Mapping Optimal Prehistoric Clay Sources: Adapting Watson's Method to GIS Technology

Elissa Hulit

Abstract: One of the basic problems in the study of prehistoric North American ceramics is clay sourcing. In a 1992 paper, Robert Watson proposed a method of predicting optimal clay sources using a combination of United States Department of Agriculture soils maps, knowledge of landscape formations, and ethnographic data to predict optimal locations for raw clay acquisition in Jefferson County, WI. These optimal locations were then compared to the site data from the Southeastern Wisconsin Archaeological Project. In this paper I discuss the results of my attempt to adapt his methodology by creating a digital model which could predict the optimal clay sources of Walworth County, WI, located just to the southeast of Jefferson County. The results of this project point out several weaknesses in the proposed model, but also highlight the benefits of using GIS for analyzing the patterns in the site data.
Key words: GIS, clay sourcing, Crawfish River, optimal resource model

Introduction

Geographic Information Science (GIS) provides a powerful tool in the study of the past. Through the use of GIS, archaeologists can explore the ways in which geographical and cultural features interact, as well as how cultural interactions can vary geographically. One of the areas in which GIS may be especially useful to archaeologists is demonstrated in the study of prehistoric pottery. Clay sourcing in prehistoric archaeology is problematic because similar ceramic styles and decorations are commonly cited as evidence of culture contact and exchange. When studying cultural interaction through ceramics, it is not enough to demonstrate that pottery collections found in different sites appear to be similar. Similar pottery may indicate movement of resources or people across the landscape. Without knowing where the ceramics were made or where the raw materials came from, there is always the possibility that stylistic or technical similarities of pottery found in different areas are coincidental.

In order to demonstrate that a pot was transported from the area in which it was made, the researcher must demonstrate that the pot was not made in the area where it was recovered. Traditionally, non-destructive methods of analysis have been limited to the study of stylistic elements and vessel body to compare clay 'recipes'. However, the development of non-destructive methods of elemental analysis, such as X-ray fluorescence, has made it possible to attempt to differentiate vessels on the basis of the elemental composition of the clay from which the vessel was made (Potts 2008). Clays are the result of physical and chemical weathering processes, and it is natural that there should be regional variations in composition based on parent materials (Perkins 2002). Once clay source compositions can be compared, the next step in this process is to try to locate potential clay sources and document elemental clay 'signatures' across a region. Use of GIS workspace to narrow the search for potential sources of raw clay may allow researchers to establish a map of regional compositions.

Field Notes: A Journal of Collegiate Anthropology 3(1): 63-72. (May 2012)

United States Department of Agriculture soils maps have the potential to greatly add to our understanding of clay source locations, however, the scale and methodology involved in creating these maps results in several problems. Soils maps are compiled with agricultural and engineering purposes in mind. The units defined are mixtures of soil series in varying concentrations and do not form abrupt boundaries. There is a natural gradation between units and the boundaries are not as clear as they appear on the soils maps(Holliday 2004). Furthermore, the maps generalize the landscape to a degree which makes them unsuitable for archaeological survey at the site level (Holliday 2004). As a result of these issues, soils maps can only be used for predicting resource location when combined with other types of data.

GIS technology has no immediate solution for these problems. GIS soils shapefiles are digitized from the paper maps, so while allowing more detail, issues of scale and boundaries remain in the digital data. While the Soil Survey Geographic Database (SSURGO) maintains a free database of soils, the data is stored by county. The series boundaries match at the county lines, but frequently the descriptions do not. One way around this problem is to go back to the soil survey manuals and read the descriptions of the main soil associations. In this way, use of the GIS shapefiles may lead to slightly better accuracy but it does not reduce the amount of work needed to isolate potential clay bearing soils in the GIS model.

In the early 1990s Robert Watson introduced a promising predictive model using soils data. Focusing on Jefferson County, Wisconsin, he based his model on soils maps, slope estimates, drainage, and accessibility as criteria for identifying the most likely areas from which useable clay might be retrieved. Watson isolated soils formed primarily from lacustrine deposits formed during the retreat of the last major glaciation around 13,000 years ago (Watson 1992) . These are secondary clays formed as clay minerals transported by wind and water are deposited in thick layers resulting in abundant and uniform sources of raw clay for pottery manufacture (Rice 1987). Also, since well drained soils along a creek or water body are more subject to slumping and erosion, they would provide good access points for a prehistoric potter. Therefore, Watson predicted creek banks that cut through lacustrine soils would be the optimal prehistoric clay sources (Watson 1992).

