20): 8151-8156.

Weaver, T.D. and K.L. Steudel-Numbers, 2005. Does climate or mobility explain the differences in body proportions between Neandertals and their Upper Palaeolithic successors? *Evolutionary Anthropology* 14 (6): 218-223.

Webb III, T., 1986. Is vegetation in equilibrium with climate? How to interpret late-Quaternary pollen data. *Vegetatio* 67 (2): 75-91.

Welker, F., M. Hajdinjak, S. Talamo, K. Jaouen, M. Dannemann, F. David, M. Julien, M. Meyer, J. Kelso, I. Barnes, S. Brace, P. Kamminga, R. Fischer, B.M. Kessler, J.R. Stewart, S. Pääbo, M.J. Collins and J.-J. Hublin, 2016. Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences* 113 (40): 11162.

Wenban-Smith, F.F., M.R. Bates and J.-L. Schwenninger, 2010. Early Devensian (MIS 5d-5b) occupation at Dartford, southeast England. *Journal of Quaternary Science* 25 (8): 1193-1199.

Wenban-Smith, F., 2010. M25 roadworks reveal earliest UK Neanderthal occupation at Dartford. *Geology Today* 26 (5): 175-179.

Wheatley, D., 2004. Making space for an archaeology of place. *Internet archaeology* 15.

White, A.A., 2012. The socal networks of early hunter-gatherers in midcontinental North America, The University of Michigan.

White, A.A., 2013. Subsistence economics, family size, and the emergence of social complexity in hunter–gatherer systems in eastern North America. *Journal of Anthropological Archaeology* 32 (1): 122-163.

White, M., P. Pettitt and D. Schreve, 2016. Shoot first, ask questions later: Interpretative narratives of Neanderthal hunting. *Quaternary Science Reviews* 140 (Supplement C): 1-20.

White, M.J., 2006. Things to do in Doggerland when you're dead: surviving OIS3 at the northwestern-most fringe of Middle Palaeolithic Europe. *World Archaeology* 38 (4): 547-575.

White, M.J. and P.B. Pettitt, 2011. The British Late Middle Palaeolithic: An Interpretative Synthesis of Neanderthal Occupation at the Northwestern Edge of the Pleistocene World. *Journal of World Prehistory* 24: 25-97.

White, T.E., 1953. A method of calculating the dietary percentage of various food animals utilized by aboriginal peoples. *American Antiquity* 18 (4): 396-398.

Whitner, R.B. and O. Balci, 1989. Guidelines for selecting and using simulation model verification techniques. Proceedings of the 21st conference on Winter simulation, at Washington, D.C., USA.

Wiegand, T., F. Knauer, P. Kaczensky and J. Naves, 2004. Expansion of Brown Bears (Ursus arctos) into the Eastern Alps: A Spatially Explicit Population Model. *Biodiversity & Conservation* 13 (1): 79-114.

Will, O., 2009. Resolving a replication that failed: news on the Macy & Sato model. *Journal of Artificial Societies and Social Simulation* 12 (4): 11.

Williams, A.N., S. Ulm, A.R. Cook, M.C. Langley and M. Collard, 2013. Human refugia in Australia during the Last Glacial Maximum and Terminal Pleistocene: A geospatial analysis of the 25-12ka Australian archaeological record. *Journal of Archaeological Science,* 40(12), 4612-4625.

Wilson, G., D.A. Aruliah, C.T. Brown, N.P.C. Hong, M. Davis, R.T. Guy, S.H. Haddock, K.D. Huff, I.M. Mitchell and M.D. Plumbley, 2014. Best practices for scientific computing. *PLoS biology* 12 (1): e1001745.

Wimsatt, W.C., 1987. False models as means to truer theories. In *Neutral models in biology*, edited by M. N. a. A. Hoffman. London: Oxford University Press, 23-55.

Wisiol, K., 1984. Estimating Grazingland Yield from Commonly Available Data. *Journal of Range Management* 37 (5): 471-475.

Wobst, H.M., 1974. Boundary Conditions for Paleolithic Social Systems: A Simulation Approach. *American Antiquity* 39 (2): 147-178.

Wolfram, S., 1983. Statistical mechanics of cellular automata. *Reviews of modern physics* 55 (3): 601.

Wood, B. and B.G. Richmond, 2000. Human evolution: taxonomy and paleobiology. *Journal of anatomy* 196: 19-60.

Wragg Sykes, R., 2010. Neanderthals in Britain: Late Mousterian Archaeology in Landscape Context, University of Sheffield, Department of Archaeology.

Wragg Sykes, R.M., 2017. Neanderthals in the Outermost West: Technological adaptation in the Late Middle Palaeolithic (re)-colonization of Britain, Marine Isotope Stage 4/3. *Quaternary International* 433: 4-32.

Wylie, A., 2002. *Thinking from Things: Essays in the Philosophy of Archaeology*: University of California Press.

Yao, X., 1993. An empirical study of genetic operators in genetic algorithms. *Microprocessing and Microprogramming* 38 (1): 707-714.

Yellen, J.E., 1977. *Archaeological approaches to the present: models for reconstructing the past*. Vol. 1: Academic Press New York.

Yeniay, Ö., 2005. Penalty function methods for constrained optimization with genetic algorithms. *Mathematical and computational Applications* 10 (1): 45-56.

Yokoyama, Y., K. Lambeck, P. De Deckker, P. Johnston and L.K. Fifield, 2000. Timing of the Last Glacial Maximum from observed sea-level minima. *Nature* 406 (6797): 713-716.

Young, D.A., 2002. A New Space-Time Computer Simulation Method for Human Migration. *American Anthropologist* 104 (1): 138-158.

Zhou, W.-X., D. Sornette, R.A. Hill and R.I. Dunbar, 2005. Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences* 272 (1561): 439-444.

Zilhão, J., 2000. The Ebro frontier: a model for the late extinction of Iberian Neanderthals. In *Neanderthals on the edge: 150th anniversary conference of the Forbes' Quarry discovery, Gibraltar*, edited by C. Stringer, R. N. E. Barton and C. Finlayson. Oxford: Oxbow Books, 111-121.

