



UNIVERSITY OF NEWCASTLE
SCHOOL OF MATHEMATICS AND STATISTICS

MMATH PROJECT REPORT

Prehistoric Population Dynamics and the Neolithic in Europe

by Kieren Charles

Supervised by
Prof. Anvar Shukurov & Dr. Andrew Baggaley

May 1, 2015

Abstract

From about 7000 to 4000 BC in Europe, the transition from the Mesolithic to the Neolithic era occurred. This was a key part of human history and saw the change from hunting and gathering to agriculture and stock breeding. This allowed for the support of a much larger population. Radio carbon dates and archaeological evidence suggest the source of the spread of farming was in the Near East. Processing of said radiocarbon dates has allowed us to produce a mathematical model which is presented here. All computations were performed with MATLAB.

Contents

1	Introduction	2
1.1	The Stone Age	2
1.1.1	Paleolithic	3
1.1.2	Mesolithic	3
1.2	Neolithic	3
2	Modelling	5
2.1	Reaction - Diffusion	6
2.2	Our Model	7
2.2.1	The FKPP Equation	7
2.2.2	Simplistic 1 Dimensional Model	7
2.2.3	Generalising to 2 Dimensions	9
2.2.4	Checking the Propagation Speed	11
2.2.5	Implementing the Landscape	12
3	Radiocarbon Dates	19
3.1	Selection of radiocarbon dates	19
3.2	Comparison with our model	21
4	Conclusion	25

Chapter 1

Introduction

1.1 The Stone Age

The Stone Age itself contains three main time periods; the Paleolithic, the Mesolithic and of course, the Neolithic. This covers a broad period of time from around 2.5 million years ago where *Homo habilis* started using the earliest stone tools (William *et al.*, 2012) extending up until around 3000BC in Europe with the start of the Bronze Age and the advent of metalworking (Gimbutas , 1973). Figure 1.1 shows the timeline of the three eras of the Stone Age.

Very recent research actually suggests the earliest stone tools came around 3.3 million years ago, this date pre dating the first members of the *Homo* genus (Callaway , 2015). This has proved shocking and leaves a lot of questions open about that time period; if the Stone Age is to continue to meet its definition then it is possible the beginning of the Stone Age could be moved back.

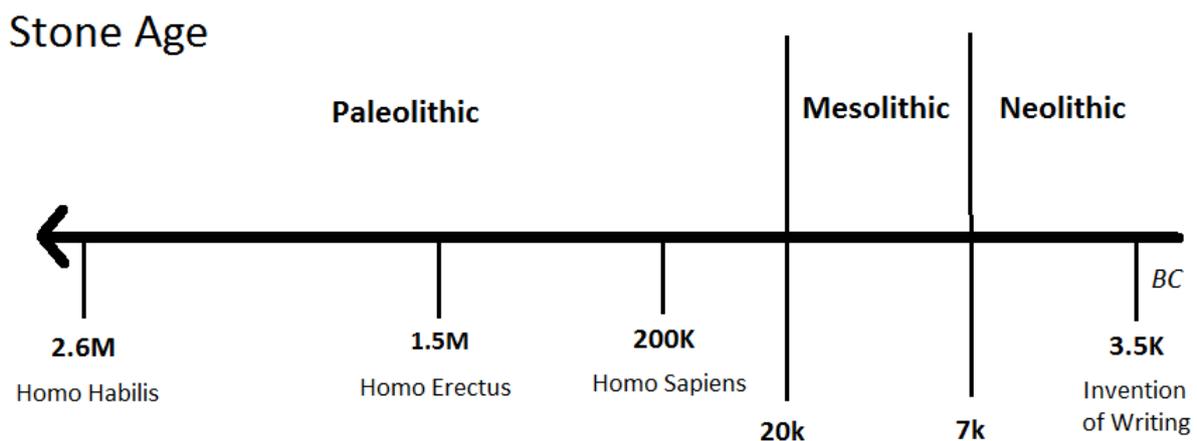


Figure 1.1: Timeline of the Stone Age showing some major events including the invention of writing, which signalled the beginning of history, before which, we call prehistory.

1.1.1 Paleolithic

The Paleolithic era spans from the beginning of the Stone Age about 2.5 million years ago to around 20,000 BC. Many significant advancements were made throughout this period, starting with homo habilis, by the end of this era, early homo sapiens had evolved. Humans generally lived as nomadic hunter and gatherers in only very small groups with no leader, though groups with more advanced food storage techniques were larger and occasionally had leadership structures (Nettle and Romaine , 2002). This clearly highlights the link to having a stable food supply being necessary for a larger population.

1.1.2 Mesolithic

The Mesolithic is the period up until the advent of agriculture, and hence varies considerably in different locations. The start of this era is associated with the end of the last ice age. Technologies had advanced considerably by this point, including the first small composite flint tools and fishing tackle. For Europe, where this report is concentrating on, the transition from the Mesolithic began at around 7000 BC.

1.2 Neolithic

The arrival of the Neolithic to Europe was a key period in the history of the European population, specifically the introduction of agro-pastoral farming (the growing of crops and the raising of livestock). This allows for a much larger population to be supported, in-fact 50 times larger (Davison *et al.*, 2009). The Neolithic era was the last part of the Stone Age. There were points where the Neolithic can be pointed to originating from all over Earth however there are variations in advances associated with different Neolithic societies. As well as agro-pastoral farming, the advent of pottery making and making use of polished stone and bone tools and large-scale manufacture of ceramic ware. Different regions had some of these advancements as can be told by the radiocarbon data. An example being the Neolithic culture in the Levant, Upper Mesopotamia, Anatolia (9800-7500 BC) where evidence shows that wheat and barley were cultivated there and permanent houses with stone foundations were used. At these sites however there is no evidence of pottery (Perles , 2001).

Figure 1.1 below shows some of the known source sites for the Neolithic but for this paper in particular we are interested in Europe and so attempt to create a model based upon the source in the Neart East at Jericho. Davison *et al.* (2009) suggests there is another source that has an impact on Europe in the Ural mountains at 8200 BC. While keeping this in mind, we are only looking here at the Near East source.

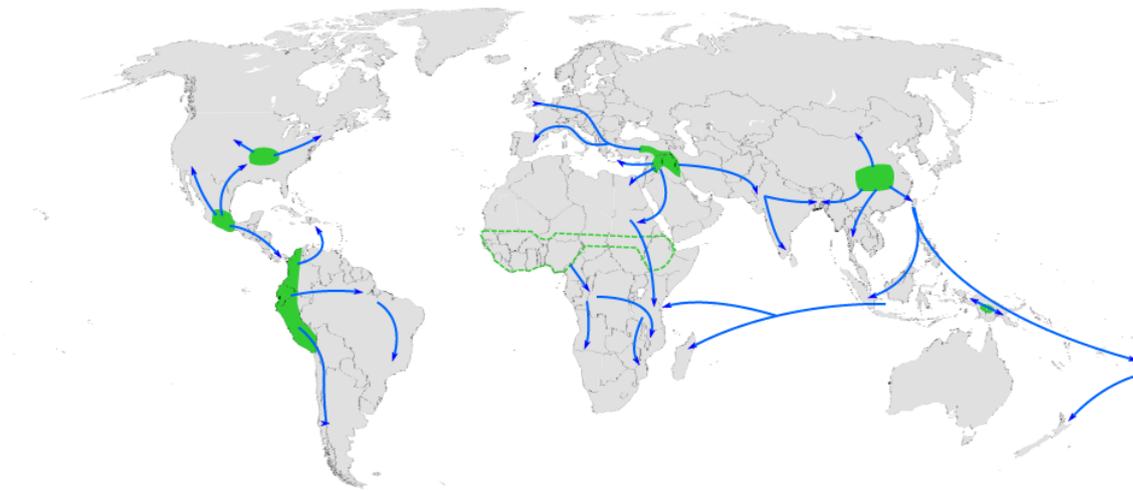


Figure 1.2: Shown in green are some of the known source sites for the Neolithic. The blue arrows indicate which direction the Neolithic population spread from its related source. The dotted green area in Sub-Saharan Africa is such because the exact source location is unknown. (Diamond J., Bellwood P., (2003). *Farmers and Their Languages: The First Expansions*, *Science* **300** 5619).

One interesting point is to look at the arrival of the Neolithic into Spain. There are two possibilities here, those being whether it arrives first from throughout Europe or from across the Mediterranean in Africa. Essentially this would cause two spreads to be occurring in Western Europe which can change results considerably. The comparison to radiocarbon dates allows us to make a conclusion.



Figure 1.3: Various different Neolithic artifacts, including axe heads, bracelets, chisels and polishing tools. Stone implements are identified as Neolithic as they are polished.

Chapter 2

Modelling

The underlying process in which the Neolithic population spreads through Europe can be considered as a 'random walk' process. This is irrespective to the real world mechanism of the spread however there is considerable interest in said mechanism. This remains controversial and there is no agreed conclusion here. Early on, the idea of direct migration of the farming population was used to create a wave of advance (Ammerman and Cavalli-Sforza, 1973). This could be viewed as the farmers either overwhelming the hunter-gatherer population or simply converting them (Price, 2000). Alternative views include the indigenous hunter-gatherers adopting Neolithic attributes through the diffusion of cultural novelties through intermarriages and assimilation (Thomas, 1996). Work recently with genetic evidence seems to favour cultural transmission (Haak *et al.*, 2005).

The 'random walk' process is such that at each point in time, each individual will take a step in any given direction with equal probability. Basically they are equally as likely to take a step forwards as they are to take a step backwards, left or right. This leads to an isotropic random walk. Figure 2.1 demonstrates a simple example.

