ABSTRACT. This article introduces a method of exploratory analysis of the geographical factors influencing large-scale innovation diffusion, and illustrates its application to the case of early pottery dispersal in the Old World. Regression techniques are used to identify broad-scale spatiotemporal trends in the innovation’s first occurrence, and regression residuals are then analyzed to identify geographical variation (climate, biomes) that may have influenced local rates of diffusion. The boundaries between the modeled diffusion zones segregate the western half of the map into a Eurasian hunter-gatherer pottery-using zone affiliated by cultural descent to the Siberian center of innovation, and a lower-latitude farming and pastoralist zone affiliated by cultural descent to the north African center of innovation. However, the correlation coefficients suggest that this baseline model has limited explanatory power, with geographical patterning in the residuals indicating that habitat also greatly affected rates of spread of the new technology. Earlier-than-predicted ages for early pottery tend to occur in locations with mean annual temperature in the range approximately 0–15°C. This favorable temperature range typically includes Mediterranean, grassland, and temperate forest biome types, but of these, the Mediterranean and the temperate deciduous forest biomes are the only ones on which regression residuals indicate earlier-than-predicted first observed pottery dates.

INTRODUCTION

In archaeology, a standard way of reconstructing dispersal routes and dispersal chronology for radiations such as those of the early farming cultures is to compile archaeological radiocarbon dates for their first observed occurrences throughout the geographical region of interest, and to look for spatial gradients in arrival times. Statistically, such trends can be evaluated using regression techniques, typically by bivariate analyses of the relationship between observed arrival time and distance from some origin point. This approach has been used to study the spread of farming into Europe across the prehistoric Neolithic transition (e.g. Gkiasta et al. 2003; Pinhasi et al. 2005). These techniques can be used to analyze innovation adoption diffusion as well as population expansion; in each case, all that is required is that the spread dynamic was influenced by geographical distance from the center or centers of origination, that the rate of spread was slow enough to be resolvable by 

14

C dating, and that a sufficient sample of well-dated occurrences is available for analysis.

In recent work of this kind, attention has focused on developing geographically explicit methods to evaluate the effects of terrain, drainage, and biome type on dispersal rates and “arrival times” across different landscapes (e.g. Davison et al. 2006; Ackland et al. 2007; Lemmen et al. 2011; Russell et al. 2014), as well as on evaluating the congruence of such archaeological models with estimates of dispersal paths and timescales derived independently from other kinds of data (e.g. genetics, in relevant cases; Ammerman and Cavalli-Sforza 1984; van Etten and Hijmans 2010).

Geographically explicit methods can be of two main kinds. Forward modeling involves setting up explicit, theory-driven geographical hypotheses about the spread dynamic, modeling them in one or two dimensions, and simulating a comprehensive set of local arrival times that can then be compared to the archaeological record. Inverse modeling involves analysis of patterns in the distribution of archaeologically observed arrival times, in a way that enables geographical hypotheses to be in-
ferred inductively from observed patterns. Each approach has its merits and its limitations. Forward models are powerful ways of exploring the implications of verbal hypotheses but limited by the plausibility of the theoretical assumptions they embody. Inverse approaches are powerful ways of making sense of pattern in noisy data, but may be limited in their predictive power by data gaps and uncertainties, theoretical uncertainties about the underlying process, and the nonuniqueness of any one model solution.

This article focuses on the inverse approach, and applies these methods to the origins and dispersals of pottery in the Old World using a data set of \(^{14}\)C dates from early pottery sites in Africa and Eurasia. There is now abundant evidence that in two areas, north Africa and east Asia, pottery was being made by hunter-gatherer groups earlier than 10,000 cal BP. In east Asia, it has even been suggested that hunter-gatherers may have invented pottery independently in southern China, Japan, and far eastern Russia (Kuzmin 2013). In a preliminary attempt at a pan-Eurasian synthesis of available dates and materials, Jordan and Zvelebil (2009), based on data collected by Hommel (2009), generated a descriptive time-space mapping of early pottery \(^{14}\)C dates from across Eurasia that appeared to lend preliminary support to a general east to west Eurasian dispersal event, in addition to possible contributions to the Near Eastern Neolithic from early pottery innovation in north Africa (Figure 1). This article will use regression analysis to explore such patterns, and to identify geographical factors that may have influenced local rates of spread.

