

**Jill of all trades and master of none?:
the challenges of doing soil chemistry analysis as an archaeologist**

Eva L. Hulse
Social Systems GIS Lab
Anthropology Department, University at Buffalo – SUNY
380 MFAC, Ellicott Complex
Buffalo, NY 14261
evahulse@buffalo.edu

Society for American Archaeology Annual Meeting 2010
St. Louis, MO

Since 2004, I have been reconstructing the spatial patterning of anthropogenic chemical enrichment at archaeological sites in northern Finland. The Kierikki archaeological sites (Figure 1) range in age from 8000 to 4000 years old, and lie in heavily leached soils classified as podzols. Conducting archaeological soil chemistry analysis in northern Finland presents a variety of challenges to standard methodology. A few issues affect spatial analysis in particular. These are: 1) Understanding the spatial characteristics of chemical enrichment in acid-leached soils, so that I could tease out the archaeologically-relevant information, 2) locating the correct tools for visualizing three-dimensional sets of chemical data, and judging whether interpolation was a benefit or a detriment to interpretation, 3) determining the best statistical methods for analyzing the dataset, and 4) linking the fairly quantitative and methodologically-heavy results to traditional anthropological understandings of how people use domestic space. In this paper I give examples of these issues as they relate to my research, and propose that we incorporate elements of another discipline which has done a better job of integrating spatial data into the research process.

The spatial characteristics of enrichment

Reconstruction of prehistoric living floors is imperfect, because post-depositional weathering of the soil moves anthropogenic enrichment deeper into the soil profile (Figure 2). The approach used by Okkonen (Petäjä-Ronkainen and Okkonen 1995; Okkonen and Petäjä-Ronkainen 1996; and personal communication), which I have tested and also subscribe to, is to focus on the darkest portion of the B horizon. This results in a two-dimensional map which roughly approximates the peak of enrichment across an archaeological site. However, under certain circumstances leaching is not as pronounced – oily hearth material counteracts the leaching process, and construction activities alter the soil texture to promote leaching in some areas and discourage it in others (Figure 3). To complicate the issue, it is not safe to assume that a single occupation surface is represented, and layered occupation surfaces may lead to layered episodes of enrichment and leaching.

The literature (for example, Petäjä-Ronkainen and Okkonen 1995; Okkonen and Petäjä-Ronkainen 1996; Kouki 2000) regarding anthropogenic enrichment in leached soils does not explain the mechanisms of leaching in a generalizable way. This is because the traditional approach, of testing either a single sample from each horizon or a single sample from only the B horizon, provides sufficiently useful results that the methodology is not explored in detail. For my own research, I want to characterize both anthropogenic enrichment and the background composition of the soil from the surface to the bottom of the core. This quest for background knowledge led me into the soil science literature, which is dominated by agricultural, ecological, and organic chemistry topics. I found myself working backward to the simplest introductory soil chemistry principles so that I could properly understand the literature on podzols and forest soils. I appreciated this side-trip into another discipline, but the more I read, the more I realized that as an amateur soil chemist, I ran the risk of grossly misunderstanding the basic concepts that I was reading about. An inevitable pitfall of interdisciplinary research is that your area of expertise, which took years to arrive at, is crossing over into an area of expertise which should take several more years of study to properly understand, and yet you need to get up to speed in a year or less. I strongly believe in the benefits of truly interdisciplinary research teams, made of experts from a number of different disciplines. I recently worked on a project in close collaboration with geologists in Russia, and having their knowledge and advice available on the spot has saved me at least a year of intensive reading.

In the event that archaeologists do not have experts from other disciplines on hand whenever they need them, it is important that we not have to re-learn methods from square one. By building up

the body of archaeological literature dealing with methodological issues (and not simply results) in soil chemistry analysis, archaeologists can get the information they need to deal with ambiguous soils and make educated assumptions. One problem is that archaeologists working in the sub-Arctic forest zone around the world may be dealing with identical issues, but are not publishing in the same languages. For example, Finnish archaeologists have in fact been conducting archaeological studies in podzols for a number of years, and have published the results in Finnish. I read and utilize this literature, but not the archaeological literature in Swedish, Norwegian, or Russian. There are many practical difficulties in cross-linguistic literature reviews, such as learning several different systems for cataloging and reporting archaeological data, not to mention learning the appropriate keywords to run searches in each language. For example, it was only through conversation with Finnish archaeologists that I learned that Finns often use the German term “ortstein” to refer to hardpan/ironpan formations.

