Determination of Vessel Morphology From Sherd Weight: A Statistical Model for Mississippian Funerary Ceramics

By: Landon D. Smith
USDA Forest Service, Region 3
517 Gold Avenue, S.W.
Albuquerque, New Mexico 87102

and

Dee F. Green
USDA Forest Service, Region 3
517 Gold Avenue, S.W.
Albuquerque, New Mexico 87102

Abstract

The number of whole or potentially reconstructable vessels from archeological context is disappointingly small. This fact seriously hampers studies of variables such as vessel capacity. If these variables were known, studies could potentially provide better understandings of social and cultural dynamics within and between prehistoric populations. Using the stochastic methods of correlation and Model II linear regression, a series of relationships were isolated that allowed the mathematical prediction of culturally meaningful measures for partial or broken vessels such as volume, diameter and height using the known, or even the estimated, value of weight.

Introduction

In a recent article on ceramic classification, Ericson and Stickel (1973) suggest the applicability of regression models for use in establishing the relationships between morphological variables in a body of ceramics (specifically weight and volume). Using the produced regression formula against the modeled data, they were able to predict capacities from known weights whose sum was within error parameters of less than one percent of the known total volume.

While demonstrating the essential viability of such an approach, the method was seriously hampered in a number of areas. For example, the sample from which the model was constructed was very small (only 25 vessels) and, furthermore, was composed only of modern Mexican ceramics--these pots being probably wheel-made or even mould-formed. Also noted were the choice of an inappropriate statistic and the incomplete assessment of the results, all of which seem to indicate a general lack of understanding of the potential of the stochastic tools used.

The possibility cannot be ignored that these factors may have produced results that are somewhat misleading when applied to most archeological data sets, particularly those within the continental United States. For example, how may these results be related to any archeological data set where manufacturing techniques were such as to result in greater variability of the morphology? Sampling theory is quite emphatic that predictions based on a sample are most suitably applied only to the sampled (or target) population (Cochran 1963:6). In the case of this example, only modern Mexican ceramics are sampled; therefore, the analytical results are most suitably extended only to modern Mexican ceramics, assuming that proper sampling procedures have been used.

Furthermore, low sample sizes are always a problem since they tend to introduce the factor of sampling error. Certainly 25 vessels seems a somewhat inadequate sample for our purposes. For these reasons and others, it was felt that another test was clearly in order--one utilizing a large sample of vessels, all of which would be of prehistoric context. Such a body of data was available, and will be discussed below.

In addition to testing the relationship between weight and volume using a ceramic sample from archeological context, we were also interested in the following:

1. Isolating those morphological variables showing particularly high correlations with each other.

2. Establishing those relationships in a quantitative manner such that prediction of one variable based on the other is possible within acceptable limits.

3. Establishing the general parameters of these highly correlated variables in order that the nature of the variables is well understood in a descriptive sense.

Ericson and Stickel (1973:366) suggested two general methods for undertaking further studies along their proposed research lines. One involved the study of a collection of whole vessels representing some general cultural area, and the second required the reconstruction of variation within a ceramic population represented by the sherds at one site. We were fortunate in having the former readily available, and suggest that the second may introduce so much uncontrolled error as to make results difficult or impossible to interpret.

Data Body

The ceramic material used in this study consist of 656 whole vessels from several sources. Two groups of vessels of known context were recovered from prehistoric graveyards at the Hazel (3PO6, 462 vessels) and Togo (3CS24, 87 vessels) sites in eastern Arkansas. The pottery was excavated during the 1930's under the Works Projects Administration by S. C. Dellinger. The final group of ceramics, consisting of 107 vessels, is drawn from a donated collection called the McPherson Collection. The consensus of opinion is that most of the ceramics in the collection came from the graves at the Togo site (e.g. Green 1974a:27, Green 1974b:1). It is not felt that these 107 vessels will introduce unacceptable variation into the study. It should be recognized that these data are constrained both spatially and temporally.

All observations were made by Green during the 1964-1965 school year, financed by a National Science Foundation grant to Dr. Charles R. McGimsey III (McGimsey and Green 1965, Green 1974a, Green 1974b, Green 1975). The attributes and variables chosen were based in part on Shepard (1961). The observations were encoded on 80-column punch cards. A discussion of the original attribute selection and observation is contained in the previous papers mentioned and the interested reader is referred to them. For our study, only those observations that utilized a measurement level of <u>interval scaling</u> were selected, some 13 in number (Figure 1). All measurements were in millimeters, liters (in the case of volume), and kilograms (in the case of weight).

Statistical Tests

Our concern, as has been mentioned, was with checking the results of Ericson and Stickel, but on a body of data we felt to be more appropriate--that is, of archeological context and of larger size. The statistical methods used are correlation and linear regression analysis.

