11

Numerical Taxonomy in Fields Other than Biological Systematics

In this chapter we give an account of the application of taxometric and closely related methods to fields outside systematic biology. We have been guided by two considerations: to show the range of subjects in which such methods are useful, and to describe the special problems that arise when they are applied to these other fields.

There is some overlap in the subject matter of the various sections, and the reader may find it useful to consult several of these for borderline topics.

11.1 ECOLOGY AND BIOGEOGRAPHY

In the last few years, there has been a swift growth of taxometric methods in ecology. Notable early work was that of Sørensen (1948), who employed both coefficients of association and cluster analysis. The modern phase can be said to have been initiated with the development of the monothetic method of Goodall (1953). In recent years W. T. Williams and his colleagues have contributed significantly to this field and there has been a considerable interchange of ideas with taxonomy. Reviews and critical discussion are given by Grieg-Smith (1964) and Gimingham (1969) and Whittaker (1972). These methods have had a powerful impact on descriptive and community ecology, especially in areas where ecology and systematics overlap (see Sections 7.4 and 7.5). Numerical taxonomy is widely used in plant ecology for vegetation studies, in which "taxa" such as woodland, prairie, or moorland can be distinguished and analyzed.

These types of vegetation are sometimes suitably represented by hierarchies analogous to those in systematics. For OTU's the basic data matrix has stands or quadrats; the species found in these units are used as characters. Those species are recorded as present or absent, or by some measure of their abundance. As well as grouping together areas (stands, quadrats) similar in their species composition, one can also group those species that are most usually associated, and Williams and Lambert (1961a) refer to these, respectively, as "normal" and "inverse" analyses, which are, of course, the Q and R analyses of the basic data matrix identified by terms less likely to cause confusion. Williams and Lambert also suggest simultaneous normal and inverse analysis, with the object of producing "noda," that is, joint clusters of both quadrats and species (Williams and Lambert, 1961b; Lambert and Williams, 1962; see also Webb et al., 1967a; Moore, Benninghoff, and Dwyer, 1967).

Because of the continuous nature of the variation in most vegetation, ordination methods are much used. The earlier ordination methods are now being superseded by various types of factor analysis (e.g., Austin and Orloci, 1966; Ivimey-Cook and Proctor, 1967; and Whittaker, 1972), but multidimensional scaling might be worth trial, particularly for such difficult studies as those on cyclic changes in vegetation (e.g., Anderson, 1967). Grieg-Smith, Austin, and Whitmore (1967) found some support for the view that cluster analysis would be most appropriate for studying major vegetation types, because it is here that sharp discontinuities are most likely.

Monothetic clustering methods have been used a good deal in ecology, particularly the association analysis of Williams and Lambert (1959). In this monothetic method a quadrat may be occasionally far removed from quadrats that are highly similar in overall floristic content, because of the chance presence or absence of the species that served to define the monothetic group. Such misplacements are evidently not uncommon (Lange, Stenhouse, and Offler, 1965; Lambert and Williams, 1966; Grieg-Smith, Austin, and Whitmore, 1967), both for normal and inverse analysis. Polythetic methods are better if detailed ecological relationships are to be studied. Nevertheless, because of their speed and ability to handle large numbers of OTU's, monothetic methods may be convenient for preliminary grouping (e.g., Crawford and Wishart, 1967; Crawford, Wishart, and Campbell, 1970).

A number of comparisons have been made in ecology between different polythetic methods, particularly between the usual clustering methods and those employing information statistics (Lambert and Williams, 1966; Williams, Lambert, and Lance, 1966; Hall, 1967b; Webb et al., 1967a; Orloci, 1968a,c, 1969a; Kikkawa, 1968). Information statistics appear promising for ecology, though they seem best suited to nonprobabilistic studies on unique sets of OTU's (Hall, 1967b). Webb et al. (1967a) found association analysis and information analysis superior to factor analytic methods in isolating the kinds of association that ecologists recognize. Williams and Lambert (1966) reported that centroid clustering with their "nonmetric coefficient" was almost as good as information analysis, but in ecology taxonomic distances were unsatisfactory, and standardization of 0,1 data was disadvantageous. Criteria of optimality present difficulties in ecological classification as in biological systematics, and "informed opinion" has most often been the test of success. However, independent criteria are sometimes available from the environment (e.g., clusters may agree closely with soil type or rainfall); thus Williams and Lambert (1960) found agreement between agricultural drainage history and subtle differences of vegetation type, and Sundman (1970) showed congruence of bacterial populations and soil type.

A considerable body of new knowledge is coming from the numerical studies of Williams and his colleagues on tropical rain forests (Webb et al., 1967a,b; Williams and Webb, 1968; Williams et al., 1969a,b; Webb et al., 1970), much of which has wide implications for ecology. This includes investigations on the extent to which certain classes of plants reflect the total floristic composition, the relation between floristic composition and microclimate, the detection of small scale vegetational pattern, and the rate and direction (in ecological hyperspace) of plant successions during the regeneration of forest.