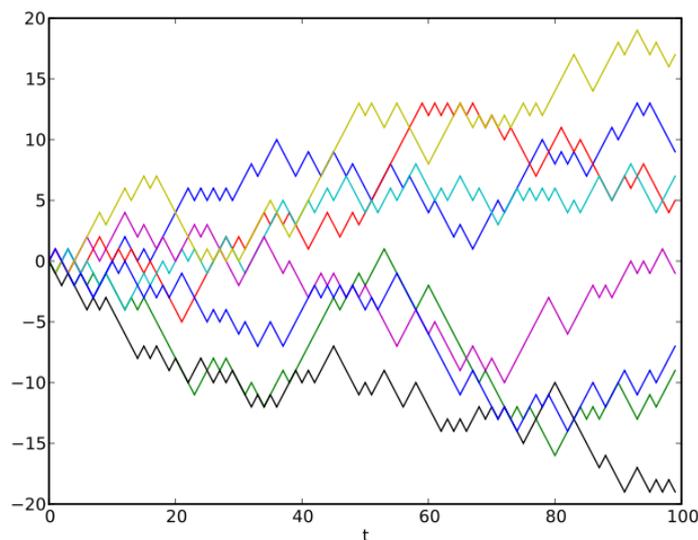


Figure 2.1: A simple example demonstrating the 'random walk' process. Eight different paths are shown progressing in time. At each time step they have an equal probability of increasing by one or decreasing by one.

Considering the spread as a 'random walk' process, mathematical modelling of it can be based on a reaction-diffusion equation.

2.1 Reaction - Diffusion

Reaction-diffusion equations are mathematical models used to explain the concentration of some variable (say human population or gas) and how it changes over time with respect to two process. Firstly, a reaction process, commonly a chemical reaction in which one substance is being created or in our case, the reproducing Neolithic population and secondly, a diffusion process. The diffusion process is what causes the concentration to spread out; a high concentration of gas in one corner of a room gradually spreads throughout the room. Again, in our case, it can be thought of as the Neolithic population migrating gradually through Europe.

Solutions of such equations can give a travelling wave, i.e. a propagation front which advances at a constant speed (in a homogeneous and one-dimensional habitat) with the relevant variable equal to a constant value (Murray, 1993). This is ideal as this type of spread of agriculture has been confirmed previously (Ammerman and Cavalli-Sforza, 1973).

Davison *et al.* (2009) looked at the European Neolithic sites in the radiocarbon dates dataset (that are associated with the source at Jericho) and compared the distance from the source to the date associated with each site; to which you would expect to see a linear relationship if a constant spreading speed is present. Figure 2.2 shows said relationship.

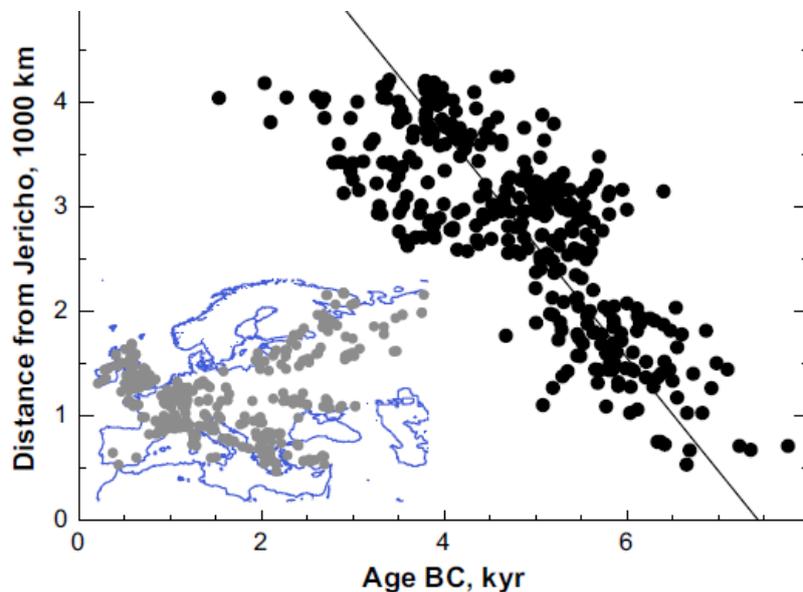


Figure 2.2: Sites associated with the Near-Eastern source at Jericho show a reasonable correlation of $C = -0.77$ with a mean speed of $V = 1.1 \pm 0.1$ km/year. (Davison *et al.*, 2009)

While not a very tight relationship, this is what is expected. The landscape in Europe is by no means homogeneous, so we would certainly expect varying propagation speeds in different locations. Take for example, mountainous regions like The Alps, here you have much decreased accessibility to humans and therefore spreading here would be very slow in

comparison to low altitude terrain. Major rivers would act in the opposite way, and increase the accessibility down the river, as well as being a very suitable place for settlement (Davison *et al.*, 2006).

This leads us on to use the FKPP equation and variants upon it to describe the intricacies detailed above.

2.2 Our Model

2.2.1 The FKPP Equation

The Fisher-Kolmogorov-Petrovskii-Piscounov equation (named after the four mathematicians who did early work with this equation) was first considered by Ronald Fisher in 1937 as a way to describe the spatial spread of an advantageous allele and explored its travelling wave solutions (Fisher, 1937). The uses for this equation have obviously far extended beyond just looking at alleles; these equations occur in many different fields including ecology, combustion and of course population dynamics. It is an example of a reaction-diffusion equation and the one used for our model. In its basic form it is given in Equation 2.1.

$$\frac{\partial u}{\partial t} = ru(1 - u) + D \frac{\partial^2 u}{\partial x^2}, \quad (2.1)$$

u is the value of the quantity which is spreading over the one spatial dimension, x , in time, t . D represents the diffusivity, that is, how fast the spread takes place. Lastly r simply represents some parameter.

As mentioned earlier, the solution in the form of a travelling wave which propagates at constant speed, V . This is a well known result when working with these equations and is given by Equation 2.2 to a first order approximation.

$$V = 2\sqrt{rD}, \quad (2.2)$$

2.2.2 Simplistic 1 Dimensional Model

To begin with, we create a model in 1 spatial dimension only. This is important to allow us to have a good understanding of our model before generalising it to a 2nd dimension. The FKPP equation (Equation 2.1) is altered slightly for our needs, introducing a new parameter that we require.

$$\frac{\partial N}{\partial t} = \gamma N \left(1 - \frac{N}{K}\right) + D \frac{\partial^2 N}{\partial x^2}, \quad (2.3)$$

Firstly we are using N to represent the Neolithic human population (just more inherently logical variable than u for our case), we now have two parameters, γ and K . These represent the intrinsic growth rate (reproducing population) and the carrying capacity of the land

respectively. That first ('reaction') term collectively models the increasing population with respect to fact there is a limit in the carrying capacity.

If we look at the case where N reaches the same value as K , we have

$$\gamma N \left(1 - \frac{K}{K}\right)$$

which is clearly going to have a value of 0. At this point this term would no longer contribute, which is as exactly as we would like. If the value of N is somehow larger than the carrying capacity K , then we would have

$$\gamma N \left(1 - \frac{K + a}{K}\right)$$

where a is a positive number representing how much larger N is. Clearly we have a negative value here, hence this term would decrease the population until it is at the carrying capacity value. Again, this is exactly what we would like.

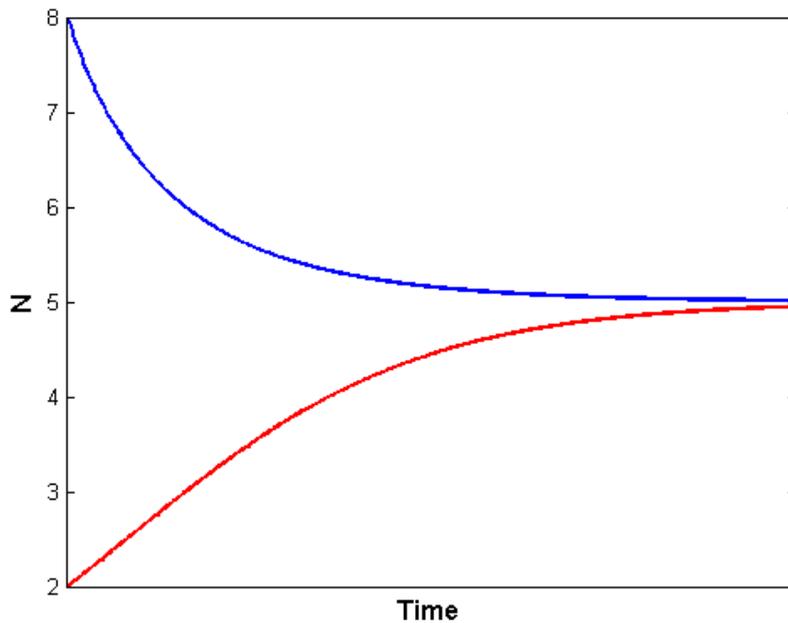


Figure 2.3: Here, $K = 5$ and two lines are shown, the red line representing the initial value of $N = 2$ and the blue line representing the initial value of $N = 8$. Over time, they increase and decrease respectively to the carrying capacity in accordance with the first term on the RHS of Equation 2.3.

Implementing this carrying capacity parameter does not affect the speed of propagation, V , so under our variables, we have

$$V = 2\sqrt{\gamma D}, \tag{2.4}$$

The second ('diffusion') term represents the migration of the Neolithic population. The parameter D is the diffusivity and can be altered to affect the speed of propagation.

For our model, the initial distribution of N is a truncated Gaussian and we adopt a numerical scheme with centred differences in space and evolves using an explicit Euler time step with forward differences in time.