Some of the issures involved in the use of correlation and regression may be found in standard statistical texts

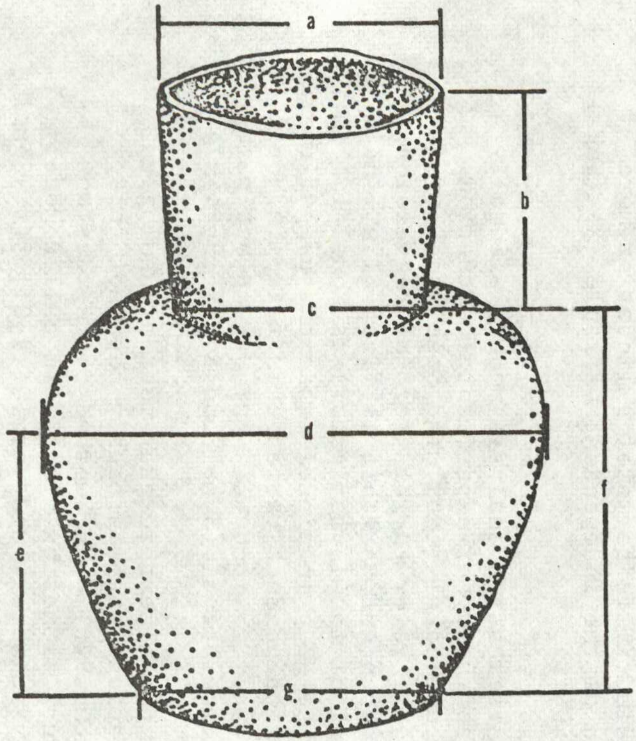


FIGURE 1. Location of the points used in measurement of the variables in this study. a) Orifice diameter. b) Distance from throat to orifice. c) Throat diameter. d) Body diameter. e) Shoulder height. f) Body height. g) Diameter of vessel at the base equator. (Drawn by Chad T. Phinney.)

(Sokal and Rohlf 1969, Steel and Torrie 1960), and are discussed in relation to this particular body of data elsewhere (Smith 1976). These concerns will be treated only cursorily here.

Correlation analysis is a method of estimating the degree to which two variables vary together. Regression analysis, on the other hand, provides an estimate of one variable's relationship with another, expressing the dependent variable (Y) in terms of a linear or more complex function of the independent variable (X) (Sokal and Rohlf 1969:404ff).

The common regression analysis model is ideally suited to the laboratory or experimental situation. Following Rohlf and Sokal (1969) and Steel and Torrie (1960:161), we will refer to this form of regression in which the independent variable is fixed (i.e., held under the control of the investigator) as Model I. An important point to keep in mind is that the use of correlation when the independent variable is fixed (a Model I regression context),"...is not in any way an estimate of parametric correlation" (Sokal and Rohlf 1969:497).

We have argued elsewhere that most archeological analytical situations would probably not meet this constraint (Smith 1976). The method of regression suggested for such situations is a form called Model II regression analysis (Steel and Torrie 1960:165ff, Sokal and Rohlf 1969:481-6). Generally speaking, it would appear that cases which fall under this form of analysis are probably better analyzed using correlation, which is completely appropriate in this context (Sokal and Rohlf 1969). But it is often also the case that a quantified expression of the relationship is desirable; thus a Model II regression approach would be utilized.

The simplest statement of the main difference between a Model I and a Model II method is that observational error must be accounted for in both the dependent and independent variables. The details of this area may be found in Sokal and Rohlf (1969:484-5) and are summarized in Smith (1976). The mathematical method we have used for our Model II regression analysis in this paper is the Bartlett's Three-group Method (Sokal and Rohlf 1969:481-6).

In summary, while similarities exist between correlation, Model I and Model II regression analysis, the underlying assumptions place constraints on their appropriate application. These must be taken into consideration by the researcher.

Results

Since our first goal was to test Ericson and Stickel's assertion of a relationship between volume and weight of a high order, and since they had employed a Model I regression approach, we felt constrained to used a Model I for the sake of comparability. Our results produced the following mathematical expression of the relationship of these two variables:

$$\text{Volume} = 1.557\ (\text{Weight}) + 23.453$$

A comparison of the sum of the actual volumes of the vessels (86238 liters) to the sum of the estimated volumes obtained using the Model I regression formula (86664.85 liters) reveals an overestimation of approximately 425 liters. This is negligible (0.49 percent) and of the same order of magnitude as that reported in Ericson and Stickel (1973:64-5). Considering the differences in manufacturing techniques and control between the two bodies of data, this is remarkable.

It is clear that, even using what we suggest are less than appropriate statistical techniques, their method works and works well, even for prehistoric ceramics. Given the basic validity of the approach, let us now expand the scope of the research and improve on the statistical tools used.