Some special problems of ecological classifications may be noted. The number of characters may be unavoidably small; thus, there may be only ten species of flowering plants in the area studied. Williams, Dale, and McNaughton-Smith (1963) have suggested a method of effectively weighting characters in proportion to their average correlation with all other characters if the number of characters is small. The significance of the absence of a plant may be uncertain: the plant may not be able to grow, or it may not have had the opportunity to colonize the area studied, or it was by chance not in the sampled quadrat. Therefore the association coefficients employed in ecology commonly exclude negative matches. Mountford (1962) considers coefficients that are independent of quadrat size (see Section 4.4). Mixed qualitative and quantitative data are discussed by Williams and Dale (1962, 1965) and Williams and Lance (1965). Goodall (1969) proposes a method for testing the significance of rare floristic associations. A list of similarity coefficients used in ecology is given by Peters (1968). Trend surface analysis has been employed by Gittins (1968) and Fisher (1968).

Most of the papers cited above have emphasized methodology. Some descriptive studies are those of West (1966) on clusters of vegetation types in forests, Schmid (1965) on underwater vegetation, and Gyllenberg (1965a), Gyllenberg and Rauramaa (1966), and Sundman and Carlberg (1967) in bacterial ecology. Rasnitsyn (1965) used Smirnov's similarity coefficient in describing the ecology of aquatic larvae. Hurlbutt (1964, 1968), by phenetic studies of mites, found some evidence for

Gause's law that species with identical ecological requirements do not occupy the same spatial niche. Ebeling et al. (1970) were unable to find support for this law in a marine habitat, but consider that they may not have had suitable material. There is considerable promise in using taxometric methods to cluster similar environments and correlate these with phenetic patterns in populations of organisms. The whole area of species diversity versus environmental diversity in which so much recent work has been done (see our discussion in Sections 7.4 and 10.3) is based on conventional interpretation of species. It is here that a quantitative phenetic interpretation of diversity could be of great heuristic value.

Studies on a larger geographic scale include those of Hagmeier and Stults (1964), Hagmeier (1966), and Kikkawa and Pearse (1969), in which zoogeographic provinces were constructed by cluster analysis. Holloway and Jardine (1968) and Holloway (1969) also distinguished faunal elements with different distribution patterns. Sakai (1971) obtained faunal regions of Dermaptera from cluster analysis. Fisher (1968), studying amphibia, reptiles, and mammals, noted good correspondence between environment (rainfall, vegetation) and trends in factors summarizing the occurrence of these species. Proctor (1967) made a principal component analysis of the liverwort flora of the British Isles. The major division was into highlands and lowlands, but the richness of the flora and oceanic influence were also reflected in the factors. Heatwole and MacKenzie (1967) noted that insular faunal similarity decreased evenly with increasing geographic distance. Sneath (1967c) applied principal component analysis to the world distribution of conifers in relation to evidence of continental drift. Holloway and Jardine (1968) applied multidimensional scaling to relate faunal distributions to geography in the Indo-Australian area, and clearly showed the different boundaries of zoogeographic regions based on different taxonomic groups (birds, butterflies, and bats). Indications of past routes of spread were also obtained. Tobler, Mielke, and Detwyler (1970) have carried this type of work further by partitioning the variation into factors representing latitude, longitude, and mobility of organisms.

Taxometric methods are also used in hydrobiology (e.g., Williamson, 1961; Fager and McGowan, 1963; Mello and Buzas, 1968) to characterize ecological associations of plankton. Neushal (1967) analyzed the effect of depth of water on the associations and Howarth and Murray (1969) found species clusters that were strongly related to environmental variables, such as salinity and ionic composition. A point raised by the work of Roback, Cairns, and Kaesler (1969) is whether multivariate methods are more sensitive indicators of pollution than the usual techniques based on a few indicator species. Valentine (1966) observed excellent concordance between geographic latitude and faunal similarity on the west coast of North America; he found that the faunal provinces given by the dendrogram were more informative than the conventional ones.

Similar ecological work has been done on fossil communities (for reviews see Buzas, 1970). Thus Johnson (1962) used linkage diagrams for clustering species of fossils in different localities and also demonstrated satisfactory results upon modern marine habitats. Kaesler (1966) made both normal and inverse studies on ostracods and foraminifers employing S_{SM}, S_J, and average linkage clustering; he found substantial agreement with earlier groups based on intuitive analyses. Maddocks (1966) and Rucker (1967) have made similar studies. Valentine and Peddicord (1967) used cluster analysis on fossil mollusk assemblages; they noted subsidiary resemblances in the similarity matrix between localities that were not evident in the dendrogram. These subsidiary resemblances are similar to ones noted by Rayner (1966) in his study on soil, though their significance is not yet clear. Park (1968) and others have studied fossil assemblages by ordination and interpreted main vectors as salinity, depth of water, and rate of sedimentation.

