Ecological Modelling 307 (2015) 22–31

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Modeling initial Neolithic dispersal. The first agricultural groups in West Mediterranean

Joan Bernabeu Aubán[a,*], C. Michael Barton[b,c], Salvador Pardo Gordó[a], Sean M. Bergin[c]

[a] Department ofPrehistoria i Arqueologia, Universitat de Valencia, Spain
[b] Center for Social Dynamics & Complexity, Arizona State University, USA
[c] School of Human Evolution & Social Change, Arizona State University, USA

ARTICLE INFO

Article history:
Received 18 December 2014
Received in revised form 25 March 2015
Accepted 26 March 2015
Available online 11 April 2015

Keywords:
Neolithic spread
West Mediterranean Neolithic
Agent based models
Socio-ecological modeling
Complex Adaptive Systems

ABSTRACT

In previous research, the SE-NW time-trend in the age of the earliest Neolithic sites across Europe has been treated as a signal of a global-scale process that brought farming/herding economies to the continent. Residual variation from this global time-trend is generally treated as 'noise'. A Complex Adaptive Systems perspective views this empirical record differently. The apparent time-trend is treated as an emergent consequence of the interactions of individuals and groups of different scale.

Here, we examine the dynamics of agricultural dispersals, using the rich body evidence available from the Iberian Peninsula as a case study. We integrate two complementary approaches: (1) creating a high resolution Agent Based Modeling environment to simulate different processes that may have driven the spread of farming; (2) collecting and synthesizing empirical archeological data for the earliest Neolithic settlements that we use to evaluate our models results.

Our results suggest that, (a) the source of radiocarbon data used to evaluate alternative hypotheses play an important role in the results; and (b) the model scenario that produces de best fit with archeological data implies a dispersal via northwestern and southern routes; a preference for leap-frog movement; an influence of ecological conditions (selecting most favorable agricultural land) and demographic factors (avoiding settled regions).

This work represents a first attempt at high-resolution bottom-up modeling of this important dynamic in human prehistory. While we recognize that other social and environmental drivers could have also affected the dispersal of agropastoral systems, those considered here include many that have been widely considered important in prior research and so warrant inclusion.

1. Introduction

The transformation of subsistence systems from hunting and gathering to farming involved fundamental changes in the relationship between humans and the environment which involved all levels of human society. The consequences of this transformation extend to the present day and, perhaps for this reason, the issue of the origin and spread of Neolithic economies remains a major topic in archeological and anthropological literature. This is certainly the case for Europe where the subject of the origin of farming societies centers on the dissemination of agricultural systems. Material culture, chronology, the absence of pre-existing wild species from which cultivated species may have been derived, and analyses of DNA from domestic animals confirm that agriculture and livestock rearing were introduced to Europe from the southwest Asia. There remains, however, a lack of consensus about the mechanisms by which this transition occurred. Did the change involve movements of people that either displaced or mixed with hunter-gatherer groups? Or was it merely material and information that traveled, such as domesticated animals, knowledge and material culture including pottery (the so-called Neolithic Package)? The latter process is commonly referred to as cultural diffusion and the former as demic diffusion. Despite a long history of diverse efforts by many scholars, the relative importance of demic vs. cultural diffusion has not yet been resolved.

A number of quantitative models have been proposed to help improve our understanding of the dynamics of the origins of agriculture in Europe. The majority and best known of these models

---

* Corresponding author. Tel.: +34 963864242.
E-mail addresses: jbauban@uv.es (J. Bernabeu Aubán), michael.barton@asu.edu (C. Michael Barton), salvador.pardo@uv.es (S. Pardo Gordó), sean.bergin@asu.edu (S.M. Bergin).

http://dx.doi.org/10.1016/j.ecolmodel.2015.03.015
0304-3800/

have been formulated on a continental scale, seeking to describe and predict a Neolithic expansion front expanding from a center of origin located in southwest Asia. Most of these models are based on versions of reaction–diffusion equations originally proposed by Fisher (1937) to represent gene flow. Ammerman and Cavalli-Sforza (1984) were the first to apply such a model to the expansion of Neolithic across Europe in order to explain an observed SE to NW trending chronological gradient of the earliest farming sites. This gradient has been subsequently re-evaluated using larger sets of dated sites (Gkiasta et al., 2003; Pinhasi et al., 2005).

This reaction–diffusion approach has recently been applied to the spread of farming in several different ways (Steele, 2009). Several authors have modeled the effect of single and multi-time delay between individual birth and dispersal (Fort et al., 2004; Isern and Fort, 2010; Pinhasi et al., 2005); the effect of anisotropic diffusion (Ackland et al., 2007; Davison et al., 2006; Isern and Fort, 2010); age-dependent mortality, fertility and dispersal persistence (Pérez-Losada and Fort, 2011); discrete and continuous dispersion kernel (Ackland et al., 2007; Isern and Fort, 2012); the effect of cohabitation of parents and their children (Isern and Fort, 2012); cultural hitchhiking of neutral traits (Ackland et al., 2007); and the effect of interaction between foragers and farmers (Ackland et al., 2007; Aoki et al., 1996; Fedotov et al., 2008).

A consistent feature of these models is that they are top-down models with aggregate approximations of locally diverse processes expressed at continental-scales, although there have been some efforts to represent environmental diversity, at least at coarse, continental scales (e.g., Ackland et al., 2007; Davison et al., 2006). While this can be a useful approach to some phenomena, it has been pointed out that such aggregate models necessarily ignore a great deal of regional and local variability in socio-ecological conditions and processes that can reduce their explanatory power and predictive utility (Bentley et al., 2009).

In addition to the mathematical models, a variety of narrative conceptual models also have been proposed by archeologists to account for the diversity of the empirical record (e.g., Bernabeu Aubán, 2002; Guilaine, 2001; Zilhao, 2001; Zvelebil and Lillie, 2000). In a non-Mediterranean context, this is the reason for Bogucki's call for agent based models to help explain the Neolithic spread to central and northern Europe (Bogucki, 2000). In general, these models invoke local to regional-scale processes that have not been included in mathematical models. Also, the data used to test the global models (e.g., radiocarbon dates for sites) have generally been of insufficiently high-resolution to test these more detailed local models.

For example, the interaction between different groups, giving rise to forms of cultural hybridization or transfers, as proposed by some models require mobilizing cultural variables such as technology, style, or networks of interaction and modeling their patterns of spatio-temporal change. Additionally, models of expansion dynamics such as leap-frog or maritime pioneer colonization, require high-resolution chronologies to identify and differentiate temporal and spatial expansion patterns at local and regional scales. This, in turn, requires detailed review of published dates in order to avoid (or at least to evaluate) problems such as the old wood effect or other taphonomic filters (Barton et al., 2001; Zilhao, 2001).

To better understand the dynamic processes that drove this socio-ecological transformation of human society, we can benefit from an approach that combines a more detailed regional scale empirical record, computational modeling, and bottom-up conceptual approach to modeling that can represent the individual and household decisions and practices that ultimately spread the agricultural way of life throughout the world. We see this as complementary to prior research summarized above, and an important way to integrate formal and narrative modeling. Conceptually, a Complex Adaptive Systems (CAS) perspective can help us to realize this objective (Miller and Page, 2007).

In previous research, the SE-NW time-trend in the age of the earliest Neolithic sites across Europe has been treated as a signal of a global-scale process that brought farming/herding economies to the continent, and even included in formal models as an a priori directionality in Neolithic dispersal. Residual variation from this global time-trend is generally treated as 'noise' due to poor quality data and/or local conditions. A CAS perspective views this empirical record differently. In the work we present here, we do not assume any geographic time-trend at the outset. Rather any apparent general trends and deviations are treated as emergent consequences of the interactions of individuals and groups of different scale.

Our goal is to explore the emergent phenomenon seen globally as the spread of the Neolithic from the perspective of local-scale processes. We do so by using an agent-based modeling (ABM) environment to systematically test the potential for different local social and ecological drivers to generate patterns at regional-scales that match the empirical data for the emergence of agricultural systems across the Iberian Peninsula. We see this study as complementary to the continental-scale modeling to provide a richer, more nuanced insight into the transition to agriculture in Europe.

2. West Mediterranean and the Iberian Peninsula

The Western Mediterranean, extending from southern Italy to Portugal and northern Africa, can be considered a single archeological unit with respect to the beginning of farming, where Early Neolithic archeological contexts share a number of common elements, exemplified by the pan-regional presence of Cardium-Impressed ceramic wares. Some consensus exists regarding the origin of these wares in southern Italy, but the debate surrounding the spread of these elements to the west remains open. Perhaps, as noted by Zeder, the processes responsible for the expansion of agricultural systems "... involved elements of demic diffusion, local adoption, and independent domestication." (2008, p. 11603, p. 11603), but the cultural contexts of relevant dispersal, routes, and tempo have not yet been resolved.

The Iberian Peninsula is a particularly good region to study the process of agricultural dispersals (Zilhao, 2003), due to its geography and the evidence for populations of foragers during the final Mesolithic (*post quem* c. 6000 BC) (Bernabeu et al., 2014; Utrilla Miranda and Montes Ramírez, 2009). Situated at the western extreme of Mediterranean, and serving as a bridge between Africa and Europe, Iberia is a subcontinent where it is possible to encounter great socio-ecological diversity at local scales important for the transition to the agriculture. For example, Iberia is the best place to evaluate whether or not the Neolithic reached westernmost Europe via dual expansion routes following the northern (i.e., Italy and France) and southern (north Africa) Mediterranean coasts (Bernabeu Auban et al., 2008; Cortés Sánchez et al., 2012; Isern et al., 2014; Linstädter et al., 2012).

Here, we examine the dynamics of agricultural dispersals using the rich body evidence available from the Iberian Peninsula. To do this we integrate two complementary approaches: (1) creating an ABM modeling environment to simulate and evaluate different processes that may have driven the spread of farming; (2) collecting and synthesizing empirical archeological data for the earliest Neolithic settlements in the peninsula that we use to evaluate our models (rather than to create the models).

2.1. The Iberian dataset

The last two decades have witnessed significant improvement in the empirical database of sites and radiocarbon dates for the

Iberian Mesolithic and Neolithic. Nevertheless, this rapidly expanding dataset has not yet been incorporated into continental-scale models of Neolithic dispersals (e.g., Fort, 2012; Pinhasi et al., 2005). Although a few very recent works have begun to remedy this problem (Bocquet-Appel et al., 2009; Fort et al., 2012; Pinhasi et al., 2005) the dataset used here represents the most complete compilation of relevant radiocarbon dates for the Neolithic available at the time of writing (see Supplementary Information Table 1). Importantly, we systematically evaluate the reliability of the dates used, including their contexts and the materials sampled, in response to recent calls for such assessment and evidence that these factors influence age estimates (Bernabeu et al., 2001; Zilhao, 2001; Zilhão, 2011). Such validation has been difficult at the scale of the entirety of Europe because empirical datasets used for top-down model calibration derive from decades of research by multiple archeological teams working in regions with different research traditions (Gkiasta et al., 2003). However, more detailed assessment of the quality of the empirical data for Neolithic dispersals is becoming more accessible at the regional scale that is the focus of the work presented here.