Selecting statistical methods

Soil chemistry interpretation is a very spatial activity, and depends heavily on statistical analysis, so it is somewhat counterintuitive that it is not much aided by explicitly spatial statistical methods. This is because on a single site level, the relationship between chemical enrichment and known archaeological features is apparent using the fuzzy logic of our own brains, while on a between-sites comparative level, the variability in size and position of archaeological features is great enough that comparisons are qualitative rather than quantitative in nature. It is possible that with a large sample of, say, 50 archaeological sites with extremely similar architecture, spatial statistics would streamline the interpretation process.

Multiple samples from each core are needed to identify archaeological features and multiple episodes of occupation. I need a way to analyze and interpret a dataset which is semi-regular in the horizontal dimensions, but highly irregular in the vertical dimension because of the variable depth of the leached and enriched layers. The spatial irregularity of the sampled area at each site made it difficult to directly compare different soil cores with one another, much less compare one site with another site. For this reason I use statistical methods which do not depend on spatial information. I simply want to know if one soil sample has similar composition to another, and if there are larger patterns visible when I map similar samples.

The basic assumption that I make about soil chemistry is that while each soil sample certainly has spatial characteristics which I know in advance, these do not necessarily influence the composition of that sample¹. In fact, statistical methods used for soil chemistry analysis do not account for qualities of neighboring samples or neighboring features. This is because a peak in enrichment is no less meaningful if it has no enriched neighbors – it may simply represent an isolated activity area, or perhaps the spacing between samples was too great to account for the size of the feature. We do not expect soil composition to vary smoothly across a site. Human activities such as pit excavation and hearth construction increase the variability of enrichment. At the same time, sites are not often large enough that we need to assign a statistical description of the relative irregularity of enrichment, such as patchiness in ecological studies (as described in Kotliar and Wiens 1990). There is no universal threshold of chemical variability that makes one area archaeologically-relevant and another not. Archaeological soil chemistry interpretation is an extremely relativistic activity.

Statistical analysis of soil composition data requires a method which results in a mappable value for each soil sample. One commonly used method is principal components analysis (Hutson and Terry

1 Except, obviously, that enrichment tends to vary predictably within the vertical profile of a podzol, so that I can delete leached samples from my dataset.

2006; Kouki 2000), which reduces complex datasets to a few groups of co-varying elements, and assigns each sample point a score reflecting how closely its composition matches the composition of each component. This works nicely for spatial data as well as non-spatial data. The resulting component scores for each datapoint can be mapped using associated x/y/z coordinates, and interpolated or not as you see fit (Figure 4). Classification analyses such as discriminant analysis (Wilson et al. 2008; Hulse 2008) and cluster analysis (Middleton 2004; Terry et al. 2004) are more predictive than descriptive, and classify soil samples based on their similarity to a set of pre-defined “type-cases.” In discriminant analysis, each sample is given a score reflecting its similarity to each of the type-cases, and these scores can be mapped according to the coordinates of each sample, and interpolated or not (Figure 5).

Selecting visualization methods

Visualization of soil chemical composition over the area of an archaeological site can be done via point markers of varying sizes, reflecting varying enrichments, or via interpolation between sample points and using a color gradient to reflect continuous changes. The interpolated approach is visually more appealing, although it makes the assumption that variation between points is predictable. Some people use the spatially-weighted statistical interpolation method known as kriging to improve on standard interpolation methods. Kriged estimates of enrichment may improve the accuracy of enrichment maps, although I have not yet tested the predictive potential of kriging at the sites where I have collected data. I have avoided this and other interpolation methods because of the difficulty in displaying layered interpolations for a multi-layer set of chemical data. Layering semi-transparent interpolated maps for each level results in a confusing image. In my case, the need to display enrichment at multiple depths in a single 2D map led me to use overlapping point markers of varying sizes instead. Volume interpolation based on voxels (volumetric/3D pixels) is also possible using GRASS GIS, and results in 3D subterranean blobs representing zones of enrichment which can be visualized in NVIZ, and rotated and sliced in various ways.

