Palaeont. afr., **19**, 1—20 (1976)

NUMERICAL METHODS IN THE DEFINITION OF PALYNOLOGICAL ASSEMBLAGE ZONES IN THE LOWER KARROO (GONDWANA) OF RHODESIA

by

Rosemary M. S. Falcon

ABSTRACT

The aim of this paper is to explore the possibility of using numerical methods in the sub-division of the palynological assemblages encountered upwards in the stratigraphic sequence in one borehole core situated in a central position in the Mid-Zambezi Basin, Rhodesia. Due to its apparently conformable thick succession and the variable micro-floras already encountered within it, this borehole has been proposed as a palynological type section for the Zambezi Basin, Rhodesia.

The numerical methods used to delineate zones of similar assemblages were basically very simple, using only presence-absence data of a set range of species. The formulae employed were those of Jaccard, Simpson and Sokal (the only linear method), and the methods of classification were simplified forms of divisive or shaded half matrices, agglomerative using single linkage clustering, and spatial ordination techniques.

The results support the concept of distinct changes of microfloras up the stratigraphic sequence. Groups of assemblages bearing similar microfloras appear stratigraphically adjacent to one another, barring certain minor instances of transition floras in the younger sequences. Measurable delineations or boundaries between these assemblage zones are illustrated. These numerically defined zones coincide very closely with the assemblage zones already proposed on a visual (quantitative and qualitative) basis.

CONTENTS

INTRODUCTION

The apparently conformable stratigraphic sequence in the Matabola Flats borehole, which penetrates twelve hundred feet or about 400 metres of Lower Karroo (Gondwana) sediments and is situated centrally in the Mid-Zambezi Basin, Rhodesia, was regarded as an excellent means of testing the concept of applying detailed biostratigraphy to the region, i.e. to establish whether, from the palynological content, sufficiently clear assemblage changes occur up the stratigraphic column to warrant delineation of zones or assemblages which may be useful in future correlation. (N.B.— The use of feet rather than metres is retained due to its practical application in Rhodesia.)

Previous statistical work, summarised by Hart (1970), has been initiated elsewhere in this field. Later Hart and Fiehler (1971) created a computer-aided storage and retrieval system for Lower Karroo stratigraphic palynology, taking for a first analysis a sequence of rocks in the Mchuchuma River Valley, Tanzania. The eventual aim of these authors is to es-

tablish a complete taxonomic and stratigraphic analysis of samples from any section or borehole. Methods, results and conclusions are as yet to be fully published, and the basic taxonomy, whether coded or linear in type, has not been established. It is therefore primarily a retrieval system at present.

Christopher and Hart (1971) have shown statistically the effects of extraneous sources of variation on miospore analysis. They establish a method of measurement "by defining their operating level in the geological maze", and further erect a simplified method of zonation based on frequency of genera, independent of lithological bias. The same sequence of rocks as used by Hart and Fiehler (above) was used, and 25 genera were analysed relative to their distance above basal glacial deposits and the rock-type of the respective samples. Of the 25 genera analysed only three showed significant and lithologically independent changes in their frequency of occurrence. The distribution of the remaining 22 genera could not be used in establishing biostratigraphic florizones due to their large fluctuations in occurrence related to factors other than lithology.

The conclusions reached by Christopher and Hart (op. cit.) on this Tanzanian sequence were that two major biostratigraphic zones with one narrow intermediate transition zone could be recognised, based upon regression equations and curves for the genera *Cordaitina* and *Protohaploxypinus*.

In the transition zone a major change in distribution pattern of these two genera occur.

These statistical subdivisions are shown to be similar to the visual zones published earlier by Hart (op. cit.) based on conventional quantitative and qualitative methods.

The conclusions reached by Hart (op. cit.) confirm that major macro- and microfloral changes were occurring during this period in the Lower Karroo, and that the relative frequencies of occurrence of certain selected genera do in fact represent good parameters for distinguishing distance or time above the glacials, i.e. biostratigraphic zones. This method may well prove to be more than adequate in defining relatively narrow stratigraphic units and in periods of time when climate change and evolution are known to have been active (Falcon, 1975), but it may prove too cumbersome or ineffective when attempting to define biostratigraphic units over a large span of time and in much thicker stratigraphic sequences. Hence the attempt to apply less sophisticated numerical techniques as seen in this paper.

In Rhodesia, to date, biostratigraphic subdivision has involved the application of visual methods including the standard use of qualitative and quantitative microfloral data from samples taken at intervals up the stratigraphic column in one type section. Delineation of miospore assemblage boundaries was based primarily on the qualitative and quantitative assemblage characteristics (assemblage zones) and secondly on the finite stratigraphic ranges of zonal or restricted species. Results of this analysis have appeared elsewhere (Falcon, 1975).

In terms of numerical and statistical techniques this visual and somewhat generalised concept of zonation is regarded as inadequate. Modern methods require that subjective and intuitive influences be superseded and that, with the advent of computerised techniques, information be objective and concise in order to obtain repeatable results.

Therefore, in order to obtain an unbiased method of grouping samples from the borehole core into groups exhibiting similar microfloral characteristics, certain numerical methods of analysis have been applied. It is the purpose of this paper to outline the methods and coefficients used, and to compare the microfloral assemblage changes thus illustrated with the zonal boundaries already proposed in the visual analysis. If it is shown that results from both methods coincide and complement one another, then the existence of objective boundaries and biostratigraphic units can be accepted, at least within the context of this one borehole core.

More detailed analysis of the data using computer techniques might yield a basis for consistently recognising finer sub-divisions of the units proposed here, which would be of use in local correlation.

PROBLEMS IN METHODS OF SUB-DIVISION OF THE KARROO

The initial problem in sub-division of Karroo sequences, as summarised by Hart (1970) and Falcon (1975), is their continental origin. The Karroo in Africa is made up of thick horizontal and originally widespread terrestrial deposits of sandstone, shale, mudstone, carbonaceous and coal deposits, some of which are economically important. These sequences are now encountered in isolated basins in central and southern Africa, the largest of which is the South African Karroo Basin. There is a very low macrofossil yield in the Lower Karroo with the result that sub-division and correlation has depended to a large extent on lithological criteria and to a very small extent on classical biostratigraphic techniques.

The fundamental problems of visual biostratigraphic zonation in the Karroo using palynology is a taxonomy which is unsteady and highly variable in interpretation, and thus methods in Gondwana palynological sub-division are basically subjective, with zonal boundaries drawn at arbitrary levels dependent on visual quantitative and qualitative microfloral characteristics and from samples taken over wide geographic areas. The biostratigraphic units themselves have been described in terms of microfloral assemblages found typically in them and none of these has been clearly related to any one section or series of specified sections.