2.2. *Selecting dating samples for model evaluation*

In selecting samples for radiocarbon dating that can inform us about the spread of agriculture, initial considerations are defining relevant Neolithic archeological contexts and the appropriate chronological range for analyzing processes that drove Neolithic dispersal. While the first of these considerations applies to sites and dates throughout the Iberian Peninsula, the second varies regionally.

To estimate a chronological range sufficient to encompass the spread of agriculture over much of the peninsula, we first identified the oldest widely accepted date for the use of domesticates in the peninsula: a date of 7569 ± 48 cal BP. (All dates used here are expressed as calibrated years BP.) We then extended this range up to 6000 cal BP to encompass the latest evidence for the initial establishment of agropastoral systems across the peninsula. For any region in the Iberian Peninsula, we selected sites representing the earliest dated evidence for domestic plants and/or animals.

For each site selected, we only use dates clearly associated with archeological contexts where the remains of domestic taxa (plant or animal) have been found, excluding dates from uncertain associational contexts. To the extent possible, we prioritize AMS dates derived from single fragments of organic material to avoid the possibility of mixing samples of different ages (Bernabeu et al., 2001). For some sites, such higher-resolution dates are not yet available, but in all cases, we restrict our sample to dates with a standard deviation of $\leq$100 years. We use the oldest radiocarbon dates that meet these criteria from the selected sites. These procedures produced a sample of 111 dates from early Neolithic sites across the Iberian Peninsula, shown in Fig. 1 (see Supplementary Information Table 1 for details).

We also classified all dates according to the material dated to better assess the quality of their age estimates. This classification distinguishes dated material as being remains of domestic plants or animals, from other short-lived taxa (e.g., unburned animal bones or shrubs), and from long-lived taxa (mainly tree wood) and bulk charcoal samples. Thus, 38 sites are dated from carbon from domestic taxa; an additional 37 sites have dates from short-lived taxa; the remaining 36 sites have dates on wood charcoal or bulk charcoal dates (but still with low standard deviations). The sites in Fig. 1 are colored according to the material dated. This allows us to compile an even larger and higher quality dataset for evaluating model performance than has been used in other recent work (e.g., Fort et al., 2012).

3. The agent based modeling environment

Many previous mathematical models have been designed to account for the empirical record of sites and their dates. While some models incorporate processes like birth and death rates derived from the ethnographic record (e.g., Isern and Fort, 2012), in others the parameters are tuned to fit the empirical record (e.g., Fort et al., 2012). Here, we utilize first principles, rather than the characteristics of the archeological record to drive model behavior. By ‘first principles’ we refer to some kind of generative process, derived from understanding human behavior, that drives model behavior. This is sometimes termed a deductive approach. This differs from the more common inductive practice in social science of deriving a mathematical relationship between empirically observed data and then applying that relationship in a model. In such cases, models are more ‘descriptive’ than generative. These are sometimes called “empirical generalizations” (Dunnell, 1982). It is also very different from the normal archeological practice of inductively inferring something about past societies from detailed analysis of the archeological record.

We treat each model scenario—a particular combination of model algorithms and parameter values—as a complex hypothesis, and carry out simulations as a set of digital experiments. Then we test the results of these experiments against the empirical record. Our goal is not to design a model that fits the empirical data as closely as possible, but to assess which scenario ‘hypothesis’ *better* fits the empirical data than the alternatives. This procedure has been inspired by the ‘modeling as experiment’ approach of Bankes and colleagues (2002) and the ‘pattern oriented modeling’ approach of Grimm and colleagues (2005; see also, Lake, 2015; van der Leeuw, 2004).

This approach allows us to systematically explore the effects of different combinations of potential drivers of Neolithic dispersals. We do not attempt to ‘reconstruct’ the past *in silico*, but instead systematically assess the ability of different driver combinations to produce model results that match (or do not match) the empirical archeological record, in the form of the earliest dated Neolithic sites in the Iberian Peninsula. Specifically, in the experiments reported here, we examine the following processes.

1. *Continuous or punctuated (‘leap frog’) movement*. We compare the results of dispersals where farming spreads to adjacent areas, analogous to wave of advance models, with dispersals in which farming can spread to non-adjacent areas at different distances from existing agricultural settlements. The latter is the kind of movement proposed in ‘maritime pioneers’ and Leapfrogging models (e.g., Zilhao, 2001).
2. *Direction of agricultural dispersal*. We compare the results of starting agricultural dispersals from a single or multiple points that, with one exception, encompass all possible regions from which the spread of farming into Iberia could have started. We identified 16 potential starting locations; 15 are located around the perimeter of the Peninsula, and one is located at the center (Madrid) as a sort of null model for comparison with the others. The sites were chosen to provide roughly regularly space, geographically reasonable starting points for dispersal into the peninsula. Except for Malaga and Gibraltar, the sites around the perimeter, are located near the mouths of rivers that would have served as convenient routes for initial dispersals of farmer/herders—especially given the generally mountainous terrain bordering the peninsula and its interior plateau, the Meseta (see Fig. 1).
3. *Ecological context*. Although, it is widely recognized that different parts of a diverse landscape are more or less suitable for early Neolithic agriculture, with few exceptions (e.g., Ackland et al., 2007; Davison et al., 2006) the ecological context of

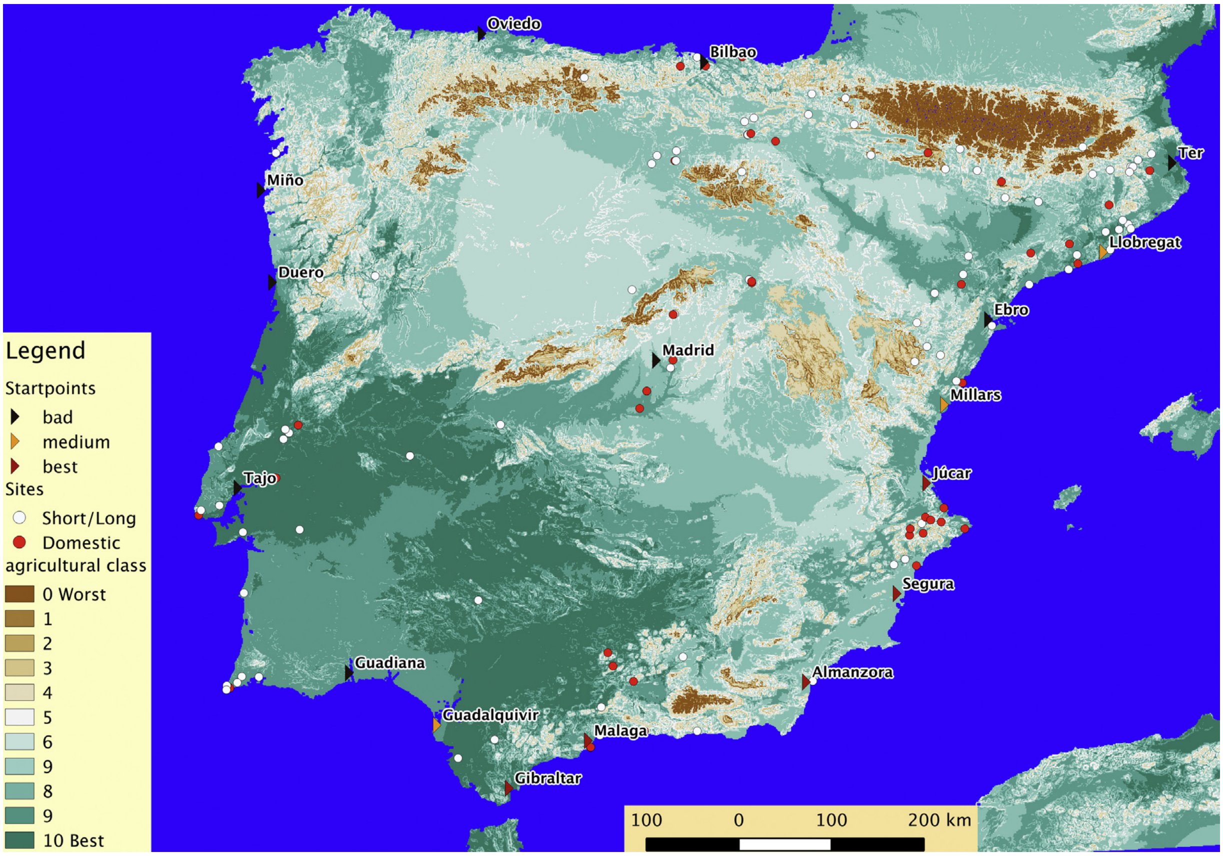


Fig. 1. Map of the Iberian Peninsula colored by values of ecological suitability index for wheat cultivation. Early Neolithic sites with radiocarbon dates used for model evaluation are shown as colored circles, colored by materials dated (see text). Starting points for spread models shown as triangles with colors indication their performance in modeling, with red indicating dates from domestic taxa and white indicating dates from other materials (see text, and Figs. 3 and 4).

agricultural dispersals has not been considered in prior modeling work, and never at a regional scale that takes into account differences in potential agricultural productivity. We examine a combination of climate and terrain in terms of its suitability for growing primitive cereals (Fig. 1).

4. *Demographic effects and anthropogenic environmental impacts.* Following up on suggestions by McClure and colleagues (McClure et al., 2006, 2009) and Shennan (2009, p. 345) about the potential impacts of socially mediated access to resources during the Neolithic, we examine the effects of population density and/or anthropogenic degradation on the tempo and pattern of agricultural dispersals. We implement a version of an Ideal Despotic Distribution algorithm from human behavioral ecology (Fretwell and Lucas, 1970; Kennett et al., 2006; McClure et al., 2006). Agriculture spreads to the neighboring cells with the highest suitability values, and suitability values decline each time agriculture "spreads" to a cell in which it is already present. In this way, agriculture will not disperse to an uncolonized "frontier" cell until the suitability of that cell for farming is equal to or greater than that of land already farmed.

This work represents a first attempt at bottom-up modeling, at the scale and resolution used here. We recognize that other social and environmental drivers could have also affected the dispersal of agropastoral systems. The factors considered here, however, do include many that have been widely considered important in prior research and so warrant inclusion. This also helps make this work complementary to other research on the spread of agriculture in Europe.