After creating the model, he then tested it against site data compiled in the Southeastern Wisconsin Archaeological Program (SEWAP). This was a project which produced an unbiased 15% stratified random sample of 170 forty acre units along the Crawfish and Rock Rivers in Southeastern Wisconsin (Goldstein 1987; Watson 1992). Of those units, 17 fell within his survey area in Jefferson County (see Figure 1), and 76% of the sites containing pottery within those survey units fell within one kilometer of an optimal clay deposit. This is important because the Exploitable Threshold Model, established by the ethnographic data compiled by Dean Arnold (1985), indicates that the majority of potters utilize clay sources within one kilometer of their habitation sites.

Watson created his model without the help of GIS and computer programs. As Figure 1 demonstrates, the hand drawn maps from Watson's article are of limited use on their own. The concept is good and the use of GIS mapping tools could make this model both more accurate and more time-efficient. Optimal areas identified in this way would then narrow the search area for the sampling of source clays. GIS also provides the researcher with the ability to statistically analyze the spatial distribution of site data. However, trends noted during the course of this project demonstrate that such straight conversion of the original predictive model to GIS was overly simplistic.

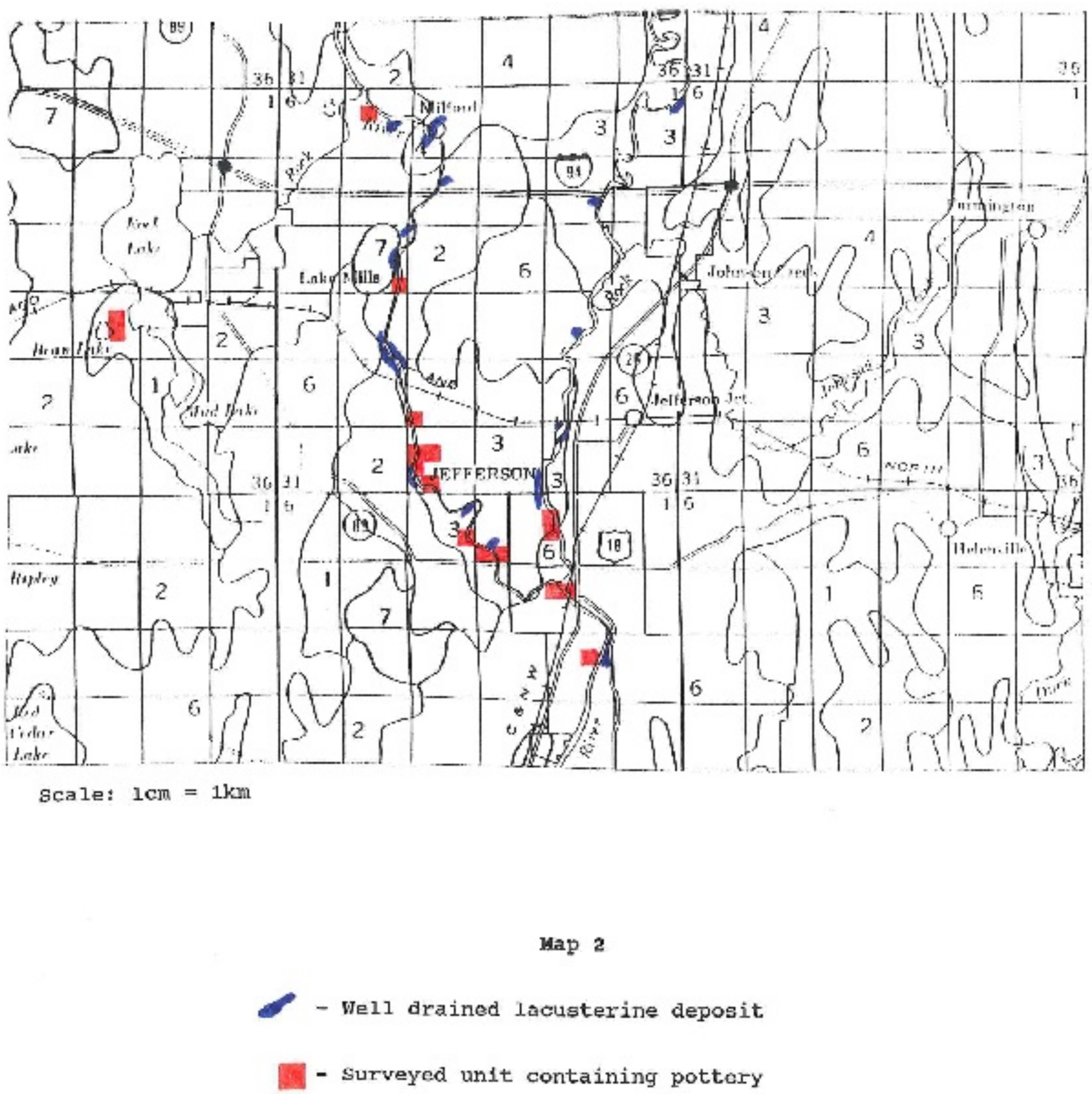

Figure 1. Watson’s survey area (Watson 1992)

Purpose

The original purpose of this project was to replicate Watson’s model in Walworth County and compare the resulting optimal sources to the location of known historic brickyards. If a geographical similarity could be demonstrated between the historically exploited clays and the optimal model, the historic clay sources would be likely to be compositionally similar to prehistoric clay sources. Clay from these early historic brickyards could then be used to fill in gaps in regional composition studies, even for areas where the original sources no longer exist or are inaccessible. Despite these initial goals, the main accomplishment of this project lies in demonstrating the difficulties of adapting predictive models that look good on paper to the digital GIS technology. The purpose was then modified to determine whether or not patterns noted in Watson's original paper stood up to GIS analysis. Artificial datasets were created and basic geostatistical analysis methods were used to test these relationships.