**MATERIALS AND METHODS**

**Site and Date Selection**

The underlying database used consists of 960 \(^{14}\)C-dated archaeological sites associated with the earliest pottery technologies in Africa and northern Eurasia (the database will be published in a future paper). Information about the dates was gleaned from site reports, academic publications, \(^{14}\)C laboratory lists, and existing databases both in print and online. Contextual data from the reports was examined carefully to ensure that all dates had clear associations with early pottery.

For the present analysis, dates were selected, for sites with multiple dates, based on their age (retaining either the oldest, or the average of two or more oldest dates if they passed Ward and Wilson’s [1978] chi-square test). Calibration of the raw \(^{14}\)C data was performed using OxCal v 4.1 and the IntCal09 calibration curve (Bronk Ramsey 2009; Reimer et al. 2009). To improve the accuracy of the regression estimation by reducing the “noise” provided by sites that do not correspond to first arrivals of pottery in the neighborhood, the data set was then further filtered using an iterative two-dimensional binning technique (to select the oldest site in a given neighborhood). This methodology is fully elaborated in Russell et al. (2014, see its Supplementary Information file): the neighborhood radius was set at 100 km, which was considered to be a minimum spatial separation required for identifying by \(^{14}\)C dating the passage of a population or innovation front. This reduced the initial sample of 960 dated sites to 400 retained for further analysis (see Figure 2a).

**Origin Centers and Distances Between Centers and Other Sites**

As a proxy for the innovation centers, the oldest sites in the database were used from three of the four regions where hunter-gatherer pottery may have independently originated (China: Xianrendong Cave, 17,460 \(\pm\) 210 BP, mean calibrated age 20,817 cal BP – the oldest date from the eastern section in the lowest pottery-bearing stratum, Wu et al. 2012; Siberia: Gromatukha, 13,269 \(\pm\) 65 BP, 16,211 cal BP – the average of the two oldest dates on organic temper in sherds from the lower layer, O’Malley et al. 1999; and north Africa: Saggai, 10,060 \(\pm\) 150 BP, 11,663 cal BP – the oldest date from the pottery-bearing layer, on *Pila* shell, although also an outlier within that stratum, Caneva 1983). Although early pottery in Japan is of similar age to that in Siberia (~16,000 cal BP, Craig...
et al. 2013), it was not included it in the modeling as we were primarily interested in the technology’s westward diffusion from east Asia. The ages and locations of the Siberian and Chinese sites effectively “block” significant independent westward diffusion of Japanese pottery innovations, in our modeling setup. We explore a scenario of three Asian origins here without prejudice to the possibility that, in the Siberian and perhaps also in the Japanese cases, pottery may in fact have diffused from an even older center in the region. We shall return to that more parsimonious scenario in a future paper, but model the multiple centers scenario here. The distance between these putative origin centers and other archaeological sites was calculated as the lengths of terrestrial least-cost paths on
a projected base map, using the Lambert Conformal Conic map projection (centered at 42.5°N, 62°E and with reference parallels 30°N and 55°N). This is an appropriate projection for the domain of interest, which has a predominant east-west orientation. To define land/sea boundaries and impose a sea mask for the distance calculation, a present-day world coastlines map was used, but extending those coastlines to 40 km offshore using a GIS buffering algorithm, to allow for dispersal and interaction by inshore and close-offshore maritime transport. This allows contact between locations separated by up to 80 km of water, and enables diffusing traits to cross most marine channels including the English Channel, the Strait of Gibraltar, and the Strait of Hormuz.

