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

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

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.

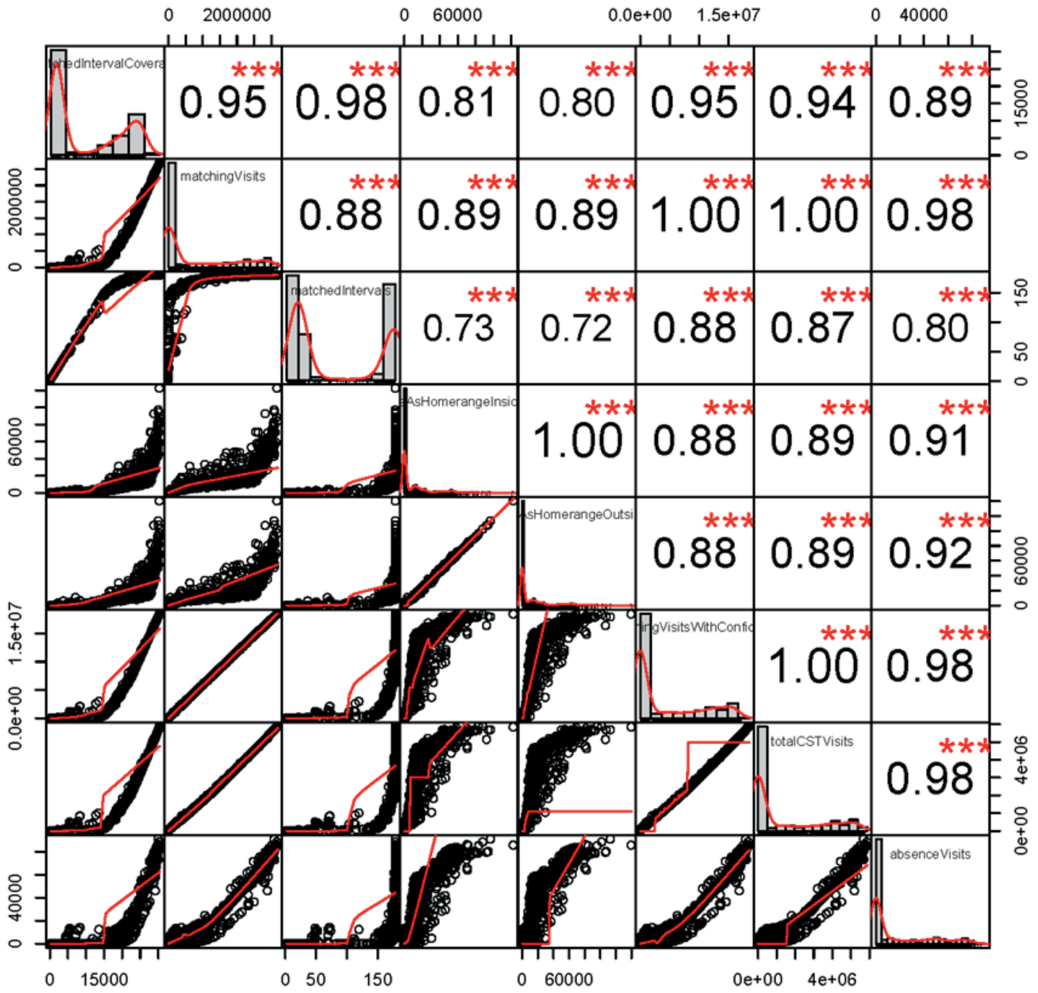

Figure 86: Correlations between all visit counting output variables for scenario Habitat-A. For instance, the correlation between matchedIntervalCoverage (simulation score) and matchedIntervals (count of all matched intervals) is 0.98, and very significant (three red stars, as have all correlations).