3.1. Modeling experiments

We created the modeling laboratory for the experiments reported here in the widely used NetLogo platform (Wilensky, 1999). We imported a georeferenced map of the Iberian Peninsula into the NetLogo environment and divided it into 5 km × 5 km cellular agents, reasonable-sized farming catchments for small, early Neolithic communities (e.g., Barton et al., 2010) and a pragmatic compromise between geographical resolution and computational requirements (Fig. 2). Although computational agents can be entities that move across a virtual world, they can also be conceptualized as stationary entities or localities (e.g., trees in a forest or farms on a landscape) that can propagate conditions or information to other localities. Models that use such cellular agents are also sometimes called *cellular automata*, or CA (Mitchell, 2009). Our modeling environment is designed in this way; each 5 km × 5 km cell represents a patch of the Iberian landscape that may or may not have the property of agriculture (i.e., represents a region occupied by farmer/herders), as well as environmental properties. In our simulations, agriculture can spread from cell to cell on the basis of decision rules outlined below, rather than representing dispersal by mobile agents that move from cell to cell (Fig. 2). In this context, we set up a series of experiments in which all of the following decision/transition rules and contextual parameters were varied systematically through a range of values.

- We varied *dispersal methods* from simple neighborhood spread (where agriculture spreads from a cell with agriculture to all adjacent cells without agriculture) to leap-frog spreading where agriculture spreads to another cell within a specified distance

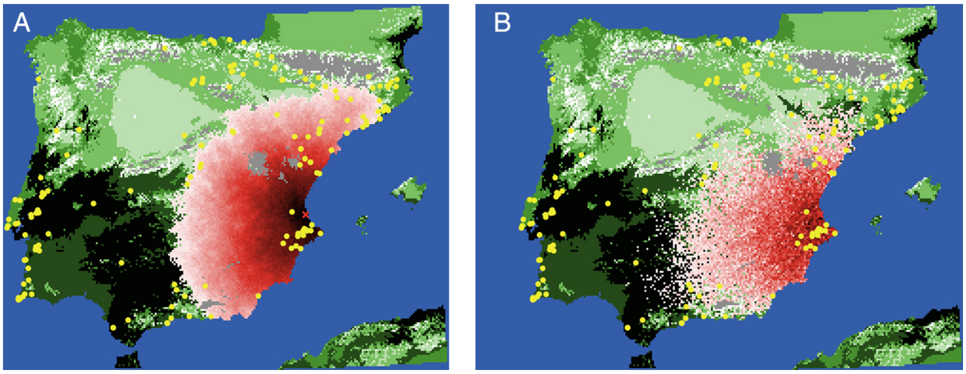


Fig. 2. Examples of spread models in action. (A) Shows neighborhood dispersal algorithm with minimum ecological suitability for spreading set to 5. (B) Shows leapfrog dispersal algorithm with maximum leap distance set to 10 cells and minimum ecological suitability for spreading set to 5. The "X" marks the starting point for the spread, and yellow dots show the location of dated Neolithic sites. The colors indicate the relative time of arrival of agriculture: darkest red is the oldest arrival time and lightest pink is the most recent arrival time. Underlying green shades show ecological suitability of cereal farming (see Fig. 1).

from the originating cell (Fig. 2). We varied the leap-frog distance from 1 to 50 cells (5-250 km).

- We started each experiment from a single 5 km × 5 km cell at one or more of the 15 starting locations around the perimeter of the peninsula mentioned above (Fig. 1). We also ran a set of control experiments in which the starting cell was located in the center of the peninsula (i.e., at Madrid) so that we also could compare the results of different possible *dispersal directions* with an impossible one (barring Neolithic air travel).
- For *ecological context*, we created a GIS base map of values representing suitability for cereal agriculture, combining climate parameters and topography, and imported it into our simulation. Each cell could assess suitability for agriculture and use that value to limit or encourage the spread of farming to another cell. Topography is derived from SRTM digital elevation models (http://www2.jpl.nasa.gov/srtm/dataprod.htm) and climate data are from the WorldClim database (Hijmans et al., 2005). In spite of pre-modern Holocene climate shifts (e.g., the Little Ice Age), Holocene climate has been comparatively stable compared with that of the Upper Pleistocene. Also paleoclimate information for this region is only available at very low spatial resolutions, especially compared to modern data such as WorldClim. And, as noted below, the computed value used for ecological suitability depends on topography as well as climate. For the experiments here, then, modern climate serves as a reasonable proxy for suitability for cereal agriculture. However, it may be informative to incorporate modeled paleoclimate information into future work if it can be suitably downscaled to the resolutions needed for all of Iberia. The procedure used to create this suitability map is detailed in the Supplemental Information. In brief, we assigned an index of agricultural suitability to each 5 km × 5 km cell that combined maximum spring temperatures, minimum March temperatures, total spring precipitation, and topographic slope. For all dispersal methods, we varied the threshold index value below which agriculture would not spread to a cell—from 0 (ecological context being unimportant) to a maximum of 10 (agriculture spreading only to the most suitable cells) (Fig. 1).
- To simulate an ideal despotic distribution type of spread to evaluate *demographic effects and anthropogenic landscape impacts*, we allowed agriculture to spread again to cells where agriculture already was present, but lowered the ecological suitability index of the cell by a percentage (varied across multiple experiments) for each time agriculture dispersed to that cell again (i.e., after the initial spread to the cell). In this way, increasing human population (represented by repeatedly spreading to a cell) caused decreasing suitability for subsequent agricultural spread. Similar to leapfrog, IDD dispersal could evaluate suitability within different sized radii around the initial cell. For the experiments we report here, a radius of 5 cells for evaluating suitability performed best and was used.

Stochasticity is embedded in the modeling algorithms to represent variation among farming communities in the application of decision/transition rules (e.g., due to local factors or differences in cultural knowledge). As agriculture begins to spread, the order by which cells with agriculture have the opportunity to spread farming to other cells varies randomly each model cycle. For leap-frog spread, a cell spreads farming to another cell chosen randomly within a specified maximum radius of spreading. When ecological context is considered, the suitability for agriculture—above a threshold which is deemed unsuitable for agriculture—affects the probability that agriculture will spread to that cell (i.e., it is not deterministic). The more ecologically suited for agriculture, the higher the probability that agriculture will spread to that cell.

This stochasticity means that each model run will produce somewhat different results, even with no change in any parameter settings. It is important, then, for this kind of modeling to capture a representative sample of model variability resulting from stochastic effects; a single run may or may not be representative. In order to assess how many repetitions of a scenario are needed to do this, we conducted a series of sensitivity tests (see Supplemental Information for details) with different model scenarios. We found that results from 10 to 20 repetitions of a particular scenario are statistically equivalent to 100 repetitions of the same scenario, and that 30-40 repetitions produced distributions of results that could not be distinguished from the results of 100 runs. For the experiments reported here, then, we repeated every scenario (i.e., every combination of parameter values) 50 times.

3.2. Evaluating model results

As discussed above, we carried out many experiments for multiple combinations of parameter values, and then compared the results to the empirical record of dated Neolithic sites. Our method of empirical validation is similar to that used by other modeling studies, but adapted for the CA modeling approach we used. Although we model the dispersion of agropastoral systems over space (i.e., across the GIS representation of the Iberian Peninsula) and through time, we have not attempted to scale our model time steps to a calendric time scale. Rather, we record the number of modeling time steps it takes for agriculture to arrive at the dated Neolithic sites, located on the GIS map of the peninsula. We then calculate the correlation coefficient (R) for the site radiocarbon dates and the arrival time of agriculture in model steps. A higher *negative* correlation indicates that a model scenario better fits the empirical dataset. (Modeled arrival times count *up* for increasingly later arrivals while radiocarbon dates count *down* for increasingly younger sites.) Below, we present the modeling results as three groups of experiments.

4. Results

4.1. Experiment Group 1

The primary goals of this set of experiments were to evaluate different starting points for the spread of farming in the Iberian Peninsula and the impacts of evaluating modeling results against radiocarbon dates on different materials. The details of the experimental protocol are shown in Table 1. Three different movement strategies and the effects considering environmental suitability for cereal agriculture were modeled and the results evaluated against four different sets of radiocarbon dates for early Neolithic sites in the Iberian Peninsula. Two dating sets limited model evaluation to sites meeting strict criteria for radiocarbon date reliability. One set of dated sites was limited to only those sites with dates on remains of domestic taxa. A second set was limited to only those sites with dates on any short-lived taxa, using mean radiocarbon dates on domestic taxa where available and dates on other short-lived taxa where there has been no direct dating of domestic taxa remains. Model results also were evaluated against all early Neolithic sites, as described above, using dating sets we term *best* and *oldest*. *Best* dates refer to mean radiocarbon dates on remains of domestic taxa where available, dates on short-lived taxa where no domestic taxa were dated, and dates on other materials where no dates on short-lived taxa were available. *Oldest* dates refer the oldest mean radiocarbon date at a site, regardless of the material dated. While the majority of the dates in the *domestic* group displayed a more restricted temporal dispersion than the other two, larger sets, it did not restrict the overall range of dates used to evaluate model results (see Supplemental Information).

This experimental protocol produced 340 different model scenarios (20 for each of the 16 perimeter starting points plus Madrid) each of which were repeated 50 times, for a total of 17,000 individual model runs. Simple Pearson correlation coefficients (R) and the probability that R is due to chance (p) were calculated using the R statistical package (R Core Team, 2014) to compare the performance of different experiment scenarios. Because many of the hypothetical starting points are very unlikely given archeological knowledge of the European Neolthic and Neolithic technology, it is not surprising that results from the majority of the scenarios exhibited a poor fit with the empirical datasets of dated Neolithic sites. However, 57 scenarios produced negative correlations with the empirical data with $p \leq 0.05$, indicating a low probability that the associations are due to chance alone. These are shown in Fig. 3. Several insights can be gained from these experiments.

The best correlations between model results and dated Neolithic sites occurred when the evaluation dating set was limited to the 39 sites with radiocarbon dating of remains of domesticates, with $R = -0.38$ in the best case. When this was expanded with the addition of 45 sites with dates on short-lived taxa, no model scenario matched the empirical data sufficiently well for a correlation to have $p \leq 0.05$. When the evaluation set was further expanded to include an additional 40 sites with other taxa (*best* dates), many other scenarios had correlations with $p \leq 0.05$, though none had R values as high as when the evaluation dataset consisted of sites with date on domesticates only. Finally, focusing on the oldest dates at a site does not seem to be a useful way to select an evaluation dataset. While several scenarios had correlations with $p \leq 0.05$, none had strong measures of association, with $R \geq -0.25$ (i.e., *worse* than -0.25, since *negative* correlations indicate good matches between the model and the empirical dates) in all cases and $R > -0.20$ in most cases. In fact, several impossible scenarios in which Madrid was the starting point for the spread of agriculture had correlations with $p \leq 0.05$, but all had $R \geq 0.21$ further suggesting that evaluating model results against this set of dates is not useful.