Zimov, S., N.S. Zimov and F.S. Chapin, III, 2012. The Past and Future of the Mammoth Steppe Ecosystem. In *Paleontology in Ecology and Conservation*, edited by J. Louys: Springer Berlin Heidelberg, 193-225.

Zimov, S.A., 2005. Pleistocene Park: Return of the Mammoth's Ecosystem. *Science* 308 (5723): 796-798.

Zobler, L., 1986. *A world soil file for global climate modeling*: National Aeronautics and Space Administration, Goddard Space Flight Center, Institute for Space Studies.

Zubrow, E., 1989. The demographic modelling of Neanderthal extinction. In *The Human Revolution. Behavioural and Biological Perspectives on the Origins of Modern Humans*, edited by P. Mellars and C. B. Stringer. Edinburgh: Edinburgh University Press, 212-231.

Zuk, M., 2013. *Paleofantasy: What Evolution Really Tells Us about Sex, Diet, and How We Live*: WW Norton & Company.

# APPENDIX 1: ABBREVIATIONS

### Abbreviations

| Abbreviation | Meaning |
|---|---|
| **ABMS** | Agent-Based Modelling and Simulation |
| **ABM** | Agent-Based Model(ling), can also be found under the terms Agent-Based Computational Modelling, Agent-Based Social Simulation, Multi-Agent systems, Distributed Artificial Intelligence, and Swarm Intelligence. |
| **ABS** | Agent-Based System. |
| **AMH** | Anatomically Modern Human(s). |
| **BP** | Before Present, where present is defined as 1950. |
| **CLIMAP** | Climate Map. |
| **FR** | Foraging range: The total area of land over which a hominin group moves and from which resource materials are taken in one year. This is equal to the *annual range*, the area that is covered for subsistence purposes in a complete year (MacDonell 1995). |
| **GARP** | Genetic Algorithm for Rule-Set Prediction. |
| **HS** | HomininSpace: The *Hominini* tribal level is today the scientific identification of humans and their ancestors as separated from the great apes, with which they form the family taxonomic unit of *Hominidae* (Wood and Richmond 2000). In popular language this is anglicized to hominins. Space is the natural living dimension for individuals. Both the simulation research methods as well as archaeological excavations add their own temporal dimension to the usage of space. The generalized nature of the implementation of the system (without the specialized datasets that mark the case study for Neanderthals) suggests the name HomininSpace (no italics) for a simulation system that models hominin behaviour in the landscape and that is validated against archaeological data concerning hominin presence and absence. HomininSpace is the designation for the complete modelling and simulation system with associated documentation that has been developed in the context of the study described in this thesis. |
| **IBM** | Individual-Based Modelling. |
| **ka** | Kilo-Annum (thousand years) ago. |
| **kya** | Kilo (thousand) years ago. |
| **LP** | Late Pleistocene (versus Middle Palaeolithic). The Late Pleistocene is defined as the period from 126 ka up to 11.7 ka. It starts with the Eemian (MIS 5e) and ends with the Holocene (http://quaternary.stratigraphy.org/majordivisions/, http://www.stratigraphy.org/index.php/ics-chart-timescale, accessed 15 January 2018). Since the simulations in this study all end at 50 ka, the final part of the Late Pleistocene is not included in this research. |
| **MAS** | Multi-Agent Systems. |
| **MAXENT** | Maximum Entropy. |
| **MH** | Mid-Holocene point in time (~6 ka), reconstructed as very warm compared to current day temperatures. |
| **My** | Million Years. |
| **Mya** | Million Years Ago. |
| **LGM** | Last Glacial Maximum (~21 ka). |
| **LMP** | Late Middle Palaeolithic. |
| **LUP** | Late Upper Palaeolithic. |
| **PMIP** | Paleoclimate Modeling Intercomparison Project. |
| **SAM** | Spatial Analysis in Macroecology. |

### Bioclimatic parameters

| **Bioclimatic parameters** | BIO1 = Annual Mean Temperature<br>BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))<br>BIO3 = Isothermality (BIO2/BIO7) (* 100)<br>BIO4 = Temperature Seasonality (standard deviation *100)<br>BIO5 = Max Temperature of Warmest Month<br>BIO6 = Min Temperature of Coldest Month<br>BIO7 = Temperature Annual Range (BIO5-BIO6)<br>BIO8 = Mean Temperature of Wettest Quarter<br>BIO9 = Mean Temperature of Driest Quarter<br>BIO10 = Mean Temperature of Warmest Quarter<br>BIO11 = Mean Temperature of Coldest Quarter<br>BIO12 = Annual Precipitation<br>BIO13 = Precipitation of Wettest Month<br>BIO14 = Precipitation of Driest Month<br>BIO15 = Precipitation Seasonality (Coefficient of Variation)<br>BIO16 = Precipitation of Wettest Quarter<br>BIO17 = Precipitation of Driest Quarter<br>BIO18 = Precipitation of Warmest Quarter<br>BIO19 = Precipitation of Coldest Quarter |
|---|---|

### Biome (Mega) classifications

| **Original biome classification** | **Mega-biome classification** |
|---|---|
| Tropical evergreen broadleaf forest<br>Tropical semi-evergreen broadleaf forest<br>Tropical deciduous broadleaf forest and woodland | Tropical forest |
| Warm-temperate evergreen broadleaf and mixed forest<br>Warm-temperate evergreen broadleaf forest<br>Warm-temperate rainforest<br>Wet sclerophyll forest | Warm-temperate forest |
| Cool evergreen needle leaf forest<br>Cool mixed forest<br>Cool-temperate rainforest<br>Cool-temperate evergreen needle leaf and mixed forest<br>Temperate evergreen needle leaf forest<br>Temperate deciduous broadleaf forest | Temperate forest |
| Cold deciduous forest<br>Cold evergreen needle leaf forest | Boreal forest |
| Temperate sclerophyll woodland and scrubland<br>Temperate evergreen needle leaf open woodland<br>Tropical savanna<br>Temperate deciduous broadleaf savanna | Savanna and dry woodland |
| Tropical xerophytic scrubland<br>Temperate xerophytic scrubland<br>Tropical grassland<br>Temperate grassland<br>Steppe<br>Xerophytic woods/scrub<br>Temperate grassland and xerophytic scrubland | Grassland and dry scrubland |
| Desert | Desert |
| Graminoid and forb tundra | Dry tundra |
| Cushion-forb tundra (cushion forb, lichen and moss tundra)<br>Erect dwarf-shrub tundra | Tundra |

| | |
|---|---|
| Low and high shrub tundra<br>Prostrate dwarf-shrub tundra<br>Tundra<br>Alpine grassland | |

(taken from http://pmip2.lsce.ipsl.fr/, accessed 24 May 2013 (Harrison and Prentice 2003)).