Regression Methods

The statistical methodology used to estimate trends in earliest pottery dates as a function of distance from origin locations in east Asia and in north Africa, involved fitting regression models (reduced major axis [RMA]; Ammerman and Cavalli-Sforza 1971; Steele 2010) to sets of paired values of site dates (mean calibrated 14C ages) and site distances from a possible origin. This enabled to estimate the mean speed of spread of the new technology (using the regression slope coefficient), and the proportion of the variation in arrival times that was accounted for by that trend (using the correlation coefficient). Because it was not known a priori which center of innovation was ancestral to the pottery found at locations very distant from either source, speeds of spread were estimated from each possible origin center by fitting regression models to dated sites within a core region subject to some cut-off distance from the source location. (The cut-off distance was determined empirically for each center by maximizing the correlation coefficient for dates and distance, subject also to the requirement that the correlation was statistically significant at α = 0.001.) Finally, having estimated the average speeds of pottery’s diffusion from the innovation centers in north Africa and east Asia, this study also modeled (using techniques developed for this purpose [Silva and Steele 2011] and
setting the onset times as the ages of the sites at the origin locations to predict the boundary effectively) the boundary between the competing spread zones as they converged.

**Analysis of Residuals**

The methods for estimating spread rates and for partitioning the map into competing spread zones enables deriving a map of predicted first arrival times, based on an analysis that allows only for the geographical effect of terrestrial and coastal-maritime mobility. By saving the RMA residuals, one can then analyze their correlation with other geographical variables. These can be ratio scale (e.g. temperature, rainfall), nominal scale (e.g. biome type), or anything in between. However, the type of data will condition the type of statistical test that can be used. The present study used ratio-scale data on mean annual temperature and on mean annual rainfall (both present day and Last Glacial Maximum from WorldClim, Hijmans et al. 2005; LGM data based on PMIP2, Braconnot et al. 2007), and have analyzed the association with residuals in empirical pottery arrival times using standard linear and nonlinear regression. In this case, ordinary least-squares (OLS) regression techniques were used because most of the error is expected to be contained in the archaeological residual site age variable. Sites with Late Glacial ages were assigned climate values based on the PMIP2 LGM model, while sites with Holocene ages were assigned climate values based on the present day. This study also used the distribution of major modern biomes (Olson et al. 2001) and treated them as nominal-scale data; median values of the archaeological residuals for each biome type are compared using boxplots and notch tests (the notch test estimates whether the median values of two samples are significantly different at $\alpha = 0.05$; Chambers 1983).
RESULTS

Diffusion Rates and Zones of Influence in the Almost-Geography-Free Model

The best-fitting linear trends for each of the three origin locations on the continental mainland are shown in Table 1. The results for China suggest both a slower mean diffusion rate (0.32 km/yr) and a smaller core region (cut-off 4790 km) than those estimated for the other two origin locations. The boundaries between the modeled diffusion zones segregate the western half of the map into a Eurasian hunter-gatherer pottery-using zone affiliated by cultural descent to the Siberian center of innovation, and a lower-latitude farming and pastoralist zone affiliated by cultural descent to the north African center of innovation (see Figure 2b). However, in all cases the coefficients of determination ($r^2$) are low, particularly so for the north African origin ($r^2 = 0.091$), and indicate that about 80% of the variation in archaeological dates in the binned and filtered data set is not explained by a model of a constant linear spread of the technology from these three competing centers.

Table 1 Parameters for the regression-based trend analysis with three competing continental mainland centers of innovation.

<table>
<thead>
<tr>
<th>Region</th>
<th>$r^2$</th>
<th>Speed (km/yr)</th>
<th>Cut-off length (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E Asia 1 (China)</td>
<td>0.186</td>
<td>0.32</td>
<td>4790</td>
</tr>
<tr>
<td>E Asia 2 (Siberia)</td>
<td>0.192</td>
<td>0.67</td>
<td>6426</td>
</tr>
<tr>
<td>N Africa</td>
<td>0.091</td>
<td>0.80</td>
<td>6336</td>
</tr>
</tbody>
</table>

The unexplained variance may reflect many things: variable diffusion rates across different habitat types; the effects of physical geographical corridors and barriers; archaeological sampling biases; and/or mistaken modeling assumptions (for example, with regard to the number, age, and location of centers of technological innovation). A major additional potential source of error is mistaken age estimation due to sample contamination or to stratigraphic misattribution: this applies to the time of appearance of pottery in our source regions as well as to the time of its appearance at other sites in the database to which, it is assumed, pottery then diffused. The first of these additional sources of variation in observed arrival times will be examined.