First of all, other variables may also show strong relationships to weight, and these could be of importance to the researcher. The best method for quickly isolating the correlated variable pairs is through a correlation matrix. Since our linear regression model will be of type II, correlation will be entirely appropriate. We can also select a level of positive or negative correlation of r = .71 because of the relationship of the r^2 as an estimator of the percent of variance accounted for by the regression formula. An r of .71 means that 50 percent of the variance present would be accounted for and this serves as what we felt would be an acceptable lower limit for rejection or use of a given variable pair.

Five pairs of variables were isolated when all vessel forms were examined as a group (see Smith 1976, Table, p. 8 for the complete correlation matrix for the data). The results of the regression analysis are presented in Table 1 and the relational network is shown in Figure 2.

Though the tabled results are self-explanatory, several general observations are in order. The first of these involves a comparison of the Model I results with the Model II results for weight and volume. While the differences are not marked, they are present. It is interesting that in this case all estimates using the Model II approach contained much less error than Model I results.

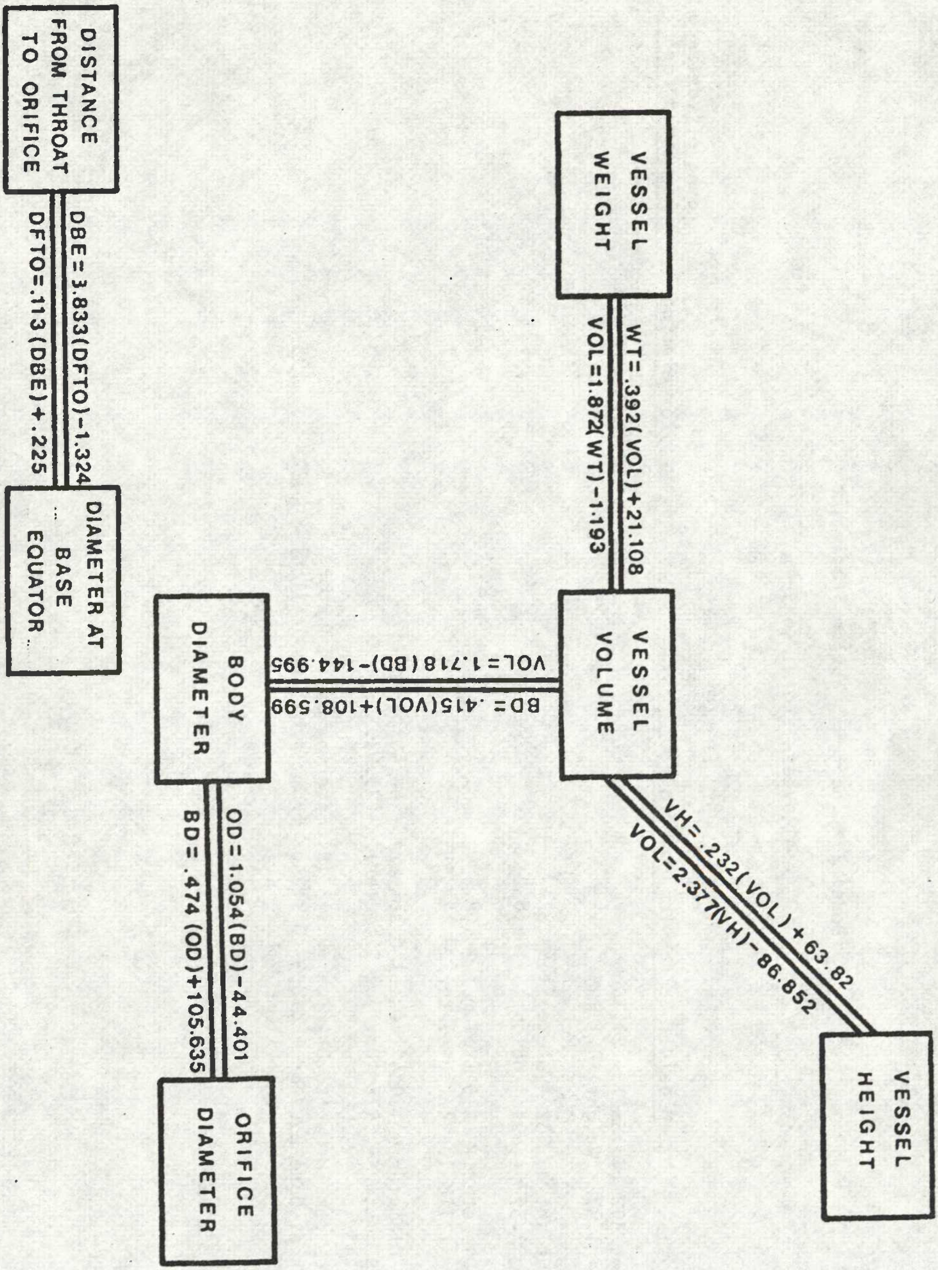
By separating the body of ceramics into three groups based on form (i.e., bowls, jars and necked vessels), it was possible to increase the r^2 values (amount of variance accounted for by the regression formula). As an example, regression formulae for the relationship of weight and volume raised the r^2 values for jars to 81 percent of the variance accounted for and for bowls, 78 percent was accounted for (see Smith 1976, Tables 3-5).