# APPENDIX 2: SUPPLEMENTARY MATERIALS

HomininSpace consist of a computer application, associated input files, simulation results and documentation. This thesis describes the study in which the HomininSpace simulation system was developed and with this thesis come Supplementary Materials, both in the appendices and in separate files collected on a single DVD disk. The content of this DVD is listed in Table 56. Furthermore this appendix contains the posters that were presented at conferences and which are referred to in this work. Such materials are normally not easily accessible and therefore included here. All materials available on the DVD are also accessible online, via a service made available by the Data Archiving and Networked Services (DANS): https://doi.org/10.17026/dans-28h-bysx. Data has been submitted using the online archiving system EASY, upon which data is archived according to the guidelines of the international Data Seal of Approval, the ICSU-WDS, and the NESTOR seal for Trustworthy Digital Archives. Moreover, the metadata fields in EASY comply with the guidelines of the Dublin Core standard[49] ensuring future availability.

### Contents of the HomininSpace data disk

The Table 56 lists the items on the disk that accompanies this study. Bold face headers are the names of directories on the disk, starting at the root. The actual filenames as included on the disk are given in italics.

**Table 56: Contents of the HomininSpace disk.**

| | | |
|---|---|---|
| **\Documentation** | | |
| | *HomininSpace ODD.docx / .pdf* | The Overview, Design and Details document for the HomininSpace model and simulation system. |
| | *HomininSpace – checkpoint database.docx* | Document describing the construction of the checkpoint database. Provides also reasons to include or exclude individual sites. |
| | *HomininSpace – thesis.docx / .pdf* | This thesis. |
| | *Climate reconstruction - Fulco Scherjon - labreport.docx* | Documentation on the use of SAM and related files. |
| | *HomininSpace_GE.kmz* | Google Earth input file with all checkpoints. |
| | *HomininSpace - user manual 1.6.docx / .pdf* | User manual describing the most common actions within the simulation system, including input and output data formats. |
| | | |
| **\Analysis** | | |
| | *Analysis results plots [scenario].pdf* | Graphical output of the statistical analysis, including correlation figures. For each scenario. |

[49] See http://easy.dans.knaw.nl/doc/pdf/DEP_All other disciplinesUK.pdf, verified 26 March 2018.

| | | |
|---|---|---|
| | *AllData.csv* | File with all simulations from all scenarios. |
| | *Batchresults_[scenario].txt* | Results for all simulations from a scenario. One file for each scenario. |
| | *[Correlation output]* | For each scenario files with correlation data for the Standard set, the Evolved set and the combined set with all simulations. |
| | *Top10 parametervalues.csv* | Overview of the best 10 simulations for each scenario. |
| | | |
| **\Model Data** | | |
| | *Neandertal sites – north west Europe.xlsx* | All checkpoints with associated intervals. Database with dating information for Neandertal sites for the simulation area. Also includes a list of MIS 6 sites (used for starting populations) and pure information points (that just register presence, but since no exact dating information is known, these are not used in validation of the results). Contains all checkpoints, monitoring checkpoints and starting locations of initial population groups. |
| | *SAM Grid a_1_12.txt* | Input grid cells. |
| | *Bintanja2008.txt* | Reconstructed temperature and sea level values. |
| | *HomininSpace Climate Grid.txt* | Climate data per grid cell. |
| | *HomininSpace Checkpoint List.txt* | The actual input file with the checkpoint information. Derived from the Neandertal sites Excel file. |
| | | |
| | **SAM input files** | |
| | *Etopo1_bedrock.zip* | Bedrock height values for selected area, XYZ format input file. |
| | *alt_2-5m_bil.zip* | Worldclim.org input files, ecological data for current and past climates in high resolution (2.5 arc minutes). |
| | *bio_2-5m_bil.zip* | |
| | *wc_2_5m_CCSM_21k_bio.zip* | |
| | *wc_2_5m_MIROC3.2_21k_bio.zip* | |
| | | |
| | | |
| **\Source code** | | |
| | *Workspace\HomininSpace* | Source Code which provides the Java source code of the HomininSpace modelling framework. This is a Repast workspace and contains next to the source also administrative files. |
| | | |
| | | |
| **\Scripts for statistical analysis** | | |
| | *Batch analysis HS results.R* | R script for statistical analysis of the combined simulation result files. |
| | *Batch print HS results.R* | R script to generate images that are used in the statistical analysis. |
| | *Analysis draw faces.R* | R script to generate the Chernoff faces of the top 10 scoring individuals. |

| | | |
|---|---|---|
| | *Climate calculations.xlsx* | Example calculations of secondary biomass. |
| | *HomininSpace Population Calculator.xlsx*. | Interactive tool to see the effect on population structure of different parameters. |
| | | |
| | | |
| **\Output data**<br>Files with batch simulation results | | |
| | *Summed data only.xlsx* | Excel sheet containing the summed results for all output variables for all relevant simulations for all scenarios. |
| | *HS Logfile 2018-11-29 0045.txt* | Example log file. |
| | | |
| | | |
| Movies with screen captures of running simulations | | |
| | *1.888.579.450 8-10-2013 50.avi* | |
| | *1.888.579.450 9-10-2013 50.avi* | |
| | *1.888.579.450 22-10-2013 50.avi* | |
| | *1.888.579.450 23-10-2013 50.avi* | |
| | | |
| | | |
| **\Review by Iza Romanowska** | | |
| | *HomininSpaceComments.docx* | Word document with the source code review results. |
| | *Re HomininSpace et al (email 2016-12-11).txt* | Verbatim email text with the overall impression and review results. |
| | *Earlier emails.txt* | Email exchange about the review. |
| | | |
| **\Applications**<br>Tools used in the framework or during development | | |
| | *Quickview* | Installation file for Quickview (needed to capture and view simulation movies). |
| | *Repast* | Installation files for Repast 2.2. |
| | *SAM* | Installation files for SAM 4.0 + documentation. |
| | | |