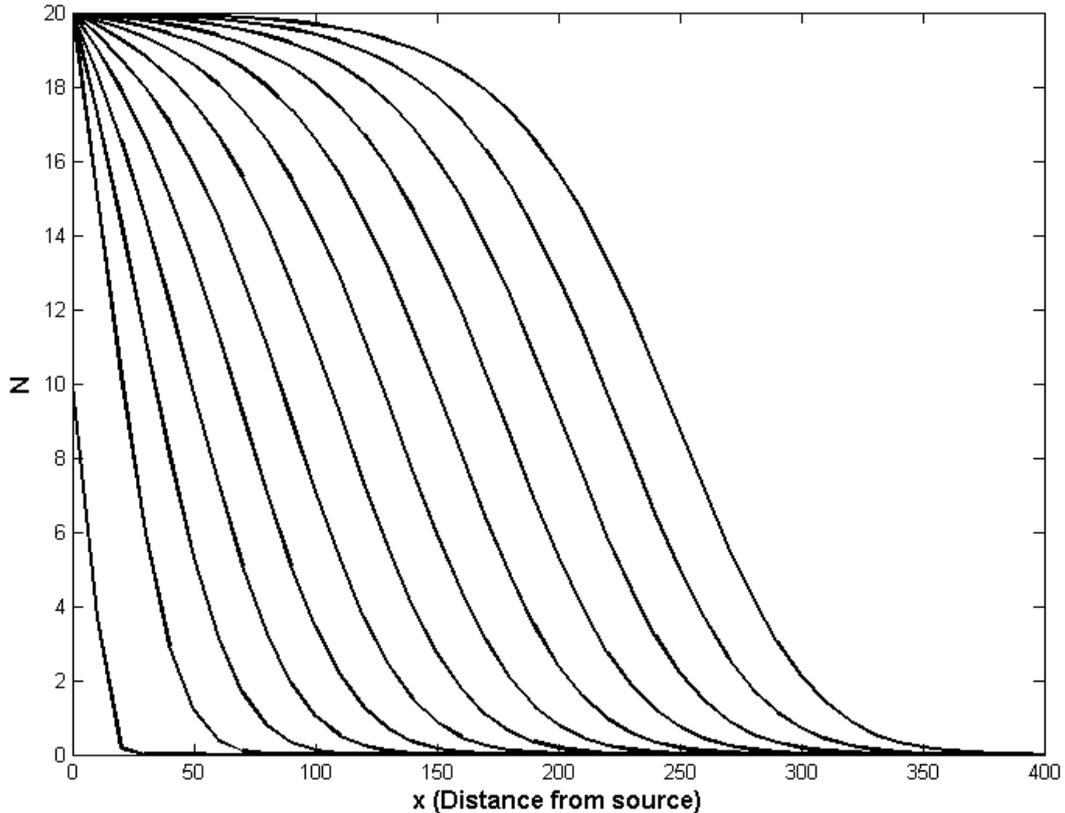


Figure 2.4: The wave front is shown here progressing forward over time, and is shown at equal time intervals. The initial distribution is shown at the far left, and the farther right, the later in time. A carrying capacity of $K = 20$ was used here with $\gamma = 0.02$ and $D = 6.25$. As time goes on, a plateau is forming at the carrying capacity. It looks as though the wave front is moving at roughly constant speed.

In Figure 2.4 we plot at equal time intervals, the advancing wave front. The value of N increases at each point until it reaches the carrying capacity, which is as expected and explained above. Due to this we start to see a plateau forming over time at this value, and as time goes on, this plateau would increase in size. Basically this is showing that between that point in space and the source, the population has reached the maximum allowed. By eyeballing the distance between each plot, we can see that the wave front is moving at roughly a constant speed. We will explore the propagation speed in more detail in Section 2.2.4.

At this stage, we understand the model and what each term contributes to the equation. We have seen what we expect when running our model in the simple 1D case and are thus confident to now generalise this to 2D.

2.2.3 Generalising to 2 Dimensions

At this stage, we now implement zero flux boundary conditions. That is $\partial N / \partial n = 0$, where n is the normal to the boundary. This does not affect our results in any significant way because of the fact the the boundary is in the sea. The spread of the Neolithic population should then not reach this boundary anyway.

Our equation is now as follows

$$\frac{\partial N}{\partial t} = \gamma N \left(1 - \frac{N}{K}\right) + D \nabla^2 N \quad (2.5)$$

Again here, before adding further complexities, we want to ensure we understand exactly what is happening with the model in order to be able to implement any future changes.

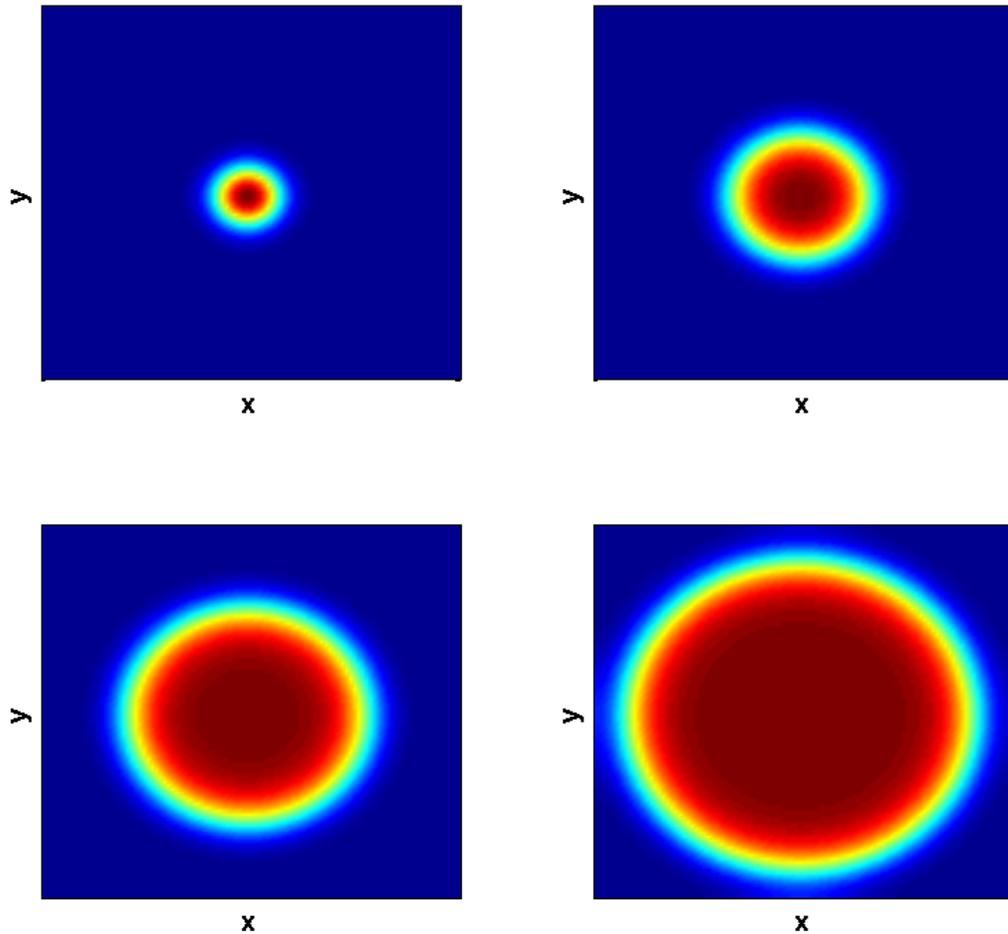


Figure 2.5: The model here running in 2D. It is shown at 4 different points in time, with the source at the centre. The plateau is shown forming similar to the 1D case and it is apparent that the wave front keeps a constant width. Here blue represents 0 and dark red represents the carrying capacity value, k .

Now at this stage in the model, the source has been made so that its location can be changed (whereas in the simple 1D model, it was simply at $x = 0$). This will enable us to set it to be at Jericho later. As hoped, we see similar features as we did previously, in Figure 2.5 we can see a similar plateau forming and increasing in size through time as N reaches the carrying capacity value, K . This gives us a good indication that the model is working properly based on our understanding from the simple 1D case. Another feature that now is clear, is that the width of the wave front looks to be constant (in Figure 2.5, that being the radial distance from the dark blue edge, to the edge of the dark red).

One thing that is no longer so clear, is whether the wave is propagating at a constant speed. It therefore makes sense to look at this in more detail, to see if this is the case.

2.2.4 Checking the Propagation Speed

Before going on to implement the landscape, and bring in all the extra detail to our model, we need to ensure it is working as expected. Checking that propagation speed is what we would expect it to be based on our approximation using Equation 2.2. This speed is a fundamental part of forming our model, as it is one key quantity that can be taken from the radiocarbon data and implemented through the values we choose for our growth rate, γ and our diffusivity D .

At this stage, it is not accurate to just eyeball the plots and must look at this in a more refined way. At each time step, we look at all the values of N where $y = 0$ and $x > 0$ (due to the fact this spread is taking place uniformly in all directions), and then select the point that falls in a small range centred around $N = 0.1K$. We then know this point will be on the wave front and exactly the distance, r , that we are from the source. We can then look at how r changes over time where we should expect a linear relationship.

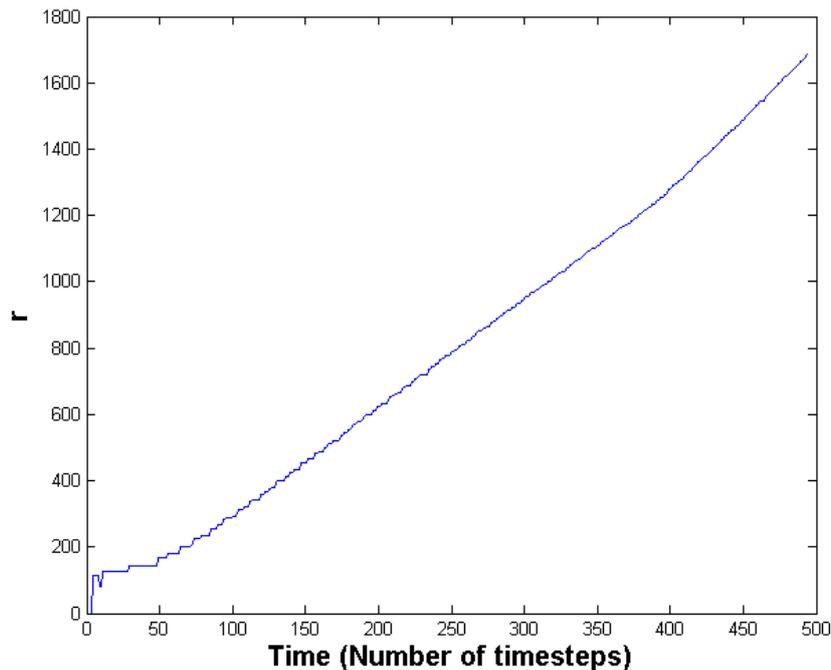


Figure 2.6: The distance of the wave front from the source is plotted here as it increases over time. Excluding some noise, a straight line is clearly visible. The gradient of this line, $m = 3.31$.