Methods

The original study area for this project was Walworth County, WI, located southeast of Jefferson County, WI (Figure 2). Walworth County was selected for the present study to adapt Watson's methods without simply duplicating his results. Walworth County is very similar to Jefferson County in size as well as geological and hydrological patterns. Selection of Walworth County also allowed for comparison of the locations of the optimal clays to historic clay resources. However various issues required reevaluation of Jefferson County and the basic principles of the predictive model.

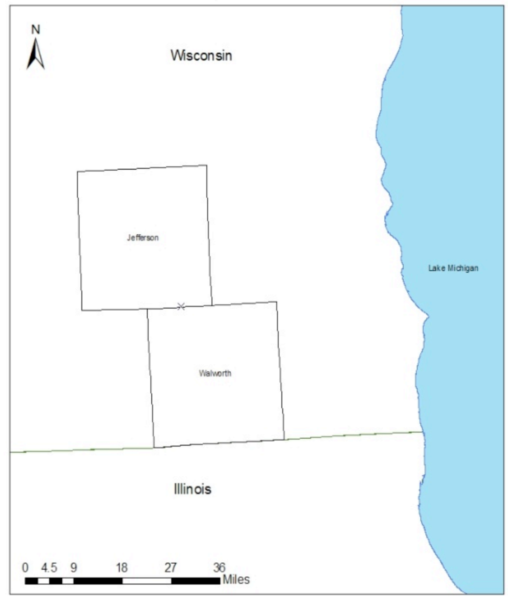

Figure 2. Jefferson and Walworth counties

The first step in creating the GIS base map was acquisition of basic layers such as shapefiles of county boundaries, water bodies, and water lines from the Wisconsin Department of Natural Resources (WIDNR) (Wisconsin DNR 1998; Wisconsin DNR 2010). Also, four digital elevation model (1degree) datasets were downloaded from the US Geological Survey (U.S. Geological Survey). Shapefiles for soils series of Walworth County and Jefferson County were downloaded from the Soil Survey Geography database files (Soils Data Mart 2009).

Once these base maps were assembled, the next step was to create a set of specialized shapefiles for use in the analysis. First, a 30m buffer was placed around water bodies and streams in Walworth County to encompass most slopes leading down to the water bodies, as well as small erosional gullies. This created a polygon suitable for use in subsequent analyses. Next, soil polygons characteristic of well drained lake deposits were extracted from the soils shapefile, including soils from the Plano-Warsaw and Casco-Fox soil associations (Figure 3) (Glocker 1971; USDA 1971). These soil types included several soil series described and utilized by Watson, such as the Hebron and Saylesville soil series.