**Poster: *Neandertals on the move. Or not.* – ESHE 2015**

# NEANDERTALS ON THE MOVE. OR NOT.

## A multi-parameter investigation of mobility types in the deep past



Fulco Scherjon
Faculty of Archeology
Leiden University, The Netherlands
f.scherjon@arch.leidenuniv.nl



Figure 1 - HomininSpace: deconstructing the Neandertal model

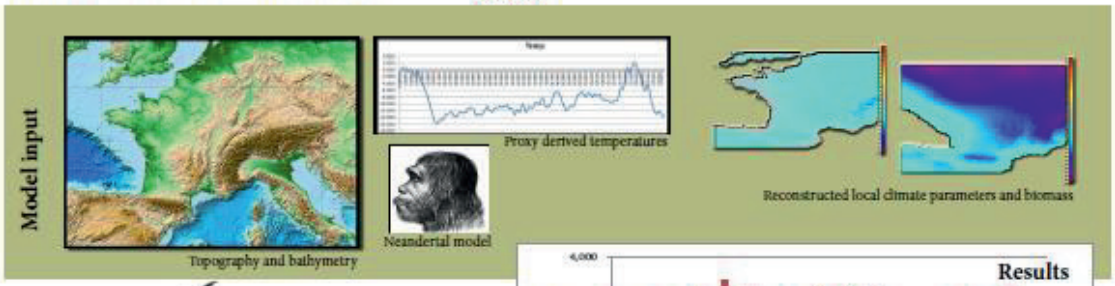
### HomininSpace

The aim of the HomininSpace modelling and simulation system is to model, as realistically as possible, past hominin movement and persistence through time, in a changing ecological and geographical environment. Key characteristics are:

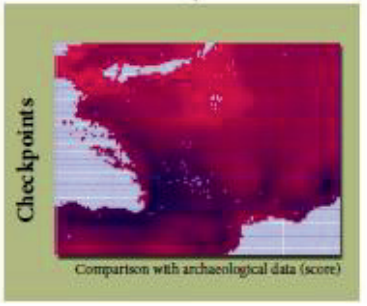
- Agent-based model (ABM);
- Fluctuating carrying capacity as key attractor;
- Parameterized and spatially explicit reconstructed paleo-environment;
- Year-by-year demographic model;
- Comprehensive database of archaeological checkpoints in space and time.

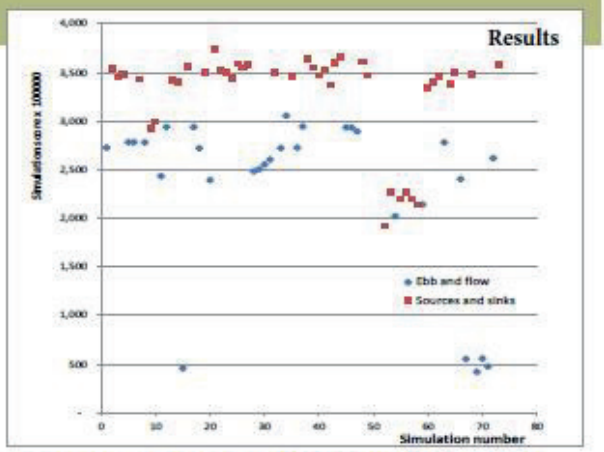
Hypothesis testing is done by comparing simulation results of hominin group presence in different scenario's against the archaeological record (simulation score). Results illustrate the effects of different parameters when reconstructing hominin behavior in the past (Figure 1). Both "ebb and flow" and "sources and sinks" mobility types have been used to explain the Neandertal record of presence and absence (Roebroeks *et al.* 2011) and have been implemented within the HomininSpace simulation system.

### Model




### Simulations




### Conclusion: Sources and sinks mobility matches archaeology best

Comparing the archaeology with the results for the two mobility types suggest that the "sources and sinks" model matches the archaeological record best. Competition for available resources is a major factor influencing mobility strategies. Possible explanations why the "sources and sinks" model scores higher and increases population levels consistently:

- When resources become scarce, hominins can sometimes survive, albeit at very low densities;
- During ebb and flow movement, groups will all converge on areas with high productivity;
- Sink populations will utilize the environment optimally, by being forced to use resources from areas that flow groups do not enter or leave when conditions become less favorable.

The implemented "sources and sinks" mobility model explains how repeated regional extinction could have been an important factor in the demography of Neandertals. The underlying results also illustrate some of the advantages of ABM over other types of modelling.