Soil classification has received increased interest lately. Leeper (1954) criticized the use of postulated origins of the soil rather than their observed properties, and this trend toward phenetic classifications is reviewed by Mulcahey and Humphries (1967), Webster (1968), and Rayner (1969).

Soils are similar to vegetation in that it is difficult to know the relative merits of ordination and clustering (see Russell and Moore, 1967; Cuanalo and Webster, 1970). But in addition there are problems of homology. It is not easy to decide which soil horizons are comparable when the resemblance between soil profiles is to be estimated. There is an arbitrary element in defining the conventional horizons. Rayner (1966, 1969) has shown that there is in general higher resemblance between the same conventional horizons than between different ones, but exceptions are not unusual. The best results (as judged against the traditional soil classifications) were obtained by first homologizing the horizons on the basis of their greatest resemblances from one soil to another and computing an average resemblance of the homologies thus determined. Quite acceptable results were obtained by employing only one horizon from each of a pair of soils, choosing this pair of horizons on the basis of the greatest degree of homology. Ordination of separate horizons was particularly instructive. Similar work has been done by Moore and Russell (1966), who examined in more detail the degree of homology between samples from different depths of a single soil profile. They noted difficulty in finding sharp distinctions between horizons, and in one case the analysis indicated subzones that were not evident before. Russell and Moore (1968) suggest a method of summing resemblances between set depths in two soil profiles to obtain an overall resemblance between the profiles. Lance and Williams (1967c) make a similar suggestion to average the character values over different depths.

Bidwell and his colleagues (Bidwell and Hole, 1964a,b; Cipra, Bidwell, and Rohlf, 1970) have shown that numerical taxonomy of soils is quite workable. The polythetic method placed together some soils that were very similar overall, but which were in separate "Great Groups" based monothetically on color. Moore and Russell (1967) note that conventional soil classifications appear to contain a substantial "general size" factor. Earlier work is that of Hughes and Lindley (1955) and

Hole and Hironaka (1960). Sarkar, Bidwell, and Marcus (1966) have investigated the minimum number of characters required to give satisfactory classifications, but their method of retaining only uncorrelated characters has serious drawbacks (Sneath, 1967b), and factor analysis sensu lato would seem preferable if the effect of correlation is to be excluded. Grigal and Arneman (1970) have compared soils and their associated vegetation by numerical methods. Webster and Wong (1969) report that soil boundaries in soil survey maps are more readily found from changes in the first vector of factor analyses than from changes in any single property.

11.2 MEDICINE

A main emphasis in taxometric work in medicine has been computer diagnosis of disease. It is thus aimed at problems of identification, and an extensive literature exists on this (for a review see Anderson and Boyle, 1968). But before discussing this it is necessary to consider the earlier stages in the classificatory process. For diagnosis one must assume prior knowledge of the disease entities and the validity of these is of critical importance. If disease taxa are unsound, the diagnostic schemes will inevitably be unsatisfactory. It is now being realized that the classification of diseases (nosology) warrants more attention than it has received; that also, like any taxonomic process, it can be carried out by taxometric methods (Sokal, 1964b; Sneath, 1965, 1966a; Baron and Fraser, 1965, 1968). Disease entities—clusters of similar individual cases of disease—can be made by Q studies, and syndromes (clusters of signs and symptoms) by R studies, although the distinction between Q and R clusters is not always made clear. Indeed diseases are in this respect analogous to the noda of ecology (see the previous section).

It is, of course, very probable that most well-known diseases are valid groupings in some sense, but there has been little study on this point; it is doubtful whether most of the obscurer diseases have been satisfactorily classified. It has been conventional to define a disease by its cause (etiology). Yet, despite the great influence of etiology on treatment, it may not always be suitable for classification. The argument here harks back to the phenetic-phyletic controversy in biological taxonomy. Some clinical entities may have a varied etiology: for example, meningitis caused by very different bacteria can be clinically indistinguishable; conversely, one etiology may produce varied signs, symptoms, and pathology, for example, syphilis. It is by no means clear how one should define health or what attributes should be used to classify diseases. It is true that there is an external criterion against which to test many disease classifications—all cases belonging to the same disease should respond to an appropriate remedy. But such a concept can permit disease groupings that may be unnatural in other respects. There is, in fact, no unifying concept of disease (see Engle, 1963; Engle and Davis, 1963; Scadding, 1963). Diseases may be defined on etiological, clinical, or anatomical bases, and many diseases are polythetic groupings.

The primary data are the individual patient and his signs and symptoms: case histories that are initially used to construct disease classifications and are the logical OTU's. Engle (1963) notes the disturbing tendency to ignore atypical cases, and the repeated redefinition of a disease monothetically according to the latest diagnostic test causes much confusion. Atypical cases point to another important reason for proper classifications: unrecognized subgroups of diseases might be found that respond to different treatments, thereby improving the overall success of treatment.