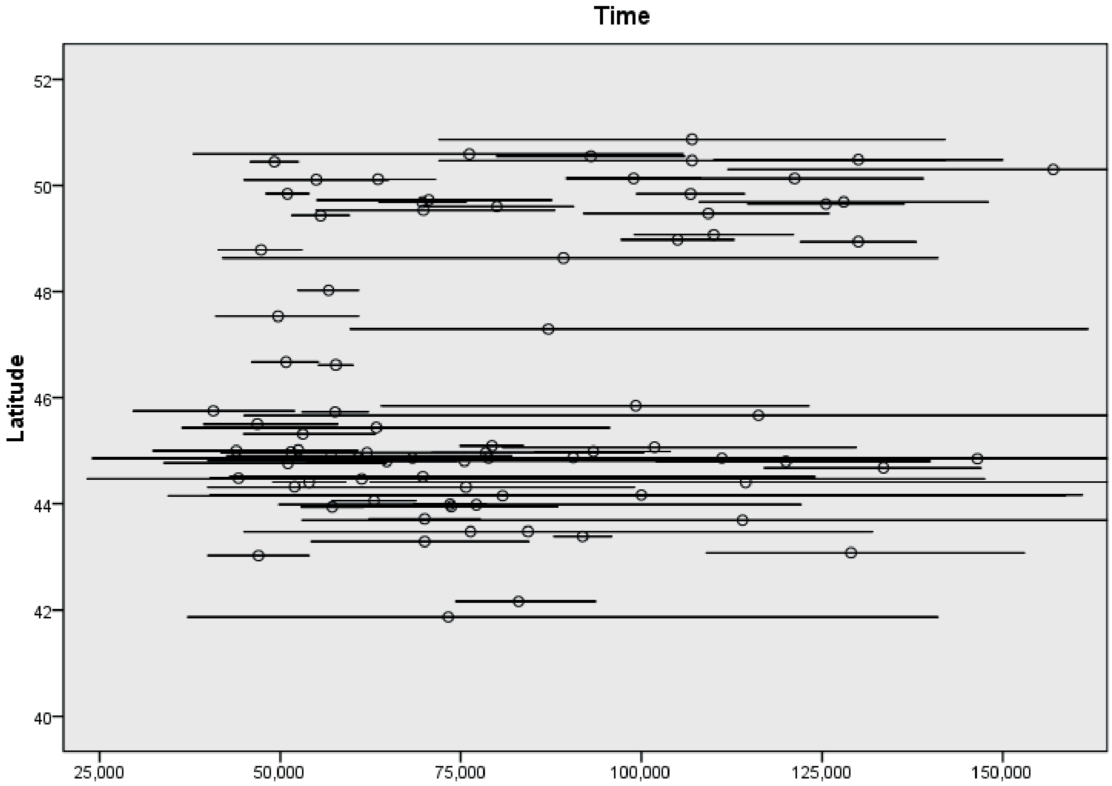

Figure 87: Illustrating the bias in the checkpoint dataset: overview of the used intervals in the simulations. X-axis gives Time, with each interval defined by starting point and end point in years ago. Vertical axis is the latitude of the site.

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.