Shown in Figure 2.6 is a plot over time of the distance from the source to the wave front. There is substantial noise right at the start, this is expected as the initial distribution for the population gradually changes into the expected shape of the travelling wave. After this and until the spread starts to reach the boundary, there is a clear linear relationship, which indicates the constant speed we are looking for. Briefly looking at Figure 2.7 where dr/dt is plotted over time, we can see this matches what we have in Figure 2.6. There is a large

peak right at the beginning and this noise decreases gradually.

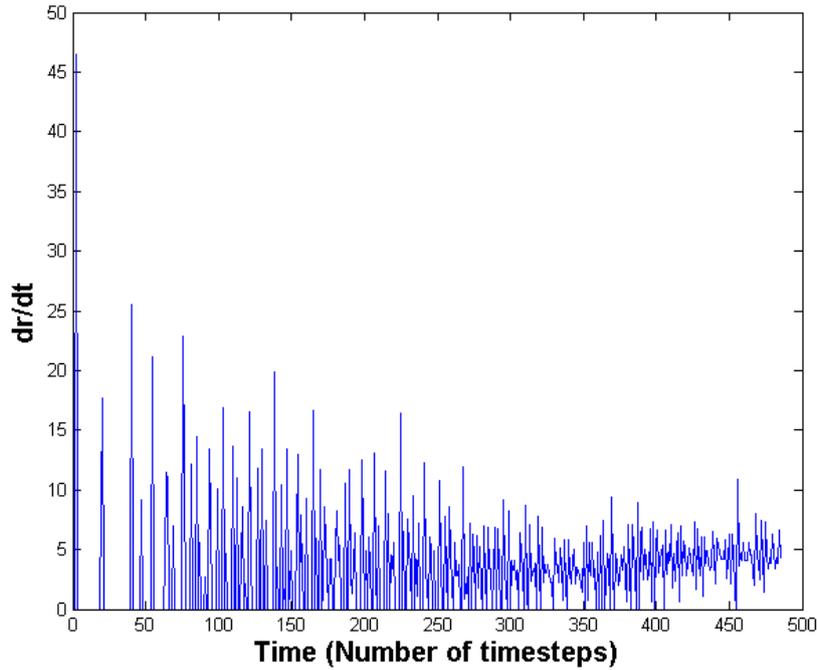


Figure 2.7: This plot shows how dr/dt varies over time. There is a lot of noise which gets smaller over time, but dr/dt stays around one value as expected.

Now by using MATLAB's linear regression tool, we obtain a value for that constant speed of propagation. This returned

$$V = m = 3.31$$

In this 2D model, we set $\gamma = 0.02$ and $D = 100$. Now using Equation 2.4:

$$V = 2\sqrt{\gamma D} = \sqrt{0.02 \cdot 100} = 1.41$$

While this value is not exact, it is reasonable and is correct to order unity as expected under the approximation. Confident our model is working as expected, the next step is to start to implement real world features and apply it to the spread of the Neolithic.

2.2.5 Implementing the Landscape

Due to the inhomogeneity of the land in Europe, we know that two things must vary, the diffusivity D , and the carrying capacity K . Certain regions would prevent a population from spreading or settling at all, an obvious such location being the seas. To do this, we need to alter Equation 2.5 slightly.

$$\frac{\partial N}{\partial t} = \gamma N \left(1 - \frac{N}{K}\right) + \nabla \cdot (D \nabla N), \quad (2.6)$$

This is a simple change to our model, as it actually only involves the addition of an extra term. By a simple vector identity, the last term becomes

$$\nabla \cdot (D\nabla N) = (\nabla D \cdot \nabla)N + D\nabla^2 N$$

The last term here is the same term we had previously in Equation 2.5 hence we only need to add $(\nabla D \cdot \nabla)N$ into our model to have the same result.

It is important to note that clearly, K and D are now functions of position rather than simply constants. Before implementing the landscape of Europe, it makes sense to do one final test with a varying diffusivity to see if the desired effect is achieved. To test this, we implement a simple function for D , with a maximum value at one corner of the domain, and minimum at the opposite corner and a gradually reducing slope between as shown in Figure 2.8. That is

$$D = D_0\left(1 - \frac{x}{L_x}\right) + D_0\left(1 - \frac{y}{L_y}\right)$$

where L_x and L_y are the widths of the domain in the x and y directions respectively.

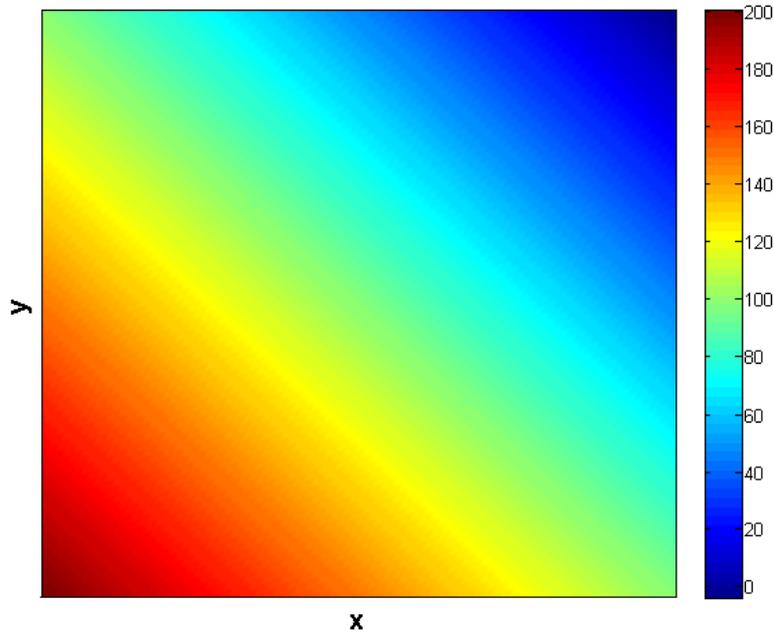


Figure 2.8: A simple example of a function for the diffusivity D with maximum value at the lower left corner of the domain, and minimum value at the upper right corner.

Now as before, we run the model and plot at several different points in time (shown in Figure 2.9). The spread is now clearly non uniform and there is an increased speed of propagation in the lower left corner of the domain, and a slowed speed of propagation in the upper right. This is as expected and shows we can easily alter the speed of spread through different regions of Europe by simply increasing or decreasing the value of D at each location.

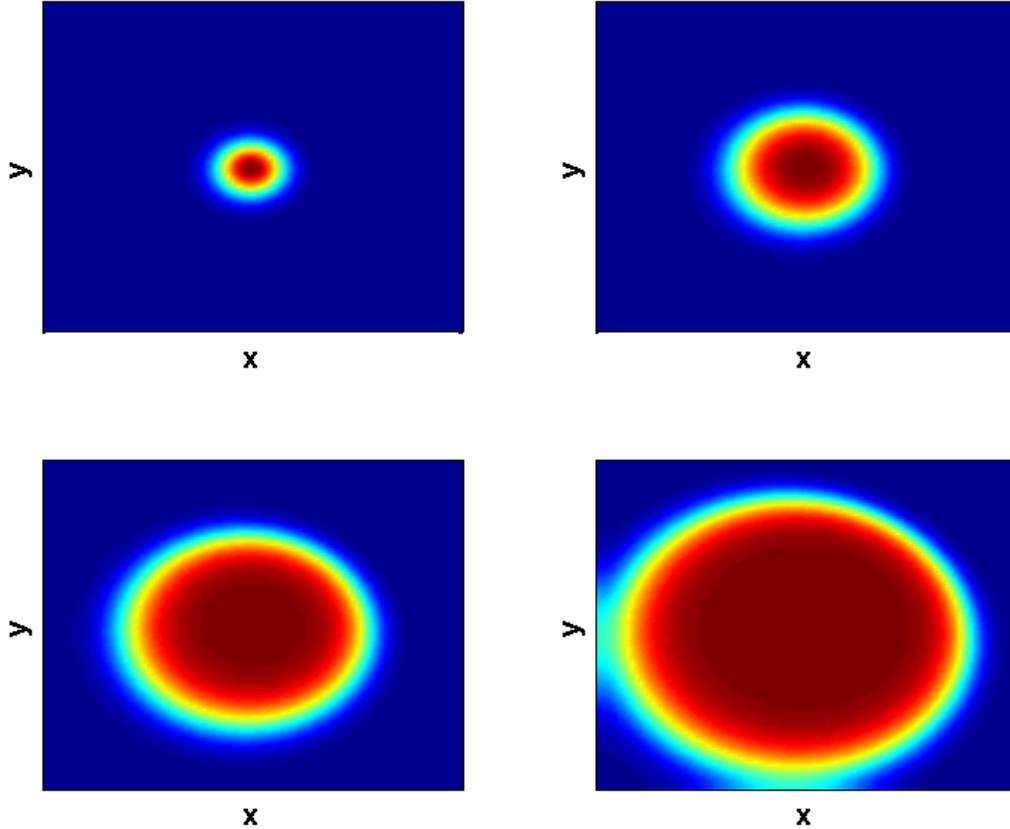


Figure 2.9: 2D model running with variable D defined as above. The spread clearly takes place faster in the lower left quarter of the domain, where the values of D are higher. This is clearly visible in the 4th image where the spread has reached the boundary but has not elsewhere.