Hayhoe, Quaglino, and Doll (1964) have studied cases of acute leukemia numerically. The classification of these has been much disputed. The authors found four groups with some overlap and nearly all the cases fitted well into these. One group had distinctive clinical features. Although the authors weighted the characters in proportion to their rarity, a reexamination without such weighting showed almost identical results apart from a scaling factor (Sneath, 1965). A similar study by Leech (1963) on milk fever in cattle showed one homogeneous cluster; Zinsser (1964) found four rather indistinct subgroups in pyelonephritis patients. Manning and Watson (1966) classified heart disease using taxonomic distance and average linkage clustering; they found three main phenons that were in reasonable agreement with clinical assessments. Goldstein and Mackay (1967) were able to distinguish two main clusters in lupoid hepatitis reflecting different pathologies. Jones et al. (1970) were able to separate, by cluster analysis, most cases of two colonic diseases with a very variable and overlapping symptomatology. Zinsser (1964) has used both factor and cluster analysis on cases of pyelonephritis.

Classification in psychiatry has many of the attributes of classification in medicine. Although some of the data employed in psychiatric classification are the same as or similar to those in psychology, the aims of psychiatry, like medicine, are etiology, therapy, and prognosis. Because psychiatry is related to both medicine and psychology, psychiatric classification encounters the taxonomic problems of both disciplines, including difficulties in validation, reliability of data, and selection of variables, which are discussed in some detail in Section 11.3, dealing with classification in psychology. Classification in psychiatry is often unreliable. Sandifer, Green, and Carr-Harris (1966) investigated the reliability of diagnosis in mental disorders by tests on case histories that were examined by different clinicians on several occasions. The cases were also classified to give a phenogram that showed six main groups of mental disorder (such as depression or schizophrenia). They found that in diagnosis the clinicians nearly always agreed as to the main group, but within these groups they very frequently disagreed about the variety or subgroup. In addition, the same clinician frequently placed a case in a different subgroup on a second occasion. This paper is an interesting contrast to that of Overall (1963), who used discriminant analysis, in which much sharper groupings were assumed.

Because of the unreliability and limited validity of present day psychiatric classification (see also Bannister, 1968), there has recently been renewed interest in evaluating current diagnostic systems and developing new ones. Numerical taxonomy has been used to evaluate depressive syndromes (Pilowski, Levine, and Boulton, 1969; Paykel, 1971), borderline psychotic states (Grinker, Werble, and Drye, 1968), the general area of psychosis (Lorr, 1966), and the severe psychiatric disorders in general (Strauss et al., 1972). In some instances classical diagnostic concepts have been upheld by these methods; in other instances new classifications have been suggested (as by Lorr). The increasing attention to more careful data collection in psychiatry and clarification of the properties of clustering algorithms as they relate to the underlying structure of the input data promise more fruitful classification systems in this field.

There are, however, a number of special problems in disease classification, as Jacquez (1964b) points out. Manifestations of a disease change during its course (a point recently taken up by Thompson and Woodbury, 1970, but one we believe needs much deeper investigation) and may become increasingly variable and difficult to summarize by classification. The normal grades into the abnormal. Again we have the parallel with biological taxonomy: the disease must be considered through its entire ontogeny (as in Hennig's "holomorph"). There are also problems in making classifications based on signs and symptoms agree with those based on etiology and other clinical evidence. This point has been emphasized in particular by Baron and Fraser (1968) and Fraser and Baron (1968) in studies of liver disease. These authors also found considerable chaining during clustering; so did Wishart (1969d) in a study of thyroid disease. For these reasons cluster analysis may be less suitable than ordination methods, though these too may present difficulties (e.g., Freer and Adkins, 1968, studying dental malocclusion, found abnormals difficult to differentiate from normals). Capon and Jellett (1968) have pointed out, in a study of polycytemia, that the effect of different coding schemes can be quite large, particularly on clusters derived from correlation coefficients. This is of more than commonplace importance in medicine, where appropriate coding and scaling is not always obvious. Capon and Jellett found that taxonomic distances gave clusters in better accord with clinical diagnoses than did correlation coefficients, so their work raises the question of what part size factors (usually representing clinical severity) should play in taxometric work in medicine. Another area requiring investigation is the observational error of clinical findings and its effect on resemblance values; this introduces problems similar to those discussed in Section 4.10.

Current views on diagnosis are well summarized in the volume edited by Jacquez (1964a). As in other fields both Bayesian and discriminant function methods are popular, varying from simple mechanical devices (for example, Nash, 1960) to those that are on-line to a computer. In the more thorough studies there is now evidence that when a straightforward question is posed (such as the chance of survival or the choice between two diagnoses), and adequate data are furnished, a good statistical method can be even more accurate than the clinical specialist. Thus, Warner et al. (1961) were able by computer to discriminate at least as well as experienced clinicians between congenital heart cases, and the discriminant function of Hughes et al. (1963) was better at prognosis than the best clinicians. Discriminant methods are now being actively applied in electrocardiography (Caceres, 1963; Klingeman and Pipberger, 1967).