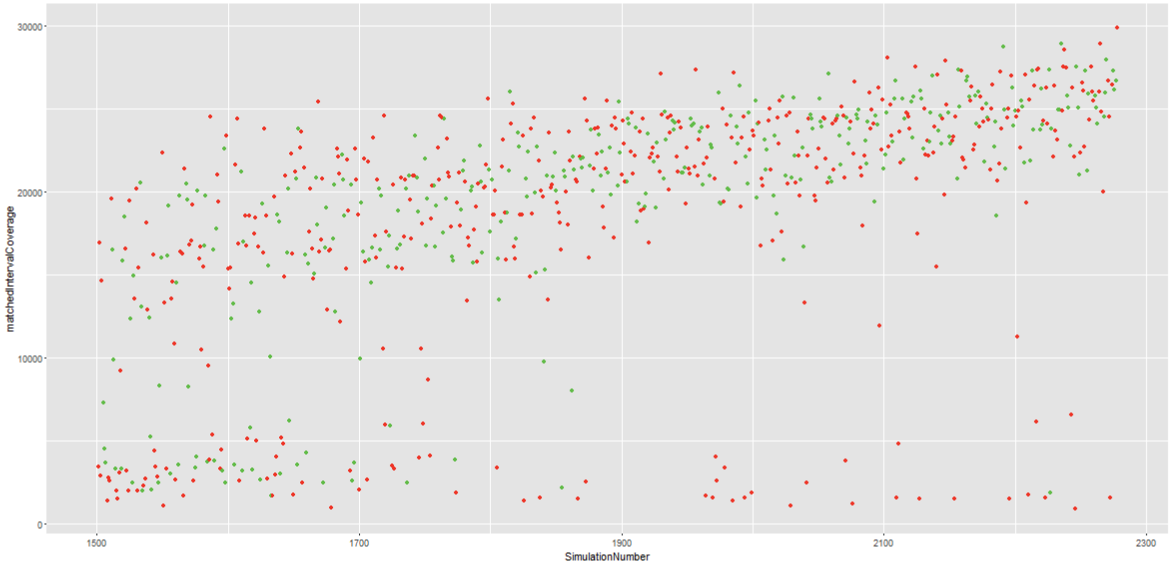

Figure 88: Mapping models resulting from mutation (in green) and combination (in red). Horizontal axis the simulation number (from 1500 onwards), the Y-axis presents the simulation score. Data from the Evolved set in scenario Energy-B.

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 HomininSpace

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 HomininSpace. 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.

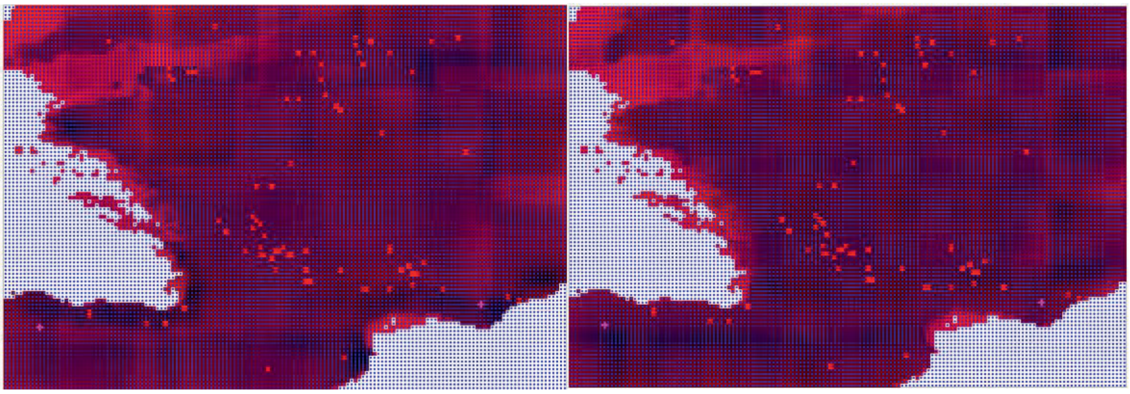

Figure 89: Illustrating the effects of using coastal resources. Left simulation 1384 with coastal resources turned on (scenario Habitat-J). On the right the same simulation with coastal resources disabled (Habitat-I).

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.