Now we are prepared to set up our model for the Neolithic spread in Europe. We first need to choose our grid spacing in two dimensions on the Earth's (approximately) spherical surface. We choose a grid spacing of 1/12 degree (which varies depending on latitude between 2-8km) as done by Davison *et al.* (2009) with boundaries of the computational domain set at 25°N and 75°N, and 15°W and 60°E. This domain, along with the European coastline is shown in Figure 2.10.

The environmental factors taken into account in our model are the altitude and coastlines. Both will be taken into account by altering the diffusivity values at certain grid points. Before detailing that, we need to get some initial values for our parameters. As mentioned in Section 2.2.4, the propagation speed is important and hence it is the focus of our model due to the fact this quantity can be easily linked to the radiocarbon age used to date the arrival of the wave of advance. A first order approximation of this speed of advance, as given in Equation 2.4:

$$V = 2\sqrt{\gamma D}$$

As this depends upon only the intrinsic growth rate and the diffusivity, it is possible to calculate values of these parameters.

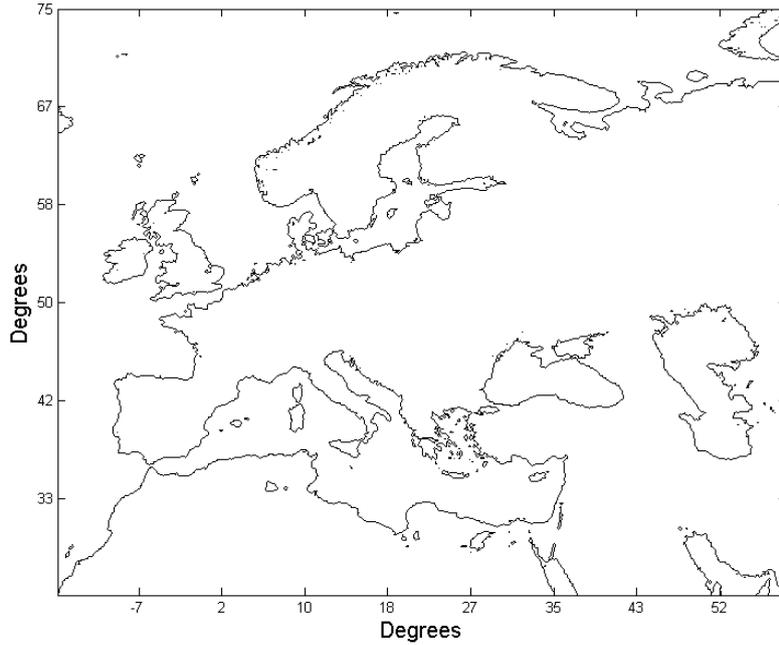


Figure 2.10: Plot of the European coastline in the computational domain. Coastlines defined as when Altitude is zero.

Estimations of these parameters has been done in many past works and we will use these values in our model. We take the intrinsic growth rate of a Neolithic population as $\gamma = 0.02 \text{ year}^{-1}$ (Birdsell, 1957). Then the mean speed of the wave front propagating through Europe of $V \approx 1 \text{ km/year}$ gives the background value of the diffusivity $D = 12.5 \text{ km}^2/\text{year}$ (Davison *et al.*, 2006). The background value for the carrying capacity K , is adopted as 3.5 persons/km^2 (Ammerman and Cavalli-Sforza, 1973).

Now there are two scenarios where the diffusivity and carrying capacity values are affected. That is at high altitudes, where farming is impractical and hence a farming Neolithic population cannot be supported and at coastlines heading into the sea which are obviously incapable of supporting the human population. D and K need to tend smoothly to zero for all of these locations as one would reasonable expect (and especially to avoid instabilities in the numerical model). We take altitudes of 1 km and above to be land where farming is impractical, and start smoothly tending to zero within 200 m of this altitude. This is done by making D and K proportional to

$$\frac{1}{2} - \frac{1}{2} \tanh\left(\frac{a - 1000 \text{ m}}{200 \text{ m}}\right)$$

where a is the altitude in metres as shown in Figure 2.12.

From the coastlines, the diffusivity tails off exponentially which allows the population to go out a short distance from the shore and this gives the model a way for the spread to reach Britain and Scandinavia from the continent. Davison *et al.* (2009) fine-tuned this exponential tail to reproduce the delay of the spread crossing the seas into Britain. This exponential tail is shown in Figure 2.13.

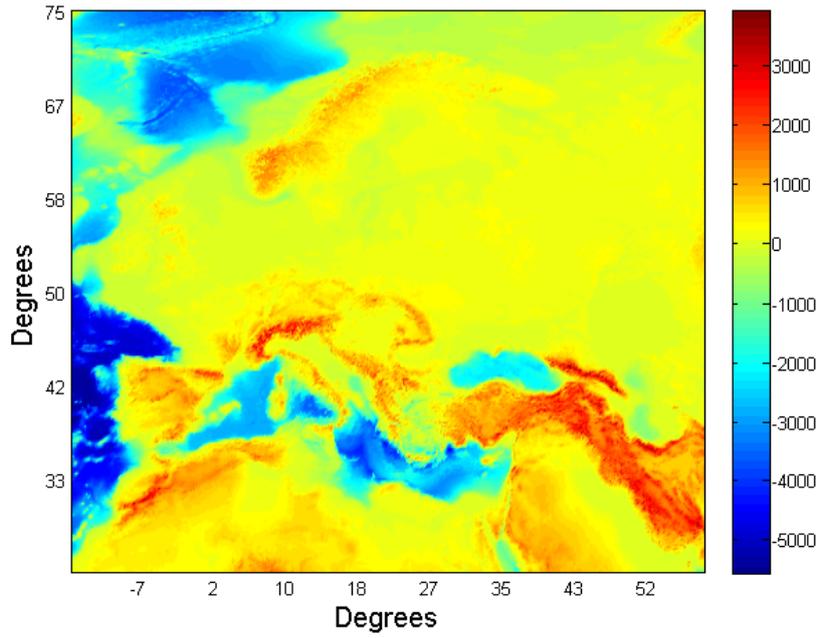


Figure 2.11: Visual representation of the altitude data used, altitude in metres.

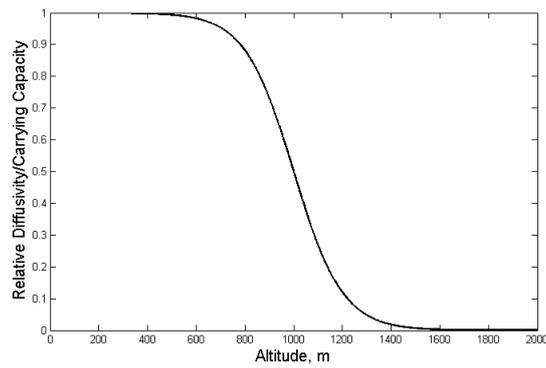


Figure 2.12: The change in diffusivity and carrying capacity based upon altitude.

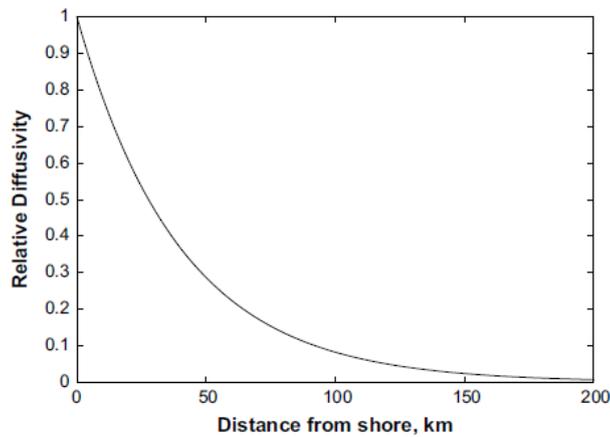


Figure 2.13: The change in diffusivity based on distance from the coast.

After this, we have a set of diffusivity and carrying capacity values for each particular grid

square representative of the landscape of Europe which can be used in our model. These are both shown in Figure 2.14.

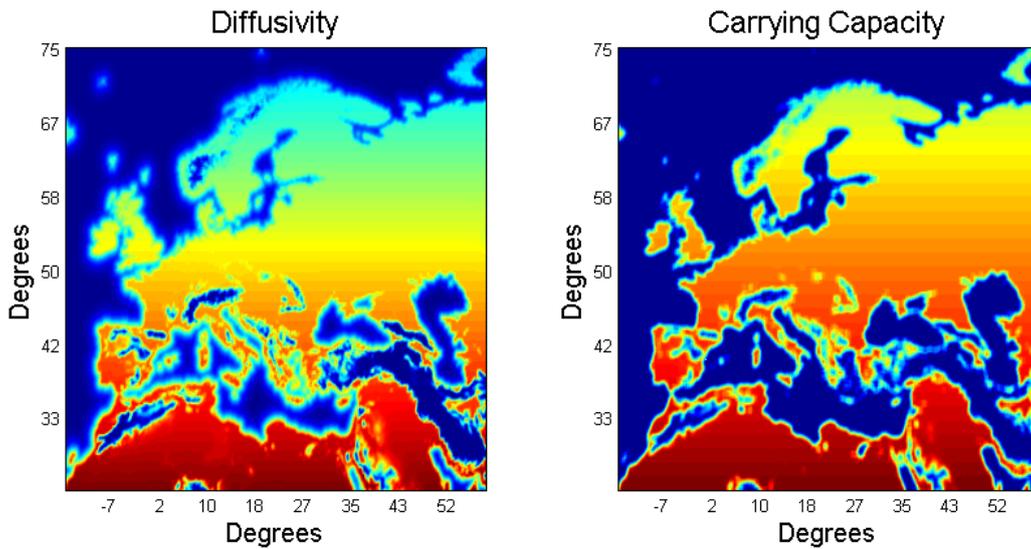


Figure 2.14: Visual representations of the values for diffusivity and carrying capacity throughout Europe. The exponential tail from the coastlines can be seen and here it is visually clear that the population will be able to spread from the continent to Britain with these values. It is also clear the spread can cross from Africa to Spain. Red represents large values, and blue small.