However, particularly in pre-Tertiary times, the variables affecting spore/pollen production in the living plant communities (ecological, depositional climatic influences and geographic separation (Dale, 1971)) and the conditions encountered after fossilisation (diagenesis, erosion, re-deposition, etc.) are now well acknowledged and biostratigraphic unit sub-division based on visual qualitative and quantitative characteristics is approached with caution (Dale and Walker, 1970; Hughes and Moody-Stewart, 1969).

For regional and local sub-division and correlation, an independent numerical approach based purely on presence/absence data is examined and analysed in this paper in the hope that this may prove more reliable than visual methods. The numerical approach in Gondwana provinces is still in its infancy, the only previous attempt known to the author being that of Hart (1970) who reports the existence of a computerised miospore retrieval system for African Karroo sequences. This, however, has a basically taxonomic emphasis and is inapplicable for the present purpose of finely delineating the Lower Karoo sequences.

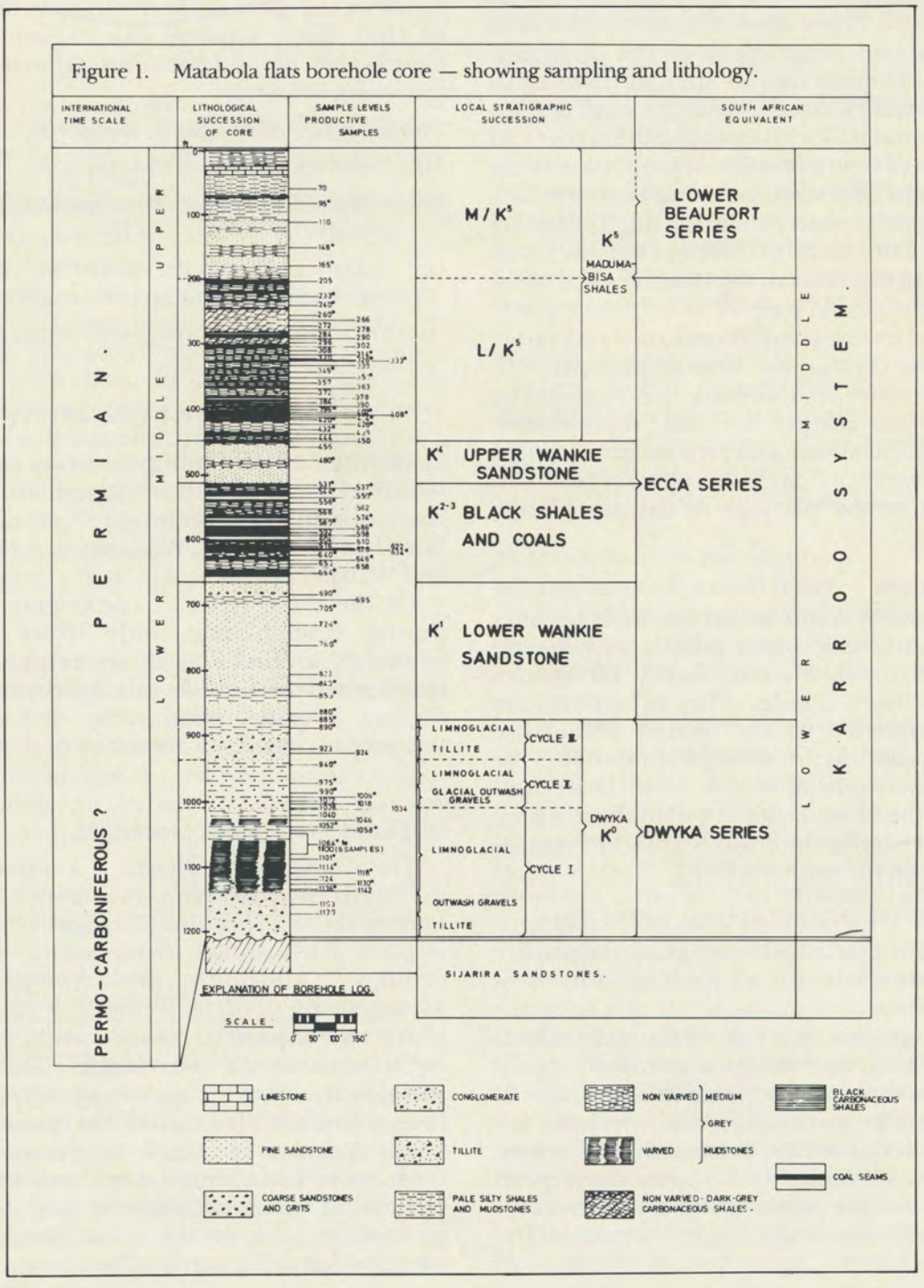


Figure 1. Matabola flats borehole core — showing sampling and lithology.

PREVIOUS SUB-DIVISION OF THE MATABOLA FLATS BOREHOLE CORE

This is discussed in detail (Falcon, 1975) and is only summarised here for the purpose of comparison (see Figure 1).

Assemblage zones

These are based on characteristic quantities of suprageneric taxa, and contain distinctive assemblages of notable genera. The boundaries are placed at points in the stratigraphic column where marked changes occur in qualitative and quantitative characteristics.

Four assemblage zones are described with boundaries at about 432 feet, at about 556 feet and at about 645 feet. The major quantitative differences lie in the change in dominance from Trilete-Monosaccate miospores in zone I to the Disaccate-Striatiti forms in zone IV. In zone II, Disaccites and Monosaccites are approximately equal in proportion (Striatiti forms comprising a small percentage of the Disaccites), whilst Triletes are still dominant. In zone III, Aletes represent a prominent proportion of the assemblages, whilst Disaccites, including Striatiti (which now represent approximately a third of this Turma) are about equal in abundance to Triletes. Striatiti forms become dominant members of the Disaccites and of the total supra-generic taxa in zone IV.

Generic content varies with dominance and variation of the major taxa up the stratigraphic column. The assemblage zones are therefore characterised by associated genera related to the suprageneric groups. Their associations and stratigraphic ranges (including appearances and disappearances) add more emphasis to the placings of the assemblage zone boundaries.