The intersection of these two shapefiles, the 30 km water buffer and the lacustrine deposits, created a new shapefile of optimal clay sources, analogous, though not identical, to those described by Watson. A one kilometer buffer was established around these hypothesized optimal resources illustrating Arnold's exploitable threshold model (Arnold 1985) (Figure 4). Approximately 77% of the area within the Walworth County falls within one kilometer of an optimal source. This indicated that there was a potential problem with the model. Watson had indicated that 76% of the surveyed units in his study area in Jefferson County fell within one kilometer of an optimal source indicating a significant relationship between ceramic bearing sites and potential clay sources. However, it was never demonstrated that the percentage of ceramic bearing sites which fell within this one kilometer buffer is significantly different from the percentage of non-ceramic bearing sites within the buffer nor from the total number of sites of either type which fall within the one kilometer buffer.

In order to reevaluate Watson's conclusion, the methods employed above were repeated for Jefferson County (Figures 5 and 6). Watson's results could have been caused by an independent random process (CSR/IRP), or a spatially random pattern of ceramic bearing sites. Geographical phenomena are rarely randomly distributed. Spatial autocorrelation is the tendency for geographical features located closer to one another to be more similar than features located farther away from one another (O'Sullivan and Unwin 2010). Spatial autocorrelation is characteristic of archaeological sites as well. Real site locations are not randomly distributed across a landscape, for example, people tend to build habitation sites in the same locations time after time and avoid the same obvious deterrents such as standing water and extreme slope. In this analysis ceramic bearing site locations are evaluated for randomness in relation to optimal clay source locations. An artificial set of randomly distributed points was created to represent a distribution of sites that are spatially random.