Another feature of the results from this group of experiments is that the best performing scenarios were those in which agriculture spread to adjacent cells (i.e., neighborhood and IDD spreading algorithms) and agricultural dispersal responded to ecological suitability for primitive wheat, calculated from climate and terrain values. With the best performing scenarios, agriculture was not allowed to spread to the least suitable land for wheat farming (ecological suitability index < 3) and the probability of spread increased with higher suitability (from ecological suitability index = 3 to 10).

Finally, the starting points with the best fit are all in eastern and southeastern Spain (Figs. 1 and 3). The lack of fit between the archeological data and model scenarios in which agricultural dispersals initiate from northwestern Spain or Portugal is not surprising. However, given that most quantitative and narrative models assume that agriculture spread to the Iberian Peninsula along the Mediterranean coast, from Italy to France to Spain, the best fit for a southeastern origins for farming and a lack of fit for any northeastern origin point are unexpected. However, as we discuss in more detail below, this is not unreasonable given the land/sea configuration of the western Mediterranean region more generally and that the earliest dates for Impressed Ware ceramics are from southern Italy (Manen, 2014).

4.2. Experiment Group 2

Experiment Groups 2 and 3 were designed to drill down into the results of Experiment Group 1, focusing on the geography of agricultural dispersals in the Iberian Peninsula (Experiment Groups 2) and exploring the space of varying movement strategies and varying sensitivity to ecological conditions (Experiment Group 3). The fact that the best fitting models in Experiment Group 1 began the spread of farming at start points in southeastern Spain, combined with a modest measure of association led us to ask if there might have been more than one starting point for Neolithic dispersals—and that the single point best fits represented the closest *single average* between *multiple initial locales* for the spread of farming.

To assess this question, we systematically evaluated the performance of scenarios with different combinations of starting points. We used the same five movement strategies used in Experiment Group 1 (see Table 1 and Fig. 3). We then created a scenario in which all 16 perimeter start locales were used to initiate agricultural dispersal. Next, we dropped out each start point, one at a time, evaluating the resulting model scenario against sites with radiocarbon dates on domesticates. If dropping the start point improved the fit with the empirical data (measured by the correlation coefficient), we kept it out; if the fit was worse without the start point, we returned it to the model scenario. The results can be seen in Fig. 4.

One combination performed considerably better than any single start point in the scenarios of Experiment Group 1, and also much better than any other combination of start points: one starting point from the northeast perimeter of the Peninsula (Rio Llobregat), one from the eastern perimeter (Rio Jucar), and one from the south (near the modern city of Málaga). Several other scenarios with paired start points performed slightly better than any single point in Experiment Group 1: three with start points from the east/southeast and northeast (R. Almanzora and R. Llobregat, R. Jucar and R. Llobregat, R. Segura and R. Llobregat), two with both start points from the east/southeast (R. Jucar and R. Segura, R. Jucar and R. Almanzora), and one with start points from the east and

Table 1
Parameters for Experiment Group 1.

Movement strategy	Neighborhood spread	Leap-frog spread	Ideal despotic distribution
Movement distance	All adjacent neighbors without agriculture	A cell without agriculture within 5 cells (25 km)	Most suitable adjacent cell(s)
Ecological threshold	0: ecology unimportant *or* 3: no spread to values less than 3 and increasing probability of spread from 3 to 10	0: ecology unimportant *or* 3: no spread to values less than 3 and increasing probability of spread from 3 to 10	3: no spread to values less than 3 and increasing probability of spread from 3 to 10. But this declines by the cost of increasing population
Demography	Not considered	Not considered	5% decrease in ecological suitability index for each time a agriculture re-spreads to a cell with agriculture

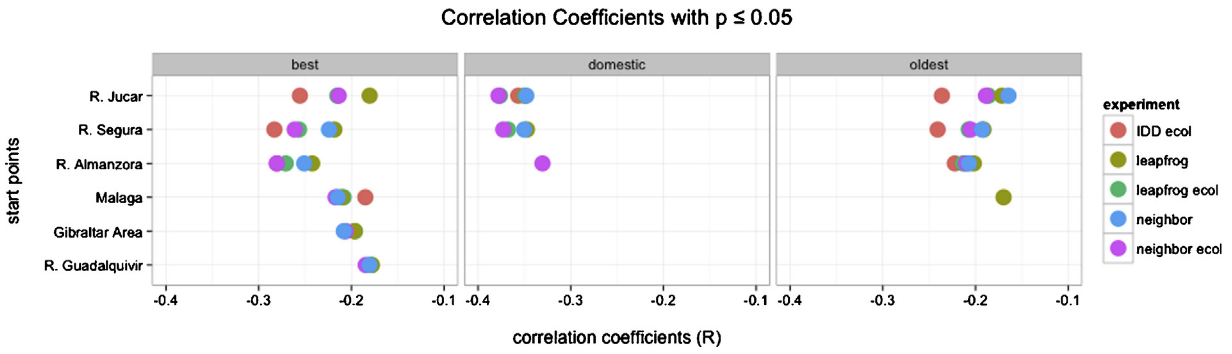


Fig. 3. Correlation coefficients for results of Experiment Group 1 with $p \leq 0.05$ evaluated against different sets of radiocarbon dates. Colors indicate different dispersal strategies employed by agents. Positive correlations and models starting from Madrid are excluded. Leap distance = 5 for leapfrog and the radius of evaluation for IDD dispersal.

south (R. Jucar and Malaga). Notably, the Rio Jucar start point is present in five of the best seven combinations of locales for the start of agricultural dispersals in the Iberian Peninsula. The worst combination of start points, of those for correlations with $p \leq 0.05$ is one from the north and one from the south (R. Llobregat and Malaga). Taken together, the central eastern perimeter remains the locale that figures consistently in most of the best performing model scenarios for agricultural dispersals, but the models better fit the empirical data when dispersals starting from both the northeast and the south are also included.

With multiple start points, movement strategies performed differently than in Experiment Group 1, with leapfrog in particular performing better with multiple start points. Investigating this difference in more detail was the basis for Experiment Group 3.

4.3. Experiment Group 3

For this group of experiments, we systematically varied all decision/transition rules for the spread of agriculture. For leapfrog spreading, we modeled the maximum dispersal distance at 2, 5,

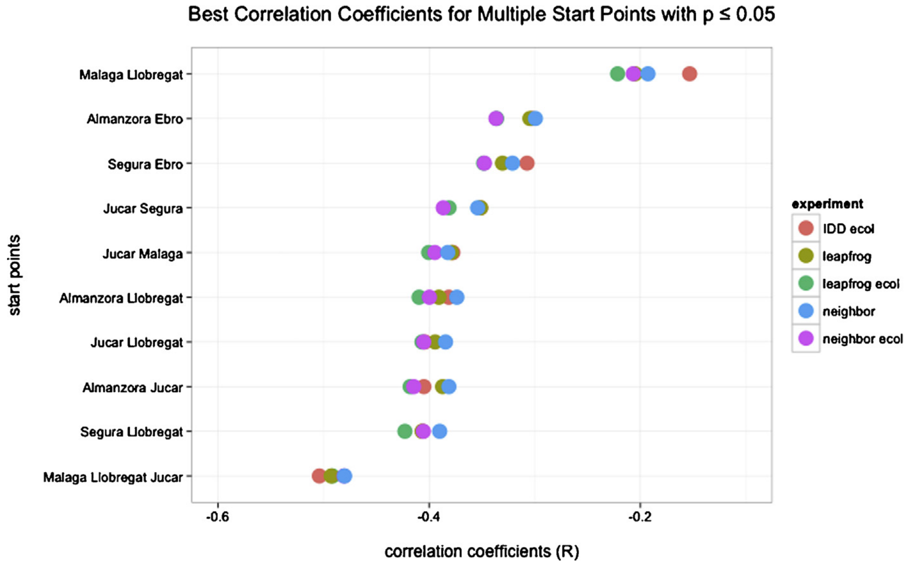


Fig. 4. Correlation coefficients for results of Experiment Group 2 with $p \leq 0.05$ for different combinations of starting points for agricultural dispersals. Colors indicate different dispersal strategies employed by agents. Only dates on domestic taxa used for model evaluation. Positive correlations and models starting from Madrid are excluded. Leap distance = 5 for leapfrog and the radius of evaluation for IDD dispersal.

10, and 20 cells from the originating cell (equal to 10, 25, 50, and 100 km). For Ideal Despotic Distribution dispersal, we varied the cost of increasing population by decreasing the ecological suitability by 2%, 5%, 10%, 20%, and 50% for each time agriculture spread to a cell that already has a farming population. (A decrease of 100% is equivalent to only spreading to cells that do not yet have agriculture, something already modeled in both neighborhood and leapfrog spreading algorithms.) We also varied the sensitivity to environmental context by changing the threshold for the minimum level of suitability for cereal farming that is acceptable for agricultural dispersals from 0 to 6 (out of a maximum possible suitability index value of 10). We used the best combination of starting points from Experiment Group 2. Again, we evaluated model results against the ages of Neolithic sites with radiocarbon dates on domestic taxa.

Results from this set of experiments can be seen in Fig. 5. It is clear that combinations of decision rules different from those used in Experiment Groups 1 and 2 produce models that better fit the empirical archeological data. Beyond this, it is difficult to characterize the many rule combinations that result in correlation coefficients between −0.50 and −0.56. One clear feature is that the worst performing models all had low values for the cost of increasing population in cells with farming populations. In other words, models in which population increase is unimportant were a worse fit to the archeological record than models in which even a slight increase in population significantly reduced a cell's desirability.

For models with $R < -0.50$, there are slight jumps in R between the best five and the rest, and the top three and the rest. These are convenient breaks to assess which combination of local decision rules produced dispersal models that better fit the empirical data. Of the top five, all have moderate to high costs for population increase: occupy cost = 20%, 50%, and 100% (the leapfrog algorithm, which does not spread to cells with agriculture, is equivalent to a cost of 100%). Additionally, three of the top five use a leapfrog algorithm, meaning that some of the best models are those in which a simple dispersal to adjacent cells—a wave of advance—does not fit the data as well as a dispersal in which new farming settlements are founded at a considerable distance (25-100 km for the models tested here) from an originating settlement. This is consistent with a tendency for new agricultural settlements to be established in cells with few or no farming populations (the other two top performing scenarios). In other words, the local decision rules for the dispersal of agriculture in the Iberian Peninsula favored pioneering strategies where new settlements would be located at a some distance (more than a day's walk) from existing settlements or otherwise avoided occupied areas. Finally, in all of the top five performing models, the minimum acceptable land for establishing new farming settlements has an index value of 5-6, the midpoint of the range of ecological suitability index values. This suggests that a scenario in which farmers who sought out only highly suitable land for cereal farming best fits the archeological evidence.