Assemblage sub-zones

Within the context of the major assemblage zones (with their characteristic supra-generic and generic contents), further sub-division based on species variations has been made. The sub-zones are therefore characterised by the relative abundance and group associations of selected species and the diversification in evolving major taxa (Disaccites-Striatiti) whilst the boundaries are placed according to the ranges of stratigraphically significant species particularly within the evolving flora.

METHODS OF NUMERICAL ANALYSIS

The aim of this numerical analysis is to compare the degrees of similarity for all pairs of samples in order to establish:

(i) groups of samples bearing similar microfloral assemblages, i.e. assemblage zones, and

(ii) levels in the stratigraphic column where significant changes in assemblage characteristics occur, i.e. boundaries between assemblage zones. The degrees of similarity vary and the greater the difference (or *dissimilarity*) between compared assemblages so the larger the microfloral change.

PREPARATION OF INFORMATION

The numerical methods chosen for this study are those employing only qualitative species information, i.e. presence/absence data. The presence of species not seen in samples located within their known stratigraphic range is not presumed.

Many methods employing quantitative information have been successfully tried on Tertiary and younger microfloral assemblage schemes (Dale and Walker, 1970; Williams and Dale, 1965; Williams *et al.* 1966), but due to the many variable factors affecting pre-Tertiary fossils the presence/absence of species is preferred at present.

The initial information was obtained by tabulating the sample numbers (43 in all) against a list of all the species taken into consideration (a total of 105). Every possible pair of samples was compared and the information tabulated under the following headings:

(i) Number of species in sample A,

(ii) Number of species in sample B,

(iii) Number of species common to both (matched, positive),

(iv) Number of species occurring in only one or other sample (unmatched, negative),

(v) Total number of species in samples A and B.

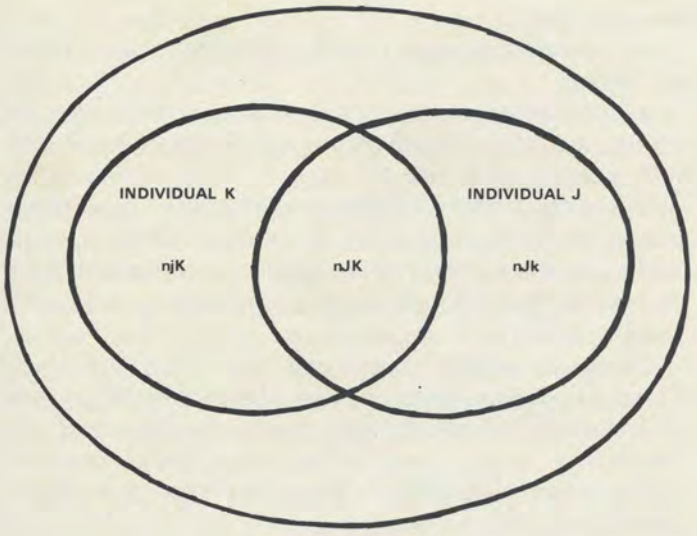
CHOICE OF COEFFICIENT

The methods of finding measures of resemblances between pairs of entities based on a number of characteristics is summarised in detail by Sokal and Sneath (1963, p. 123), Williams *et al.* (1966) and Dale and Walker (1970).

Of the three types of coefficients used in computing resemblance, only those of *association (similarity)* and of *distance* are employed expressing direct similarity, whilst the coefficient of distance implies a spatial relationship and a quantitative measure of degrees of similarity or dissimilarity.

Coefficient of association (similarity)

The similarity coefficient is a numerical value of the difference between two entities or, as used herein, the miospore assemblages in two samples. It is particularly suited to measuring associations in ecology (Goldsmith, pers. comm.). The basic arrangement of data is the 2 × 2 table, with n characters (species) scored for two individuals (samples or sample assemblages). The characters are sub-divided into presence (positive) or absence (negative) classes for each of the two individuals. The count is then taken for those characters positive to both individuals, negative to both individuals (but present in other individuals) and for those only positive in one or the other. In two imaginary samples, J and K, one then has:



n = no. of characters (species) in the whole sequence, of which J and K possess a certain number. Therefore
JK = shows positive presence of one character in both samples, and
jk = negative presence;
nJK = no. of characters *positive* to both individuals;
njk = no. of characters negative to both; and
nJk, njK = no. of characters positive to only one or other individual.

The number of characters with common presence or absence in both samples are termed "matched", and their sum is termed "m". The number of characters occurring only in one or other sample is termed "unmatched", and their sum is termed "u". The total number of characters, n, is therefore:

$$n = nJk + njk + nJk + njK$$
$$= nJ + nj = nK + nk$$
$$= m + u$$

The fundamental formula as defined by Sokal and Sneath (op. cit.) consists of the number of matches (or total number of common species) — the numerator, divided by a term implying the possible number of comparisons — the denominator. The larger the coefficient value for the comparison between two samples, the greater the similarity between them.

The formula or coefficient of similarity employed in this study is that of *Simpson*[m2] in Mortimer (1969).

$$\text{i.e. } C = \frac{nJK}{\min (nJ, nK)}$$
$$= \frac{\text{total no. of spp. common to both}}{\text{no. of spp. in the less diverse assemblage}}$$

This coefficient allows for wide variation in numbers of species per sample, from rich to poor. An assemblage poor in species variation may therefore possess a high degree of similarity (up to 100%) when compared with another rich in species.

The *coefficient of distance* is a measure of similarity utilising distance or dissimilarity based on Pythagoras's theorem. The techniques employed are therefore geometric and are here only being considered in two dimensions. (See Sokaland Sneath, op. cit., p. 143.)

The measure of distance is estimated on the basis of quantitative values of negative matches of species between two samples or miospore assemblages. These values may be plotted on single or rectangular co-ordinates. If the two samples under consideration are identical, their spatial position will coincide, and the distance between them will be zero. The greater the dissimilarity, the greater the distance between them. Therefore distance complements similarity.

The coefficient of distance employed in this study is that of Sokal (Sokal and Sneath, 1963, p. 147).

$$Djk = \sqrt{\sum_{i}^{n} (Xij - Xik)^2}$$

where Xij = 0 if the species i is absent from sample j;
= 1 if species i is present in sample j.

This coefficient of distance (D) generalises the concept of Euclidean distance between two points in an n-dimensional space, i.e. it embraces information regarding the presence or absence of the whole range of species in a two-dimensional setting, relative to two standard points or samples.

Methods of classification

There are a variety of ways of constructing classification systems for grouping assemblages with similar characteristics. The system is basically hierarchical or non-hierarchical (see Dale and Walker, 1970; Gordon and Birks, 1972).

