JOURNAL OF MULTIVARIATE ANALYSIS 4, 308-326 (1974)

A Survey of Statistical Problems in Archaeological Dating

R. M. CLARK

Department of Mathematics, Monash University, Clayton, Victoria, Australia 3168

Communicated by M. Rosenblatt

This expository paper gives a survey of statistical problems arising in two important and widely used scientific methods of dating archaeological deposits, namely tree-ring-calibrated radiocarbon dates and seriation.

1. INTRODUCTION

Archaeologists are increasingly relying on scientific methods for both relative and absolute dating. These methods include numerical techniques such as seriation by multidimensional scaling [29, 31], physical methods such as potassium-argon dating, thermoluminescent dating, radiocarbon dating [1], and biological methods such as tree-ring dating [17] and varve chronology [61].

All these methods involve the measurement of various physical quantities with possible errors of a random nature. Thus, in a trivial sense, all these methods of dating involve some sort of statistical analysis. But in most cases, the statistical problems are overshadowed by the technical difficulties of the method.

Two important exceptions are radiocarbon dating and seriation; here the statistical problems are paramount. The aim of this paper is to give an expository survey of various statistical problems and proposed solutions in connection with these two methods of dating.

2. RADIOCARBON DATING

2.1 *Background*

Radiocarbon or carbon-14 is produced in the atmosphere by the action of cosmic rays at the rate of two radiocarbon atoms per square centimetre of the

Received April 10, 1974.

AMS 1970 Subject Classification: Primary 62P99, 62-02.

Key words and phrases: Archaeology, radiocarbon dates, seriation, polynomial regression, incidence matrices.

earth's surface per second, or equivalently, 7.5 kg globally per year (see [1]). Although these isotopes of carbon are radioactive, they behave chemically like ordinary carbon, and so, e.g., they combine with oxygen to produce carbon dioxide. This radioactive carbon dioxide is absorbed by plants during photosynthesis, and by animals which eat the plants. While a plant or animal is alive, any of its radiocarbon atoms which decay are immediately replaced by "fresh" radiocarbon from the atmosphere, and so the concentration of radiocarbon in all living matter must be the same as that in the atmosphere, currently 1 in 10^{12} relative to the most common isotope, carbon-12.

However, once an organism dies, its store of radiocarbon is not replaced but decays away exponentially at a known rate. Thus in principle, the age of any sample of organic matter (such as fragments of wood or charcoal from an archaeological site) can be determined simply by measuring the current concentration of radiocarbon in that sample.

The steady exponential decay of radiocarbon may be represented by the equation

$$A_m = A(x)\, e^{-\lambda x}, \tag{1}$$

where $A(x)$ denotes the concentration of radiocarbon in the sample when it "died" x years ago, A_m the measured concentration of radiocarbon in the sample *now*, and λ denotes the decay constant related to the half-life T by

$$\lambda = \log 2/T. \tag{2}$$

In practice, both x and $A(x)$ are unknown. The radiocarbon age, y, of the sample is determined by *assuming* that $A(x) = A_0$, the concentration of radiocarbon in living material *now*, and solving the equation

$$A_m = A_0 e^{-\lambda y}$$

giving

$$y = (1/\lambda) \log(A_0/A_m). \tag{3}$$

Further details may be found in the nonspecialist articles by Libby [34], Michael and Ralph ([38], Ch. 1), Renfrew ([46], Ch. 3 and Appendix), while [1, 5, 41] discuss the complications which arise in practice.

In practice, the quantities A_m and A_0 in (3) are determined by counting for several days the emitted beta-particles arising from the disintegration of carbon-14 atoms in the sample. Thus every radiocarbon date is subject to unavoidable random errors of measurement, due to the random nature of this emission, and to the continual random fluctuations in the background radiation, which cannot be eliminated entirely, despite the use of anti-coincidence counters [5, 34]. The resultant lack of precision due to "counting errors" of any radiocarbon date is expressed in terms of its standard deviation, which may range from 40 to 120 yr [14, 37, 58]. However, most radiocarbon laboratories ignore the possibility of additional noncounting errors [41].

In order to date samples from widely different geographical locations, it is necessary to assume that the concentration of carbon-14 in living matter at any time in the past has been the same all over the world. That is, $A(x)$ is to be thought of as this global concentration in all living matter x years ago. Recent work [33] confirms this assumption, often called Libby's principle of simultaneity.

2.2. *Tree-Ring Calibration of Radiocarbon Dates*

Initially, the validity of the radiocarbon dating method was checked by dating samples of known age from archaeological sites in Europe and the Middle East. Subsequently, more accurate comparisons [16, 65] indicated that in the past 1300 yr, the atmospheric concentration of carbon-14 had fluctuated by up to 1.5% (equivalent to 120 yr), but there had been no systematic deviation from recent levels, apart from changes in the past century caused by the burning of fossil fuels and the explosion of nuclear weapons.

The situation was transformed by the very long dendrochronology obtained by Ferguson [17, 18] using the bristlecone pine *Pinus aristata*. These trees, which grow high in the White Mountains of California, live for up to 4,600 yr [52], and it is possible to date fragments which are even older than the oldest living tree, by "cross-dating" tree-rings (see [19, 21, 54] for further details). In this way, Ferguson was able to date fragments of bristlecone pine up to 8,200 yr old, with an error claimed to be no greater than 10 yr.