Residuals: Effects of Climate and Biome Type

Local climate has a clear association with residual advancement or delay in observed dates for early pottery use as shown in Figure 3 ($r^2 = 0.25$, indicating that climate explains a quarter of the variance unexplained by the simple linear spread models; the climate data used for the Late Glacial sites are LGM values and may not be good reflections of conditions at the time these sites were occupied, since using only present-day temperature values for all sites yielded a higher $r^2$ value of 0.32). Earlier-than-expected ages for early pottery tend to occur in locations with mean annual temperature of approximately 0–15°C. Sites that are much colder or warmer than this tend also to have relatively late first observed pottery dates. This favorable temperature range typically includes Mediterranean, grassland, and temperate deciduous and coniferous forest biome types, and boreal forest at the colder extreme of this range (Woodward et al. 2004). Variation in annual mean precipitation has no separate or additional effect on the residuals (results not shown), suggesting that this correlation with temperature extends across the precipitation regimes of each of the aforementioned biome types.

A similar pattern emerges from an analysis of residuals grouped by major modern biome type (Figure 4). The Mediterranean and the temperate deciduous forest biomes are the only ones on which median values for residuals are significantly positive, indicating earlier-than-predicted first observed pottery dates. However, and subject to the limitation of our only using modern biome dis-
Modeling Spatial Innovation Diffusion from $^{14}$C Dates

Figure 3  Effects of climate (mean annual temperature, °C) on residually unexplained variance in first observed dates for pottery in the entire studied domain. The choice of the oldest reliable date for pottery at each center of origin as the onset time for the modeled spread explains the preponderance of negative values. Third-order polynomial fitted curve.

Figure 4  Residuals by major modern biome type. The triangles show the upper and lower 95% confidence interval for the medians. The choice of the oldest reliable date for pottery at each center of origin as the onset time for the modeled spread explains the preponderance of negative values.
tributions, there is no such effect for the temperate grassland, temperate coniferous or boreal forest sites, enabling us to further constrain the set of “favorable habitats” beyond that obtainable just from the annual temperature/residual archaeological age association. Given the high density of European early Neolithic dates in our sample (Figure 2a; Table 2) found within these biomes, it is likely that this reflects a distinctive diffusion dynamic for the spread of early agriculture.

Table 2  Median and notch range values for the residuals according to biome type, with sample sizes.

<table>
<thead>
<tr>
<th>Biome</th>
<th>Lo Notch</th>
<th>Median</th>
<th>Hi Notch</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mediterranean Forests, Woodlands &amp; Scrub</td>
<td>97.7</td>
<td>193.9</td>
<td>290.0</td>
<td>32</td>
</tr>
<tr>
<td>Temperate Broadleaf &amp; Mixed Forests</td>
<td>56.7</td>
<td>91.4</td>
<td>126.0</td>
<td>128</td>
</tr>
<tr>
<td>Flooded Grasslands &amp; Savannas</td>
<td>−236.2</td>
<td>82.4</td>
<td>400.9</td>
<td>5</td>
</tr>
<tr>
<td>Temperate Conifer Forests</td>
<td>−71.4</td>
<td>66.1</td>
<td>203.7</td>
<td>7</td>
</tr>
<tr>
<td>Boreal Forests and Taiga</td>
<td>−46.2</td>
<td>56.6</td>
<td>159.3</td>
<td>47</td>
</tr>
<tr>
<td>Deserts &amp; Xeric Shrublands</td>
<td>−138.1</td>
<td>−75.7</td>
<td>−13.4</td>
<td>34</td>
</tr>
<tr>
<td>Temperate Grasslands, Savannas &amp; Shrublands</td>
<td>−238.7</td>
<td>−85.3</td>
<td>68.0</td>
<td>34</td>
</tr>
<tr>
<td>Tropical &amp; Subtropical Grasslands, Savannas &amp; Shrublands</td>
<td>−353.2</td>
<td>−186.7</td>
<td>−20.1</td>
<td>17</td>
</tr>
<tr>
<td>Tropical &amp; Subtropical Dry Broadleaf Forests</td>
<td>−375.1</td>
<td>−194.5</td>
<td>−13.8</td>
<td>5</td>
</tr>
<tr>
<td>Tundra</td>
<td>−675.5</td>
<td>−267.6</td>
<td>−140.4</td>
<td>6</td>
</tr>
<tr>
<td>Montane Grasslands &amp; Shrublands</td>
<td>−1104.2</td>
<td>−553.1</td>
<td>−205.9</td>
<td>3</td>
</tr>
<tr>
<td>Mangroves</td>
<td>−897.4</td>
<td>−745.8</td>
<td>−594.2</td>
<td>2</td>
</tr>
<tr>
<td>Tropical &amp; Subtropical Moist Broadleaf Forests</td>
<td>−1109.2</td>
<td>−785.5</td>
<td>−461.9</td>
<td>16</td>
</tr>
</tbody>
</table>