I. *Hierarchical methods* involve:
(a) agglomerative techniques, a monothetic form of singly clustering together similar entities; and
(b) divisive techniques, dividing up the total number of entities into smaller and smaller groups.

In order to make significant hierarchies clumped distribution of samples is required, i.e. nests within the hierarchies. Random distribution of samples may give some nested groups, but due to too many intermediate or transitional members, the hierarchies or divisions may not be as clear cut as those in which the samples are retained in stratigraphic order.

In making clusters of similar groups, the dividing lines or assemblage zone boundaries must be placed where there are significant breaks in microfloral composition. The *level of termination* (or *rank*) at which the hierarchical schemes are divided (in order to establish the groups of similar assemblages and the degrees of dissimilarity between them) is somewhat arbitrary. The level of rank adopted in this paper is that which illustrates most clearly the maximum number of assemblage sub-zones without omitting transitional single members and with the least amount of overlapping between nests or clusters.

It is an accepted fact (Sokal and Sneath, 1963; Goldsmith, pers. comm.) that in ecological analysis of extant organisms, overlapping of assemblages or communities is inevitable. However, due to the long time span represented by this borehole sequence, microfloral assemblage changes are expected to be more marked. Climatic and ecological conditions are the accepted causes of community changes, but other factors to be considered when covering a long period of time are those of evolution, disconformities in the stratigraphic sequence, long erosional periods, deposition of reworked sediments bearing older microfloral elements and repetition of climatic cycles. These conditions will influence the final interpretation of the assemblages analysis, or indeed may show up in the course of the analyses.

The methods of illustration or graphic representation of the hierarchical methods applied in this paper are (a) shaded half matrices and (b) single linkage clusters. These are considered useful preliminary techniques in terms of analysis of the Matabola Flats borehole core, and were drawn using the Simpson coefficient.

II. *Non-hierarchical methods* include ordination, a further method of sub-division of assemblages into similar groups based on spatial linear relationships. This technique is based on the distance-measure of all sample assemblages relative to two reference assemblages on a standard axis. *Sub-division* of assemblage zones is by means of magnitude of distance between two nests or groups of assemblages; the larger the measure between them, the greater the microfloral dissimilarity.

Illustration of this method of classification adopted in this paper includes (a) single axis ordination, and (b) spatial ordination (clustering). In single axis ordination the intercept of the perpendicular from the point at which the arcs meet (i.e. the measured distance of each sample from the two standard samples) and the axis gives the relative distance of dissimilarity along the axis of the sample assemblage from the two reference samples. The closer the sample assemblages, the greater their similarity. Ideally assemblages from samples closely adjacent would fall into neat clustered groups or nests to form assemblage zones or biostratigraphic units (if all members are stratigraphically adjacent).

Spatial ordination comprises the composite spatial relationship of two single axis ordinations. The sample assemblages relative to four basal axial samples appear as points where the perpendiculars meet. Those samples most closely grouped are recognised to be most similar. Ideally, nested or clustered groups of samples would represent biostratigraphic units if stratigraphically discrete, or "associated assemblages" if randomly scattered in the stratigraphic sequence. The larger the gaps between the nests, the greater the microfloral change.

Nomenclature

A few relevant terms used in this paper are defined below:

Sample denotes the small rock unit at a specific depth in the borehole core from which the miospores are obtained.

Assemblage refers to the characteristic qualitative microfloral content of each sample. Both sample and assemblage refer to a single entity. Samples are named in feet (depth) and assemblages numbered 1–43 as in Figure 7.

Visual assemblage zones and sub-zones (VAZ) are biostratigraphic units, having distinct stratigraphic delineations (boundaries) and characterised by qualitative and quantitative microfloral content. Analysis and definition is based on inspection alone. (See Figure 12).

Numerical assemblage zones (NAZ) are units consisting of assemblages (=samples) grouped together on the basis of similarity or dissimilarity of the microfloral content, irrespective of the stratigraphic positions. The microfloral analysis is based solely on presence-absence data while delineations of the NAZs are based on the magnitudes of spatial breaks between them.

Results

These will be discussed under the headings of classification methods and comparisons made between the two coefficients involved in this study.

(a) Divisive Techniques in Classification

The technique applied is that of *differential shading of a similarity matrix*. The coefficient used was that of Simpson. This is the simplest method of illustrating the sub-division among the assemblages, as groupings may be seen at a glance. (See Figure 2.)

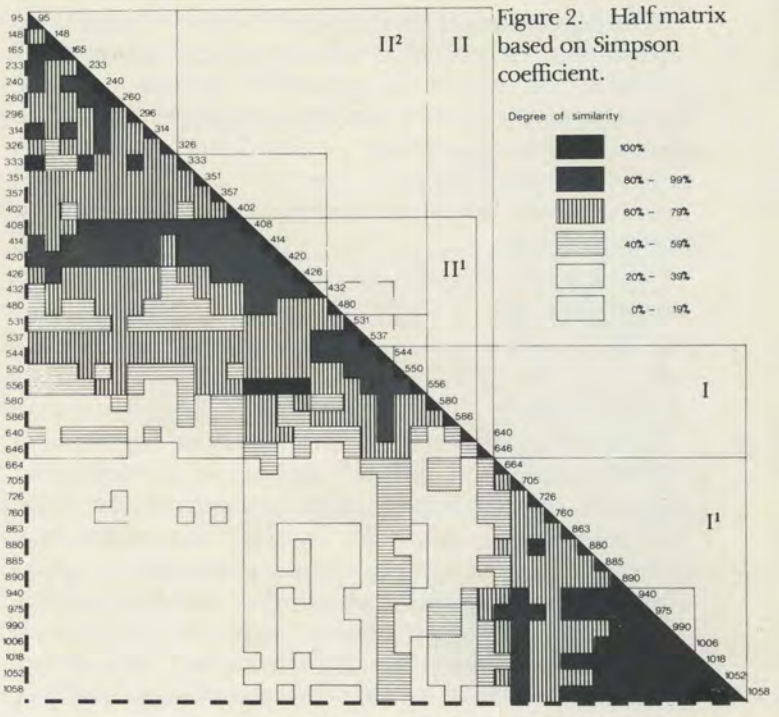


Figure 2. Half matrix based on Simpson coefficient.

The coefficient values for all possible pairs of assemblages are placed into a half matrix. These values are divided into five equal classes of magnitude between 0 and 100%, the sixth indicating absolute similarity or 100%. Each class possesses a different degree of shading proportional to its magnitude and when applied to the values in the squares of the half matrix should show dark areas of high value and light areas of low value.