Nearly 600 tree-ring-dated samples of bristlecone pine have now been radiocarbon-dated at three laboratories [14, 37, 57, 58]. These results indicate major fluctuations in the concentration of radiocarbon prior to 1,000 B.C., the radiocarbon dates being systematically younger (i.e., more recent) than the corresponding tree-ring or calendar dates by up to 700 yr. Therefore it is necessary to correct or calibrate radiocarbon dates, at least for material which is more than 3,000 yr old.

Libby's principle of simultaneity implies that all samples of age x should have the same radiocarbon age $F(x)$, which is related to the radiocarbon concentration $A(x)$ by

$$\begin{aligned} F(x) &= (1/\lambda)\log[A_0/A(x)\, e^{-\lambda x}] \\ &= x + (1/\lambda)\log[A_0/A(x)]. \end{aligned} \tag{4}$$

This "calibration function" F thus defines the theoretical relationship between radiocarbon dates and tree-ring or calendar dates. Given the radiocarbon age, y_0, of a particular archaeological sample, its real age, or the calibrated radiocarbon age, x_0, is given by

$$y_0 = F(x_0). \tag{5}$$

Since we have no independent information concerning past variations in $A(\cdot)$, the radiocarbon concentration of the atmosphere, F is unknown and must be estimated. The data in [14, 37, 57, 58] satisfy the equations

$$y_i = F(x_i) + e_i, \tag{6}$$

where x_i denotes the tree-ring or calendar age of a typical bristlecone pine sample, y_i its corresponding radiocarbon age with its attendant measurement error e_i, and $F(\cdot)$ denotes the calibration function (4). Since the errors in the tree-ring dates are so small, the x's may be regarded as constants. The problem is then that of estimating the regression function F from the approximate values of its ordinates at known abscissae $\{x_i\}$.

Suess [56] was the first to publish a tree-ring calibration curve for radiocarbon dates, based on data from bristlecone pine samples. His revised curve [58], extending back to 5,300 B.C., has been widely used by archaeologists; the consequent corrections to existing radiocarbon dates from archaeological sites have had a dramatic effect on the interpretation and understanding of European prehistory (see Renfrew [44–46]).

Suess's calibration curve contains a large number of irregular undulations or "kinks," implying that several calendar dates may correspond to the same radiocarbon date, i.e., Eq. (5) need not have a unique solution for x_0. There has been much argument concerning the magnitude, location and even the reality of these kinks (see discussion following [58]). There is no intrinsic geophysical reason why there should not be kinks in the calibration function, since the rate of production of carbon-14 could have changed rapidly in the past, due to changes in the cosmic-ray flux, the earth's magnetic field and/or solar activity [7, 13]. However, to a statistician, Suess's curve is most unsatisfactory, since it was obtained by freehand smoothing [58, p. 310].

The first calibration curve derived by explicit statistical methods [63] tacitly assumed that the data satisfied the equations

$$x_i = G(y_i) + e_i', \tag{7}$$

where x_i and y_i have the same meanings as above, and $G(\cdot)$ denotes an "inverse calibration function," assumed to be a low-order polynomial whose parameters were then estimated by least-squares. Equation (7) implies that there can be only *one* calendar date corresponding to any given radiocarbon date, but it does not rule out the possibility of several radiocarbon dates corresponding to the same calendar date, which does not make sense geophysically. If the x_i's are regarded as constants and the y_i's as random variables, it does not make sense to estimate G by least-squares either.

Alternative calibration curves have been produced [15, 39, 60] with the deliberate aim of smoothing out any kinks and so facilitating the calibration of dates. These curves, obtained by a combination of polynomial regression and moving-average techniques, are not entirely satisfactory from the statistical viewpoint, as noted in [47]. In particular, there is a fundamental error in the formula given in [15] for the standard error of a calibrated radiocarbon date.

There is broad agreement amongst these various calibration curves, although there are differences in detail. The discrepancy between the tree-ring age x and the radiocarbon age $F(x)$ implies that the atmospheric concentration of carbon-14 in the past ($A(x)$) has differed from current levels (see Eq. (4)). Various causes of this past variation in carbon-14 concentration have been suggested (see Sect. 2.4 of [1] for a review); it is thought that the long-term fluctuation is related to changes in the earth's magnetic field while the short-term fluctuations reflect solar activity. As yet, there is no definitive explanation, and so the form of the calibration function F is largely unknown. Nonparametric estimates [43] of F would therefore seem more appropriate than estimates based on an assumed parametric form which could be incorrect.

2.3. *Validity of the Bristlecone Pine Calibration*

The validity of the bristlecone pine calibration of radiocarbon dates has been questioned on both geophysical and archaeological [35] grounds. For example, Berger [5] has suggested that the carbon-14 concentration of bristlecone pine wood could be abnormally high due to *in situ* production of carbon-14 by cosmic rays at the high altitude at which the bristlecone pine grows. Also, it is possible that, because of the very narrow rings of the bristlecone pine (typically 0.1 mm), the inner rings of a tree may be enriched in radiocarbon by sap from the outer (younger) rings.

The historical calendar of ancient Egypt provides an independent chronology for testing the bristecone pine calibration prior to 1500 B.C. Qualitative comparisons by several workers [37, 51, 57] of radiocarbon dates of Egyptian samples and the accepted Egyptian chronology show general agreement with the bristlecone pine results, within the uncertainties of archaeological context. Clark and Renfrew [10] were the first to conduct a statistical comparison of radiocarbon dates of bristlecone pine samples and Egyptian samples, assuming that over the time-period considered, namely 3100–1800 B.C., the calibration curve could be represented by either a polynomial or a continuous piecewise linear curve. This analysis showed that there was no reason to doubt the validity of the bristlecone pine calibration; this conclusion should be regarded as provisional to some extent, because of the uncertainties regarding the calendar dates of the Egyptian samples. Incidentally, an examination of the radiocarbon dates of paired samples indicated that the real errors of measurement could be 40% greater than those reported by the laboratories.