Successful diagnostic techniques would seem to need provision for (1) deciding on what signs, symptoms, and diseases are pertinent (and conversely, what is considered "normal" or "healthy"); (2) the satisfactory construction of disease entities or the corroboration thereof; (3) the weighting of attributes for diagnosis and discrimination; (4) the questioning of the physician to confirm signs he may have misread or overlooked; and (5) the correction of the disease entities by new data. There will be many opportunities for fruitful collaboration between those working on medical diagnosis and those working on taxonomic identification methods; Card (1967) presents a good discussion of this. Many of the problems may be beyond our present capabilities, but attempts to solve them may yield important byproducts. Thus it is already clear that better clinical data are needed, which may lead to improved medical textbooks and methods for earlier diagnosis.

In conclusion, some miscellaneous studies may be briefly mentioned. Gyllenberg et al. (1963b) applied numerical taxonomy to food hygiene in delimiting bacterial flora associated with good and bad keeping qualities of milk. Sneath (1966d) studied the relationship between chemical structure and biological activity in some peptides of pharmacological interest, the results suggesting that one might make some predictions on activity from the chemical structure. Izzo and Coles (1962) described a polythetic optical recognition method for identifying abnormal blood cells, and recent taxometric applications to cytopathology are discussed in some detail by Bartels, Bahr, and Wied (1968) and Wied, Bahr, and Bartels (1970). Some numerical taxonomic work was done by Ornstein (1965) on serum proteins, and by James (1964) on nutrition. Selwood (1969) used the expected similarity value of leucocyte antisera on cell panels to identify allelic antigens in tissue typing.

11.3 THE SOCIAL SCIENCES

Methods similar to those in numerical taxonomy have been used for many years in the social sciences. Factor analysis and clustering methods were pioneered in psychometrics, and in social anthropology some farsighted work was done by Boas at the end of the last century and by Kroeber in the first half of this century (reviewed by Driver, 1962). Early quantitative classifications in political science were carried out by Rice and Beyle (see Grumm, 1965). General reviews of taxometric methods include those of Driver (1965) on social and physical anthropology, Cattell (1966a) in psychology, and Cowgill (1967) and Clarke (1968) in archaeology. Alker's (1969) review in social science is particularly useful because of its wide range.

Classification work in psychology has been dominated by factor analytic methods, and there is an extensive literature on this. Much less has been done on cluster analysis (this is reviewed by Cattell and Coulter, 1966, and Bromley, 1966). This emphasis on ordination rather than class making has the empirical justification that psychological traits show continuous distributions rather than discrete ones. Cluster analyses have nevertheless proved interesting in the study of emotions and verbal concepts (Stringer, 1967; Miller, 1969).

A serious problem, one constantly argued about since the first attempts to measure intelligence, is the validity of the character sets on which resemblances are to be based. Thus tests that discriminate between the sexes are excluded from intelligence tests because intelligence is intended to be uncorrelated with sex; but tests that discriminate between races may not be excluded. Many tests are greatly influenced by social environment. In addition, tests administered to subjects on successive occasions do not elicit exactly the same responses, so that workers in psychology are much concerned with the reliability of the variates used in their computations. These questions have been discussed at length by Hawkins (1964) and it is clear that it is difficult to choose objective sets of characters for psychometrics.

Numerical taxonomic tests of the ease and accuracy by which intuitive classifications and ordinations can be made by eye (see Section 5.10) relate to the psychology of perception and should have interest for workers in this field (see, for example, Micko and Fischer, 1970; Marr, 1970).

In social anthropology and linguistics, problems of evolution have attracted much interest; this has been well discussed by Kroeber (1960). Such studies are particularly difficult because of the reticulate nature of cultural evolution. Again we meet with the problem of the validity of sets of characters. Some early papers employing special resemblance coefficients are those of Klimek (1935) and Clements (1954). Milke (1949) and Howells (1966) made important observations on the numerical relation between increasing geographic distance and increasing cultural dissimilarity. Driver and Schuessler (1967) have followed up the suggestion of Leach (1962) and examined a world ethnographic sample of cultural traits by factor analysis to see what support is given to traditional ways of dividing up cultures. This is evidently a complex matter, for only two of the five main factors were readily interpreted. These two were evidently related to patricentered and agricultural societies, respectively. Hudson (1967) made a numerical taxonomy of social attitudes to artists and scientists.

Clausen (1967) reviews numerical taxonomy in political science. In this field Grumm (1965) studied the voting behavior of legislators by average linkage clustering; he found, besides the main political parties, subgroups of a few individuals who followed a distinctive voting policy. Alker (1969) has used numerical cladistics to make an interesting study of the evolution of political systems.