Linking data to spatial behavior

Archaeologists examining archaeological soil chemistry must make the leap from numbers in a spreadsheet to human behavior. This activity enters the notoriously-slippery territory of middle-range theory (Raab and Goodyear 1984), where all archaeologists eventually find themselves wrestling with the correspondence between things in the dirt and people in the past. Once a suitable mapping solution is reached, the chemical patterns need to be interpreted in light of human behavior. This involves a combination of guesses as to enrichment source based on chemical composition, and comparison of the patterns of enrichment with macro-features of the site – does enrichment follow the walls of a structure, for example. The perceived relationship between enrichment and architecture has somewhat more interpretive power than estimates of enrichment source.

There are a number of obstacles to making a confident identification of the materials which actually decayed into the soil to enrich it in the first place. Organic sources of enrichment tend to enrich in the same elements – phosphorus, sodium, potassium. This means that simple assessment of the levels of each element will not tell the difference between seal butchering waste and fish butchering waste. The power of chemical analysis is as an indicator of presence or absence, when overlaid with other patterns. Even the successful studies such as (Terry et al. 2004) indicate areas of organic decay vs. areas enriched with trace elements but devoid of organic decay (e.g. hypothetical craft areas) vs. cleaned areas (e.g. patio and plaza space). Samuel Vaneeckhout's extensive shovel-testing of sampled

areas fills out the pattern of inorganic waste, so in Kierikki as well, areas of lithics mixed with household waste contrast with other areas containing lithic waste but no organic enrichment (Figure 6). The way patterns of organic waste, inorganic waste, and deliberate cleaning overlap and relate to architecture is where multi-element analysis displays the most utility. There are other compositional analyses which go farther towards identifying specific sources of enrichment, such as gas-chromatographic identification of unique lipids (as in Hjulstrom et al. 2006).

In principle, archaeological soil chemists have it easy. All you have to do is measure the chemical composition of residues left by living people, make a note of the activities that left those residues, and then compare to the composition of your archaeological soils (perhaps by means of discriminant analysis or cluster analysis). The most successful ethnoarchaeological studies do just this (for example, Terry et al. 2004, who compare the activities and use-of-space of modern Maya villagers with their not-so-distant Post-Classic ancestors). Of course, it is often difficult to find living people who engage in subsistence-level activities using Neolithic technology, based on the same plant and animal resources as in the past. And soil weathering tends to shift elements around a bit, which means that a single sample may not actually represent all of the composition of the activity that took place there (see Figure 2). My foray into soil science did not leave me with the tools necessary to rewind the motion of each of the elements back to their original location, but I hope that future researchers will be able to accomplish this.

When the most salient categories of human behavior are “cleaned” vs. “not-cleaned”, you run the risk of telling an overly-simplified story of humans in the past. Fortunately, trash-disposal behavior has been recorded by Western ethnographers for centuries, and a quick dip into eHRAF (the electronic Human Relations Area Files²) gave me as many examples of the range of human waste-disposal behavior as I could possibly synthesize. Unfortunately, human behavior is never as simple as cleaning or not-cleaning, and the subtleties of waste categorization go way beyond organic and inorganic. However, some common themes emerged which may be applicable to my archaeological sites. People who live near water often use that water to dispose of trash. Indeed, increased enrichment on the downhill slope from residential sites may indicate beach middens. People surrounded by scavenging animals worry less about formal disposal of organic trash, which may lead to overall buildup of organic enrichment. The presence of scavengers among the Kierikki people is debatable – a probable dog bone was among the excavated materials in 2004 (Costopoulos, personal communication), but there is no hard evidence for dog-keeping, and they kept no pigs or chickens, the other usual garbage disposals.