### Selected references

[illegible]

HUMAN ORIGINS GROUP
FACULTY OF ARCHAEOLOGY - LEIDEN UNIVERSITY
www.humanoriginsleiden.org

**Poster: *How (not) to model Neandertal extinction*. –ESHE 2016**

# How (not) to model Neandertal extinction

Fulco Scherjon
Faculty of Archeology
Leiden University, The Netherlands
f.scherjon@arch.leidenuniv.nl

Krist Vaesen
Faculty of Archeology
Leiden University, The Netherlands
Eindhoven University of Technology
k.vaesen@tue.nl

Alexander Verpoorte
Faculty of Archeology
Leiden University, The Netherlands
a.verpoorte@arch.leidenuniv.nl

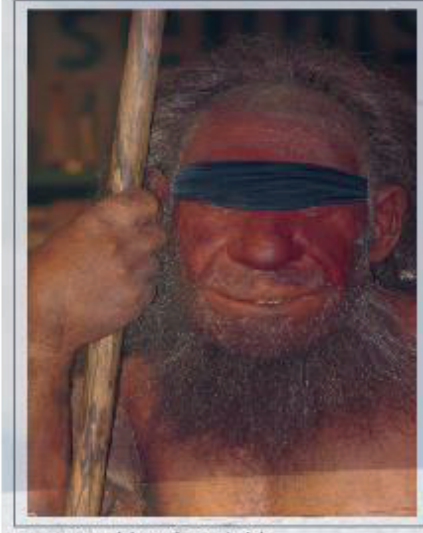


Figure 1 - Unequal chances for Neanderthals.

## If you handicap the competition, do not be surprised to win...

We analysed existing simulation models for Neanderthal demise. Each of the models assumes a large difference between Neanderthals and AMH (Fig. 1). Evidently, under such conditions the replacement of Neanderthals by AMH is inevitable (Fig. 2). Each model we looked at implements an unsubstantiated (relative) handicap for Neanderthals:

- increasing mortality rates for Neanderthals [6], and decreasing rates for AMH [1];
- decreasing growth rates for Neanderthals, but increasing rates for AMH (due to differences in technological development [3,4, 5, 7] or in social organization [2]).

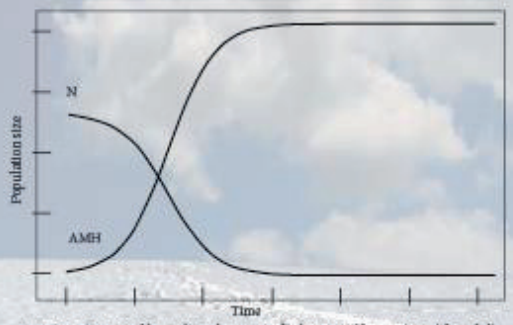



Figure 2 - The inevitable result: replacement of N by AMH (figure adapted from [7]).

| Reference | Handicap (assumptions) | Results | Model type |
|---|---|---|---|
| Zubrow 1989 [1] | Seemingly small but effectively large demographic differences, with lower birthrate for N. | Large discrepancies in mortality and life expectancy, with an increment in N mortality of 1–2 % leading to extinction within as little as 30 generations. | population model |
| Horan et al. 2005 [2] | Increasing social organisation for AMH. | Unavoidable replacement of N by AMH after varying amount of years. | mathematical |
| Flores 2011 [3]<br>Flores 1998 [7]<br>Murray 2004 [5] | Higher mortality rate for N. | N are replaced by AMH (model implemented by Murray 2004). Flores 2011 uses the computed similitude from [7] to calculate diffusion coefficients. | nonlinear mathematical model |
| Gilpin et al. 2016 [4] | Lower cultural level interacting with population size. | Replacement of larger N population by AMH. | mathematical |
| Sorensen 2011 [6] | 1% reduction in death during childbirth and in death by hunting incidences between AMH and N. | Population growth amongst AMH despite adverse climatic conditions ~45 kya. | mathematical |

## Conclusion: How to model Neanderthals?

We introduce a conservative model which assumes no difference between N and AMH and uses only population size. This null model suggests that stochasticity and allee effects (i.e., populations at low numbers are affected by decreasing growth rates) might have been sufficient for the Neanderthals to have gone extinct (Fig. 3).

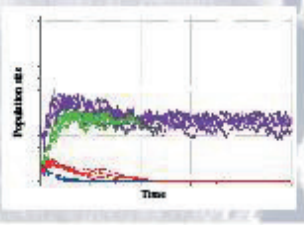



Figure 3 - Preliminary results for different population sizes of N, including allee effects. Green and Blue illustrate critical population sizes, where Purple represents survival and Red extinction levels.

References

[1] Zubrow 1989
[2] Horan et al. 2005
[3] Flores 2011
[4] Gilpin et al. 2016
[5] Murray 2004
[6] Sorensen 2011
[7] Flores 1998

Human Origins Group

The research was supported by the Netherlands Organisation for Scientific Research (N.W.O. VIDI Grant 016.144312) (KV) and the N.W.O. Spinoza program (FS).

### Comparing past climate models: CCSM versus MIROC

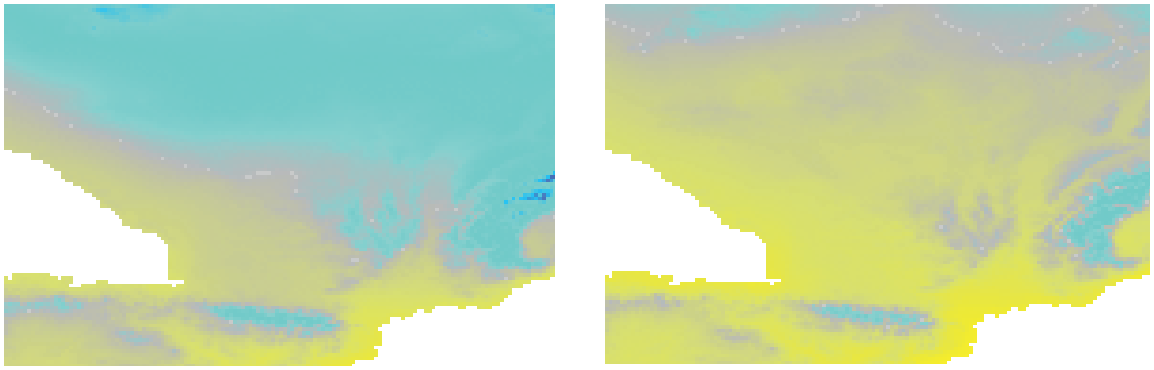
The LGM period is part of the focus of the Paleoclimate Modeling Intercomparison Project (PMIP, Joussaume and Taylor 1995) which is an international effort to evaluate models using widely varying climate conditions. Climate for this defining period in the past has been extensively documented using a diverse array of paleo-climate proxy data sets and is used to calibrate (new) climate models. The LGM climate differs significantly from today and can be interpreted as an extreme climate anomaly for which there is no present day analogue and which is difficult to reconstruct in simulations, but for which the major forcing factors are relatively well known. The LGM is a severe glacial period with changes in greenhouse gases, sea level and ice sheets. As such, climate parameters can be used to model extreme climate changes (Prentice *et al.* 2000). Due to the fact that conditions for the LGM are well studied, results from simulations including this period can be compared against other modelling results (Otto-Bliesner *et al.* 2006) and thus the dataset as a whole gains in confidence.