We use the source location and start date as in Davison *et al.* (2009). That places the start location at 35°N , 39°E with propagation starting at 6700 BC. Our model is shown running in Figure 2.15 where four points in time are shown. The first point to mention here, is the bottleneck through Turkey. Much of Turkey is high altitude terrain, leaving only a small bottleneck along the Southern coast for the spread to pass through. This proved an issue in actually getting the model in a situation where it would run and produce results. The values for D have to be sufficiently smoothed out in order to not cause instabilities but not so smooth that the bottleneck was not present.

The second point is the propagation wave reaching Spain through Africa prior to the wave front arriving from Turkey. Not too much is known about the spread of the Neolithic in Northern Africa and it is possible the diffusivity values associated with this region are too high. This is something that we will consider when comparing our results with the radiocarbon dates.

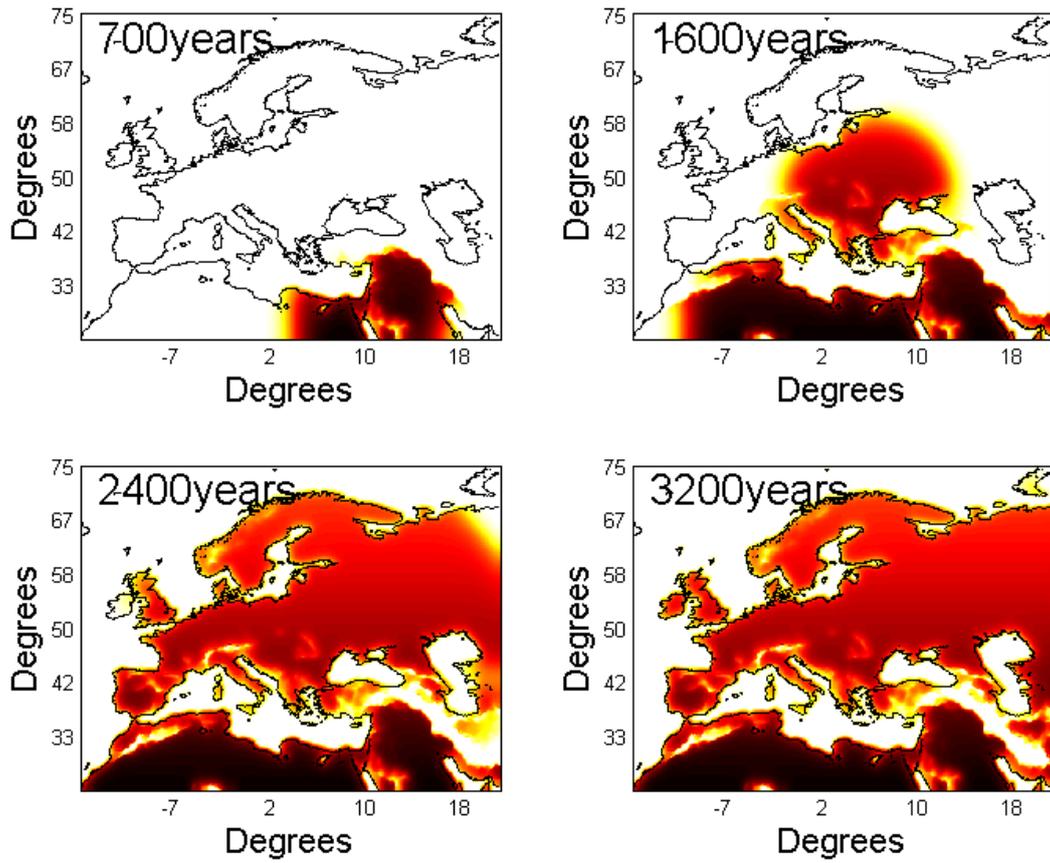


Figure 2.15: Visual representations of the values for diffusivity and carrying capacity throughout Europe. The exponential tail from the coastlines can be seen and here it is visually clear that the population will be able to spread from the continent to Britain with these values. It is also clear the spread can cross from Africa to Spain. Red represents large values, and blue small.

Chapter 3

Radiocarbon Dates

3.1 Selection of radiocarbon dates

The selection of dates used in this study were taken with permission from Davison *et al.* (2009). There is a large number of dates available from all over Europe which will allow us to look at our model's accuracy throughout different regions. The data for Southern, Central and Western Europe was originally from Shennan and Steele (2000), Gkiasta *et al.* (2003) and Thissen *et al.* (2006). The Eastern Europe dates were taken from Timofeev *et al.* (2004) and Dolukhanov *et al.* (2005). The treatment of the dates was completed by Davison *et al.* (2009) and is described in more detail there but is summarised here.

The main problem in selecting the dates is that many of the archaeological sites considered have many radiocarbon dates; some up to 30-50. Each individual radiocarbon measurement has an associated laboratory error, which was then converted into a calibration error σ_i , however this laboratory error is not really appropriate as we want to characterise the accuracy of the date compared to the true age of the archaeological site. This meant it was logical to establish an estimate of a minimum error and this was used for treating sites with multiple dates. An global minimum error of $\sigma_{min} = 160$ years was obtained from well explored, archaeologically homogeneous sites with a large number of tightly clustered dates.

There were then several strategies used to reduce each site to one or two dates and give each an associated error value. Firstly and most simply, if a geographical location had only one radiocarbon measurement associated with early Neolithic activity, then that date is taken to be the most likely date for the arrival of the Neolithic. The error associated with this date is taken to be the maximum of the calibration error obtained at the 99.7% confidence level and the global minimum error just discussed. Next if there are only a few (less than 8) dates measurements for a particular site, and all those dates are all within the calibration error, then their weighted mean value is taken to be the most likely date. Uncertainty in this scenario is characterised with an error equal to the maximum of the calibrated errors σ_i , the standard deviation of the dates involved, and the global minimum error.

In the circumstance where there are a series of dates that cluster in time but do not agree within the calibration error, a different approach is needed depending on the number of dates available. If there are only a small number of dates (again, less than 8), then the mean of the dates is taken. Reasoning being that any more sophisticated statistical technique is not appropriate with such a small sample size. The error is chosen in the same was as previous.

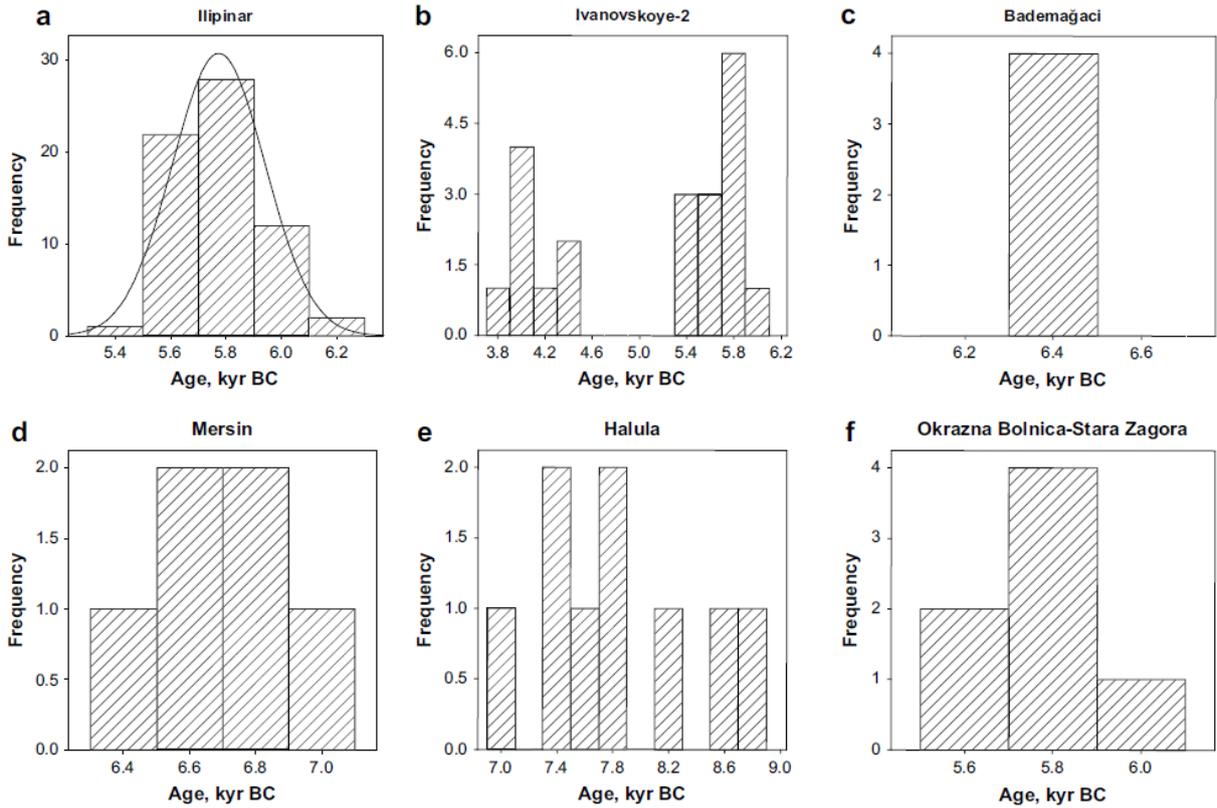


Figure 3.1: Examples of histograms from many Neolithic sites and how their date was determined. (a) There are a large number of dates from Ilipinar which are approximately normally distributed. They are not within the calibration error so the χ^2 criterion is used to calculate the age of this site. (b) There is clearly two distinct peaks here, suggesting multiple waves of Neolithic settlers through this site. Each peak is treat independently via the χ^2 criterion. (c) This site has 4 dates which very tightly cluster around a single date. The weighted mean is taken here. (d) There is no clear peak in the dates from Mersin and are spread out over a wide period of time. The oldest date is taken here to represent the arrival of the Neolithic. (e) Similar to (d), no particular peak, with a wide spread of dates. (f) There are only 7 dates from this site, which means there are not a sufficient number to apply the χ^2 criterion therefore the mean date is used.