In the study of languages we may mention the work of Dyen (1962, 1965) and Carroll and Dyen (1962) on computer classification of languages, and Ross (1950) on philology. When vocabularies are used as the character sets of language studies it is usual to measure resemblance as the proportion of "plausible cognates" in the word list. Thus "decem" (Latin) and "deka" (Greek) are obviously cognate, but are these cognate with "ten"? Homologies may therefore be difficult to decide. There may be incongruence between different parts of the vocabulary, or between vocabulary and grammar, when languages of mixed origin like English are studied. It may not be clear whether low resemblance is due to interdependence (see Ellegård, 1959, on this and the significance of negative matches). Computing problems also arise with the very big matrices of word lists. Dyen, James, and Cole (1967) have attempted numerical cladistics based on the rate of replacement of words in different branches of a family of languages, but reticulation (due to word borrowing) poses formidable problems. Some work on grammar has been done by Svartvik (1966) and Carvell and Svartvik (1968), in which clusters of prepositional phrases and idioms were found. An important study relating language to culture is the taxometric study of the dialects of Salish Indians of the American Northwest by Jorgensen (1969).

In physical anthropology taxometric methods are being slowly introduced, though the application of multivariate statistics in this field dates back several decades. Mention has been made in Sections 4.9 and 8.5 of studies on growth and on discriminant functions in man. There have been a few attempts to measure resemblance between human populations (e.g., Rao, 1952; Cavalli-Sforza and Edwards, 1964; Vyas et al., 1958; Neel and Ward, 1970; Jardine, 1971), though Huizinga (1962, 1965) has some pertinent comments on resemblance coefficients. Much disagreement exists on the pattern of variation in humans, despite the large body of data obtained by earlier methods (Bielicki, 1962; Wierciński, 1962), and there is urgent need for investigating racial differences by the study of individuals as the OTU's. Numerical taxonomic methods have recently been applied to social geography (Goddard, 1968; Krueckeberg, 1969; Willmott and Grimshaw, 1969).

In archaeology the OTU's can be of many kinds, from individual artifacts like axes to collections of many kinds of artifacts that together represent a cultural assemblage. The OTU's may be from many sites and periods. The choice of appropriate OTU's is therefore a critical step, depending on the sort of study planned (well discussed by Tugby, 1965). Material may however be scanty, and with simple objects (e.g., flint arrowheads) there may be very few characters that can be considered meaningful. There are a number of special resemblance coefficients used for cultural assemblages and Driver (1965) reviews these. Neither factor analysis nor cluster analysis has been employed widely, though clusters of artifact types or traits have been recognized by eye from Q or R kinds of resemblance matrices (see Tugby, 1958; Lewis and Kneberg, 1959; Clarke, 1962). Polythetic methods appear to be superior to monothetic ones (Doran and Hodson, 1966). A detailed study of Iron Age brooches (Hodson, Sneath, and Doran, 1966) using details of style and manufacture as characters showed that cluster analysis was satisfactory in some respects: near-duplicates were clustered together and atypical brooches were confirmed as such. In this study no large clusters were found, evidently because the designs belonged to a single cultural tradition, though as noted below fashions slowly changed with time. Hodson (1969) has shown the power of taxometrics to reveal structure in a collection of bronze tools analyzed for trace elements. True and Matson (1970) report on cluster analysis of archaeological sites in Chile.

The problem of finding time trends in fashions is one of special importance to archaeology and has its parallel in evolutionary biology. Frequently no independent dating is available, so that the trends must be discovered from changes in pattern of the artifacts and assemblages; the aim is to arrange these into chronological sequence. The classic conventional study is that of predynastic Egypt by Flinders Petrie (see Kendall, 1963). This was first attempted numerically by Brainerd (1951) and Robinson (1951) and has been discussed in Section 5.6 under the heading of seriation. Freeman and Brown (1964) illustrate with a statistical study the danger of assuming temporal changes on slender evidence. A study with a nonlinear scaling method has been made by Hodson et al. (1966), and this showed a clear correlation between the major axis of variation and the date (from independent criteria). This trend was quite difficult to detect by eye. The problem is one of ordination, and, if nonlinear methods are not essential, factor analysis sensu lato may be useful and can handle much larger numbers of OTU's, as pointed out by Cowgill (1968) and illustrated by Hodson (1970).

11.4 THE EARTH SCIENCES

Applications to the earth sciences are becoming numerous. Paleontology has already been discussed in Section 6.5, paleoecology and soil studies have been described in Section 11.1. Useful general references to numerical taxonomic methods in geology are those by Miller and Kahn (1962), Krumbein and Graybill (1965), and Harbaugh and Merriam (1968). These works also cover factor analysis and discriminant functions, topics covered in more detail in the review by Reyment (1963). Clustering methods in particular are discussed by Jizba (1964) and Parks (1966). Much work is summarized in Merriam (1966).