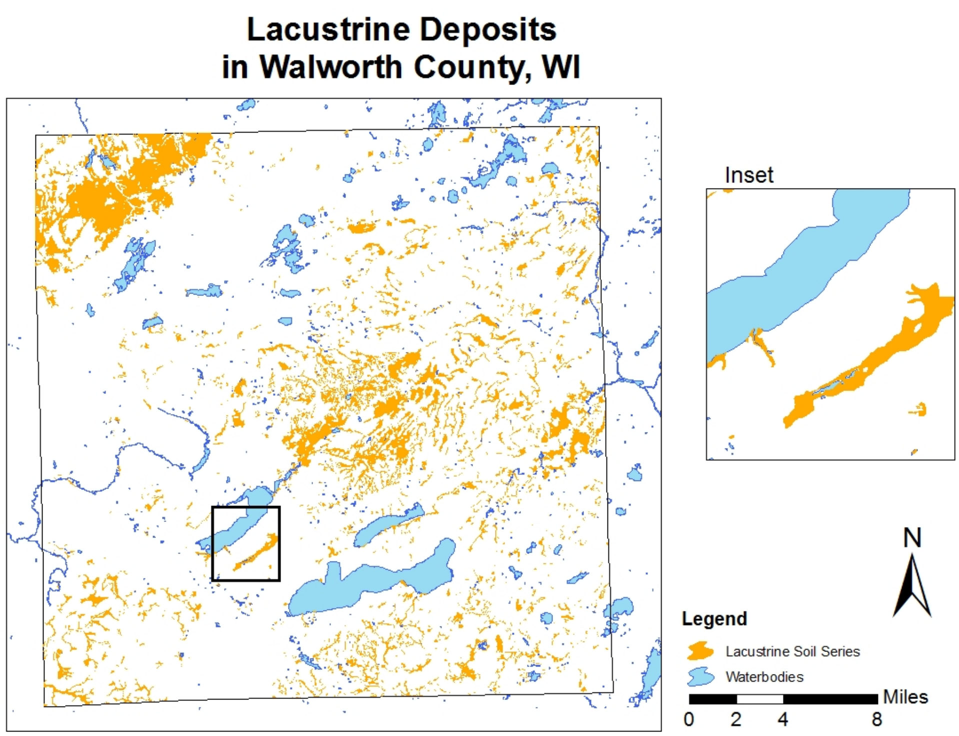

Figure 3. Lacustrine deposits in Walworth County, Wisconsin

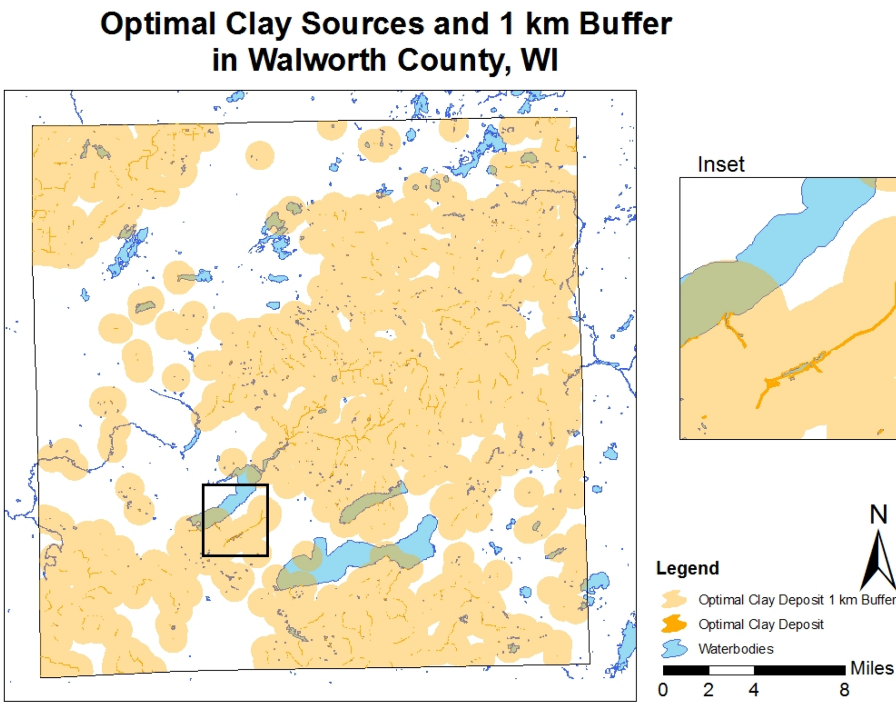

Figure 4. Optimal clay deposits and the one kilometer buffer in Walworth County, Wisconsin