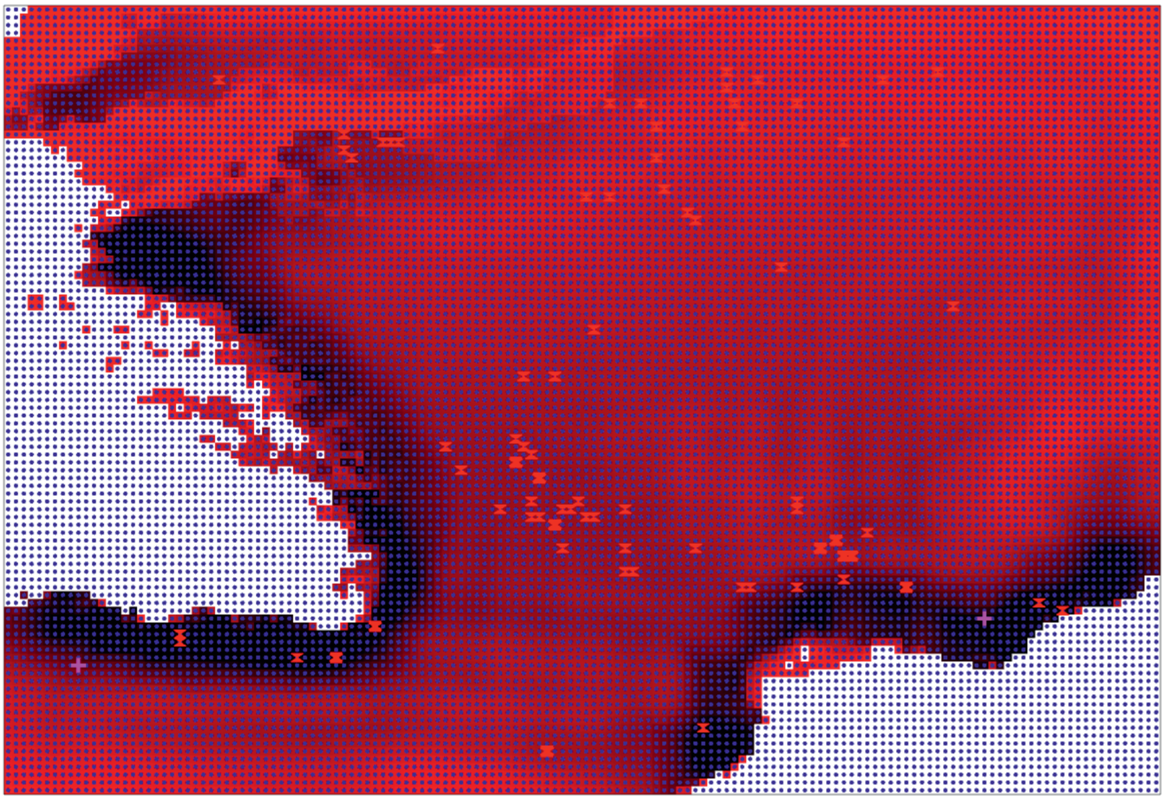

Figure 90: The attractiveness of coastal resources. Simulation 1365 from Scenario Energy-B but with dynamic hominins.

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).