**DISCUSSION**

This article has presented a regression-based analysis of trends in 14C dates for early pottery in north Africa and Eurasia, and explored a hypothesis of competing diffusion from three continental mainland centers of innovation. In any study on this scale, one must remember that patterns observed may be heavily influenced by archaeological sampling biases, while patterns may also be missed due to mistaken modeling assumptions (for example, with regard to the number, age, and location of centers of technological innovation). A major potential source of error is mistaken age estimation due to sample contamination or to stratigraphic misattribution: this applies to the time of appearance of pottery in our source regions as well as to the time of its appearance at other sites in the database to which, it is assumed, pottery technology then diffused. For example, Saggai was chosen as the source location for the “pottery Mesolithic” of Sudan-Sahara, but the date chosen (10,060 ± 150 BP) is an outlier within that site and stratum that some have queried (e.g. Mohammed-Ali and Khabir 2003; Sadig 2013), and it may be that north African hunter-gatherer pottery appeared later than this date suggests (for pottery Mesolithic sites elsewhere in north Africa with dates from the earlier 10th 14C millennium BP, see e.g. Sadig 2013:29). This still predates the earliest pottery in the Near East (e.g. early 9th 14C millennium BP at Ganj Dareh, Zagros Mountains, Iran, Smith 1978; Zeder and Hesse 2000; very late 9th/early 8th 14C millennium BP in Syria – Nieuwenhuyse et al. 2010; Early/Middle PPNB in Israel, perhaps therefore as old as late 9th 14C millennium BP, Biton et al. 2014; see also Le Mièvre and Picon 1999); however, we must emphasize the dependence of our results on the modeling assumptions about number, age, and location of independent centers of pottery innovation.

Subject to these constraints, we find that our best-fitting baseline model predicts a boundary between competing diffusion zones in western Eurasia roughly corresponding to a stable frontier between early agricultural and late hunting-and-gathering societies recognized by archaeologists on other
Modeling Spatial Innovation Diffusion from \(^{14}C\) Dates

grounds. Methodologically, this approach has been very different to that taken by Davison et al. (2009), who address a similar modeling problem. The studied approach assumed that the map of early pottery use can be segregated according to a first-arrival rule (i.e. we do not follow Davison et al. 2009 in assuming that successive waves of diffusion of technological innovation can interpenetrate each other’s core domains of influence without interference or competitive exclusion). However, the correlation coefficients suggest that our own baseline model still has limited explanatory power, while geographical patterning in the residuals indicates that habitat also greatly affected rates of spread of the new technology. Future work will therefore explore models that incorporate more geographical variation, and that allow for the possibility that each center of innovation may have had a distinctive diffusion dynamic reflecting the varying ecocultural niches in which pottery vessels were first used.

ACKNOWLEDGMENTS

We thank Yaroslav Kuzmin for the invitation to participate in this session, and for his engaging commentary on our approach at the meeting itself. Yaroslav also kindly contributed to a further revision of the Asian part of the database, as did Henny Piezonka, and we gratefully acknowledge their assistance. We also gratefully acknowledge Peter Hommel’s contribution in compiling and updating the original database. This article is an output of the project “Ceramics before Farming: prehistoric pottery dispersals in Northeast Asia,” which is funded by the UK Leverhulme Trust (2011-14; grant reference F/00 152/AM; PI Peter Jordan; research fellow Kevin Gibbs). These authors would like to thank the Leverhulme Trust for this support.

REFERENCES