5. Discussion

In this paper, we illustrate the potential of bottom-up modeling for investigating the dispersal of agropastoral economies and lifeways in Europe, focusing on the Iberian Peninsula as a case study. Additionally, we use computational modeling more as a method of formalizing and testing multiple, complex hypotheses about local-scale decision rules than as means of quantitatively characterizing agricultural dispersals at the continental scale. That is, we are not taking a modeling approach to reconstructing the past. Nor do we claim that any of our models can reveal the 'true' past. What we have shown, however, is that models that implement certain local rules and geographic locales for the initiation of agricultural dispersals better fit the empirical archeological data of the Iberian Peninsula than do models constructed with other parameters. Taking this approach offers new insights about the social and ecological drivers responsible for the spread of agriculture, as well as confirming prior interpretations of the archeological evidence.

For example, our modeling experiments confirm an intuitive, widely held assumption about the direction for the spread of farming in the Iberian Peninsula. In all model scenarios that fit the empirical data with a low probability of an association due only to chance, agricultural dispersal began somewhere along the Mediterranean perimeter of the peninsula. In spite of the discovery of a few early Neolithic sites in Portugal and the possibility of LBK dispersals around the western terminus of the Pyrenees, no scenario initialized in the west or northwest of the peninsula fits the empirical data as well as do ones initialized along the Mediterranean perimeter. Although this is an unsurprising conclusion, it bears mentioning because it shows that the bottom-up modeling approach that we take produces results in line with relatively clear-cut trends in the empirical data.

The research presented here also underscores the importance of the reliability of the archeological data used to develop and (in our case) evaluate models of any kind, narrative, mathematical, or computational. Barring contamination or laboratory error, radiocarbon dates on the remains of domestic taxa should provide unequivocal evidence of the presence of farming economies at particular times in the past. There is less certainty that dates obtained from other organic materials are associated with agricultural practices (Bernabeu et al., 2001; Zilhao, 2001). Any model that purports to represent the spread of agriculture should produce results that are at least able to account for the space–time distribution of Neolithic sites based on dates from domestic taxa. Many of the scenarios modeled here cannot account for the distribution of Neolithic sites dated in this way, but some do so rather well. However, those models that do fit the unequivocally dated Neolithic sites do not fit well with larger samples of sites where the oldest dates are used, regardless of the material dated, or even or even sites that include those dated from short-lived taxa in addition to those dated from domesticates. This strongly suggests that we need to carefully assess the chronological quality of the individual radiocarbon dates (or samples used for any other dating method) that are combined into a regional dataset used to analyze of the dynamics of agricultural dispersals (Gkiasta et al., 2003). A radiocarbon dataset that includes substantial values that do not accurately represent the age of the phenomenon of interest can lead to spurious results and inaccurate models.

This work offers insights and questions about the routes by which agropastoral systems reached and spread across the Iberian Peninsula. To date, many mathematical models have represented agricultural dispersal as spreading around the Mediterranean coast of Europe and entering the peninsula from the northeast (e.g. Ackland et al., 2007). However, the single points of origin for the Iberian Neolithic for the scenarios that best fit the empirical data are located in eastern and southern Spain (Figs. 1 and 3), suggesting that Neolithic farmers may have arrived on the Iberian coast by boat. But scenarios with multiple origin points display even better correspondence with the archeological evidence, with a best fit from simultaneous initiation of agricultural dispersals in the northeast, east, and southeast Iberian coast (Figs. 1 and 4) (also see Isern et al., 2014; Lemmen et al., 2011). Such a scenario could be produced by temporally equivalent arrivals of farmers spreading along the Mediterranean coast of Europe by land and other settlers arriving by boat from Africa. Alternatively, it could result from farming groups spreading along the Mediterranean coast by sea at several locales over a short period of initial colonization (i.e., simultaneous

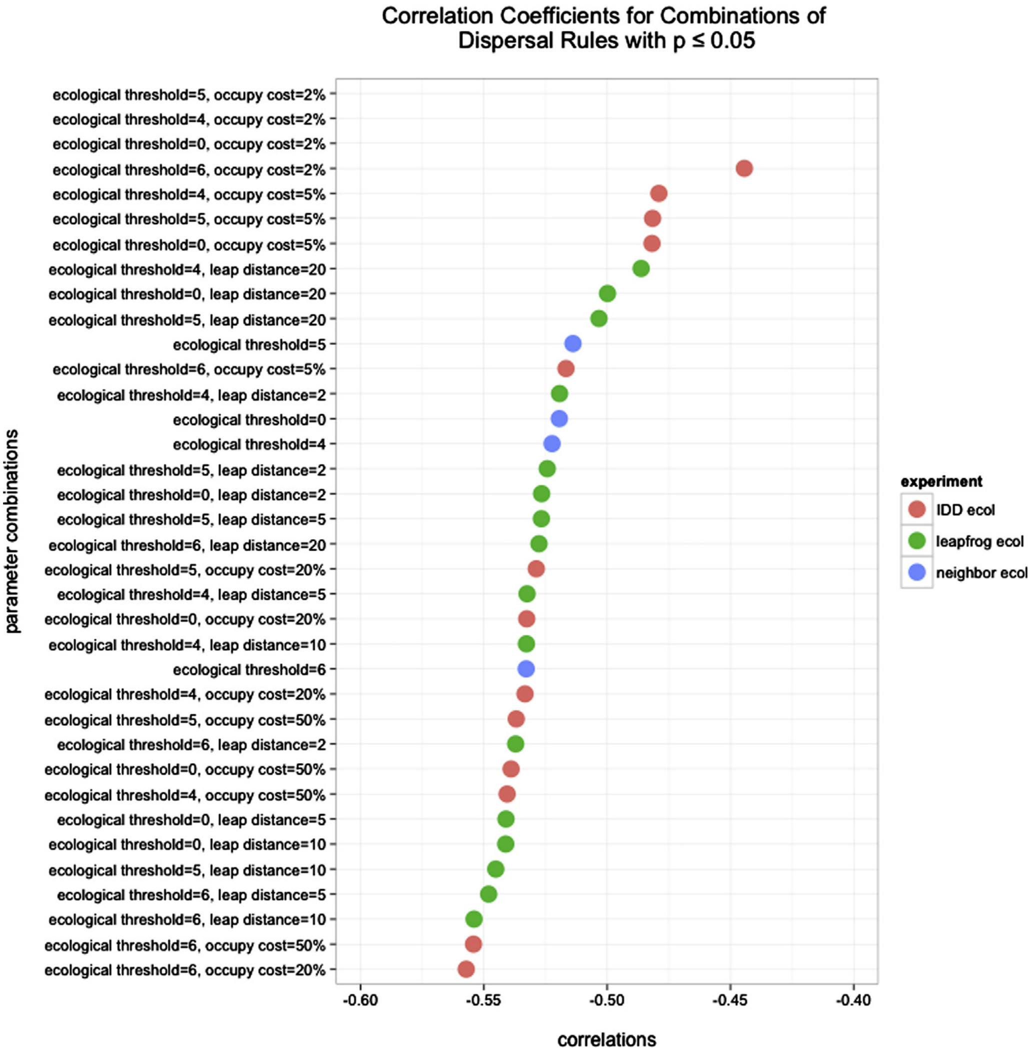


Fig. 5. Correlation coefficients for results of Experiment Group 3 with $p \leq 0.05$ for different combinations of parameters for modeled dispersal decisions. Colors indicate different dispersal strategies employed by agents. The best combination of starting points from Experiment Group 2 used. Only dates on domestic taxa used for model evaluation. Positive correlations and models starting from Madrid are excluded. IDD models used a radius = 5 cells for suitability evaluation.

within the resolution limits of radiocarbon dating), followed by a lengthier period of land-base dispersals inland from each initial coastal settlement (Zilhao, 2001). Both alternatives involve seafaring as an important component of agricultural dispersals (see also Davison et al., 2006).

Finally, this approach provides insights into the local-scale decisions by Neolithic communities that resulted in the large-scale dispersal of agriculture. The model scenarios that best fit the empirical data were driven by rules for the establishment of new farming communities that avoided already settled regions, sometimes by moving up to 50-100 km from the originating area, and preferred the highest quality land available, while completely avoiding land that was even slightly marginal for growing wheat (see Fort et al., 2012; Zilhao, 2003). This is also consistent with the nature of archeological evidence for early Neolithic settlement across the peninsula, dominated by small hamlets or farmsteads and lacking nucleated villages.

Bottom-up and top-down modeling are complementary approaches to formalizing hypotheses about the dynamics of human societies. Top-down modeling is important for describing general trends in societies across large scales and over long time periods; bottom-up modeling is key to understanding the decisions and practices of the individual members of those societies that resulted in the general trends observed. The formalization inherent in both kinds of modeling approaches is an essential step for the ability to systematically compare and test hypotheses about the space–time dynamics of past human society against a fragmentary and incomplete archeological record. Formalization also offers the potential for increased transparency and replicability in specifying hypotheses and building theory about the drivers of social change. For this potential to be realized, however, it is important that the details of formal models be accessible to other scientists, as well as the narrative describing the results of their application. To this end, we have published the NetLogo code and associated data files used in the research reported here in the CoMSES Computational Model Library (https://www.openabm.org/model/4447/version/3/view). We hope to build on this work to further investigate local decisions and routes for agricultural dispersals and also invite others to leverage our research to better understand this most important transformation in the human career.

Acknowledgements

This research was funded in part by National Science Foundation grant DEB-1313727 from the Coupled Natural and Human Systems

Program and from Spanish Ministry of Science grant HAR2012-33111 from the MesoCocina: Los últimos caza-recolectores y el paradigma de la neolitización en el mediterráneo peninsular.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2015.03.015.

References

Ackland, G.J., Signitzer, M., Stratford, K., Cohen, M.H., 2007. Cultural hitchhiking on the wave of advance of beneficial technologies. Proc. Natl. Acad. Sci. 104, 8714–8719, http://dx.doi.org/10.1073/pnas.0702469104.

Ammerman, A.J., Cavalli-Sforza, L.L., 1984. The Neolithic Transition and the Genetics of Populations in Europe. Princeton University Press, Princeton.

Aoki, K., Shida, M., Shigesada, N., 1996. Travelling wave solutions for the spread of farmers into a region occupied by hunter-gatherers. Theor. Popul. Biol. 50, 1–17.

Bankes, S.C., Lempert, R., Popper, S., 2002. Making computational social science effective: epistemology. methodology, and technology. Soc. Sci. Comput. Rev. 20, 377–388.