Ultimately, the picture that emerges is very complex in terms of site formation processes and site taphonomy, and overly simplistic in terms of human behavior – the inevitable result of viewing a few hundred years of human behavior through a screen of 5000 years of nutrient cycling, groundwater accumulation, tree-root activity, freezing, and thawing. This makes the non-human processes extremely obvious, and somewhat distracting. Soil chemistry analysis is more susceptible to this type of distraction than traditional artifact analysis, because it is difficult to draw a line between human and non-human processes at work.

Problems and solutions

On my good days, I feel like this type of research has made me into a well-rounded researcher, drawing on mathematics, chemistry, ecology, geology, computer science, history and the arts in order to tell a story about the past. On my bad days, I feel like I'm applying amateur-level knowledge of disciplines I'm not trained in, in order to pass myself off as a specialist on human prehistory. Good

² ehrafworldcultures.yale.edu

archaeology must be more than just the sum of its interdisciplinary parts. To reach this point, I have identified two areas where additional information and training would promote better research:

1. Accessibility of previous studies that could allow archaeologists to advance knowledge incrementally, rather than re-invent the wheel, such as cross-linguistic data sharing, and a centralized searchable clearinghouse of published research.
2. Standard tools which are appropriate to a broad range of archaeological data analysis and visualization problems, and basic training in these, so that researchers spend more time analyzing data than learning to use tools.

As my contribution to the discussion about solutions to archaeology's problems with spatial data, I am also offering a brief description of cognitive neuroscience. Neuroscientists have done a better job of integrating spatial data analysis with socially-relevant results than archaeology has. I am moderately qualified to comment on this as I have been a research aide in a cognitive neuroscience lab since 2004, and there are glaring differences in the process of data analysis in each discipline. There are some good reasons why brain data is easier to generalize, and therefore easier to develop tools for working with, but I suspect that much of the difference in tools and training has to do with relative funding for the two disciplines. The end result is that neuroscience labs spend less time hammering at methodological issues, and more time sticking their data into standard tools that conduct standard analyses, so that researchers can worry more about the results than the process.

Like archaeologists examining soil composition, neuroscientists don't often use spatial statistics. They use independent components analysis and discriminant analysis to identify individual active voxels, which they map over time throughout the 3D volume of the brain. However, multi-voxel pattern analysis (MVPA) utilizes spatial patterning of brain activation as well as measures of absolute brain activation in each voxel. The process is conceptually similar to discriminant analysis. The discrimination is not made at the level of points, but at the level of groups of voxels. Data are sorted into brain patterns that occur as a result of a particular stimulus, and then new data can be classified based on these known patterns. The archaeological possibilities are intriguing, although the lack of standard structure between archaeological sites may prevent easy application.

Neuroscientists have a few standard tools for working with spatial data, which could be adapted to archaeological applications:

BrainMap³ is an Internet-accessible database that allows you to search for published case studies relevant to a region of interest simply by selecting coordinates or naming a region. The now-defunct Geography Network had a similar search function for global geographic data. The practical issues of setting up such a database for archaeology are no greater, as inclusion in the database involves either manual entry by an employee of the BrainMap project, or self-submission by the author of a study. In our case, naturally, access to precise locations of archaeological sites may need to be restricted to approved account-holders.

AFNI/Analysis of Functional NeuroImages⁴ is a free, open-source collection of tools which do what GIS software does for archaeologists. As open-source software, it has more in common with GRASS GIS than with ArcGIS. One difference between AFNI and GRASS is that AFNI allows you to normalize your results based on standard brains, so that you can easily compare one person with

3 www.brainmap.org

4 afni.nimh.nih.gov

another. Statistical methods for neuroscience assume that you can compare one dataset with another one a one-to-one level. AFNI has full support for voxel-based statistical analysis and visualization. ArcGIS has barely begun to incorporate 3D visualization, much less analysis. GRASS GIS is functionally similar to AFNI, in that it permits advanced statistical analyses via a command line interface, and 2D or 3D visualization of the results. However, training in ArcGIS is far more common despite its limited 3D capabilities.

