CORRESPONDENCE ANALYSIS IN ARCHAEOLOGICAL PALYNOLOGY

In large part, archaeological palynology involves finding and recognizing palynological data patterns that can be interpreted as products of some specific paleoenvironmental condition, some form of cultural behavior or the diagnostic indices of a particular temporal horizon. Computer programs which analyze data matrices to reveal patterns that are not inherently obvious are available, but such forms of statistical analysis of archaeological site-context pollen records has been discouraged by the "cheaper is better" philosophy that has characterized CRM-sponsored archaeological palynology. Today, I'd like to discuss the results of a pilot study in the use of correspondence analysis to discover archaeologically relevant patterns of palynological data.

................................**First Slide**

This is a map of the Basin and Range physiographic zone that emphasizes a district informally known as the Arizona Transition Zone. As you can see, the Transition Zone consists of a jumble of valley basins separated by the slopes of mountains hills and ridges. Since the slopes face in all directions, and the valley floors range through a 3000 foot elevational gradient, ecological conditions in the area are not only highly variable, they are very highly localized. There are actually only four plant communities in the area: Desert Scrub, Desert Grassland, Chapparal and Woodland. But each supports a wide range of habitat and microhabitat conditions, often separated by very narrow ecotones, and a similarly wide range of plant associations. Surface sample data suggests pollen rains in the district are similarly heterogeneous. The samples produce a broad range of pollen spectrum characteristics, but no particular portion of that range seems diagnostic of any particular sort of plant community or plant association. The single product of 25 years of study of surface samples from this district is recognition that we cannot find securely evidenced diagnostic relationships between the vegetation and the pollen records it produces.

Yet this district yields an archaeological record of occupation by populations of at least four different archaeological cultures (Hohokam, Salado, Northern Sinagua and Southern Sinagua), and appears to have been one of the more densely inhabited areas of the American Southwest between 800 and 1500 A.D. Archaeologists working in the area strongly suspect that there is a real story to be told through analysis of the adaptive relationships of the occupants of Transition Zone sites and their local paleoecological contexts, and archaeological palynologists have been analyzing samples recovered from sites in the district for the past 25 years. So far, we have been unable to identify patterns of palynological data that can be interpreted as indices of paleoenvironmental conditions, so archaeological palynologists have had to content themselves with the rather mundane problem of identifying evidence of subsistence patterns that can support or amplify evidence obtained through analysis of flotation samples.

I hoped correspondence analysis might be a way to break through this impasse. Correspondence analysis is a form of principle components analysis that has been found particularly helpful for identifying patterns that occur in data sets characterized by a high degree of heterogenity. Basically, this form of analysis applies an algorithm that establishes a mathematical "profile" for each of the rows and each of the columns of a data matrix. It then plots the relationships of those profiles to an analysis of variance in a way that displays the correspondence of one set of profiles with the other. With pollen data, the profiles for the various pollen taxa are related to the profiles for the various samples of a population in a way that illustrates their correspondence as points plotted in the same part of a two dimensional graph. Samples that have more similar pollen values are plotted closer to each other on the diagram, and the pollen taxa that most highly influence that positioning of the samples are plotted to the same region of the diagram.

In practical terms, what this means is that correspondence analysis is a way to reveal which pollen taxa make primary contributions to any clusters of pollen samples that are plotted along the axes of variability in the data population. Here's an illustration of how it works.

Second slide

This slide displlays the resuls of a correspondence analysis of the pollen spectra of 43 surface samples from the Cherry Creek basin. All but a few samples were recovered from various plant associations in the Desert Scrub plant community, with some from the Desert Grassland community at higher elevations and a single sample was collected at the extreme margin of the basin's floor.

The lower graph shows that half of the samples occur in the lower right quadrant of the diagram. The upper graph documents the fact that only seven of the 22 pollen taxa influence this distribution at all, and only two have a very strong influence. Alternatively, half of the remaining samples cluster in the lower left quadrant, and it is the frequency values for six taxa that influence that distribution. The taxonomic array ranges from pollen transported long distances from much higher elevations to the pollen of plants typically observed in more arid basins at significantly lower elevations. I have highlighted those pollen types produced by plants adapted to the most warm or xeric and the most cool or mesic habitats. As you can see, the distribution of the two principle clusters of samples corresponds to the distributions of frequency values for these two classes of pollen types. When the vegetation of the sample locations was analyzed, it became clear that these samples were recovered from the most mesic and the most xeric local habitats of the Desert Scrub and Desert Grassland plant communities. So correspondence analysis provides a mechanism by which such habitat conditions could be recognized palynologically.