Alternatively, an independent chronology can be obtained by examining the annual layers or "varves" found in clays that originate in the beds of lakes dammed up by glaciers. Recent comparisons [55, 61] of radiocarbon dates and varve chronologies in Sweden and U.S.A. have confirmed the general trends of the bristlecone pine calibration.

The data used in these independent checks on the bristlecone pine calibration are not sufficiently precise to either substantiate or refute the kinks in Suess's calibration curve. There is little geophysical evidence for or against the existence of kinks. It has been shown [4] that over the past century there is a high negative correlation between radiocarbon concentration and sunspot number, implying the existence of kinks in the calibration function with a period of about 11 yr. However, these kinks are "averaged" out in the bristlecone pine data, because all the samples used contain at least 10 consecutive rings, and so each radiocarbon date is related to the *average* concentration over those 10 years.

In principle, the existence of kinks can be tested statistically by fitting a polynomial curve of sufficiently high degree to the bristlecone pine data, and comparing the Residual Sum of Squares with some independent estimate of the variance of the errors $\{e_i\}$ in (6). Clearly, the errors reported by laboratories must be regarded as an underestimate of the real errors, since the reported error relates to counting fluctuations only. Many laboratories suggest that noncounting errors are negligible, but few have published any statistical evidence to support this assertion. It seems unlikely that the question of kinks can be resolved by purely statistical methods until a more careful analysis is made of the measurement errors involved.

2.4. *Calibration of Floating Chronologies*

If the calibration curve F contains kinks, a single radiocarbon date may correspond to more than one calendar date. However, in such a case, this ambiguity regarding the calendar date of an archaeological sample may be removed if a *floating chronology* is available.

Suppose, e.g., we dig up a large log at a particular archaeological site. We may take samples from various parts of the log and determine their radiocarbon dates. In addition, we can determine the *relative* calendar ages of these samples, simply by counting the tree-rings on the log. Consequently, although we know the "spacings" between the samples along the calendar-age axis, we do not know the absolute date of the sequence of samples; the sequence or chronology may be regarded as *floating* along the calendar-age axis.

Since we know the relative ages of the samples in the floating chronology, we may construct a short "calibration curve" for the floating chronology. We may then compare not only the ordinates but also the derivative of this calibration curve with those of the "master" calibration curve derived from the bristlecone pine samples. Thus, by "matching the kinks" in these two curves, we may eliminate or reduce any ambiguity concerning the calendar age of the floating choronology.

This leads to a most interesting statistical problem which can be formulated as follows. Suppose that the master calibration curve is based on data from m bristlecone pine samples, whereas the floating chronology under consideration contains n samples. Let x_i and y_i denote respectively the tree-ring date and radiocarbon date of the ith bristlecone pine sample, $i = 1, 2, \ldots, m$; let x_{1i} denote the tree-ring date (relative to the arbitrary origin) of the ith sample in the floating chronology, and y_{1i} the corresponding radiocarbon date, $i = 1, 2, \ldots, n$. Then we have

$$\begin{aligned} y_i &= F(x_i) + e_i & i &= 1, 2, \ldots, m \\ y_{1i} &= F(A_1 + x_{1i}) + e_{1i} & i &= 1, 2, \ldots, n, \end{aligned} \tag{8}$$

where A_1 denotes the (unknown) calendar age or tree-ring date of the arbitrary origin of the floating chronology. The e's denote random errors of measurement, the x's are assumed known and error-free, and F is the unknown calibration function defined by (4). The problem is to estimate A_1, in the presence of the "infinite-dimensional" nuisance "parameter" F.

This is a generalisation in two ways of the well-known problem of inverse calibration from a fitted regression line. Firstly, F is not necessarily a linear function. Secondly, the classical problem of inverse regression corresponds to the special case where all the x_{1i} are equal. Here, the n points $\{x_{1i}\}$ may be distinct, but there is only one location parameter to estimate, namely A_1.

The use of floating chronologies to increase the precision of calibrated radiocarbon dates was pioneered by Ferguson *et al.* [20], using data from a chronology comprising 311 consecutive rings. These authors used a graphical procedure, using Suess's free-hand curve as an estimate of F in (8). The first statistical approach to this problem [9] assumed that, over the relatively short interval of interest, the calibration function F in (8) could be represented by a straight line. The problem then reduced to that of fitting parallel lines to the two groups of data, the parameter A_1 being proportional to the difference between the intercepts of these two lines.

This approach is not entirely satisfactory, because if F were really linear, only one calendar date would correspond to any given radiocarbon date, and there would be no ambiguity concerning the date of the floating chronology. In any case, the assumption of linearity may not be justified by the data.

More generally, we could assume that F in (8) can be represented by some low-order polynomial whose coefficients are to be estimated from the data. However, as noted in [11], this leads to a nonlinear model with its resulting complications. For example, even if we assume that F is a quadratic, equations (8) become

$$y_i = B_0 + B_1 x_i + B_2 x_i^2 + e_i \qquad i = 1, 2, \ldots, m$$
$$y_{1i} = (B_0 + B_1 A_1 + B_2 A_1^2) + (B_1 + 2A_1 B_2) x_{1i} + B_2 x_{1i}^2 + e_{1i} \quad i = 1, 2, \ldots, n \tag{9}$$

which are nonlinear in the parameters (A_1, B_0, B_1, B_2).