In the present application, the assemblage values in the half matrix were arranged in stratigraphic order so that the clusters of similar assemblage groups, although appearing rearranged due to their marked nestings or groupings, are in fact unbiased and merely arranged according to depth.

The half matrix shading illustrates that samples (assemblages) are grouped together in large distinct clusters in ascending or descending stratigraphic order. At a glance, four major groups of clustering are obvious but on coefficient value analysis (i.e. according to the percentage units) a different and more complex picture emerges.

At the low similarity value of 40–59%, two large overlapping groups are seen: NAZ I = 1,058–544 and II = 646–95 ft. The overlap or transition area is between samples 544 and 646 — which coincides with the proposed transition zones in the visual assemblage zones. A further sub-division at the 60–79% level shows one large microfloral unit, NAZ I^1, below 664 ft and two overlapping above 640 ft, NAZ II^1 and NAZ II^2.

NAZ I is characterised by one large cluster of samples centred around samples 940 ft, 975 ft, 990 ft, and 1 006 ft. No further assemblage breaks occur of any significance. It is of interest to note that all the samples above 890 ft and below 1 018 ft do not show any greater similarity to each other — their uniform value is 60–79%. But all these samples have greater values when compared to 940–1 006 ft. This is considered to be a function of the coefficient in that the richest and most varied miospore taxa occur in samples 940–1 006 ft; all other samples with similar species but fewer in variety therefore show extremely high values when compared with these rich samples.

NAZ II is divided into two overlapping zones with values of 60–79%: 640–408 ft (II^1) and 556–95 ft (II^2). A further three sub-zones may be sub-divided. These are arbitrary units based on careful comparison of all samples, and are not clearly seen in the visual assemblage zones. Overlapping, clustering and repetition of high similar values account for this. One small transition zone occurs. Sample assemblages down to a level of 333 ft show close comparisons between each other and samples 408–420 (80–99%), i.e. repetitive or associated assemblages. But samples 35 to 402 show closer affinity to 408–426 only. There would seem therefore to be a break of some small order in the region of sample 333. Similarly, assemblages 420–408 compare at a similarity of 80–99% not only with assemblage 480 below (and 556, a relatively weak sample), but all possible assemblages above, possibly illustrating a new flora at level 426–432 ft which continues through the younger sediments — this agrees with the *Vittatina-Lueckisporites* visual assemblage zone. Sample 480, a stratigraphically isolated sample, shows equal affinity up to 408 ft and down to 544 ft. This may therefore be termed a transition assemblage.

The "problem" assemblages in samples 640, 646 and 664 are indicated by the Simpson coefficient as being transitional, but with a major dividing point in sample 646. This assemblage bears a 60–79% similarity to the adjacent major assemblage groups above and below (537–550 and 940, 975). The younger assemblage 640 is moderately similar to 646 but below this a break occurs in older assemblages. The natural affinity is to the younger assemblage group 408–556. Alternately sample 664 ft shows a moderate break (±50%) with younger assemblages above but a close affinity to older ones below.

Figure 3 shows the numerical assemblage zones as analysed in Simpson's coefficient. Four breaks shown in Simpson's analysis coincide with four visual assemblage zones delineations at 885 ft, 664 ft,

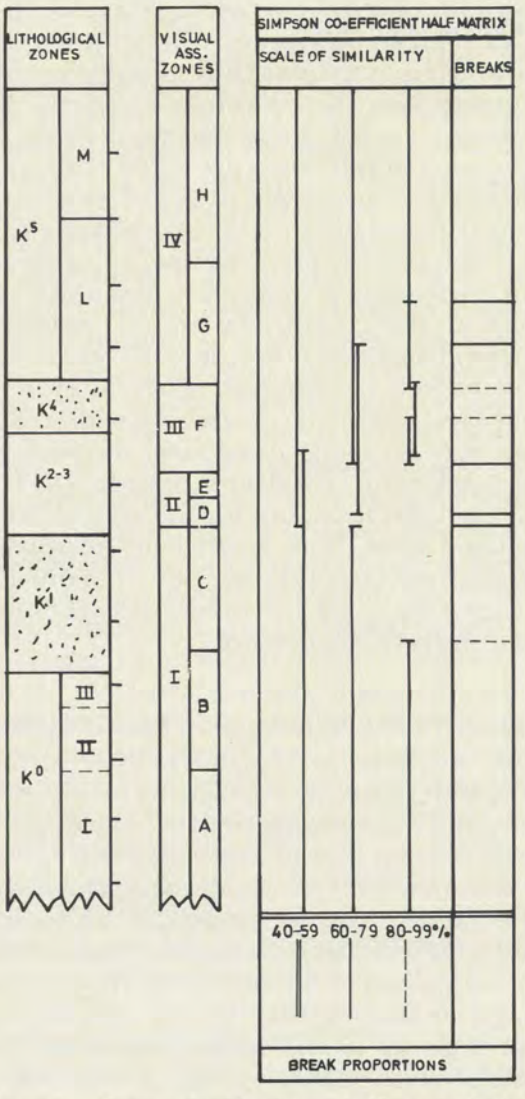


Figure 3. Divisive classification — half matrix shaded.

556 ft and 432 ft, i.e. boundaries between visual assemblage zones, B and C, C and D, E and F, F and G. A further four are possible. These numerical assemblage breaks illustrated are not necessarily the delineations of discrete and separate assemblage zones. The scale depicting the order of importance of the breaks must be taken into account when assessing the comparisons of the actual visual assemblage zones and numerical assemblage zones, and where the major microfloral vegetational changes occur.

To summarise, the shaded half matrix illustrates a major microflora below 664 ft, a transitional zone between 646 ft and 580 ft, and a series of minor microfloras overlapping one another above 556 ft. The degree to which each assemblage zone varies can be defined numerically.

The Simpson coefficient ("total common species to number of species in poorer sample") allows for the comparison of assemblages poor in variety with very rich ones. This prevents the masking of poor samples and serves the useful purpose of dividing the two major floral assemblage zones (NAZ I and II). Thereafter the focal point of the numerical assemblage zones appears to be those samples bearing the richest species variation. Adjacent to these groups of high similarity are samples less varied in content which may exhibit clear and distinct affinity to one group, or a shared affinity to several. By this means large-scale sub-zones and microfloral breaks or overlapping of sub-zones are made obvious.