Classification of rock types is now an active field. When such classifications are based on specimens with few petrographic components, scaling is likely to present problems. Procedures like standardization of the composition percentages may have marked effects on the relationships, and it is therefore important to consider with some care the purpose of these procedures. In some specimens the proportions of minor constituents may be correlated, and it may then be useful to treat the analogues of size and shape separately; Imbrie and Purdy (1962) have suggested the use of the vector angle (Section 4.11). More usually distances, correlations, and association coefficients have been used (see Harbaugh and Demirmen, 1964

Behrens, 1965; Bonham-Carter, 1965). Another problem arises when the composition is expressed as percentages that necessarily add up to 100. This introduces effects on correlations, which have been discussed by Krumbein (1962).

Studies on rocks include those of Varty and White (1964) on the classification of clays, which the authors found to form a single cluster, whose extreme members had been given separate names, although a reexamination by Rayner (1965) gave some evidence of subclusters. Silicious rocks were studied by Howd (1964), who used the method of Rogers and Tanimoto (1960). Obial (1970) applied numerical methods to classifying stream sediments from a mineralogical survey and obtained clusters that closely reflected the bedrocks.

Imbrie and Purdy (1962), Purdy (1963), and Bonham-Carter (1965) examined the same data on modern carbonate sediments by factor analysis and cluster analysis. The factors and clusters were highly congruent. Parks (1966) also found agreement between factors and clusters in Q and R studies of rocks. Behrens (1965) used correlations and cluster analysis in a study of the composition of limestones; he found four main clusters of the constituent materials, which appeared to relate to the environments in which these facies were deposited. Similar studies were made by Bonham-Carter (1967a) and Veevers (1968). Kaesler and McElroy (1966) examined clusters of different sandstones and the localities from which they came; members of the same cluster usually occurred close together on the map. Chave and Mackenzie (1961) have used linkage diagrams (Section 5.9) for work of this kind. As the result of a faunal distance study, Hecht (1969) obtained evidence that in the Miocene there was no sharp change in sea temperature between Florida and New Jersey, unlike the present day.

A special problem that is similar in many ways to finding homologous positions in protein sequences (for which similar methods are used) is that of matching up the strata in different rock sections and revealing cyclic phenomena of deposition (cyclothems). Cross-association (Sackin, Sneath and Merriam, 1965; Sackin and Merriam, 1969) has been used for such studies. Some illustrative examples and discussions of problems are given by Merriam and Sneath (1967) and Sneath (1967e, 1969c). This exploratory work suggests that the method has considerable power for finding the best matches and can tolerate a limited number of gaps in the successions of strata. Although it is also capable of demonstrating cyclothems where successive cycles of rock types occur (such as shale, limestone, sandstone, shale, limestone, sandstone,...), its sensitivity for this is less certain. A problem awaiting solution is to allow for differences in thickness of corresponding strata.

Another, more difficult subject is the measuring of resemblance between geological or geographic surfaces. Some exploratory work has been reported, most of it based on results of trend surface analysis (Merriam and Lippert, 1966; Merriam and Sneath, 1966; Sneath, 1966c). This may employ either the trend coefficients or the residuals from fitted surfaces. Comparison between maps is also pertinent in geography (see Haggett, 1965), and numerical taxonomic methods are now being actively used in several branches of geography (Ahmad, 1965; Berry, 1968; Johnston, 1968; Cole and King, 1968; Spence, 1968), and in the field of tree-ring research (e.g., Fritts, 1963). Skaggs (1969) applied cluster analysis to tornados using meteorologic variables, and demonstrated several types of tornados characteristic of different geographic regions. A review of applications to land surveys is given by Williams and Lance (1969).

11.5 OTHER SCIENCES AND TECHNOLOGY

Most of the applications of taxometric methods in other sciences have been in the fields of library science and pattern recognition; an interest common to both is information theory. The close connections between taxonomy and information theory may be seen by reading the articles of Good (1958, 1962). He discusses clumps and clusters and how to find them (under the name of "botryology") and also takes up some of the philosophical points of polythetic classes. An example is the concept of a cow cited in lectures by the Cambridge philosopher John Wisdom. A cow has four legs and gives milk—but it may have three legs and may not supply milk. No one property may be essential to its "cowness." Estrin, quoted by Good (1958), suggests information retrieval by asking for *k* out of *n* index words before selecting a document as relevant. But the group "cows" is a polythetic taxon. So are Estrin's document clusters. This is clearly similar to the philosophical considerations of Beckner (1959, 1964), discussed in Section 2.2. A perennial problem in library science is the classification of knowledge. Subject headings such as "chemistry" or "biology" are evidently polythetic concepts, with much overlap. They also change with time, as new disciplines arise. The conventional problems of classification are thus formidable. A less formidable problem is now being explored by taxometric methods. This is the construction by computer of clusters of words that occur together in documents, which is reviewed in Stevens, Giuliano, and Heilprin (1965), Needham (1967) and Sparck Jones and Jackson (1970). The aim is to find the best key words for indexing documents. Words that consistently appear together in documents are good candidates because they indicate a cluster of related concepts. For example, "magnetic," "transistor," "voltage" would point toward electronics as the subject of the document. These studies face two major problems: the matrices of word occurrences are very large, although most entries will be zeros; and the required clusters are overlapping and perhaps of rather special kinds. Methods will therefore develop along somewhat different lines from those in biological numerical taxonomy. Work on the classification of knowledge is relevant to language translation research (Parker-Rhodes, 1961). And Deutsch (1966) has attempted to formalize the performance of a classification as a way of establishing communication codes for information retrieval.