In such cases, the linear/nonlinear structure of the model can be used to construct significance tests for A_1. If A_1 were replaced by any *given* hypothetical value $A_1{}^*$, the above equations (and the equivalent equations for any higher-order polynomial) become linear in the remaining parameters. Standard theory of linear models may then be used to construct a significance test for the hypothesis $A_1 = A_1{}^*$. Since this argument holds for *any* value $A_1{}^*$, confidence limits for A_1 can be found simply as the set of hypothetical values $A_1{}^*$ *not* rejected by the chosen significance test. A modified version of this procedure was used in [11] for the simultaneous calibration of the *two* floating chronologies from Auvernier, Switzerland (see below).

Since the form of F is unknown *a priori*, it may be more appropriate to estimate F by some nonparametric method, such as those proposed by Priestley and Chao [43], Nadaraya [40] and Rosenblatt [50]. These methods make no assumptions regarding the parametric form of F but merely assume that F possesses a certain degree of smoothness. If such an estimator is used, it is possible, in principle, to test any hypothetical value of A_1 as follows. Firstly, we obtain an estimate f of F, independent of A_1, using the bristlecone pine data only. Then for any given hypothetical value $A_1{}^*$ of A_1, we examine the deviations

$$R_i = y_{1i} - f(A_1{}^* + x_{1i}) \qquad i = 1, 2, \ldots, n$$

of the radiocarbon dates of the samples in the floating chronology from the dates predicted by the estimated calibration curve. If the hypothetical value is close to the correct value, the deviations $\{R_i\}$ should be collectively "small," and vice versa. The precise formulation of this idea as a significance test is given by Clark [8], and is developed for a wide class of estimators f.

Suess [59] subsequently used his graphical procedure to calibrate simultaneously *two* floating chronologies from an archaeological site at Auvernier. The data for this analysis may be described by the following model, using an obvious extension of the previous notation.

$$
\begin{aligned}
y_i &= F(x_i) + e_i & i &= 1, 2, \ldots, m, \\
y_{1i} &= F(A_1 + x_{1i}) + B + e_{1i} & i &= 1, 2, \ldots, n_1, \\
y_{2i} &= F(A_2 + x_{2i}) + B + e_{2i} & i &= 1, 2, \ldots, n_2.
\end{aligned} \tag{10}
$$

Here, subscripts 1 and 2 refer to the first and second floating sequence, respectively, and the parameter B denotes a possible systematic difference between radiocarbon dates of bristlecone pine and Auvernier wood of the same calendar age, due to possible *in situ* production of carbon-14 in bristlecone pine. In this particular case, stratigraphic and dendrochronological evidence showed that the second floating chronology was older than the first. In other words, the parameters A_1 and A_2 must satisfy a constraint of the form

$$A_2 \geqslant A_1 + a, \tag{11}$$

where a is a known constant, depending on the choice of arbitrary origin for the two chronologies.

Least-squares solutions of this problem have been obtained [9, 11] by assuming that F could be represented by various low-order polynomials. The parameter B was either estimated from the data or given a hypothetical value derived by geophysical arguments. In the latter paper, the prior information (11) was taken into account by a suitable truncation of the ordinary unconstrained confidence region for (A_1, A_2). The various solutions in these papers showed that the constraint (11) and the value of B were more critical than the assumed degree of the polynomial form for F.

2.5. *Generalisations and Unsolved Problems*

Clearly, the simultaneous calibration of two floating chronologies can be generalised to $k \geqslant 2$ chronologies, the difficulties being technical rather than conceptual.

Alternatively, it would be interesting to adopt a Bayesian approach to the Auvernier problem. It seems plausible that the archaeologist's prior information concerning the dates A_1 and A_2 of the two sequences is made up of two components. Firstly, the difference $(A_2 - A_1)$ in the dates of the chronologies may be estimated from the relative depths at which the relevant material was uncovered at the site. Secondly, the date of, say, the younger sequence may be estimated by examining any artifacts from the site. If these sources of information are independent, we may postulate a prior density $p(A_1, A_2)$ for (A_1, A_2) of the form

$$p(A_1, A_2) = p_1(A_1)\, p_2(A_2 - A_1),$$

where p_1, p_2 are density functions. In the Auvernier case p_2 is positive only on some subset of $[a, \infty)$.

The distinguishing feature of a floating chronology is that the differences $\{(x_{1i} - x_{1j})\}$ in the calendar dates of the samples in the chronology are known *exactly*. More generally, we may know only that certain differences in the dates of our samples are nonnegative. For example, Vogel [62] considers the simultaneous dating of three samples from an archaeological site in which the stratigraphic evidence shows that the first sample is older (lower) than the second, and in turn that the second sample is older than the third.

The problem may be represented by the equations,

$$\begin{aligned} y_i &= F(x_i) + e_i \qquad & i &= 1, 2, \ldots, m, \\ y_{1i} &= F(x_{1i}) + e_{1i} \qquad & i &= 1, 2, 3, \end{aligned}$$

where x_i, y_i and F have the same meanings as in (8), and x_{1i} and y_{1i} denote the calendar age and radiocarbon age respectively of the ith archaeological sample. Again the e's denote random errors of measurement. The problem is to estimate the three parameters x_{11}, x_{12}, x_{13} subject to $x_{11} \geqslant x_{12} \geqslant x_{13}$.