Simpson's coefficient would therefore seem to be adequate at delineating major assemblages to these zones no matter how poor the species variation. In terms of correlating a random or unknown sample numerically, this coefficient would indicate the major assemblage zone or zones to which it may be most closely compared. Thereafter a more stratigraphically limited comparison within the zone or zones may be made using such methods as the Jaccard coefficient (Sokal and Sneath, 1963). This method has been used, but for the sake of brevity is not included here. Both coefficients adequately illustrate the presence of the previously defined visual assemblage zones and the breaks between them in varying degrees of importance.

(b) Agglomerative Techniques in Classification

Cluster analyses are agglomerative techniques based on high similarity coefficients for the purpose of delineating or grouping similar individuals. *Single Linkage clustering* is one of many methods of clustering. It is initiated by grouping together individuals with the highest related similarity; thereafter by lowering gradually the coefficient value of admission, more members are admitted into the cluster. The single linkage is the means whereby an individual or group of individuals (cluster) are joined. Sokal and Sneath (1963, p. 180) discuss the single linkage methods at length.

If an individual or cluster joins a second cluster on the basis of equal similarity between one member in each, then the inferred group similarity may be too generalised a concept and many members may in fact be quite removed from each other. Similarly, one individual or cluster, when representing a transition microflora, may at the same level of similarity be required to join two separate other clusters not related to each other (except having this common element). This "overlapping" effect is a disadvantage and transitional or repetitive cycles are therefore difficult to illustrate. Where this has occurred (primarily in the youngest floras) a method of cross-hatching has been adopted in this paper.

Further formulae and methods of recalculation have been devised (Sneath, *in* Sokal and Sneath, 1963, p. 181) to overcome this problem, but for the purpose of this paper only the initial single-linkage methods have been employed to test their possible validity, and for confirmation of other methods of zone delineation. In order to test two aspects, the Simpson coefficient values have been applied to samples arranged in stratigraphic order (i.e. *stratigraphically aligned*), whilst the Sokal coefficient values have been used to illustrate the clustering of *randomly distributed* samples.

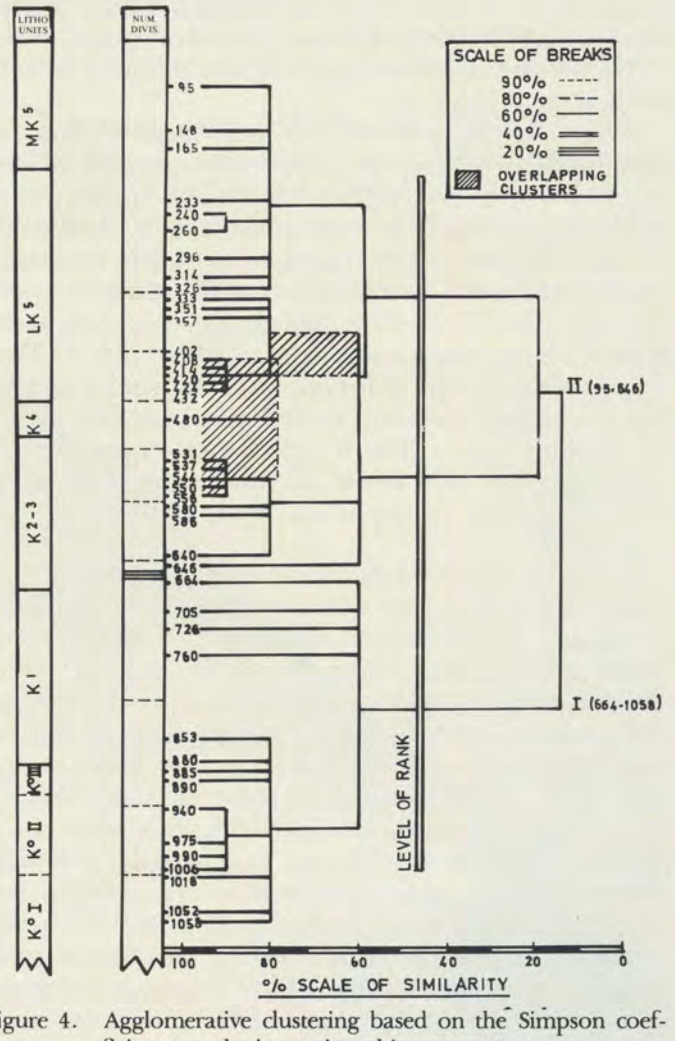


Figure 4. Agglomerative clustering based on the Simpson coefficient samples in stratigraphic sequence.

(a) *Stratigraphically-aligned assemblage clusters*

Single linkage clustering of sample assemblages placed in stratigraphic sequence is illustrated in Figure 4 and the coefficient used is the qualitative one of Simpson. The sample assemblages are placed relative to their depth in the borehole core with an adjacent log indicating the lithological units. The levels at which the clusters join represent the values of similarity which have been obtained using the Simpson coefficient and are therefore taken from the same half matrix of values as the *shaded* half matrix. The scale of similarity is again divided into five units representing 0–100% similarity.

Assemblages with the highest degree of similarity are initially joined, e.g. 90%, followed at a lower level by those sample assemblages with the next highest similarity coefficient value, e.g. 80%. Assemblages, or clusters, which compare with one or more individuals in another cluster are also joined, indicating a generalised grouping or overlapping but in these cases cross-hatching indicates the span of time involved.

The results are shown in Figure 4. There are four clusters at the highest level of similarity (90%) followed by four further clustered groups corresponding to those seen at the 80% level in Figure 2. One transition zone exists at this level between 544 ft and 408 ft and another one between 408 ft and 420 ft at the next level.

The basic *level of rank* has been drawn between 60% and 40% on the scale of similarity. This level was selected to illustrate as simply as possible the clustering of all assemblages into three major groups or floras with the younger two overlapping. The major break occurs between the lower two floras, i.e. between samples 646 and 664. This break agrees with other results already seen in the half matrix. The clusters of Simpson are generalised, wider embracing stratigraphically, and most significant at one level (80%, a relatively low degree of similarity). This clustering and series of breaks agrees with those of the half matrix, but adds a *two-dimensional* aspect of magnitude to the breaks.

(b) *Randomly distributed clusters*

Single linkage clustering by means of randomly distributed assemblages is illustrated in Figure 5. Sokal's coefficient was used in this technique, to test once more the validity of sample groupings based on both a different coefficient and method of clustering. The values are plotted on a scale of dissimilarity indicating the greater the coefficient value the larger the dissimilarity between assemblages. For the purposes of unbiased comparison, the assemblages are numbered (see Figure 6).