WorldClim provides two sets of data for the Last Glacial Maximum (LGM, around 21ka): the Community Climate System Model (CCSM) and the Model for Interdisciplinary Research On Climate (MIROC). Both are coupled general circulation models, with MIROC developed at the Center for Climate System Research (CCSR) at the University of Tokyo (Hasumi and Emori 2004) and CCSM designed at the University Corporation for Atmospheric Research (UCAR) (Collins *et al.* 2006). MIROC couples models for atmosphere, land, river, sea ice and ocean where CCSM couples models for land, sea-ice, ocean and atmosphere. MIROC is available upon request for collaborative researchers (http://ccsr.aori.u-tokyo.ac.jp/~yangpeng/source_center.htm, accessed 25 July 2013), where CCSM is actively developed into a successor named Community Earth System Model (CESM, http://www2.cesm.ucar.edu/, accessed 25 July 2013). Both are used primarily for future climate predictions, for instance by the Intergovernmental Panel on Climate Change (IPCC) (Hamilton 2007).

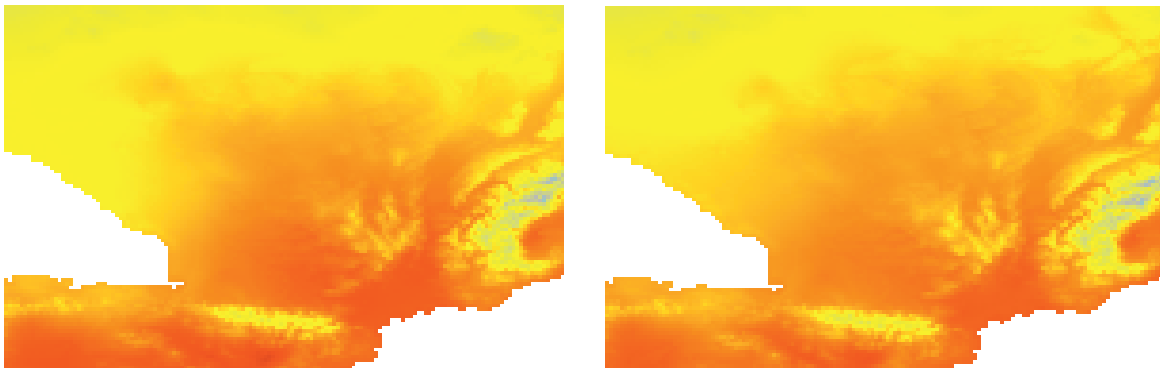
The reconstructed LGM environments from CCSM and MIROC are different. See Figure 94 for a visual comparison of the most relevant datasets. Especially temperatures as simulated by CCSM are generally lower than those from MIROC, but precipitation levels are also lower. These differences are likely to influence simulation results (Olfert *et al.* 2011). Both models are however among the best for modelling LGM circumstances (Maris *et al.* 2012; Otto-Bliesner *et al.* 2007). Instead of averaging the results of the two models as done by Waltari *et al.* (2007), the underlying research uses one single model. CCSM has

been selected due to the fact that the available data from MIROC were created using version 3.2 which contains an error in the land surface scheme. Correcting this error, as has been done in 3.2.2, would lower the temperature values (Maris *et al.* 2012, 804).

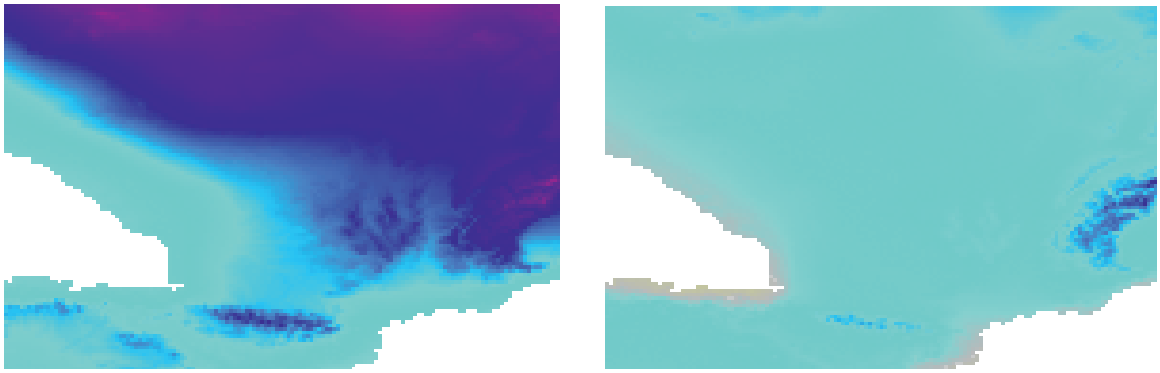
The chosen colour scheme for the temperature maps (values in Table 3) is constructed based on the colours used by The Weather Network (example in http://scsjournal.files.wordpress.com/2011/11/the-weather-network.jpg, accessed 21 July 2013). The colour scheme for the precipitation values is inspired by the Canadian weather reports (http://scsjournal.wordpress.com/, accessed 21 July 2013). Note that the temperature values in the WorldClim database must be divided by 10 to get degrees Celsius.