Vogel gives a simple graphical solution using Suess's free-hand curve, but ignoring the fact that this curve is only an estimate of F based on observations containing random error. There is as yet no satisfactory statistical solution to this problem.

The determination of the form of the calibration function F remains an important practical problem. If, as seems likely, this must be determined experimentally rather than theoretically, the magnitude of the errors associated with radiocarbon dates becomes crucial. Radiocarbon laboratories are rarely explicit as to how their reported errors of measurement are computed, and it seems likely [10, 41] that most laboratories underestimate the real error. Therefore there is an urgent need for radiocarbon laboratories to conduct suitably designed experiments just to keep a check on their measurement errors.

3. Seriation

3.1. *Background*

A common problem in archeaology is the seriation, i.e., the reconstruction of the chronological order, of a set of objects using only information defining the degree of similarity between pairs of objects. For example, suppose we have a number of graves containing different varieties of pottery. Our aim is to place these graves in temporal order, using the basic archaeological principle that graves which are close together in temporal order are more likely to have similar contents than graves which are further apart in temporal order. This principle is derived from the assumption that each variety of pottery was in existence during a unique relatively short period of time, and that the "mix" of varieties has changed gradually with time.

Such a problem was first considered in 1899 by Petrie [42], who was confronted with 900 predynastic Egyptian graves containing a total of about 800 varieties of pottery. In the past decade, the availability of high-speed computers has rekindled interest in the general problem of seriation. We need not restrict ourselves to graves and pottery, for essentially the same problem has arisen in other contexts as diverse as philology [6, 12] and epigraphy [53].

3.2. *Mathematical Formulation*

In order to formulate the above problem in mathematical terms, the following definitions [28] will be useful.

(1) An *incidence matrix* A is one whose typical element a_{ij} takes the value 1 if the jth variety of pottery is present in the ith grave, and is otherwise zero.

(2) An incidence matrix is in *Petrie* form if in *every* column there is only a single run, if any, of 1's.

(3) An incidence matrix A is *Petrifiable* if there exists a permutation matrix P such that PA is in Petrie form.

(4) A symmetric matrix is in *Robinson* form if, when going to the left or down from any position on the main diagonal, the elements never increase.

Kendall [26–29] has developed a mathematical theory of seriation from incidence matrices. This is based on the proposition that if each grave contains representatives of every variety of pottery extant at the time of internment, and the rows (graves) of the given incidence matrix A are rearranged in the correct temporal order, then A will be in Petrie form. Of course, the rows of A are initially in some arbitrary order, and so the problem of seriation is simply that of Petrifying A, i.e., finding that permutation of the rows of A which converts A to Petrie form, assuming that this can be done.

There are two practical difficulties in this approach. Firstly, if some varieties are "missing" from some graves, the incidence matrix A will be only approximately in Petrie form, even when the rows are rearranged in the correct order. Secondly, even if no varieties are missing from any graves, an incorrect ordering of the rows may still produce a matrix in Petrie form. In other words, there may be more than one way of Petrifying A.

Despite these difficulties, further examination of the idealised problem has been worthwhile, since it has suggested methods of seriation which, although not exact, are satisfactory for real-world problems.

The *possibility* of Petrifying A may be examined using results in [22]. The main result of this paper may be stated as follows.

THEOREM. *If two incidence matrices A and B have the same number of rows and columns, and if $A'A = B'B = V$, then B can be permuted into Petrie form if and only if A can.*

Thus the question of whether A can be Petrified can, in principle, be answered if we know only the matrix V derived from A, further reference to A being unnecessary. Fulkerson and Gross [22] give a graph-theoretic algorithm for identifying Petrifiable incidence matrices from their V matrices.

Having determined that A can be converted to Petrie form, the following theorem of Kendall [27] is relevant to the next problem: that of finding a particular permutation for doing this.

THEOREM. *If the incidence matrix A is Petrifiable, then the row permutations which Petrify A are exactly those which, when applied to the rows and columns of $G = AA'$ simultaneously, put G into Robinson form. Thus the required permutation may be found either from A or the derived matrix G.*

Both the derived matrices V and G contain information on the chronology which is being reconstructed. The (i, j)th component of V is equal to the number of graves in which the ith and jth varieties are both present. Thus V contains information on the chronology of the varieties, for if the temporal ranges of two varieties overlap, they are likely to appear together in a number of graves roughly proportional to the extent of their overlap.

On the other hand, the (i, j)th component of G is equal to the number of varieties which are present in both graves i and j. Thus G may be regarded as a similarity matrix for the graves, and so by the basic archaeological principle stated previously, G contains information on the chronology of the graves.

The same ideas and results hold good, with only minor modifications [30, 31], if we start with an abundance matrix rather than an incidence matrix. In an abundance matrix, the (i, j)th element is a nonnegative number specifying the frequency of occurrence of the jth variety in the ith grave.

3.3. *Proposed Methods of Solution*

A direct search for the correct row-permutation by examining all possible row permutations of A is clearly impractical, once the number of rows (graves) exceeds even 8 or 9. In any case, the incidence matrix A need not necessarily be of Petrie form even when the rows are rearranged in the correct order, because of "missing" varieties. What is needed is some possibly heuristic approach which will yield an acceptable seriation in the presence of small perturbations from the ideal mathematical model considered above.

Petrie himself used a variety of methods based on what Kendall [26] has called Petrie's *Concentration Principle*. This asserts the following.

In the best arrangement (of rows of the incidence matrix) the temporal range of each separate variety should be minimised.