The method of progressive addition involves the linking together of random sample assemblages bearing the lowest dissimilarity values, and thereafter admitting usually singly assemblages with progressively increasing dissimilarity.

The initial groupings of related assemblages occurs at a level of 2.8 and progresses by single linkage admittance of one or more clusters up the scale of dissimilarity. *Level of rank* is arbitrarily taken as being 4.4. Above this, at level 4.7, three entities emerge, one random assemblage (21) and two large and distinct clustered microfloral groups stratigraphically separated with a large break above assemblage 28. At the basic rank level, two major clusters and five single assemblages join. The two major assemblage groups represent (1) one biostratigraphic entity, NAZ I (assemblages 29–34 inclusive, or 664–1 058 ft), and (2) one near biostratigraphic entity, NAZ II (1–27, or 95–640 ft), excluding assemblages 21, 22, 23 and 27 which have been regarded elsewhere as a separate group. Assemblage 28 is quite unique and stands alone, joining above the basic rank level to the NAZ I. This assemblage has previously been illustrated as representing a lone transition position, but eventually joining the upper limit of NAZ I at a low level of similarity. The two largest breaks therefore occur below and above assemblage 28. Breaks between these assemblages are regarded as similar to one another but marginal to the major NAZ II.

Within the two large assemblage groups, NAZ I and NAZ II, seen at basic level or rank, smaller groups of equally clustered assemblages show the existence of mutually related floras, the most notable of which in NAZ I is assemblage 35–43, inclusive, i.e. 885–1 058 ft, (i.e. VAZ A and B, found in the Dwyka interglacial sediments). Lower orders of dissimilarity occur in adjacent assemblages which agrees with the transitional nature of NAZ C.

In NAZ II, grouping of assemblages is more random and shows a reorganisation of stratigraphic levels in an attempt to show similarity groupings. At 4.2 a large number of assemblages and clusters join showing a broad constant level of dissimilarity. Subgroups below this do, however, possess assemblages almost adjacent to one another and linked progressively up the stratigraphic column with minor readjustments in the stratigraphic order.

Sub-division at level 4.0 shows low dissimilarity (high similarity) between assemblages 2, 3 and 4 and 11, 12, followed closely by assemblage 7, then 8 and 9, thereby forming a group of assemblages 2, 3, 4, 7, 8, 9, 11 and 12 leaving out assemblages 1, 5 and 6 which are separately grouped together. Such separation may be related to reworking of older sediments bearing similar assemblages, or by cyclic repetition of a similar climate and ecological environment. Further sub-groupings are illustrated at 4.0, with one significant group, 25 and 26, at 3.3. These assemblages are shown to be associated again in ordination classification.

Where assemblages are linked closely, but in fact are not adjacent to one another stratigraphically, the possibility of similarity caused by reworking of sediments or climatic oscillations must be considered carefully. These linkages, although relatively infrequent, may therefore be due to a number of

FIG. 5.

SINGLE LINKAGE CLUSTERING BASES ON THE SOKAL COEFFICIENT - RANDOM SAMPLES JOINED (PROGRESSIVE ADDITION).

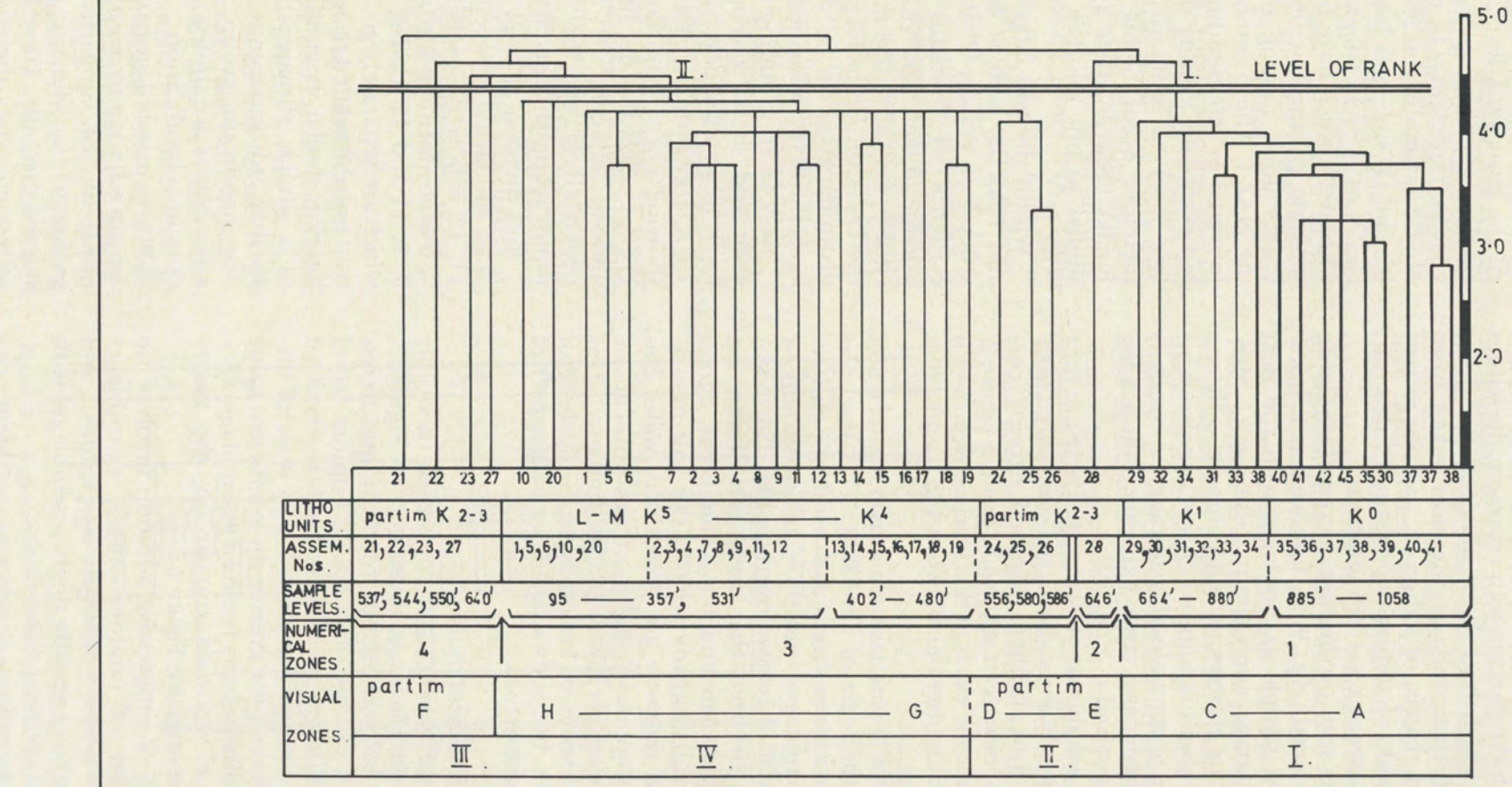


Figure 5. Single linkage clustering bases on the Sokal coefficient — random samples joined (progressive addition).

factors which must be taken into consideration in the final interpretation of this method of analysis.