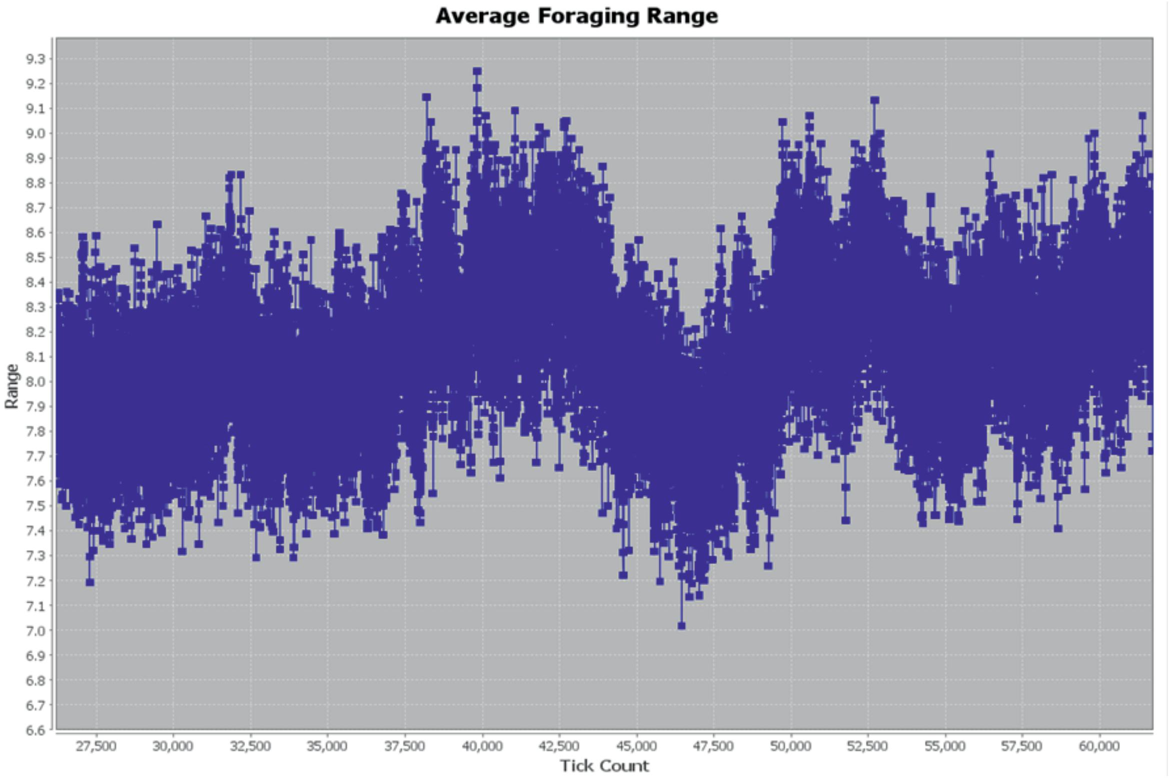

Figure 91: Overview of the average foraging range for part of the simulation period. Largest value around 9.3, minimum value around 7.0, standard deviation around 2. Figure created in HomininSpace, by calculating the average foraging range of all present groups for each time step. Simulation 1384, scenario Habitat-I, the maximum value for the foraging range is 11.

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.