If the cluster of dates is large (more than 8), the χ^2 statistical test can be used to calculate the most likely date, T , from the sample. The exact technique is described in detail by Dolukhanov *et al.* (2005). The algorithm is as follows:

- Calculate

$$T = \frac{\sum_{i=1}^n t_i / \tilde{\sigma}_i^2}{\sum_{i=1}^n 1 / \tilde{\sigma}_i^2}$$

- Calculate the statistic

$$X^2 = \sum_{i=1}^n (t_i - T)^2 / \tilde{\sigma}_i^2$$

- Compare X^2 with χ_{n-1}^2 , if $X^2 \leq \chi_{n-1}^2$ then T is taken as the most likely date for the arrival of the Neolithic. If not, then the dates that provide the largest contribution to X are discarded one at a time until they can be compared in this way.

with $\tilde{\sigma} = \max(\sigma_i, \sigma_{min})$. That is the maximum of the calibration errors of all the individual dates and the global minimum error.

Lastly, for sites where there are many radiocarbon dates that do not cluster around a single date, a histogram was analysed. There are two main possibilities, the first being that the data have a wide range with no peak. This would suggest that there may have been sustained Neolithic activity at the site and hence the oldest date was chosen as the most likely. The second possibility is that there are two or more peaks; this suggests that there may have been multiple waves of settlement at this location. Multiple dates were given to these sites and were calculated by treating each peak independently.

3.2 Comparison with our model

We are now in a position where we can compare the results of our model with the radiocarbon dates. The quality of the model is assessed by considering the difference in time between the model arrival time, T_m at each site, to said site's associated radiocarbon date T , obtained as described above. We calculate the time lag for each site:

$$\Delta T = T - T_m$$

and will use various plots to judge which sites are modelled badly and how the model could then be improved.

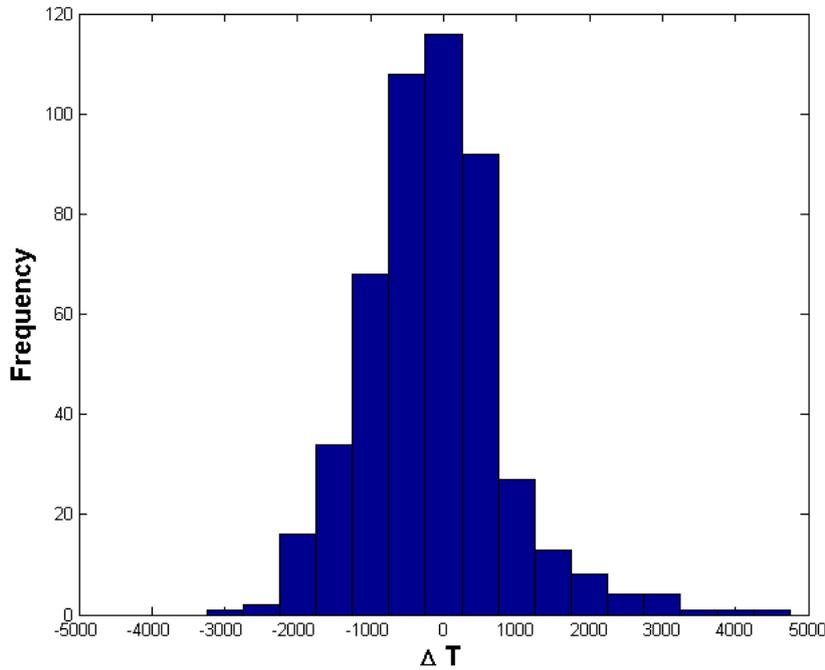


Figure 3.2: A Histogram of the time lags for each site. The ΔT values are clearly not normally distributed; there is a distinct positive skew.

Figure 3.2 shows a histogram of the ΔT values and highlights a clear positive skew. Ideally, we should have a normally distributed set of values, which would imply on average that the model was correct. The positive skew shows that there are many sites for which the model

is reaching too soon and possibly needs to be slowed down. In Figure 3.3, we look at all the archaeological sites and place them into categories depending upon how inaccurate the associated model time of arrival was. This enables us to clearly identify that there is a major problem in Great Britain, as at almost all the sites here, the model arrived too soon. Most likely this implies that there is not enough delay before the spread crosses the sea.

The sites which the model gets most incorrect however, are the sites in Eastern Europe. which is expected based upon Davison *et al.* (2009) and the work which was done on identifying a second source in Eastern Europe which many of these sites are attributed to. There are also a large number of sites near the source at Jericho where the model is arriving too late. These sites are not of much concern compared to Southern, Central and Western Europe which is where our primary interest is.

For a more quantifiable figure on how well our model represented the Neolithic spread, there are 125 archaeological sites which have $\Delta T \geq 1000$. There are 261 archaeological sites which have $\Delta T \geq 500$, which means roughly 48.9% of the sites were modelled well (Taking modelled well to mean a ΔT of less than 500).

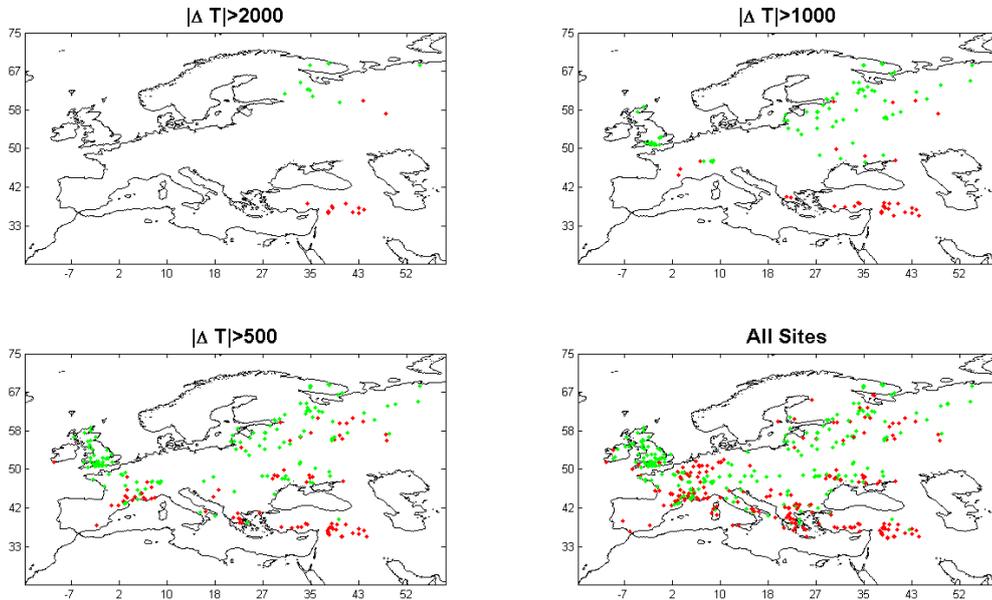


Figure 3.3: Each subplot shows all the sites which are have a sufficient enough error from the model. Green represents the model arriving too early at at that site, and red represents the model arriving too late. There are a large collection of green sites in Great Britain.

As briefly discussed at the end of Section 2.2.5, the model reaches Spain first, from Africa. This may be affecting some of the results for the sites in Western Europe and it is still a controversial topic whether the Neolithic did reach Spain in this way. It is a valid test for us to now suppress the spread through Africa by reducing the diffusivity in this region. This was simply done by selecting a set region of longitudes and latitudes and halving the diffusivity there. Figure 3.4 shows the updated diffusivity values. This will ensure the spread reaches Spain first through Europe and we will see if there is any difference in the results.

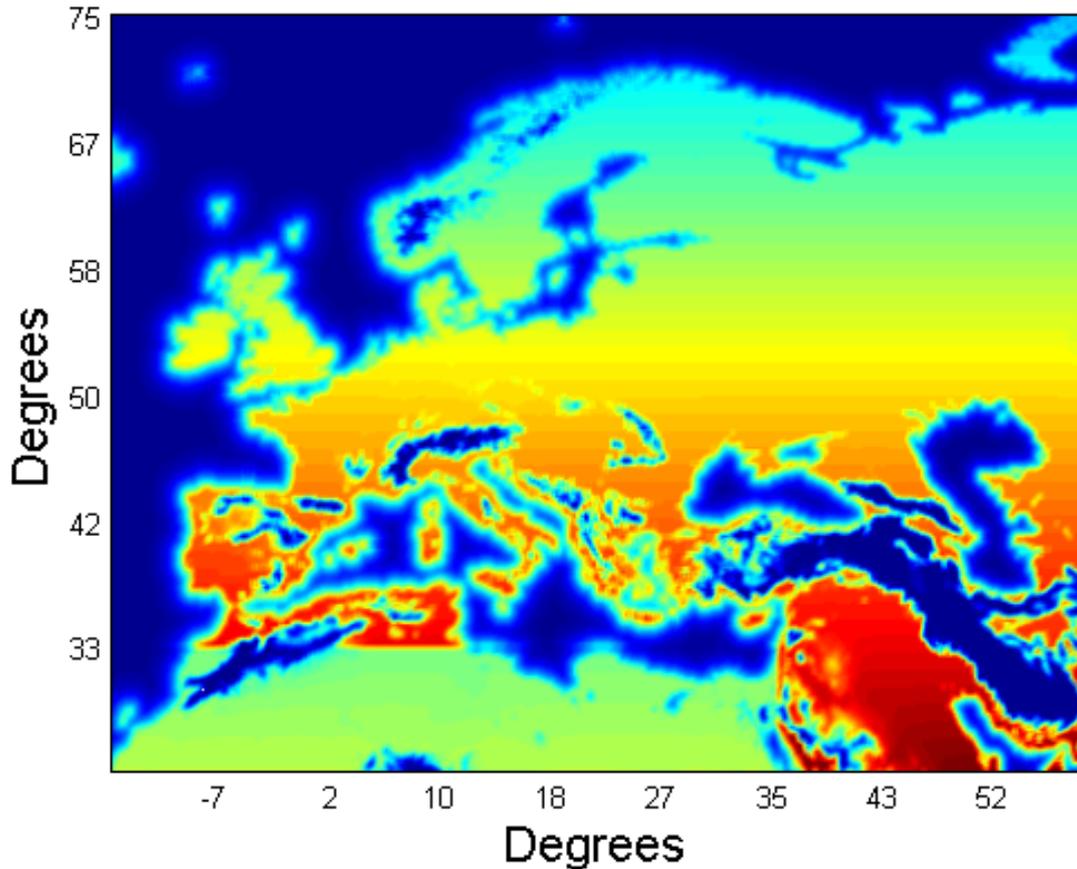


Figure 3.4: Updated diffusivity values with suppressed Africa.