This idea was formalised in [26] but the solution given there is unsatisfactory, for two reasons. Firstly, the solution still involves the examination of all possible permutations, although various techniques [25] may be used to restrict the search to selected subsets. Secondly, the mathematical model is highly artificial and specific, and recent experiments [32] with this method on real data have given unsatisfactory results.

Kendall's alternative solution [28, 29, 31] treats the derived matrix $G = AA'$ as a similarity matrix for graves, and thereby uses multidimensional scaling to represent each grave by a point in 2-dimensional space. If the graves can be seriated, then these points should lie on a straight line. This linear ordering then identifies the correct permutation to convert G to Robinson form, or equivalently, to convert A to Petrie form.

Experiments [29, 31] with artificial data show that the correct order of the "graves" can be recovered by this method, but with the surprising side-effect that the points representing the graves are plotted along a horseshoe curve. As a further demonstration of the power of this method, a map of Romania was reconstructed [29] knowing only the *rankings* of pairs of towns in terms of their distance apart. In other words, one knows only which pair of towns is the closest, which pair is the next closest, and so on, down to the most remote pair, but the actual distances between towns are unknown. The resulting map, reconstructed from this derived "similarity matrix" for towns, was remarkably accurate.

This method was applied [29] to archaeological data relating to 59 graves, from the La Tène Cemetery at Münsingen-Rain, which contain a total of 70 varieties of pottery. The computer's seriation was in good agreement with previous results [24] and with the geographical ordering of the graves along the major axis of the cemetery.

Gelfand [23] gives two well-defined algebraic algorithms for transforming, as nearly as possible, any given similarity matrix (not necessarily Kendall's G) into Robinson form. Further, he proves that if the given similarity matrix can be transformed into Robinson form exactly, then both of his algorithms provide the necessary transformation. The results obtained by using these methods on actual data are in good agreement with seriations derived previously [48, 49] by independent methods.

In a paper generalising some of Kendall's results, Wilkinson [64] notes a close connection between the problem of Petrifying incidence matrices and the familiar travelling salesman problem. In principle, the row-permutation needed to Petrify a given incidence matrix can be determined from the solution of an appropriate travelling salesman problem. However, this idea has apparently not yet been used in practice.

Previously, Robinson [49] had considered the slightly different problem of seriating a number of archaeological deposits, given only the percentages of various types of pottery in each deposit. Since some of these deposits corresponded to different levels of the same archaeological site, this stratigraphic evidence supplied partial prior information on the temporal order of the deposits. Robinson constructed a similarity matrix for deposits using as similarity coefficient for deposits i and j the quantity

$$S_{ij} = 200 - \sum_{k=1}^{v} | p_{ik} - p_{jk} |,$$

where p_{ik} and p_{jk} denote respectively the percentage of pottery at deposits i and j which is of the kth type, $k = 1, 2, \ldots, v$. He then gave a heuristic two-stage procedure for re-arranging the rows and columns of this similarity matrix so that it approximates the Robinson form subject to the restriction that the known stratigraphic relationships be preserved.

Since this procedure still involved the examination of a large number of possible permutations, several authors developed alternative graphical procedures for this type of seriation problem, using either the similarity coefficients [48] or the percentage of pottery directly [2, 36]. In general, these extremely rapid methods have produced results in close argeement with those obtained by the more elaborate methods.

Finally, although the context is philological and not archaeological, the chronological seriation [3, 12] of 7 works of Plato, based on the frequency distributions of sentence endings, is worthy of mention. Apart from [26], these papers are the only ones on seriation in which a specific probability model is used. Assuming the *Republic* to be the earliest of the seven works and the *Laws* to be the latest, Cox and Brandwood [12] used a simple parametric model to estimate the linear order of the remaining five works. The subsequent modification [3] of their method, using a less restrictive probability model, nevertheless produced the same estimated order for the 7 works. Recently, Boneva [6] produced a seriation of all 45 books of Plato based on the frequency distributions of sentence endings, using Kendall's "horseshoe" program [31] and, unlike [12], treating all 10 books of the *Republic* and all 12 books of the *Laws* as separate books. This seriation was in broad agreement with the earlier conclusions of both statisticians and philologists.

3.4. *Unsolved Problems.*

Most of the seriation methods described in the preceding section are mathematical rather than statistical, in the sense that no explicit probability model is used. Consequently, there has been little or no examination of the statistical performance of these methods in terms of sampling fluctuations.

For example, certain varieties of pottery may be "missing" from certain graves at random, due to either not being deposited in the grave in the first place or not being discovered upon excavation. Alternatively, there may be random errors in the classification of pottery after excavation. In either case, the graves-versus-varieties incidence matrix, and hence the final seriation, will be subject to random errors. In principle, the final seriation, will have a probability distribution related to the error structure of the incidence matrix. However, most of the preceding methods ignore the possibility of random errors in the incidence matrix.

Other questions, which are at present largely unresolved, are as follows.

(1) What do we mean by a good seriation? Equivalently, on what basis should we compare one possible seriation with another from the same data? Some tentative suggestions may be found in [23, 31].

(2) How can we test the assumption that seriation is possible?

(3) Most methods of seriation assume that each variety was present during a single time-interval. What happens if some varieties were in use in several disjoint time-intervals?

(4) To what extent are the graphical procedures unambiguous? If different people apply these procedures to the same data, will they necessarily obtain the same answers?

(5) If seriation is possible, do the graphical procedures always produce the correct answer?

Experience has shown that, provided the objects (graves) to be seriated can be placed in an approximate serial order, the different methods described above will produce seriations which are generally in sufficiently close agreement for practical purposes.