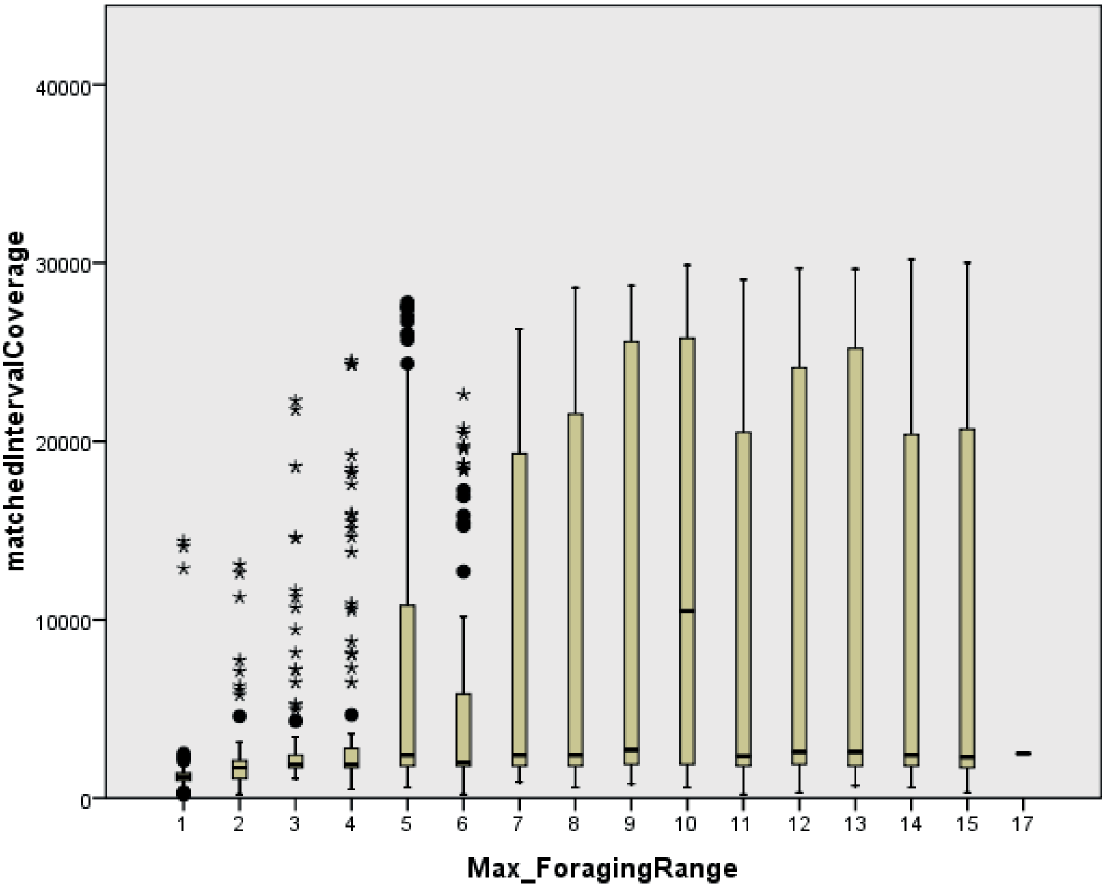

Figure 92: Simulation score versus maximum foraging range. For low values of the foraging range good scores are difficult to obtain. Increasing the maximum beyond 9 does not influence the score much. Values are from scenario Habitat-C.

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.

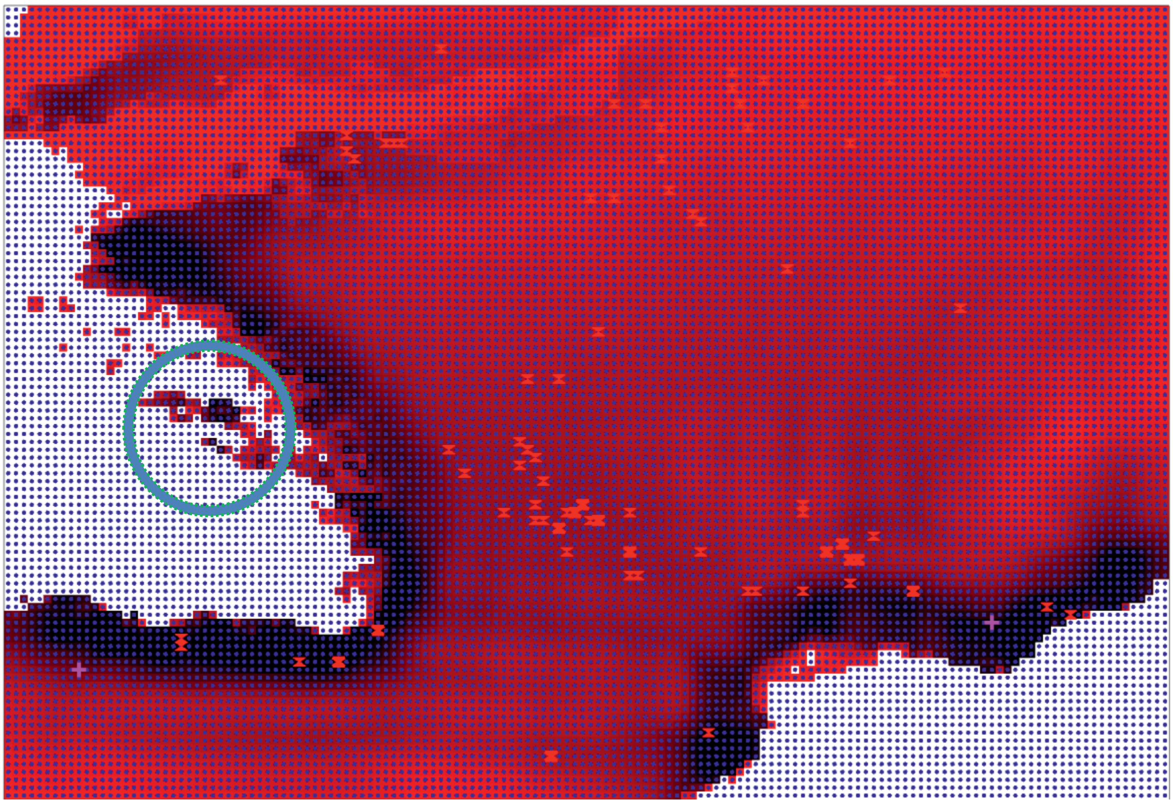

Figure 93: Illustrating the ability to cross larger waterbodies to some of the islands off the coast of France. Produced in Scenario Energy-D with dynamic hominins, simulation 1365.

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.