Figure 3.5 is similar to Figure 2.15, only with Africa suppressed. It is clear by looking at the 3rd snapshot, that the spread did indeed in this scenario, reach Spain through Europe first. We expect any improvements to the results will be marginal, as there are only a handful of sites in the region which is affected by the arrival of 'two' separate wave fronts. It will also not improve upon the fact the model arrives in Britain too early, as shown in Figure 2.15, the spread through Africa does not reach Britain first.

As expected, there is little change in the results, due to this only having a possible impact in Spain and France. Figure 3.6 shows there are a few sites which have been modelled worse in this scenario, and indeed if we consider the number of sites in each category again, we see an increase. Under this model, there are 127 sites with $\Delta T \geq 1000$ and a total of 268 sites with $\Delta T \geq 500$. This leaves us with 47.6% of sites modelled well. This is a decrease of 1.3% suggesting that it is more likely that the spread did reach Spain first From Africa. This is by no means a verifiable conclusion due to the fact we are dealing with just a handful of sites in this region and there is sufficient room for future work on this question.

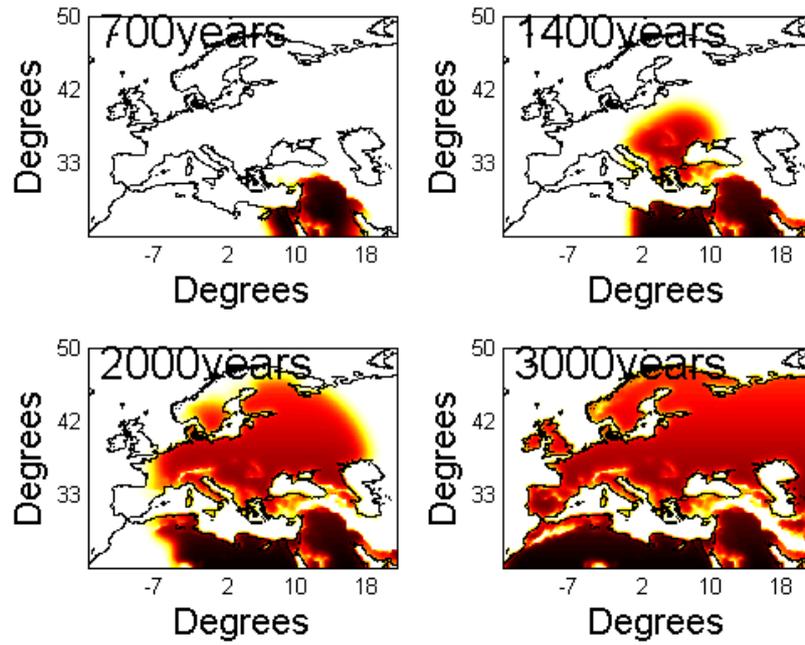


Figure 3.5: Model running shown at 4 different snapshots in time, with suppressed Africa. In this scenario, the wave front spreading through Europe reaches Spain first.

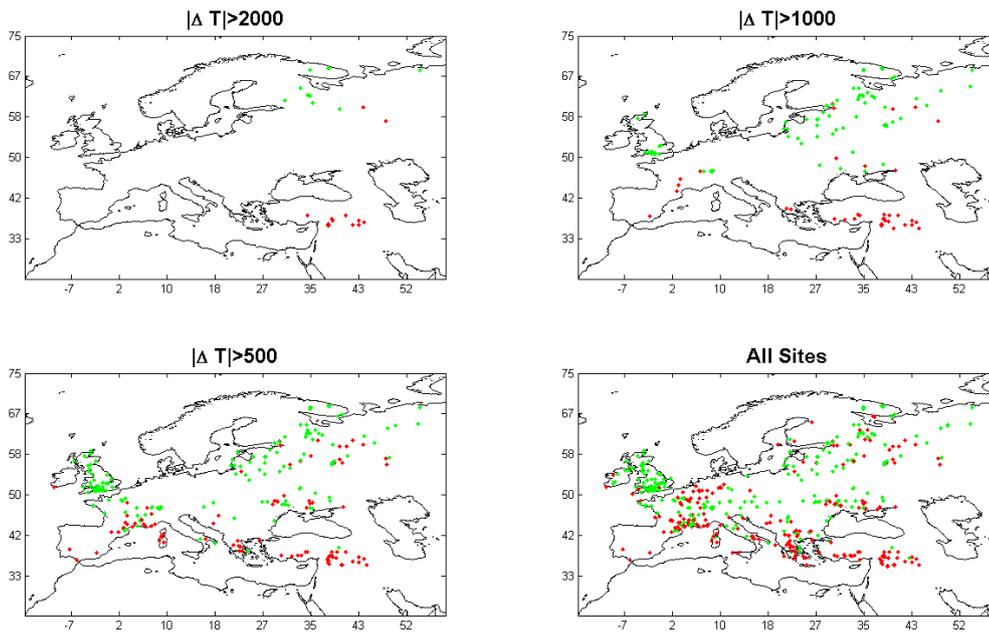


Figure 3.6: A few sites have different results in Spain and France. Mostly they have moved into the $\Delta T \geq 1000$ category from the $\Delta T \geq 500$ category. Notably, this model is providing worse results.

Chapter 4

Conclusion

The complexities of the Neolithic era causes the spread to be very difficult to model. The fact there are various types of Neolithic population who had advancements in some aspects of this era, but not others mean it is very unlikely just one or two simple spreads took place at this time. There is possibly room for more work in categorising the radiocarbon dates into more types of Neolithic population which likely spread in different waves.

Our model proved fairly good in the Southern, Central and Western areas of Europe, but had problems elsewhere, specifically Britain where the model arrived much too soon. The added complexity of Britain being over the sea is more challenging to model, given little information is known about humans seafaring ability during this period.

The question of whether the Neolithic arrived in Spain through Africa or from through Europe needs further exploration, and has potentially some interesting results if so. Little information is known on the Neolithic in Northern Africa and perhaps advents of slightly different technologies could have taken place through this route and introduced those into Europe through Spain. When suppressing Africa in our model, the results for some of the sites in Western Europe got worse, suggesting it is more likely that the Neolithic did arrive in Spain from Africa. We are dealing with such a small number of sites in this region however, so it is not possible to conclude that this is the case.

Bibliography

- Ammerman A.J., Cavalli-Sforza L.L., (1973). *A population model for the diffusion of early farming in Europe*, Duckworth, London, Pg 343-357.
- Birdsell J.B., (1957). Some population problems involving Pleistocene man, *Cold Spring Harbor Symposium on Quantitative Biology* **22**, 47–69.
- Callaway E., (2015). Oldest stone tools raise questions about their creators, *Nature* **520**, 421.
- Davison *et al.* (2006). The role of waterways in the spread of the Neolithic, *Journal of Archaeological Science* **33**, 641–652.
- Davison *et al.* (2009). Multiple sources of the European Neolithic, *Quaternary International* **203**, 10–18.
- Dolukhanov P., *et al.* (2005). The Chronology of Neolithic Dispersal in Central and Eastern Europe, *Journal of Archaeological Science* **32**, 1441-1458.
- Fisher R.A., (1937). The wave of advance of advantageous genes, *Annals of Eugenics* **7**, 355–369.
- Gimbutas M., (1973). The Beginning of the Bronze Age in Europe and the Indo- Europeans 3500-2500 BC, *Journal of Indo-European Studies* **1**, 163.
- Gkiasta *et al.* (2003). Neolithic transition in Europe: the radiocarbon record revisited, *Antiquity* **77**, 45–61.
- Haak *et al.* (2005). Ancient DNA from the first European farmers in 7500-year-old Neolithic sites, *Science* **310**, 1016–1018.
- Murray J.D., (1993). *Mathematical Biology*, Springer Verlag, Berlin.
- Perles C., (2001). *The Early Neolithic in Greece. The First Farming Communities in Europe*, Cambridge University Press, Cambridge.
- Price T.D., (2000). *Europe's First Farmers*, Cambridge University Press, Cambridge.
- Romaine S., Nettle D., (2002). *Vanishing Voices: The Extinction of the World's Languages*, Oxford University Press, Pg 102-103.
- Shennan S., Steele J., (2000). Spatial and chronological patterns in the neolithisation of Europe, Available from, http://ads.ahds.ac.uk/catalogue/resources.html#c14_meso.
- Thissen L., *et al.* (2006). CANeW 14C databases and chronocharts (2005-2006 updates), Available from, <http://www.canew.org/data.html>.

- Thomas J., (1996). *The cultural context of the first use of domesticates in Continental and Northwest Europe*, UCL Press, London, Pg 310-322.
- Timofeev V.I., *et al.* (2004). *Radiocarbon Chronology of the Neolithic in Northern Eurasia*, Tesa St, Petersburg.
- William *et al.* (2012). *The Essence of Anthropology*, 3rd edition, Cengage Learning, Boston, Pg 83.