Training is the final area in which neuroscience may be doing better than archaeology. My neuroscience lab promotes basic training for all researchers in the use of the common tools. Instead of struggling with limited training in the tools, and a fragmented user community, graduate students, research aides, and professors alike have a common baseline of familiarity with the tools which allows them to move on to more substantive topics, like the results of different experiments. Some of this expertise is the result of a slightly different career trajectory for neuroscience students. It is very common to spend one or two years before grad school working as a research aide for an established lab – in fact, it is difficult to be a competitive applicant for a PhD program unless you have co-authorship on a publication resulting from original research. Therefore, many graduate students have already been trained in the basic tools that they will need to collect and analyze data. Because of differences in funding and research between neuroscience and archaeology, it is impractical to expect incoming graduate students in archaeology to have such training. However, spatial research in archaeology will be advanced if everyone receives practical training in GIS from expert users during their first year in graduate school. This would establish a community of researchers with expertise to share with one another, and allow everyone to move beyond the tools and towards the results of research.

Acknowledgements

This research was possible thanks to the following people and organizations: Ezra Zubrow, Tina Thurston, Sarunas Milisauskas, Andre Costopoulos, Jari Okkonen, Paula Kouki, Ieva PaberzYTE, Samuel Vaneeckhout, Dustin Keeler, Milton Nuñez, Patrik Franzén, the University of Oulu Northern Cultures and Societies Field School, T. Douglas Price, James Burton, and the University of Wisconsin-Madison Laboratory for Archaeological Chemistry, the University at Buffalo Anthropology Department's Marian White Fellowship, the University at Buffalo College of Arts and Sciences Dissertation Fellowship, the American-Scandinavian Foundation, the Finlandia Foundation, the University at Buffalo Mark Diamond Research Fund, the Graduate Student Employees Union Professional Development Fund, the University at Buffalo Anthropology Department's Michael Matsumunyane Travel Grant, the Sackler Institute for Developmental Psychobiology at Weill Cornell Medical College, and National Science Foundation Grant #0631230 (Social Change and the Environment in Nordic Prehistory).

References

Hjulstrom, Bjorn, Sven Isaksson and Andreas Hennius. 2006. Organic Geochemical Evidence for Pine Tar Production in Middle Eastern Sweden During the Roman Iron Age. *Journal of Archaeological Science*, 33 (2): 283-294.

Hulse, Eva L. 2008. "The difference between dirt and other dirt: using multivariate statistical analysis to classify chemical soil enrichment at Late Stone Age archaeological sites in North Ostrobothnia,

Finland." PhD diss., Department of Anthropology, University at Buffalo – SUNY.

Hutson, Scott R and Richard E Terry. 2006. Recovering social and cultural dynamics from plaster floors: chemical analyses at ancient Chunchucmil, Yucatan, Mexico. *Journal of Archaeological Science*, 33 (3): 391.

Kotliar, Natasha B and John A Wiens. 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. *Oikos*, 59 (2): 253-260.

Kouki, Paula. 2000. Polvijärven Multavierun asuinpaikan geokemialliset analyysit (Geochemical analyses from the dwelling site of Multaviera in Polvijärvi). *Muinaistutkija*, 3: 2-14.

Middleton, W D. 2004. Identifying Chemical Activity Residues on Prehistoric House Floors: A Methodology And Rationale For Multi-Elemental Characterization of a Mild Acid Extract of Anthropogenic Sediments. *Archaeometry*, 46 (1): 47-65.

Okkonen, Jari and Anne Petäjä-Ronkainen. 1996. *Geokemiallisia havaintoja Kemijokivarren kivikautisista asumuspainanteista*. Oulu, Finland: Oulu University Archaeology Laboratory.

Petäjä-Ronkainen, Anne and Jari Okkonen. 1995. Traces of Human Prehistoric Influence on Podzol Profiles - Preliminary Remarks. *Fennoscandia archaeologica*, 12: 151.

Raab, L Mark and Albert C Goodyear. 1984. Middle-Range Theory in Archaeology: A Critical Review of Origins and Applications. *American Antiquity*, 49 (2): 255-268.

Terry, Richard E, Fabian G Fernández, J Jacob Parnell and Takeshi Inomata. 2004. The story in the floors: chemical signatures of ancient and modern Maya activities at Aguateca, Guatemala. *Journal of Archaeological Science*, 31 (9): 1237-1250.