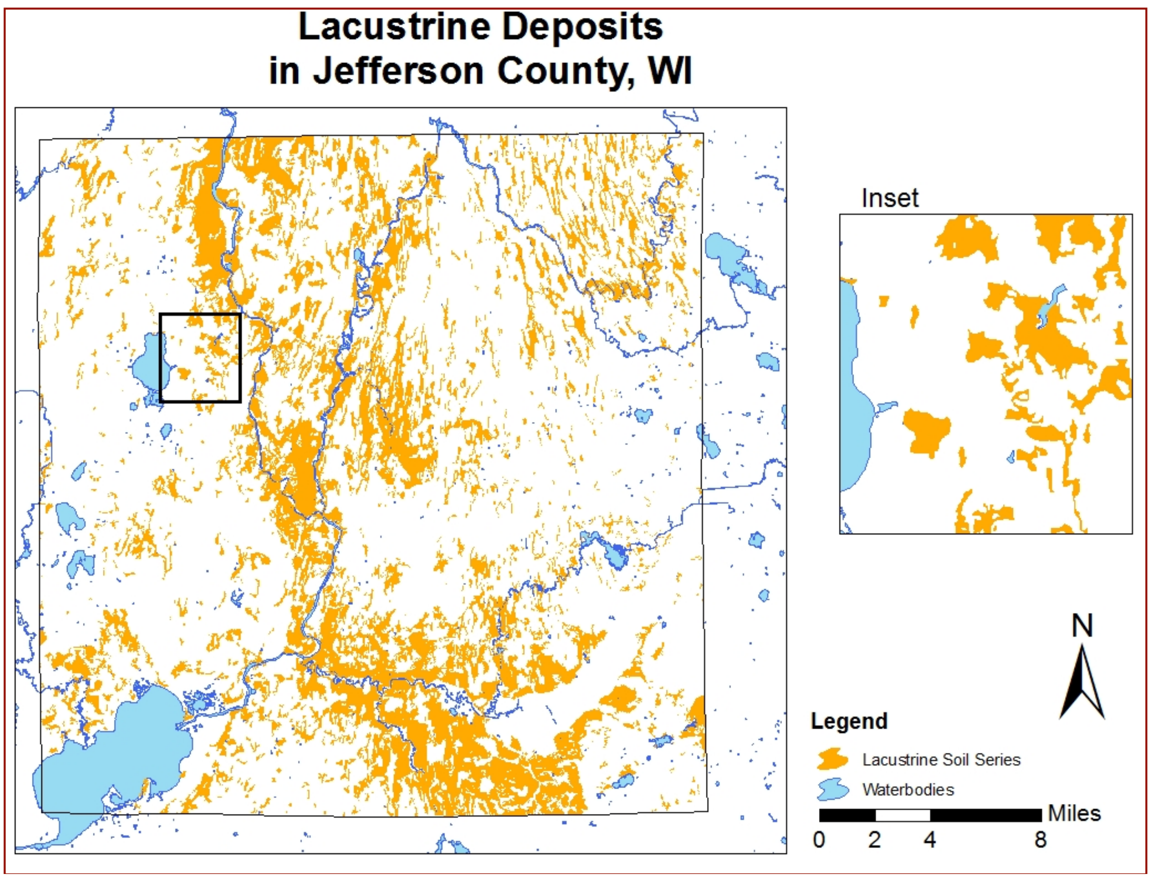

Figure 5. Lacustrine deposits in Jefferson County, WI

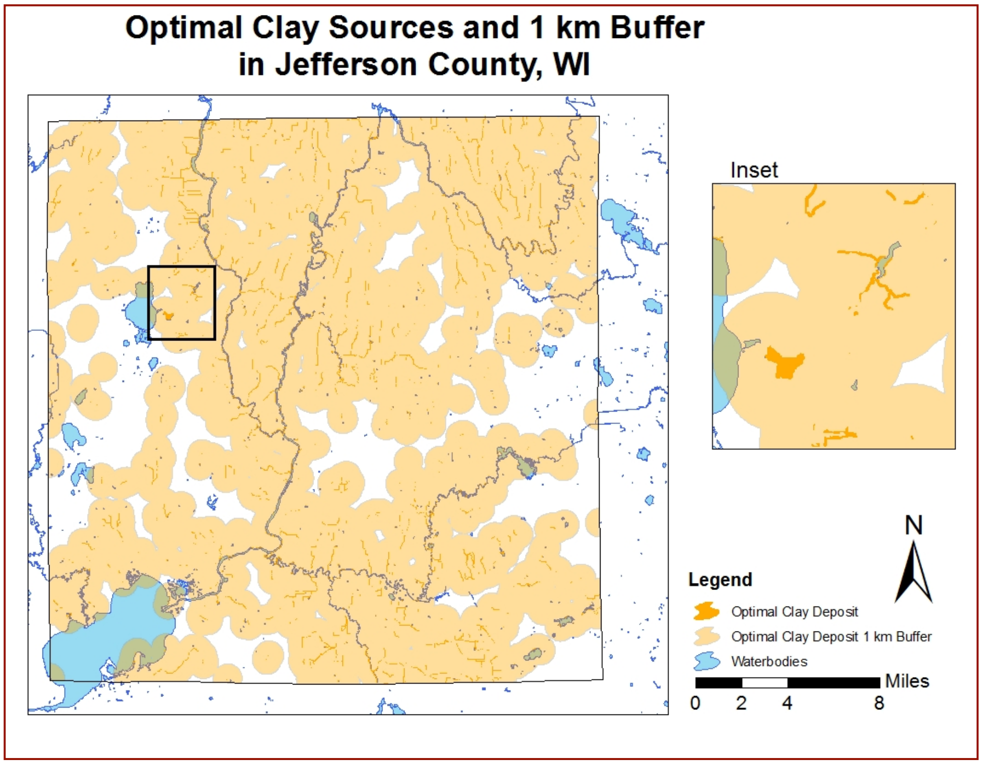

Figure 6. Optimal clay sources and one kilometer buffer in Jefferson County, Wisconsin

This process is slightly more complicated than the previous steps as Watson omitted several key details. For example, he indicated how many survey units containing pottery fell within Jefferson County, however, he did not mention how many survey units without pottery fell within the study area, nor does he indicate how many discrete site locations were included in each survey unit. Watson wrote only in regards to the survey units. Of the 17 units that fell within his study area, 13 units, or 76% of the sites, fell within one kilometer of an optimal clay source.

To test whether this same pattern could be observed under a CSR/IRP model, a random pattern of pseudo-site data was created using the spatstat package in the R statistics program to create a sample of 150 points. The survey units for the SEWAP project were 40 acre units. Assuming a site density of one site per 40 acres in Jefferson county, would require a shapefile containing 9,318 'sites', however, using this many points would overload the process. Instead, a sample of 150 points would yield roughly one site per $10km^2$ for Jefferson County, which is approximately $1508km^2$ in size. This process was done in the R statistics package rather than in ArcGIS because the spatstat package is capable of generating a spatially random point pattern of geographic coordinates, and the R program can also create a third variable, z, for each point (Baddeley and Turner 2005). This z variable is a random binary variable of zeroes and ones, coded such that zero stands for 'sites without ceramics', while site points coded with a z value of 1 can be considered 'sites with ceramics'.