Barton, C.M., Bernabeu Aubán, J., Aura Tortosa, J.E., Molina Balaguer, L., 2001. Historical Contingency, Nonlinearity, and the Neolithization of the Western Mediterranean. Presented at the 66th Annual Meeting of the Society for American Archaeology.

Barton, C.M., Ullah, I.I.T., Mitasova, H., 2010. Computational modeling and Neolithic socioecological dynamics: a case study from southwest Asia. Am. Antiq. 75, 364–386.

Bentley, R.A., Layton, R.H., Tehrani, J., 2009. Kinship, marriage, and the genetics of past human dispersals. Hum. Biol. 81, 159–179.

Bernabeu Aubán, J., 2002. The social and symbolic context of Neolithization. In: The Mediterranean Landscape of the Neolithic. Unicersidad de Valencia, Saguntum Extra-5, pp. 209–233.

Bernabeu Auban, J., Molina Balaguer, L., Orozco Köhler, T., Díaz Castillo, A., Barton, C.M., 2008. Early neolithic at the Serpis Valley, Alicante, Spain. In: Diniz, M. (Ed.), The Early Neolithic in the Iberian Peninsula. Regional and Transregional Components. Proceedings of the XV World Congress (Lisbon, 2006), International Series, BAR, Oxford, pp. 53–59.

Bernabeu, J., Barton, C.M., Perez Ripoll, M., 2001. A taphonomic perspective on Neolithic beginnings: theory, interpretation, and empirical data in the western Mediterranean. J. Archaeol. Sci. 28, 597–612.

Bernabeu, J., García Puchol, O., Pardo, S., Barton, M., McClure, S.B., 2014. Socioecological dynamics at the time of Neolithic transition in Iberia. Environ. Archaeol. 19, 214–225, http://dx.doi.org/10.1179/1749631414Y. 0000000032.

Bocquet-Appel, J.-P., Naji, S., Linden, M.V., Kozlowski, J.K., 2009. Detection of diffusion and contact zones of early farming in Europe from the space–time distribution of 14C dates. J. Archaeol. Sci. 36, 807–820, http://dx.doi.org/10.1016/j.jas.2008.11.004.

Bogucki, P., 2000. How agriculture came to north-central Europe. In: Europe's First Farmers., pp. 197–218.

Cortés Sánchez, M., Jiménez Espejo, F.J., Simón Vallejo, M.D., Gibaja Bao, J.F., Carvalho, A.F., Martinez-Ruiz, F., Gamiz, M.R., Flores, J.-A., Paytan, A., López Sáez, J.A., Peña-Chocarro, L., Carrión, J.S., Morales Muñiz, A., Roselló Izquierdo, E., Riquelme Cantal, J.A., Dean, R.M., Salgueiro, E., Martínez Sánchez, R.M., De la Rubia de Gracia, J.J., Lozano Francisco, M.C., Vera Peláez, J.L., Rodríguez, L.L., Bicho, N.F., 2012. The Mesolithic–Neolithic transition in southern Iberia. Quat. Res. 77, 221–234, http://dx.doi.org/10.1016/j.yqres.2011.12.003.

Davison, K., Dolukhanov, P., Sarson, G.R., Shukurov, A., 2006. The role of waterways in the spread of the Neolithic. J. Archaeol. Sci. 33, 641–652.

Dunnell, R.C., 1982. Science, social science, and common sense: the agonizing dilemma of modern archaeology. J. Anthropol. Res. 38, 1–25.

Fedotov, S., Moss, D., Campos, D., 2008. Stochastic model for population migration and the growth of human settlements during the Neolithic transition. Phys. Rev. E 78, 026107, http://dx.doi.org/10.1103/PhysRevE.78.026107.

Fisher, R.A., 1937. The wave of advance of advantageous genes. Ann. Eugen. 7, 355–369.

Fort, J., 2012. Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. Proc. Natl. Acad. Sci. 109, 18669–18673.

Fort, J., Jana, D., Humet, J., 2004. Multidelayed random walks: theory and application to the Neolithic transition in Europe. Phys. Rev. E 70 (031913), 1–6.

Fort, J., Pujol, T., Linden, M.V., 2012. Modelling the Neolithic transition in the near East and Europe. Am. Antiq. 77, 203–219.

Fretwell, S.D., Lucas Jr., H.L., 1970. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheor. XIX.

Gkiasta, M., Russell, T., Shennan, S., Steele, J., 2003. Neolithic transition in Europe: the radiocarbon record revisited. Antiq. Oxf. 77, 45–62.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Moolj, W., Railsback, S.F., Thulke, H., Weiner, J., Wiegand, T., DeAngelis, D., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310, 987–991.

Guilaine, J., 2001. LA diffusion de L'Agriculture en Europe: une hypothese arythmique. Zephyrvs 53–54, 267–272.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978, http://dx.doi.org/10.1002/joc.1276.

Isern, N., Fort, J., 2010. Anisotropic dispersion, space competition and the slowdown of the Neolithic transition. New J. Phys. 12, 123002.

Isern, N., Fort, J., 2012. Modelling the effect of Mesolithic populations on the slowdown of the Neolithic transition. J. Archaeol. Sci. 39, 3671–3676, http://dx.doi.org/10.1016/j.jas.2012.06.027.

Isern, N., Fort, J., Carvalho, A.F., Gibaja, J.F., Ibañez, J.J., 2014. The neolithic transition in the Iberian Peninsula: data analysis and modeling. J. Archaeol. Method Theory 21, 447–460.

Kennett, D., Anderson, A., Winterhalder, B., 2006. The ideal free distribution, food production and the colonization of Oceania. In: Kennett, D., Winterhalder, B. (Eds.), Foraging Theory and the Transition to Agriculture. Smithsonian Institution Press, Washington, DC, pp. 265–288.

Lake, M.W., 2015. Explaining the past with ABM: on modelling philosophy. In: Wurzer, G., Kowarik, K., Reschreiter, H. (Eds.), Agent-Based Modeling and Simulation in Archaeology. Advances in Geographic Information Science. Springer, Switzerland, pp. 1–35.

Lemmen, C., Gronenborn, D., Wirtz, K.W., 2011. A simulation of the Neolithic transition in Western Eurasia. J. Archaeol. Sci. 38, 3459–3470, http://dx.doi.org/10.1016/j.jas.2011.08.008.

Linstädter, J., Medved, I., Solich, M., Weniger, G.-C., 2012. Neolithisation process within the Alboran territory: models and possible African impact. Quat. Int. 274, 219–232, http://dx.doi.org/10.1016/j.quaint.2012.01.013.

Manen, C., 2014. Spatial, chronological and cultural dynamics of the Neolithization in the Western Mediterranean. In: La Transition Néolothique En Méditerranée, Archives d'Ecologie Préhistorique. Ed. Errance., pp. 405–418.

McClure, S., Jochim, M.A., Barton, C.M., 2006. Behavioral ecology, domestic animals, and land use during the transition to agriculture in Valencia, eastern Spain. In: Kennett, D., Winterhalder, B. (Eds.), Foraging Theory and the Transition to Agriculture. Smithsonian Institution Press, Washington, DC, pp. 197–216.

McClure, S.B., Barton, C.M., Jochim, M.A., 2009. Human behavioral ecology and climate change during the transition to agriculture in Valencia, eastern Spain. J. Anthropol. Res. 65.

Miller, J.H., Page, S.E., 2007. Complex Adaptive Systems: An Introduction to Computational Models of Social Life. Princeton University Press, Princeton, NJ.

Mitchell, M., 2009. Complexity: A Guided Tour. Oxford University Press, USA.

Pérez-Losada, J., Fort, J., 2011. Spatial dimensions increase the effect of cultural drift. J. Archaeol. Sci. 38, 1294–1299.

Pinhasi, R., Fort, J., Ammerman, A.J., 2005. Tracing the origin and spread of agriculture in Europe. PLoS Biol. 3, e410.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Shennan, S., 2009. Evolutionary demography and the population history of the European Early Neolithic. Hum. Biol. 81, 339–355.

Steele, J., 2009. Human dispersals: mathematical models and the archaeological record. Hum. Biol. 81, 121–140.

Utrilla Miranda, P., Montes Ramírez, L., 2009. El Mesolítico Geométrico en la Península Ibérica. Universidad de Zaragoza, Monografías Arqueológicas.

van der Leeuw, S.E., 2004. Why model? Cybern. Syst. Int. J. 35, 117–128, http://dx.doi.org/10.1080/01969720490426803.

Wilensky, U., 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University.

Zeder, M.A., 2008. Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. Proc. Natl. Acad. Sci. 105, 11597–11604, http://dx.doi.org/10.1073/pnas.0801317105.

Zilhao, J., 2001. Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. Proc. Natl. Acad. Sci. U. S. A. 98, 14180.

Zilhao, J., 2003. The Neolithic transition in Portugal and the role of demic diffusion in the spread of agriculture across West Mediterranean Europe. In: The Widening Harvest. The Neolithic Transition in Europe: Looking Back, Looking Forward. Archaeological Institute of America, pp. 207–223.

Zilhão, J., 2011. Time is on my side. . .. In: Hadjikoumis, A., Robinson, E., Viner, S. (Eds.), The Dynamics of Neolithisation in Europe. Oxbow Books, pp. 46–65.

Zvelebil, M., Lillie, M., 2000. The transition to agriculture in eastern Europe. In: Price, T.D. (Ed.), Europe's First Farmers. Cambridge University Press, Cambridge.

Supplemental Information for Modeling Initial Neolithic Dispersal. The first agricultural groups in West Mediterranean

Sites Used for Model Evaluation

Table 1 lists the sites and radiocarbon dates used for evaluating modeling experiment results. The criteria for selecting or rejecting dates is discussed in the main text. These represent all dates for the Early Neolithic in the Iberian Peninsula available at the time of writing.

Ecological Suitability Index

We classified landscape cells based on their suitability for cereal agriculture, focusing on wheat, using a combination of terrain and climate parameters (Bevan and Conolly, 2004; Lersten, 1987; López, 1991; Wardlaw et al., 1989). These are summarized in Table 2.

Terrain Classification

Terrain was derived from a 90m resolution SRTM (Shuttle Radar Topography Mission) DEM. Based on analyses by Bevan and Conolly (Bevan and Conolly, 2004), slopes ≤ 10° were considered most likely to be cultivated without terracing, and slopes ≤ 5° were considered the most preferable due to less loss of rainfall due to runoff and reduced chance for erosion. Slopes higher than 10° were classified as potentially cultivatable, given the coarse spatial resolution, but less desirable; they were divided into those between 10° and 15°, and those above 15°.