Numerical taxonomy can assist in pattern recognition, which Lusted (1960) refers to as one of the challenging problems to be found in many fields of science.

Pattern recognition usually means the identification of written or printed characters, also sounds, pictures, etc., with patterns already described and stored in the memory of a computer (see Nagy, 1968, for a general review, and Casey and Nagy, 1971, for a popular account). It is therefore mainly concerned with techniques of identification, with the special requirement that incorrect identifications must be very few. A great variety of identification methods have been proposed, but probably most are simultaneous polythetic ones and in consequence various resemblance measures have been devised, often paralleling closely those in numerical taxonomy (Unger, 1959; Bonner, 1962; Firschein and Fischler, 1963; Mattson and Dammann, 1965). A modified numerical taxonomy was used by Sneath (1964b) to solve a simple jigsaw puzzle.

A recent development is the proposal of mixed systems for classification and identification. In these each unknown is in turn incorporated into the classification and new classes of pattern are set up when examples are received that do not fit the previous classes (Sebestyen, 1962; Ornstein, 1965; Rosen, 1967). Although this work is directed mainly toward making machines that can learn to read the handwriting of different people, clearly it has potential for many fields.

Applications of numerical taxonomy in scattered fields include economics (Möller, 1964; Fisher, 1969; Goronzy, 1969; Fisher, Williams, and Lance, 1967); chemistry (Oyama and Carle, 1967); naval studies (Cattell and Coulter, 1966); market research (Joyce and Channon, 1966; Green, Frank, and Robinson, 1967; Frank and Green, 1968); and genetic analysis of mutants (Gillie and Peto, 1969). Such methods have also been used in studies on biological activity in relation to chemical structure (Sneath, 1966d; Simon, 1968). Using a clustering method, Harrison (1968) was able to obtain considerably better prediction of activity of new chemical compounds than would be expected by chance, and such techniques may well have advantages over the more usual regression analysis approaches.

11.6 THE ARTS AND HUMANITIES

Most applications of numerical taxonomic methods in fields of the arts and the humanities have been directed toward questions of disputed authorship, with a little work on literary style. These are mostly problems of identification rather than classification. A general survey is the volume edited by Leed (1966). There has been much controversy over the kinds of traits needed to reach accurate conclusions about authorship, particularly with ancient texts (for example, see Morton and Winspear, 1967). From the *known* works of specific authors, it can be determined which traits are easily imitated, or which vary much with the author's period or subject matter. Such traits are obviously undesirable for identification. Less attention has been given to the best discriminatory statistics or to the number of traits that are needed. A milestone is the work of Mosteller and Wallace (1964) on the authorship of *The Federalist* papers. Another detailed study is that by Ellegård

(1962) on the authorship of the letters, also published in the eighteenth century, signed by "Junius." Meier-Ewert and Gibbs (1970) demonstrate by cluster analysis the great sensitivity of doublets of letters in distinguishing not merely different languages (as is well known) but, less expectedly, different authors using the same language.

A different approach is illustrated by the elegant study of Blackith (1963) on Latin poets, employing ordination. He was able to isolate several components of style that appear to be reliable indicators of authorship. Of particular interest was the finding that the mental stress of Ovid's exile to Tomis was clearly reflected in stylistic features of his poetry written during his exile.

Less attention has been given in arts and humanities to quantifying relationships or making classifications. Buechley (1967) has made cluster analyses of family names in different geographic localities and Griffith (1967, 1968) has applied numerical taxonomy to medieval manuscripts by Juvenal; his ultimate aim is the more difficult task of determining their cladistic, as well as phenetic, relationships. In a similar study of New Testament manuscripts Griffith (1969) found that their dates were clearly reflected in a seriation, with the Codex Bezae (a puzzling manuscript of uncertain origin) consistently aberrant. Wishart and Leach (1970) used several methods of clustering and ordination to examine the works of Plato and to determine the probable chronology; they obtained an arrangement identical with that in an earlier independent mathematical study by Cox and Brandwood (1959). Such methods are now being applied to music; an example is the project of Bernstein (1967) to quantitate the properties of musical style. The proceedings of a recent symposium on mathematical methods in the humanities have been edited by Hodson, Kendall and Tăutu (1971).