The resulting shapefile (Figure 7) contains 150 points, 77 of which were coded as ceramic bearing, and 73 were coded as non-ceramic bearing (Table 1). ArcGIS software provides several functions that allow the researcher to evaluate the degree of autocorrelation and clustering of site data. The random model of site data could then be tested for spatial auto correlation both at the global and local levels. At both levels, the analysis should reflect the randomness of the data. There should be no spatial clustering of the z variable (ceramic vs. non-ceramic bearing sites).

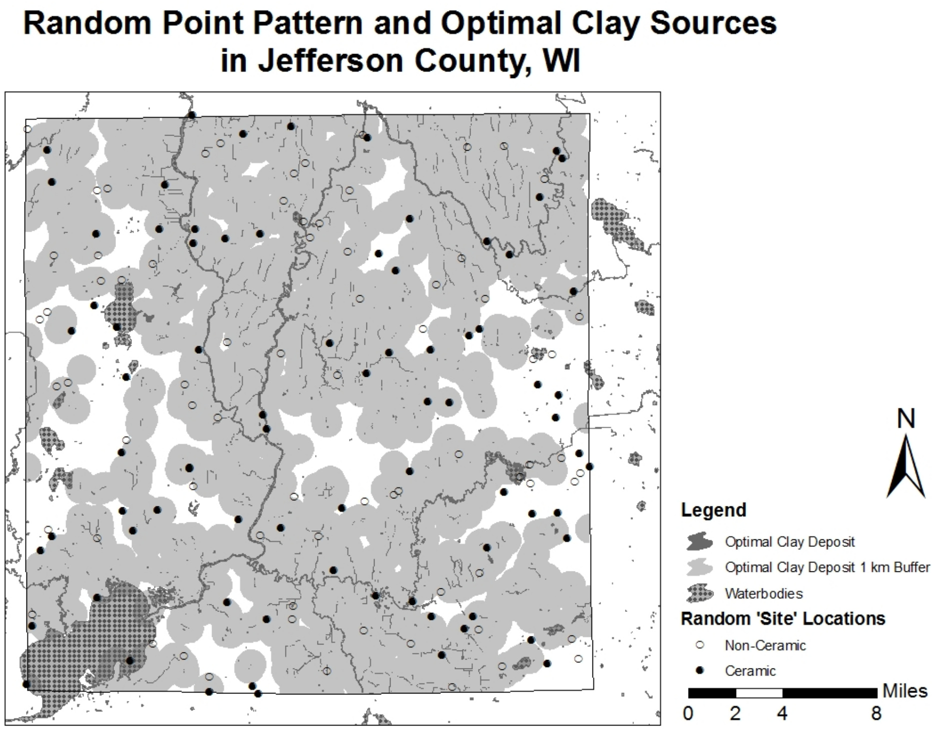

Figure 7. Random Point Pattern and Optimal Clay Sources, Jefferson County, Wisconsin

Table 1. Summary of Spatially Random Point Pattern Distribution

Z Variable	Within 1 km of an Optimal Source? No	Within 1 km of an Optimal Source? Yes	Total
Non-Ceramic Bearing (0)	16	57	73
Ceramic Bearing (1)	25	52	77
Total	41	109	**150**

Results

Global statistics of the z variable for the random dataset reveals no autocorrelation, as expected given that the dataset was randomly generated (Moran's I=0.07, Z score=0.83, Getis-Ord G=0, Z score= 0.3). After performing a spatial join between the random data points and the one kilometer buffer, the points were assigned a value based on their location (0 = outside the one kilometer buffer, 1 = inside the one kilometer buffer). The global statistics for this location variable revealed a high level of autocorrelation (Global Moran's I=0.48, Z score=5.58, very clustered). The testing of this location variable is an extension of the idea that optimal clay sources are likely to be closer to one another, since they are derived by the same process. This test demonstrates that when points fall within areas that are spatially autocorrelated, and when these points are then given values corresponding to these areas, the points will appear to mimic the autocorrelation of the area leading to spurious conclusions.