Climate

Climate parameters were derived from WorldClim database (Hijmans et al., 2005)(http://www.worldclim.org). We acquired gridded climate maps for monthly maximum temperatures, monthly minimum temperatures, and total monthly precipitation, chosen because of the impact of these climate parameters on wheat growth (Lersten, 1987; López, 1991; Wardlaw et al., 1989). Index values were assigned as shown in Table 2.

The ecological suitability index was created by summing the three climate index maps and slope index map. The resulting map was upscaled to a 5km resolution when uploaded to NetLogo (using the NetLogo GIS Extension) to provide grid cells with a sufficiently fine resolution to capture the likely catchments of simple agriculturalists while coarse-grained enough to allow for rapid simulation runs.

Sensitivity Tests for Repeated Model Iterations

Because of the inherent stochasticity in agent based models, where each agent makes decisions based on internal rules and the environmental context, including the state of neighboring agents in this study, no two model runs produce exactly the same results—even when initiated with exactly the same parameters. So it is necessary to do multiple runs for each scenario tested. But because the sensitivity of the interactions of agent decision rules and environmental parameters varies with every model created, there is no standard rule for the appropriate number repetitions needed. We estimated the appropriate number of repetitions for the modeling environment used here by a set of systematic sensitivity tests.

We created a scenario for each of the dispersal algorithms: neighborhood dispersal, leapfrog dispersal, and Ideal Free Distribution dispersal. We repeated each of these scenarios 100 times. We then randomly sampled 10, 20, 30, 40, 50, 60, 70, 80, and 90 runs from each of the 100 repetition sets and compared them statistically to the full 100 repetitions. These can be seen graphically in Figures 1-3. We also conducted *equivalence tests* using the 95% confidence intervals (Robinson and Froese, 2004) to assess the likelihood that the smaller samples differed from the large repetition set. All of the smaller number of repetitions were found to be equivalent to the large repetition set for each of the dispersal algorithms at $p \leq 0.05$. However, it is apparent from Figures 1-3 that smaller samples differed to an increasing degree from the 100 repetition set. The 50 repetition samples showed a distribution matching all larger samples, as well as being statistically equivalent to the large repetition sets. Hence, to be conservative, we decided to repeat each scenario 50 times.

Distribution of Dates on Different Materials

Figure 4 shows the distribution of means for dating subsets used to evaluate models: *domestic* (dates on domestic taxa only), *short-lived* (dates on domestic and short-lived taxa), and *best* (dates on domestic and short-lived taxa where available, plus dates on other materials when domestic and short-lived taxa are not available).

References Cited

Bevan, A., Conolly, J., 2004. GIS, Archaeological Survey, and Landscape Archaeology on the Island of Kythera, Greece. Journal of Field Archaeology 29, 123–138. doi:10.1179/jfa.2004.29.1-2.123

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965–1978. doi:10.1002/joc.1276

Lersten, N.R., 1987. Morphology and Anatomy of the Wheat Plant. American Society of Agronomy 13, 33–71.

López, L., 1991. Cultivos herbáceos. Ed. Mundi-Presa, Madrid.

Robinson, A.P., Froese, R.E., 2004. Model validation using equivalence tests. Ecological Modelling 176, 349–358. doi:10.1016/j.ecolmodel.2004.01.013

Wardlaw, I., Dawson, I., Munibi, P., 1989. The tolerance of wheat to hight temperatures during reproductive growth. 2. Grain development. Aust. J. Agric. Res. 40, 15–24.

Table 1. Sites and radiocarbon dates used to evaluate model experiment results. All dates are given as calibrated BP.

Site	**Province**	**Other Charcoal**		**Short Lived Taxa**		**Domestic Taxa**		**Oldest**	**Best**
		1 sigma range	**mean**	**1 sigma range**	**mean**	**1 sigma range**	**mean**	**mean**	**mean**
Abautz	Alava	6562 – 6673	6618					6618	6618
Abric de la Falguera	Alacant					7324 - 7488	7406	7406	7406
Almonda	Santarem			7324 - 7422	7373			7373	7373
Alto de Rodilla	Burgos			7005 - 7159	7082			7082	7082
Arenaza	Vizcaya					6788 - 6989	6889	6889	6889
Atxoste	Alava			7020 - 7244	7132			7132	7132
Barruecos	Caceres	6888 - 6999	6944					6944	6944
Benamer	Alacant			7431 - 7551	7491			7491	7491
Buraco da Pala	Braganza	6656 – 6726	6691					6691	6691
Ca Estrada	Girona			6485 - 6627	6556			6556	6556
Cabranosa	Faro			7474 - 7580	7527			7527	7527
Caldeirao	Santarem					7166 - 7412	7289	7289	7289
Can Bellsola	Barcelona	7024 - 7262	7143					7143	7143
Can Roqueta	Barcelona			7274 - 7415	7345			7345	7345
Canaleja 2	Caceres	7013 - 7169	7091					7091	7091
Cariguela	Granada			7168 - 7246	7207			7207	7207
Carrascal	Lisboa					7174 - 7252	7213	7213	7213
Casa da Moura	Leiria			6746 - 6894	6820			6820	6820
Casa Montero	Madrid	7332 - 7461	7397			7017 - 7167	7092	7397	7092
Castelo Belinho	Faro			6500 - 6662	6581			6581	6581
Cerro Virtud	Almeria			6797 - 6941	6869			6869	6869
Chaves	Huesca	7576 - 7672	7624			7435 - 7499	7467	7624	7467
Cingle del Mas Cremat	Castello			6795 - 6930	6863			6863	6863
Codella	Girona			6440 - 6625	6533			6533	6533
Costamar	Castello					6787 - 6888	6838	6838	6838
Cova Avellaner	Girona			6502 - 6739	6621			6621	6621
Cova Colomera	Lleida					6991 - 7156	7074	7074	7074
Cova de la Sarsa	Valencia					7338 - 7463	7401	7401	7401

Site	**Province**	**Other Charcoal**		**Short Lived Taxa**		**Domestic Taxa**		**Oldest**	**Best**
		1 sigma range	**mean**	**1 sigma range**	**mean**	**1 sigma range**	**mean**	**mean**	**mean**
Cova de lOr	Alacant					7332 - 7429	7381	7381	7381
Cova del Petroli	Castello	6796 - 6907	6852					6852	6852
Cova del Toll	Barcelona					7320 - 7417	7369	7369	7369
Cova del Vidre	Tarragona	6952 - 7233	7093					7093	7093
Cova den Pardo	Alacant					7463 - 7563	7513	7513	7513
Cova Font Major	Tarragona					7174 - 7270	7222	7222	7222
Cova Foradada	Tarragona			7017 - 7167	7092			7092	7092
Cova Fosca	Castello	7959 - 8154	8057					8057	8057
Cova Fosca Ebo	Alacant					7312 - 7416	7364	7364	7364
Cova Gran	Girona			6787 - 6928	6858			6858	6858
Cova Sant Marti	Alacant			6485 - 6627	6556			6556	6556
Cueva Chica de Santiago	Sevilla	6911 - 7172	7042					7042	7042
Cueva de la Dehesilla	Cadiz	7011 - 7274	7143					7143	7143
Cueva de la Higuera	Madrid			7029 - 7259	7144			7144	7144
Cueva de los Marmoles	Cordoba					7023 - 7164	7094	7094	7094
Cueva de los Murcielagos C.	Cordoba	7173 - 7262	7218			6991 - 7156	7074	7218	7074
Cueva del Gato	Zaragoza	7029 - 7252	7141					7141	7141
Cueva del Toro	Malaga	7167 - 7318	7243					7243	7243
Cueva Lobrega	Rioja			7005 - 7251	7128			7128	7128
El Barranquet	Valencia					7332 - 7478	7405	7405	7405
El Cavet	Tarragona	7476 - 7569	7523					7523	7523
El Congosto	Madrid			6791 - 6930	6861			6861	6861
El Mirador	Burgos	7851 - 7941	7896			7172 - 7267	7220	7896	7220
El Miron	Cantabria	6482 - 6715	6599			6300 - 6395	6348	6599	6348
El Tonto	Toledo					7029 - 7246	7138	7138	7138
Font de la Vena	Barcelona	6969 - 7243	7106					7106	7106
Font del Ros	Lleida	7426 - 7551	7489					7489	7489
Fuente Celada	Burgos			6943 - 7151	7047			7047	7047
Gruta do Correio-Mor	Lisboa	7177 - 7413	7295	7173 - 7317	7245			7295	7245
Hostal Guadalupe	Malaga			7175 - 7262	7219	7163 - 7246	7205	7219	7205
Huerto Raso	Huesca	7168 - 7293	7231					7231	7231

Site	**Province**	**Other Charcoal**		**Short Lived Taxa**		**Domestic Taxa**		**Oldest**	**Best**
		1 sigma range	**mean**	**1 sigma range**	**mean**	**1 sigma range**	**mean**	**mean**	**mean**
Kobaederra	Vizcaya	6301 - 6501	6401			6018 - 6280	6149	6401	6149
La Dou	Girona	6354 - 6497	6426					6426	6426
La Draga	Girona	7279 - 7417	7348			7019 - 7158	7089	7348	7089
La Lampara	Soria	7938 - 7995	7967	7294 - 7416	7355	7167 - 7259	7213	7967	7213
La Paleta	Toledo	7486 - 7581	7534			6636 - 6732	6684	7534	6684
La Revilla del Campo	Soria	7958 - 8007	7983	7256 - 7410	7333	7258 - 7314	7286	7983	7286
La Serreta	Barcelona	7328 - 7436	7382					7382	7382
La Vaquera	Segovia	7798 - 7955	7877	7324 - 7421	7373			7877	7373
Las Torrazas	Teruel	6305 - 6402	6354					6354	6354
Les Guixeres	Barcelona					7501 - 7575	7538	7538	7538
Los Cascajos	Navarra	7324 - 7419	7372			7030 - 7258	7144	7372	7144
Los Castillejos	Granada					7173 - 7271	7222	7222	7222
Los Gitanos	Cantabria			6678 - 6849	6764			6764	6764
Los Husos I	Rioja			7028 - 7253	7141			7141	7141
Los Husos II	Rioja			6803 - 6952	6878			6878	6878
Mas Is	Alacant					7498 - 7578	7538	7538	7538
Molino de Arriba	Burgos			6943 - 7151	7047			7047	7047
Monte dos Remedios	Pontevedra	6511 - 6639	6575					6575	6575
Novelda	Alacant	7269 - 7414	7342					7342	7342
Paco Pons	Zaragoza	6799 - 6952	6876					6876	6876
Padre Areso	Navarra	6016 - 6283	6150					6150	6150
Parco	Lleida	6980 - 7164	7072					7072	7072
Paternanbidea	Navarra	6891 - 7137	7014					7014	7014
Pena d'Agua	Santarem	7585 - 7666	7626					7626	7626
Penya Larga	Alava					7521 - 7617	7569	7569	7569
Penya Oviedo	Cantabria	5920 - 5986	5953					5953	5953
Pla del Serrador	Barcelona	6561 - 6667	6614					6614	6614
Plansallosa I	Girona	7002 - 7164	7083					7083	7083
Portalon	Burgos	8485 - 8607	8546	6890 - 7151	7021			8546	7021
Prazo	Guarda	6471 - 6628	6550	6353 - 6446	6400			6550	6400
Puyascada	Huesca	6673 - 6842	6758					6758	6758