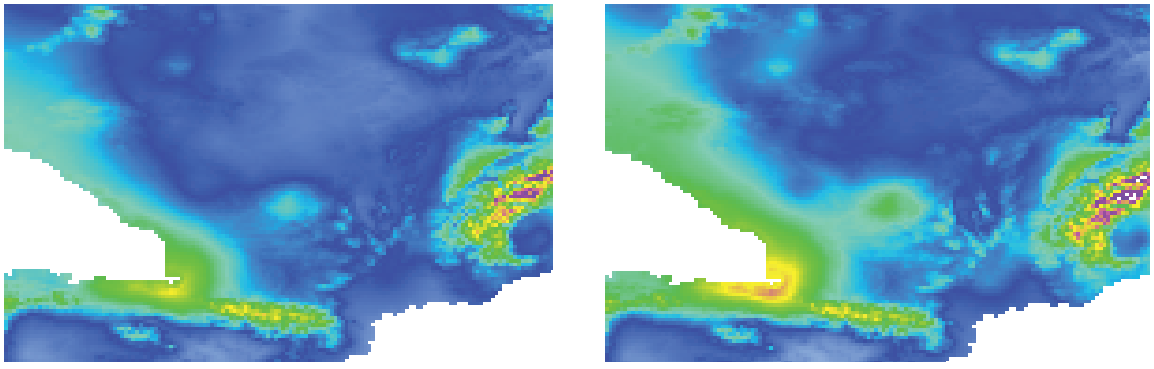
BIO1 = Annual mean temperature



BIO5 = Max temperature of the warmest month



BIO6 = Min temperature of the coldest month



BIO12 = Annual precipitation

**Figure 94: Overview of the results for the main biological values for CCSM (left) and MIROC climate models with unified colour coding schemes.**

# SUMMARY

This study presents an agent-based simulation model exploring the patterns of presence and absence of Late Pleistocene Neanderthals in western Europe. HomininSpace implements a parameterized generic demographic and social model of hominin dispersal while avoiding parameter value biases and explicitly modelled handicaps. Models are simulated through time within a high-resolution environment where reconstructed temperatures and precipitation levels influence the carrying capacity of the landscape. Model parameter values are assigned and varied automatically while optimizing the match with Neanderthal archaeology using a Genetic Algorithm (GA) inspired by the processes of natural selection. The system is able to traverse the huge parameter space that is created by the complete set of all possible parameter value combinations to find those values that will result in a simulation that matches well with archaeological data in the form of radiometrically obtained presence data.

This study further implements and explores with the HomininSpace tool ten questions that are unresolved in the current debate on Neanderthals. Modelled Neanderthal hominins were capable of using the environment to their benefit irrespective of energy distribution or dispersal modus. Adding coastal resources to the model allows a more relaxed demographic sub-model with more realistic parameter values and less physical stress for female Neanderthals. Adding the ability to cross open water does not influence the simulation scores much, but allowed hominins to reach England earlier. With the inherent flexibility of the model the hominins appear to have difficulty staying out of England in most simulations.

Almost 40,000 simulations were executed for a total of nearly 3.3 billion simulated Neanderthal years. Very short lived and fast reproducing Neanderthals score well in certain scenarios. High birth rates and low mortality rates for (pre) fertile hominins were successful in others. Individuals of post-fertile conditions were generally selected against. Subsistence requirements could vary widely but were met with large foraging ranges, increased energy extraction, or simply by survival in very low numbers. The power of this model to recover quickly after setbacks excludes the need for implementing unwarranted hypotheses about either environment or behaviour of past hominins. With HomininSpace a new and powerful tool to explain archaeological patterns and explore new and old questions is added to the toolset of archaeologists and paleoanthropologists.

# SAMENVATTING

HomininSpace is een simulatie systeem waarmee groepen van jager-verzamelaars gesimuleerd kunnen worden in een gereconstrueerd, realistisch landschap. Met behulp van deze tool heb ik gekeken naar de mogelijke oorzaken voor aan- en afwezigheid van Neanderthalers in het westen van Europa, min of meer het hedendaagse Frankrijk. De studie concentreert zich op het late Pleistoceen, onder andere omdat er voor deze periode relatief veel data voorhanden is. Het model wat is geïmplementeerd in HomininSpace is een generiek verspreidingsmodel, met demografische en sociale parameters. Het systeem is ontworpen om zelfstandig op zoek te gaan naar de meest waarschijnlijke parameter waarden, zonder dat daarbij input van de gebruiker nodig en gewenst is. Hiertoe wordt gebruik gemaakt van een genetisch algoritme. Hierdoor kunnen in de gigantische parameter ruimte die waarden gevonden worden die tot een goede match met de archeologie leiden. Simulatie resultaten worden vergeleken met een database met 470 directe dateringen van Neanderthaler aanwezigheid verspreid over 83 archeologische sites.

Met de HomininSpace tool worden een aantal relevante openstaande vragen uit het vakgebied van de Paleolithische archeologie verkend. Bijvoorbeeld over het gebruik van voedsel uit de kustgebieden, wat niet zozeer tot meer matches leidt maar wel meer realistische parameter waarden toelaat (bijvoorbeeld geboortecijfers onder de 40%). Of het kunnen oversteken van open zee waardoor vooral Engeland toegankelijker wordt. Waarbij direct moet worden aangetekend dat het model veel moeite heeft om überhaupt uit Engeland weg te blijven, waardoor de archeologisch aangetoonde perioden van afwezigheid voor dit gebied mogelijk opnieuw geanalyseerd moeten worden.

In totaal zijn ongeveer 40.000 simulaties uitgevoerd, waarbij elke simulatie een periode van 81.000 jaar beslaat. In totaal zijn er dus ongeveer 3,3 miljard Neanderthaler jaren gesimuleerd. “Live fast and die young” Neanderthalers waren succesvol in sommige simulaties, in andere waren juist hoge geboortecijfers en lage sterftecijfers belangrijk voor een goede match. Over het algemeen worden periodes van voedseltekorten overbrugd door het vergroten van de gebieden waarin voedsel gevonden wordt, door het meer efficiënt verwerken van voedsel of simpelweg door te zorgen dat er maar heel weinig individuen het voedsel moeten verdelen. Met HomininSpace hebben archeologen en paleoantropologen een flexibel gereedschap gekregen waarmee nieuwe en oude vragen beantwoord kunnen worden.