The Cherry Creek Basin surface samples had been selected for analysis because they sampled a less disturbed version of the flora which exists in the Tonto Creek Basin a few miles to the west. A great deal of archaeological work has been done there recently, and a suite of 167 pollen spectra recovered from 9 Hohokam and 17 Salado archaeological sites is available.

Third Slide

Correspondence analysis of the archaeological pollen spectra revealed, as you can see, a pattern lacking sample clusters and lacking any sort of systematic correspondence between the distribution of samples and the distribution of indices of xeric or mesic local habitat conditions. However, it also revealed other things.

You can observe here that the majority of the samples are distributed in the lower half of the graph while almost all of the pollen taxa are distributed in the upper half. Also, that a few samples are distributed in the upper right hand corner and that some taxa are also distributed in that fashion.

Fourth Slide

In the upper graph I have highlighted the distributions of three sorts of pollen taxa: (1) pollen of taxa that are products of field cultivation, including the weeds that characterize irrigated fields, and taxa that the flotation record identifies as intensively collected wild food plants; (2) pollen transported from long-distance sources; and (3) the pollen type normally interpreted as the primary product of human habitat disturbance. The lower graph illustrates the distribution of samples from the sort of specialized storage facilities observed at certain Salado sites and from the pithouse structures characteristic of Hohokam sites.

Correspondence analysis allows us to recognize that Salado granery facilities were also used to store other plant materials. It also suggests that the reason that the samples data is distributed in the lower half of the diagram is related to the effect of human disturbance on the paleoenvironment. Samples in which chenopod pollen is not overrepresented contain more long-distance transport pollen types, which skews their distribution to the left side of the diagram. Many samples in this class were recovered from the floors and fill of Hohokam pithouses. Alternatively, samples in which chenopod pollen occurs in high or very high frequency are skewed to the lower right of the diagram, and few such samples derive from subsurface contexts.

Fourth Slide

When the distributions of samples from Hohokam sites and Salado sites are compared, it becomes clear the distributions of Salado samples are skewed by the influence of chenopod pollen and the pollen of cultivated and collected taxa to a far greater degree than the Hohokam samples. But some Salado samples fall into the "middle ground" of the horizontal axis of the diagram where most Hohokam samples are found.

Fifth Slide

That distribution is partially explained on this diagram, which illustrates the distributions of the pollen records from plaza surfaces and from rooms at Salado sites. Rooms were roofed spaces, plazas were unroofed. The palynological effects of local disturbance, and of the storage and consumption of plant foods seems to be more strongly expressed in roofed contexts.

I undertook this research to discover whether or not correspondence analysis would be a more effective means of revealing the paleoenvironmental contexts of prehistoric occupations of the Arizona Transition Zone. I found that it has that potential, but it cannot be realized with the sorts of pollen records archaeological palynologists have been recovering from sites in the region for the past 25 years. Correspondence analysis of those sorts of pollen records, however, suggests both the reason this is so and the solution to the problem. Simply stated, the pollen records we have are so affected by palynological expressions of intensive agricultural production and human-induced disturbance that they mask the palynological signatures of paleoenvironmental conditions, and the palynological record of paleoenvironmental changes that were products of natural processes. If we really want to know more about the character of the natural paleoenvironment, we must observe sufficiently more pollen from each sample that this mask can be statistically removed.

Alternatively, we can take advantage of the evidence that correspondence analysis offers a good deal of potential to explore ways in which behavior patterns are palynologically expressed. This study, for example, suggests that Salado populations practised a significantly more intensive form of agriculture in the Tonto Basin than the earlier Hohokam population, and had a greater adverse impact on the local environment. Perhaps these effects were simple by-products of population size, but that can be checked by comparing the spectra of smaller and larger Salado sites. We may find that the palynological patterns revealed by this analysis suggest and help us to identify basic distinctions in the adaptive structures of the two groups.