Statistical methods have been used in archaeology in contexts other than dating. However, even in the restricted field of archaeological dating surveyed in this paper, many of the statistical problems are nontrivial and there are many interesting unsolved problems.

REFERENCES

[1] AITKEN, M. J. (1970). Physics applied to archaeology. I. Dating. *Reports on Progress in Physics* **33** 941–1000.

[2] ASCHER, M. (1959). A mathematical rationale for graphical seriation. *American Antiquity* **25** 212–214.

[3] ATKINSON, A. C. (1970). A method for discriminating between models. *J. Roy. Statist. Soc. Ser. B* **32** 323–353.

[4] BAXTER, M. S. AND WALTON, A. (1971). Fluctuations of atmospheric carbon-14 during the past century. *Proc. Roy. Soc. Ser. A* **321** 105–127.

[5] BERGER, R. (1970). The potential and limitations of radio-carbon dating in the Middle Ages: a radiochronologist's view. In *Scientific Methods in Medieval Archaeology* (R. Berger, Ed.), pp. 89–139. Univ. Calif. Press, Berkeley.

[6] BONEVA, L. I. (1971). A new approach to a problem of chronological seriation associated with the works of Plato. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, P. Tautu, Eds.), pp. 173–185. Univ. Press, Edinburgh.

[7] BUCHA, V. (1970). Influence of the earth's magnetic field on radiocarbon dating. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 501–512. Wiley, New York.

[8] CLARK, R. M. (1973). Nonparametric function-fitting. Ph.D. Thesis, University of Sheffield, U.K.

[9] CLARK, R. M. AND RENFREW, C. (1972). A statistical approach to the calibration of floating tree-ring chronologies using radiocarbon dates. *Archaeometry* **14** 5–19.

[10] CLARK, R. M. AND RENFREW, C. (1973). Tree-ring calibration of radiocarbon dates and the chronology of ancient Egypt. *Nature* **243** 266–270.

[11] CLARK, R. M. AND SOWRAY, A. (1973). Further statistical methods for the calibration of floating tree-ring chronologies. *Archaeometry* **15** 255–266.

[12] COX, D. R. AND BRANDWOOD, L. (1959). On a discriminatory problem connected with the works of Plato. *J. Roy. Statist. Soc. Ser. B* **21** 195–200.

[13] DAMON, P. E. (1970). Climatic versus magnetic perturbation of the atmospheric C-14 reservoir. In *Radiocarbon Variations and Absolute Chronology* (I.U. Olsson, Ed.), pp. 571–594. Wiley, New York.

[14] DAMON, P. E., LONG, A. AND GREY, D. C. (1970). Arizona radiocarbon dates for dendrochronologically dated samples. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 615–618. Wiley, New York.

[15] DAMON, P. E., LONG, A. AND WALLICK, E. I. (1972). Dendrochronological calibration of the carbon-14 time scale. *Proceedings of 8th International Conference on Radiocarbon Dating, Lower Hutt City* **1** 45–59.

[16] DE VRIES, H. (1958). Variations in concentration of radiocarbon with time and location on earth. *Nederl. Akad. Wetensch. Proc. Ser. B* **61** 94–102.

[17] FERGUSON, C. W. (1968). Bristlecone pine, science and esthetics. *Science* **159** 839–846.

[18] FERGUSON, C. W. (1970). Dendrochronology of bristlecone pine, *Pinus aristata*. Establishment of a 7484-year chronology in the White Mountains of eastern-central California, U.S.A. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 237–259. Wiley, New York.

[19] FERGUSON, C. W. (1970). Concepts and techniques in dendrochronology. In *Scientific Methods in Medieval Archaeology*. (R. Berger, Ed.), pp. 183–200. Univ. Calif. Press, Berkeley.

[20] Ferguson, C. W., Huber, B. and Suess, H. E. (1966). Determination of the age of Swiss lake dwellings as an example of dendrochronologically-calibrated radiocarbon dating. *Z. Naturforsch.* **21A** 1173–1177.

[21] Fritts, H. C. (1972). Tree rings and climate. *Scientific American* **266** No. 5, 92–100.

[22] Fulkerson, D. R. and Gross, O. A. (1965). Incidence matrices and interval graphs. *Pacific J. Math.* **15** 835–855.

[23] Gelfand, A. E. (1971). Rapid seriation methods with archaeological applications. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.), pp. 186–201. Univ. Press, Edinburgh.

[24] Hodson, F. R. (1968). *The La Tène Cemetery at Münsingen-Rain.* Stämpfli, Berne.

[25] Hole, F. and Shaw, M. (1967). Computer analysis of chronological seriation. *Rice University Studies* **53** 1–166.

[26] Kendall, D. G. (1963). A statistical approach to Flinders Petrie's sequence-dating. *Bull. Inst. Internat. Statist.* **40** 657–680.

[27] Kendall, D. G. (1969). Incidence matrices, interval graphs and seriation in archaeology. *Pacific J. Math.* **28** 565–570.

[28] Kendall, D. G. (1969). Some problems and methods in statistical archaeology. *World Archaeology* **1** 68–76.

[29] Kendall, D. G. (1971). A mathematical approach to seriation. *Philos. Trans. Roy. Soc. London Ser. A* **269** 125–135.

[30] Kendall, D. G. (1971). Abundance matrices and seriation in archaeology. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **17** 104–112.

[31] Kendall, D. G. (1971). Seriation from abundance matrices. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.), pp. 215–252. Univ. Press, Edinburgh.