# ACKNOWLEDGEMENTS

Major parts of the model used in this research were built during my research master and doctoral work at the Leiden University. Now that this thesis is completed there are many people to be thanked. First of all acknowledgement and gratitude to those who made this research possible, and who advised, supported and believed in me: Wil Roebroeks, Alexander Verpoorte, Kathy MacDonald, and Hans Kamermans. The research was supported by the Netherlands Organisation for Scientific Research (N.W.O.) Spinoza program awarded to prof. dr. Wil Roebroeks who enabled my employment as a researcher and PhD candidate at Leiden University.

I wish to express my genuinely profound thanks to fellow researcher and friend Alison Smith with whom many stimulating discussions helped shaping my ideas, and who kept a very kind and supporting interest towards my work. We both moved on but stay connected. Many thanks are given to my colleagues for inspiring and sometimes heated discussions, interspaced with encouraging moments of fun and insight: first of all Femke Reidsma for lending more than one ear and restarting my computer after every forceful shutdown on Mondays. In random order I would also like to thank Andy Sorensen for the simulated fire indicators, Gerrit Dusseldorp for the optimal tea talks, Marie Soressi for inspiring innovations, Morgan Roussel for deep lithic insights, Jose Joordens for reminding me of Africa, Jac Aarts for access to DNA analysis, Eduard Pop for graphical advise, Krist Vaesen for modelling insights and reviewing my hominin model, Corrie Bakels for reviewing the reconstructed environments, Iza Romanowska for scrutinizing my code, Karsten Lambers for addressing computational archaeology, Lia Hemerik for her matrices, and Shumon Hussain for all those Ahmarian dates! I thank Iza also for the implementation of HomininSpace Lite, a replication effort that provided many insights! And Andy, Marie and Morgan for providing additions to my checkpoint list. Alexander, your critical and just reviews cannot be given enough appreciation.

Mijn grootste dank gaat uit naar Marielle Jonkman, mijn vrouw en levenspartner. Zij heeft mij altijd gesteund en ontzettend veel in mij geïnvesteerd. Door haar passie en goedhartigheid kan ik mijn dromen waar maken. Dit werk is opgedragen aan mijn zoon, Thor Scherjon. Vanaf nu geen virtuele Neanderthalers meer mee op vakantie!

I alone remain responsible for any and all errors.

# CURRICULUM VITAE



I, Fulco Scherjon, was born on the second of January 1969 in Amsterdam, The Netherlands. I was originally trained as a computer scientist and information analyst at Delft University (MSc). During an extensive career in software design and development I worked in many diverse projects for different employers that included river flow simulation software, database for planning of promotional activities, and eye care screening software.

I obtained a research master (MA Archaeology, Human Origins) in 2013. With this background I am working together with other researchers in a wide range of projects where computing knowledge is useful. Examples include a database constructed for lithic analysis, principal component analysis performed on experimental data sets, and research in which genetic data from extinct hominins has been analyzed for specific base patterns. My main research subject is modelling and simulation in an archaeological context.

My teaching experience started at Delft University, where I taught as a student assistant different programming languages include C++, Scheme and Pascal. At Leiden University I participated in many courses including “Data Analysis and Interpretation”, “Palaeolithic Europe”, and “Advanced Themes in Archaeology”. I obtained a 15k€ grant from Leiden University to develop, record and teach a Small Private Online Course (SPOC) titled “Modelling and Simulation for Archaeologists”, with Iza Romanowska and Karsten Lambers. I followed the Honours Class Archaeology in 2008, was admitted to the Honours College Linguistics in 2010, and obtained the Basis Kwalificatie Onderwijs (BKO, university teaching licence) in 2017.

## Selected publications

2017 Sorensen, A.C. and F. **Scherjon**, 2017. fiReproxies: A computational model providing insight into heat-affected archaeological lithic assemblages, *PLoS ONE* 13(5):e0196777.

2016 Aarts, J.M.M.J.G., Alink, G.M., **Scherjon**, F., MacDonald, K., Smith, A.C., Nijveen, H. and W. Roebroeks, 2016. Fire Usage and Ancient Hominin Detoxification Genes: Protective Ancestral Variants Dominate While Additional Derived Risk Variants Appear in Modern Humans, *PLoS ONE* 11: e0161102.

2015 **Scherjon** F., Bakels C.C., MacDonald K. and W. Roebroeks, 2015. Burning the Land. An Ethnographic Study of Off-Site Fire Use by Current and Historically Documented Foragers and Implications for the Interpretation of Past Fire Practices in the Landscape, *Current Anthropology* 56(3): 299-326.

## Education

2014 **Research Master of Arts in Archaeology**. Thesis: *HominінSpace – Modelling and Simulating Hominins moving through Time and real geographical Space.* Supervisors: Prof. Dr. J.W.M. Roebroeks and Dr. K. MacDonald. Grade: 9. Leiden University, the Netherlands. Simulation system developed in Repast and written in Java.

1995 **Master of Science in Applied Mathematics and Informatics**. Thesis: *Timing Analysis – A Case Study.* Supervisors: Prof. Dr. Ir. J. van Katwijk, Dr. Ir. W.J. Toetenel and Drs. Ing. K. Brink. Grade: 8. Delft University of Technology, the Netherlands. Research includes an explicit model of the i86 processor and an automated source code analysis

# Virtual Neanderthals

This study presents an agent-based simulation model exploring the patterns of presence and absence of Late Pleistocene Neanderthals in western Europe. HominínSpace implements a parameterized generic demographic and social model of hominin dispersal while avoiding parameter value biases and explicitly modelled handicaps. Models are simulated through time within a high-resolution environment where reconstructed temperatures and precipitation levels influence the carrying capacity of the landscape, in the number of edible ungulates that can be hunted. Model parameter values are assigned and varied automatically while optimizing the match with Neanderthal archaeology using a Genetic Algorithm inspired by the processes of natural selection.