Wilson, Clare A, Donald A Davidson and Malcolm S Cresser. 2008. Multi-element soil analysis: an assessment of its potential as an aid to archaeological interpretation. *Journal of Archaeological Science*, 35 (2): 412-424.

Figures

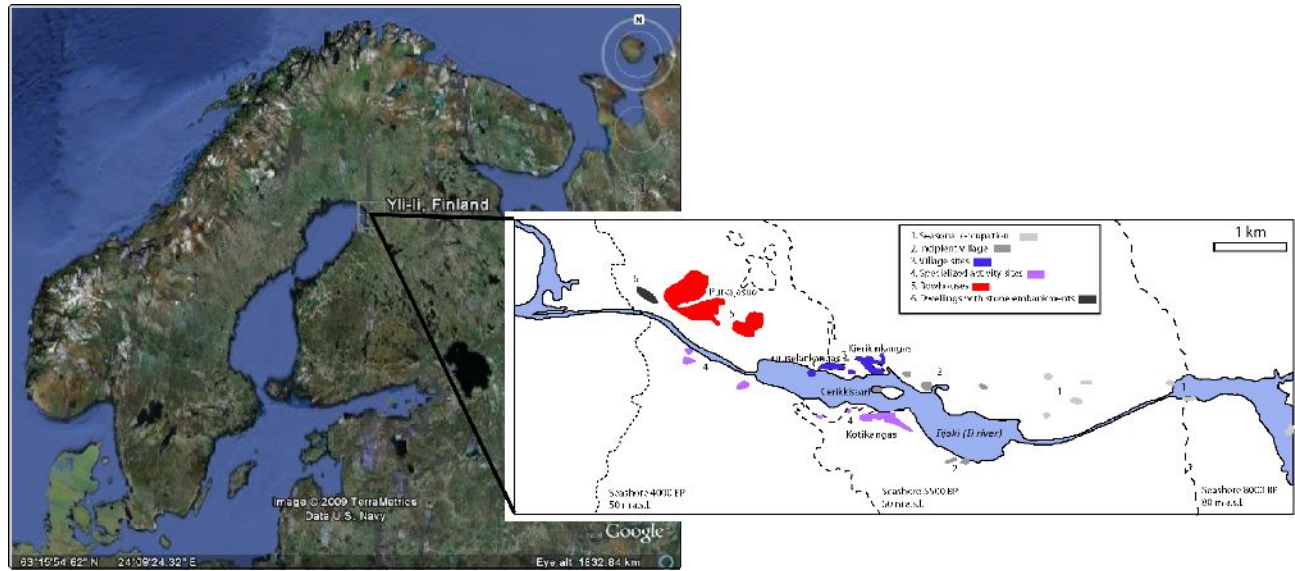


Figure 1: The location of the Kierikki sites in northern Finland.

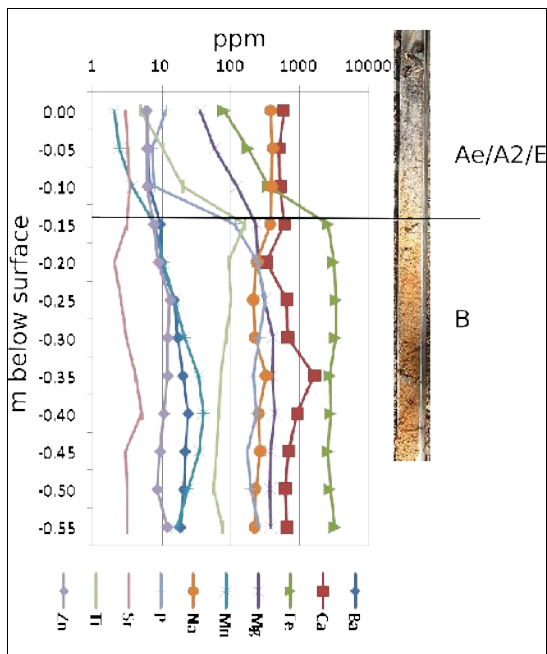


Figure 2: Chemical enrichment in a podzol varies with depth (podzol in photo is not the same core that provided chemical data)



Figure 3: Pits and postholes increase the potential for leaching and preserve the outlines of structures.