In the random dataset, 72.67% of the sites fell within one kilometer of an optimal source. Chi-square tests of homogeneity indicate that there is no evidence that the proportion of ceramic sites within one kilometer of an optimal source differs from the proportion of non-ceramic sites within one kilometer of an optimal source (χ^2=0, df=1, p-value=1). This was to be expected given that the dataset was randomly generated. However, the interesting point is that the proportion of ceramic bearing sites within one kilometer of an optimal source from the random dataset can then be compared to the proportion of real ceramic bearing sites within one kilometer of an optimal source in Watson's model. There is no evidence that these proportions differ significantly (χ^2=0.0022, df=1, p-value=0.9628). Also, approximately 79.26% of the study area lies within one kilometer of an optimal source. This indicates that Watson could have come to the same conclusions from a model that was spatially random regarding the location of so-called 'optimal sources'. This is not to say that ceramic bearing site locations are completely random, only that the locations of sites containing pottery may be independent of the hypothesized optimal clay source locations.

Conclusion

It appears that Watson's Predictive Model has the potential to help archaeologists better understand how prehistoric people utilized raw clay resources. However, in adapting the model to use with a GIS, several weaknesses have been identified. Most notable is the inability of the model to predict locations of ceramic bearing sites from the optimal source locations. The results of the statistical analysis show that Watson did not demonstrate that ceramic-bearing sites are correlated with the predicted sources. In fact, what Watson's model showed was that potential clay sources are highly correlated with one another and that any distribution of sites falling within the exploited threshold model reflects that autocorrelation. Due to the nature of geographic data and the way in which optimal sources are predicted, optimal clay sources are necessarily autocorrelated with one another. In order to make Watson's model more useful, it is necessary to demonstrate that ceramic bearing sites are more closely associated with potential sources than non-ceramic bearing sites, which Watson fails to do. Returning to the original SEWAP data might allow the researcher to determine whether ceramic bearing sites do correlate with the predicted sources.

It may also be possible to refine the approach by removing creeks and rivers which have been excessively altered by post-settlement disturbance. Furthermore, carefully planned field work and survey of the areas discussed here may provide a familiarity with the landforms of Southeastern Wisconsin in new and helpful ways not available to the researcher at this time. Future research regarding predictive models should focus on clarifying the issues discussed here and establishing this link between predicted sources and known prehistoric sites.

References

Arnold, Dean
1985 Ceramic Theory and Cultural Process. Cambridge: Cambridge University Press.
Baddeley, Adrian, and Rolf Turner
2005 spatstat: An R Package for Analyzing Spatial Point Patterns. Journal of Statistical Software 12(6).
Glocker, Carl L.
1971 Soil Survey of Jefferson County, Wisconsin: U.S. Department of Agriculture, Soil Conservation Service.
Goldstein, Lynne
1987 The Southeastern Wisconsin Archaeology Project: 1986-87 & Project Summary. University of Wisconsin-Milwaukee.
Holliday, Vance T.
2004 Soils in Archaeological Research: Oxford University Press.
O'Sullivan, David, and David Unwin
2010 Geographic Information Analysis. Hoboken, New Jersey: John Wiley & Sons, Inc. .
Perkins, Dexter
2002 Mineralogy: Prentice Hall.
Potts, Philip J.
2008 Introduction, Analytical Instrumentation and Application Overview. *In* Portable X-Ray Spectrometry; Capabilities for In Situ Analysis. P.J. Potts and M. West, eds. Cambridge, UK: The Rocal Society of Chemistry.
Rice, Prudence
1987 Pottery Analysis: A Source Book: University of Chicago Press.
Soils Data Mart
2009 SSURGO Soils file, Vol. 3/15/2010. http://SoilDataMart.nrcs.usda.gov/: U.S. Department of Agriculture Natural Resources,.
U.S. Geological Survey
USGS National Elevation Dataset. http://ned.usgs.gov/: U.S. Geological Survey.
USDA
1971 Soil Survey, Walworth County, Wisconsin: United States Department of Agriculture, Soil Conservation Service.
Watson, Robert
1992 A Predictive Model for Ceramic Site Distribution based on Exploitable Clay Resources. Milwaukee, Wisconsin: University of Wisconsin - Milwaukee.
Wisconsin DNR
1998 30m DEM, Vol. 3/11/2010. ftp://gomapout.dnr.state.wi.us/geodata/elevation/DEM_30-meter/dem_30m.zip: United States Geological Survey (USGS).
2010 Waterbody Area, Vol. 4/15/2010. ftp://gomapout.dnr.state.wi.us: Wisconsin Department of Natural Resources.