Site	Province	Other Charcoal		Short Lived Taxa		Domestic Taxa		Oldest	Best
		1 sigma range	mean	1 sigma range	mean	1 sigma range	mean	mean	mean
Retamar	Cadiz			7570 - 7687	7629			7629	7629
Riols I	Zaragoza	6746 - 7140	6943					6943	6943
Roca Chica	Malaga					7029 - 7269	7149	7149	7149
Sanavastre	Girona	6501 - 6650	6576					6576	6576
Sao Pedro de Canaferrim	Lisboa	6800 - 7138	6969					6969	6969
Senhora das Lapas	Santarem			6883 - 7155	7019			7019	7019
Serrat del Pont	Girona			7328 - 7428	7378			7378	7378
Valada do Mato	Evora	6797 - 6941	6869					6869	6869
Vale Boi	Faro			6894 - 7140	7017	6805 - 6943	6874	7017	6874
Vale Santo1	Faro			7028 - 7256	7142			7142	7142
Ventana	Madrid					7183 - 7410	7297	7297	7297
Cueva del Hoyo de la Mina	Malaga	6995 - 7171	7083					7083	7083
Padrao	Faro			7427 - 7559	7493			7493	7493
Cueva de Nerja	Malaga					7438 - 7552	7495	7495	7495
Cueva de los Murcielagos A.	Granada			6884 - 7141	7013			7013	7013
Tossal de les Basses	Alacant					6680 - 6855	6768	6768	6768
Cova de les Cendres	Alacant	7515 - 7665	7590			7335 - 7475	7405	7590	7405
Can Sadurni	Barcelona					7318 - 7417	7368	7368	7368
Sant Pau del Camp	Bacelona			7170 - 7261	7216			7216	7216
Plaza Vila de Madrid	Barcelona			7325 - 7420	7373			7373	7373
Can Xammar	Barcelona	7170 - 7249	7210					7210	7210

Table 2. Environmental parameters used to calculate Ecological Suitability Index.

Parameter	Values	Index Value
Slope	16° – 100°	1
	11° - 15°	2
	6° - 10°	3
	0° - 5°	4
	cell is ocean	NULL
Mean Maximum Spring Temperature (degrees C for March, April, and May)	< 18° or > 30°	0
	25° - 30°	1
	18° - 24°	2
Minimum March Temperature	< 0°	NULL
	0° - 4°	1
	≥ 5°	2
Total Spring Precipitation (mm for March, April, and May)	< 100mm or > 600mm	0
	100mm – 149mm	1
	301mm – 600mm	1
	150mm - 300mm	2

Figure 1

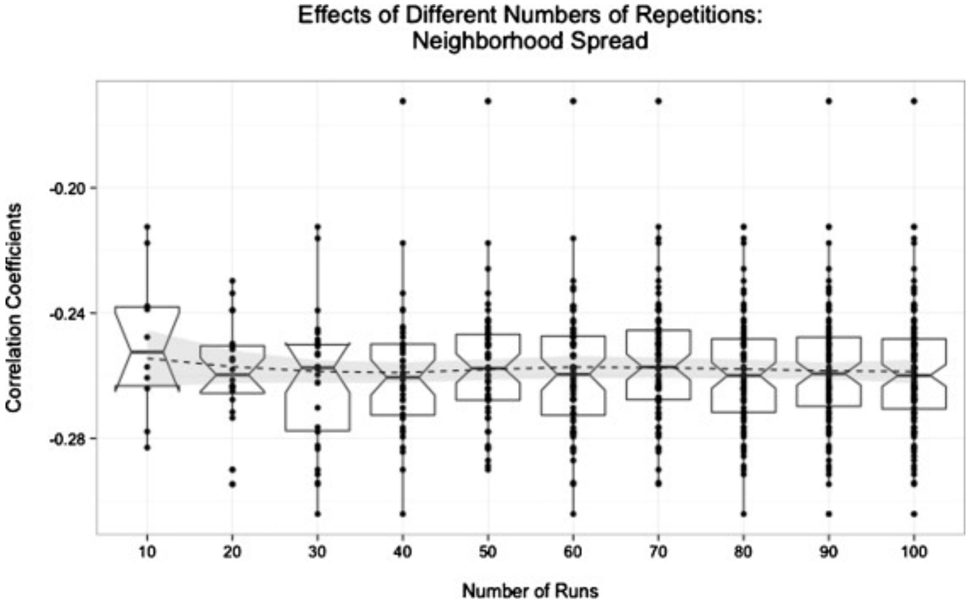
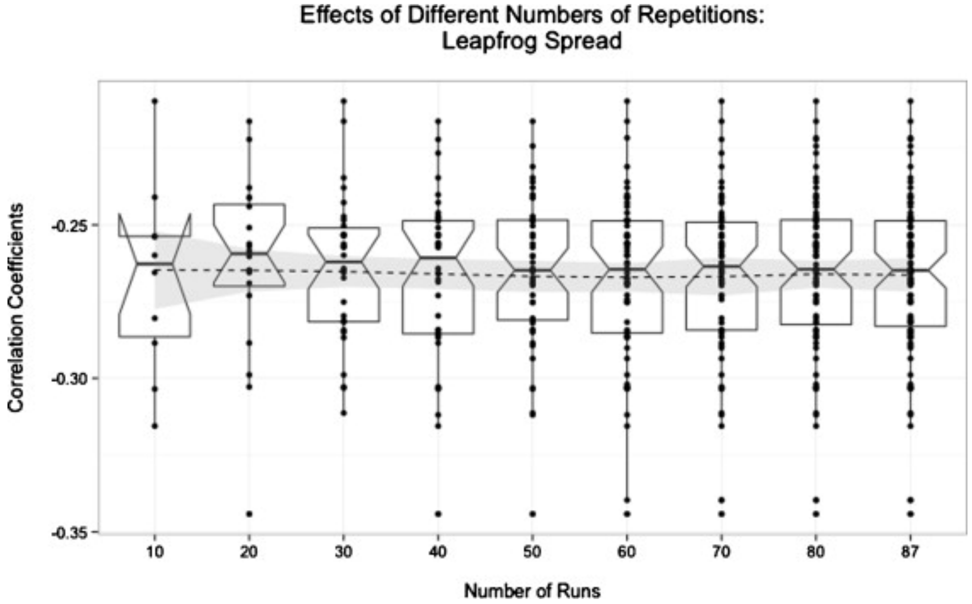

Effects of Different Numbers of Repetitions:
Neighborhood Spread
Correlation Coefficients
-0.20
-0.24
-0.28
10
20
30
40
50
60
70
80
90
100
Number of Runs

Figure 2



Effects of Different Numbers of Repetitions:
Leapfrog Spread
Correlation Coefficients
-0.25
-0.30
-0.35
10
20
30
40
50
60
70
80
87
Number of Runs

Figure 3

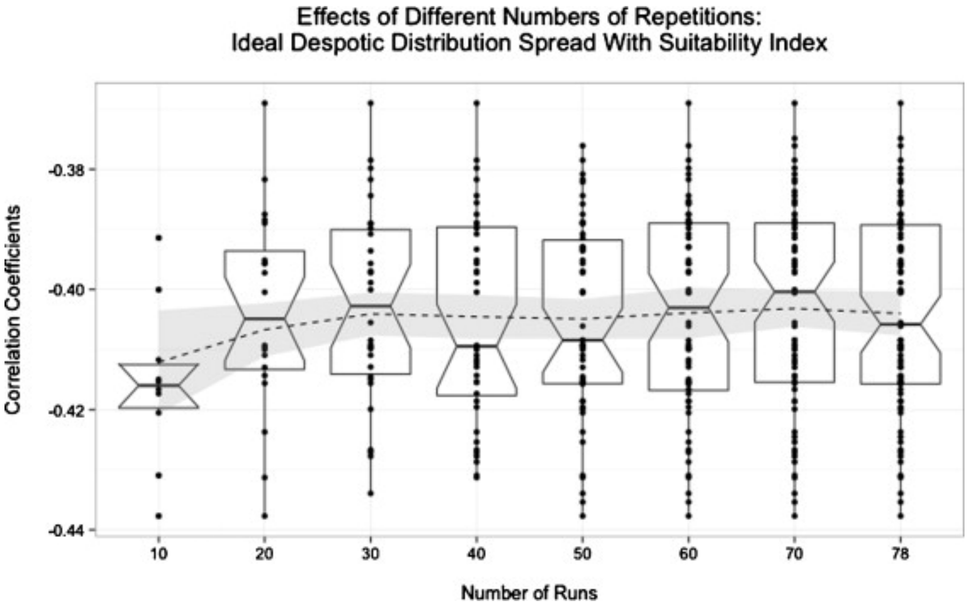
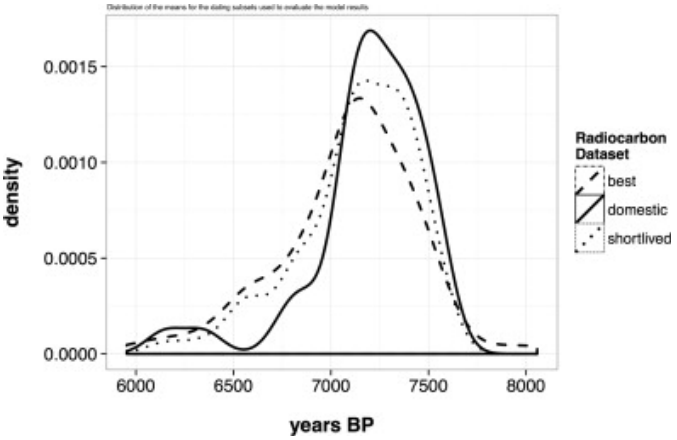

Effects of Different Numbers of Repetitions:
Ideal Despotic Distribution Spread With Suitability Index
Correlation Coefficients
-0.38
-0.40
-0.42
-0.44
10
20
30
40
50
60
70
78
Number of Runs

Figure 4



Distribution of the means for the dating subsets used to evaluate the model results
density
0.0015
0.0010
0.0005
0.0000
6000
6500
7000
7500
8000
years BP
Radiocarbon Dataset
best
domestic
shortlived