Model II Linear Regression Statistics

All Vessel Forms (n=594)

Table *1.*

Model II Regression Formula	Correlation	Percent of Variance Accounted For
Weight = .392 (Volume) + 21.108 Volume = 1.872 (Weight) - 1.193	.75	57
Vessel Height = .232 (Volume) + 63.82 Volume = 2.377 (Vessel Height) - 86.852	.72	51
Vessel Diameter = .415 (Volume) + 108.599 Volume - 1.718 (Vessel Diameter) - 144.995	.77	60
Orifice Diameter = 1.054 (Vessel Diameter) - 44.401 Vessel Diameter = .475 (Orifice Diameter) + 105.635	.75	57
Distance from Throat to Orifice = .113 (Diameter at Base Equator) + .225 Diameter at Base Equator = 8.833 (Distance from Throat to Orifice) - 1.324	.99	98


DISTANCE FROM THROAT TO ORIFICE
DBE=3.833(DFTO)-1.324
DFTO=.113(DBE)+.225
DIAMETER AT BASE EQUATOR
VESSEL WEIGHT
WT=.392(VOL)+21.108
VOL=1.872(WT)-1.193
VESSEL VOLUME
VOL=1.718(BD)-144.995
BD=.415(VOL)+108.599
BODY DIAMETER
VH=.232(VOL)+63.82
VOL=2.377(VH)-86.852
VESSEL HEIGHT
OD=1.054(BD)-44.401
BD=.474(OD)+105.635
ORIFICE DIAMETER

A variable such as volume may be estimated not only from a weight value but also from vessel height or vessel diameter. Comparison of the various results can help to reduce the error in estimation present.

Summary and Conclusions

Improving on an approach first suggested by Ericson and Stickel (1973), it has been possible to quantify the morphological relationships for a body of ceramics such that prediction of these variables given only weight, or an estimate of the weight, is easily possible within a relatively small range of error. It is felt that the approach has the obvious potential of moving from a readily obtainable datum for partial or broken vessels such as the weight, to variables that have more potential meaning in a cultural sense such as capacity.

The research, only partially reported here and discussed more fully in Smith (1976), has moved rapidly from linear models to multiple regression (which are both additive models). Work now completed has included the construction of interactive (i.e., nonadditive) multivariate models for this data body which has had the result of accounting for almost 90 percent of the variation present within the data (see Jensen 1973 for the general procedure). The importance of using a higher r^2 as the major criterion for a model or approach should be noted.

The information potential of potsherds has hardly been tapped. Just counting and typing these numerous artifacts is only a beginning, and it is felt that approaches similar to the one outlined here are a further step toward reconstructing man's past behaviors in more meaningful ways.

References Cited

Cochran, William G.
1963 Sampling Techniques, 2nd Edition. John Wiley and Sons. New York.

Green, Dee F.
1974a Cluster analysis of 549 whole Mississippi ceramic vessels. Ph.D. dissertation. Arizona State University. Tempe.

1974b A statistical analysis of decoration on Mississippian funerary ceramics from the Hazel and Togo Sites, eastern Arkansas. Newsletter of Computer Archeology, 10:1-25. Tempe.

Green, Dee F.
1975 Testing a traditional typology using cluster analysis. Computer Applications in Archeology 1975, p. 25-32. Birmingham, England.

Ericson, Jonathon E. and E. Gary Stickel
1973 A proposed classification system for ceramics. World Archaeology, 4:357-367.

Jensen, Chester E.
1973 Matchacurve-3: Multiple-component and multidimensional mathematical models for natural resource studies. USDA Forest Service Research Paper INT-146. Ogden, Utah.

McGimsey, Charles R., III and Dee F. Green
1965 Computer analysis of eastern Arkansas ceramics. Paper presented at the 1965 meeting of the Society for American Archaeology, Urbana.

Shepard, Anna O.
1961 Ceramics for the Archeologist. Carnegie Institution of Washington Publication 609.

Smith, Landon D.
1976 Regression modeling and archeological data. Newsletter of Computer Archaeology, XI (3) 1-18. Tempe.

Sokal, Robert R. and F. James Rohlf
1969 Biometry. W. H. Freeman and Company. San Francisco.

Steel, Robert G. D. and James H. Torrie
1960 Principles and Procedures of Statistics McGraw-Hill Book Co. New York.