[32] Kivu-Sculy, I. (1971). On the Hole-Shaw method of permutation search. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.), pp. 253–254. Univ. Press, Edinburgh.

[33] Lerman, J. C., Mook, W. G. and Vogel, J. C. (1970). C-14 in tree rings from different localities. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 275–301. Wiley, New York.

[34] Libby, W. F. (1970). Radiocarbon dating. *Philos. Trans. Roy. Soc. London Ser. A* **269** 1–10.

[35] McKerrell, H. (1972). On the origins of British faience beads and some aspects of the Wessex-Mycenae relationship. *Proc. Prehistoric Soc.* **38** 286–301.

[36] Meighan, C. W. (1959). A new method for the seriation of archaeological collections. *American Antiquity* **25** 203–211.

[37] Michael, H. N. and Ralph, E. K. (1970). Correction factors applied to Egyptian radiocarbon dates from the era before Christ. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 109–120. Wiley, New York.

[38] Michael, H. N. and Ralph, E. K. (1971). *Dating Techniques for the Archaeologist.* M.I.T. Press, Cambridge, Mass.

[39] Michael, H. N. and Ralph, E. K. (1972). Discussion of radiocarbon dates obtained from precisely-dated Sequoia and bristlecone pine samples. *Proceedings of 8th International Conference on Radiocarbon Dating, Lower Hutt City* **1** 28–43.

[40] Nadaraya, E. A. (1964). On estimating regression. *Theor. Probability Appl.* **9** 141–142.

[41] Neustupný, E. (1970). The accuracy of radiocarbon dating. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 23–34. Wiley, New York.

[42] PETRIE, W. M. F. (1899). Sequences in prehistoric remains. *J. Anthropol. Inst.* **29** 295–301.

[43] PRIESTLEY, M. B. AND CHAO, M. T. (1972). Non-parametric function fitting. *J. Roy. Statist. Soc. Ser. B* **34** 385–392.

[44] RENFREW, C. (1970). The tree-ring calibration of radiocarbon: an archaeological evaluation. *Proc. Prehistoric Soc.* **36** 280–311.

[45] RENFREW, C. (1971). Carbon-14 and the prehistory of Europe. *Scientific American* **225**, No. 4 63–72.

[46] RENFREW, C. (1973). *Before Civilisation: The Radiocarbon Revolution and Prehistoric Europe.* Jonathan Cape, London.

[47] RENFREW, C. AND CLARK, R. M. (1974). Problems of the radiocarbon calendar and its calibration. *Archaeometry* **16** 5–18.

[48] RENFREW, C. AND STERUD, G. (1969). Close-proximity analysis: a rapid method for the ordering of archaeological materials. *American Antiquity* **34** 265–277.

[49] ROBINSON, W. S. (1951). A method for chronologically ordering archaeological deposits. *American Antiquity* **16** 293–301.

[50] ROSENBLATT, M. (1969). Conditional probability density and regression estimators. In *Multivariate Analysis—II* (P. R. Krishnaiah, Ed.), pp. 25–31. Academic Press, New York.

[51] SÄVE-SÖDERBERGH, T. AND OLSSON, I. U. (1970). C-14 dating and Egyptian chronology. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 35–56. Wiley, New York.

[52] SCHULMAN, E. (1958). Bristlecone pine, oldest known living thing. *National Geographic Magazine* **113** 354–372.

[53] STEFAN, A. (1971). Application of mathematical methods to epigraphy. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.), pp. 267–275. Univ. Press, Edinburgh.

[54] STOKES, M. A. AND SMILEY, T. L. (1968). *An Introduction to Tree-Ring Dating.* Chicago Univ. Press, Chicago.

[55] STUIVER, M. (1970). Long-term C-14 variations. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 197–213. Wiley, New York.

[56] SUESS, H. E. (1965). Secular variations of the cosmic-ray produced carbon-14 in the atmosphere and their interpretations. *J. Geophys. Res.* **70** 5937–5952.

[57] SUESS, H. E. (1967). Bristlecone pine calibration of the radiocarbon time scale from 4100 BC to 1500 BC. In *Proceedings of Symposium on Radioactive Dating and Methods of Low-Level Counting, Monaco, March* 1967, pp. 143–151. I.A.E.A., Vienna.

[58] SUESS, H. E. (1970). Bristlecone pine calibration of the radiocarbon time scale from 5400 BC to the present. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 303–312. Wiley, New York.

[59] SUESS, H. E. (1970). The neolithic of Auvernier, Switzerland. *Antiquity* **44** 91–99.

[60] SWITSUR, R. V. (1973). The radiocarbon calendar recalibrated. *Antiquity* **47** 131–137.

[61] TAUBER, H. (1970). The Scandinavian varve chronology and C14 dating. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 173–196. Wiley, New York.

[62] VOGEL, J. C. (1970). Discussion on historical chronologies. In *Radiocarbon Variations and Absolute Chronology* (I. U. Olsson, Ed.), pp. 125. Wiley, New York.

[63] WENDLAND, W. M. AND DONLEY, D. L. (1971). Radiocarbon-calendar age relationship. *Earth Plan. Sci. Letters* **11** 135–139.

[64] WILKINSON, E. M. (1971). Archaeological seriation and the travelling salesman problem. In *Mathematics in the Archaeological and Historical Sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.), pp. 276–283. Wiley, New York.
[65] WILLIS, E. H., TAUBER, H. AND MÜNNICH, K. O. (1960). Variations in the atmospheric radiocarbon concentration over the past 1300 years. *Radiocarbon* **5** 1–4.