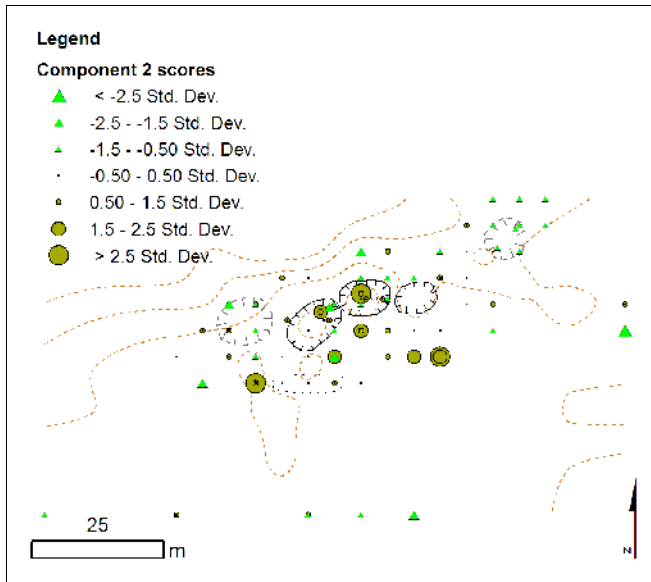


Figure 4: An example of principal components scores, mapped as point values across the site of Kotikangas 2008

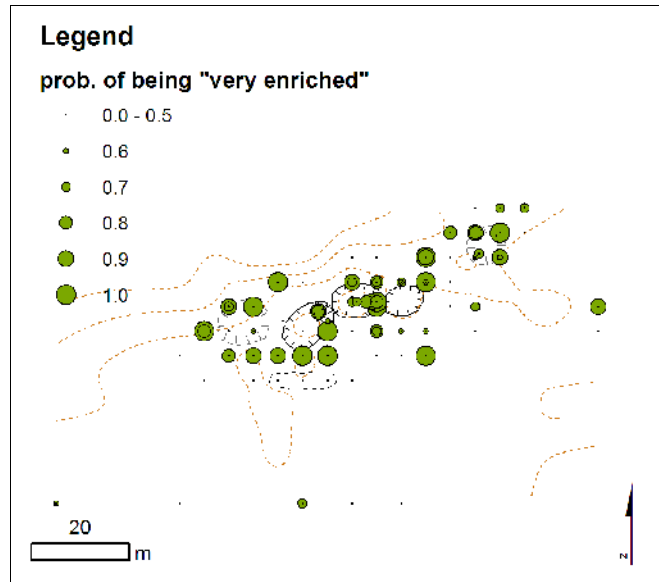


Figure 5: An example of discriminant function classification probabilities, mapped as point values across the site of Kotikangas 2008

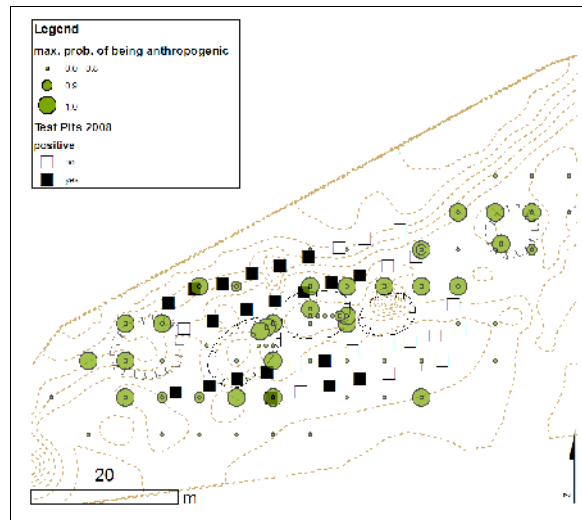


Figure 6: Anthropogenic chemical signature overlaid on locations of shovel tests at Kotikangas 2008.