(c) Summary

A summary of the assemblage boundaries using agglomerative sub-division may be seen in Figure 6.

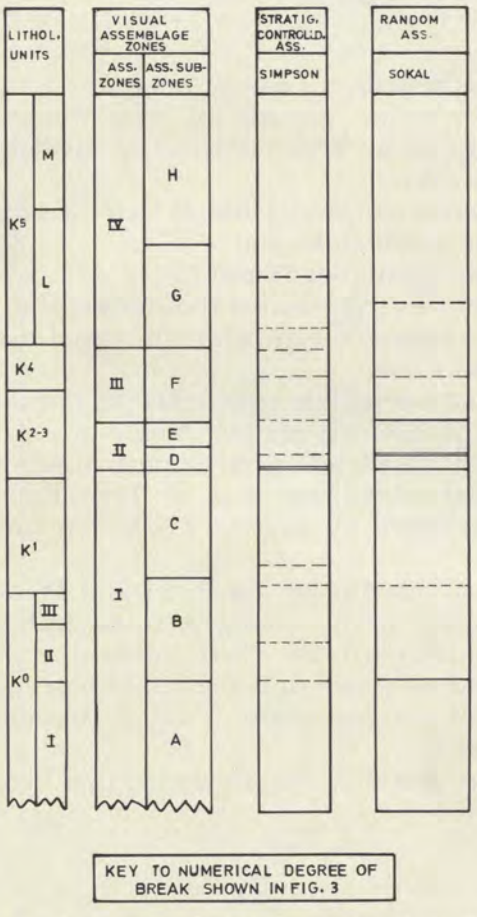


Figure 6. Agglomerative sub-division: single linkage clustering.

(1) *Stratigraphic distribution of assemblages*

The primary advantage of stratigraphically controlled clustering methods is that it allows for quick checking as to the natural groupings of the samples within the lithostratigraphic context and within a two-dimensional background. The disadvantages involve the difficulty in illustrating the distributed affinities of transitional assemblages to the major floras.

In comparison with the visual assemblage zones, the delineations of assemblage sub-zones A–H are moderately well defined by the Simpson coefficient. Further detailed definition of the sub-zones are deemed potentially possible using more refined methods in the future.

(2) *Random distribution of assemblages*

Sokal's linear coefficient used in single linkage clustering of random assemblages shows good stratigraphic grouping of similar assemblages. The four major groups seen at a standard rank level of 4.4 virtually coincide with the four major assemblage zones obtained by visual means. Breaks between these zones are noted with varying importance in the stratigraphically aligned cluster analyses. Further sub-divisions in the younger zones (above 28) are based on minor groupings of stratigraphically disorientated assemblages which may show affinity to each other due to cyclic factors as discussed previously. This re-arrangement of the assemblages is not possible as is clearly illustrated in earlier methods.

The advantage of this method is the unbiased approach to clustering, based solely on numerical levels of dissimilarity. The distinct numerical assemblage zonal breaks seen here confirm more positively the major natural visual groupings already proposed. The slightly disorientated minor groups of assemblages seen above 28 may well prove useful in ecological studies.

On summarising the agglomerative techniques of sub-division, emphasis must be placed on the random distribution method (Sokal coefficient). This gives (1) an unbiased, two-dimensional grouping of similar assemblages at high levels of dissimilarity; (2) it surpasses the clumsy nature of illustrating transitional assemblages (as in stratigraphic methods); and (3) natural groupings of randomly distributed low-level dissimilarity groups may allow for the study of palaeogeographical or palaeoecological factors common to all of them. This method warrants further study.

The advantages of the stratigraphically aligned agglomerative procedures are based on the natural microfloral changes up the stratigraphic column which may be correlated with lithological factors. This method may only be used over long time intervals, involving thick sequences. Narrow stratigraphic limits would cause undue overlapping and difficulty in the clustering of similar randomly distributed assemblages.

In terms of comparison with the visual assemblage groups, the Sokal and Simpson coefficients illustrated by these agglomerative techniques both confirm the boundaries of the major assemblage zones and most sub-zone delineations (with the possibility of even further divisions).

SPATIAL ORDINATION — A METHOD OF NON-HIERARCHICAL CLASSIFICATION

The spatial relationships between compared assemblages requires a coefficient of distance based on linear geometrical theories. For this method of classification, termed spatial ordination or clustering, Sokal's coefficient was chosen (see Methods of Analysis).

Two methods of illustrating spatial grouping amongst similar assemblages are employed:
(1) — Single axis ordination, and
(2) — Spatial or double axis ordination.

(a) Single axis ordination

This is a simple graphical method or ordination based on the relationships of all assemblages to two

selected reference samples (assemblages) placed along one axis. The assemblages chosen should possess the highest dissimilarity value when compared with each other. The axis is constructed with the length representing the distance between these reference samples. Thereafter the position of each sample with reference to the standard pair is found by geometrical construction and a perpendicular dropped from this point on to the axis. The positions of the samples along the axis are thus proportional to their dissimilarity distance from each reference sample. If the axis chosen corresponds to *time* and the reference samples are sited at opposite ends of a stratigraphic sequence, then samples falling at the same point on the same axis are indicated to be of a similar age (Mortimer, 1969). By inference this method therefore picks out stratigraphic relationships among the samples. If, however, the reference samples are chosen due to the greatest difference in their *qualitative* content (i.e. where the highest value of dissimilarity lies), then samples falling into specific groups along the axis may indicate changes in and the presence of natural floral zones. These "zones" may coincide with and even further resolve those assemblage zones already created on a time basis.

The half matrix containing the values of the Sokal coefficient may be seen in Figure 7; the samples are numbered once more for convenience. Two reference pairs of assemblages were chosen to illustrate the single axis ordination, i.e. (a) samples 1 and 39; and (b) 14 and 38. The first pair was chosen due to their *positions* at virtually oppostie ends of the stratigraphic sequence and in order to illustrate any potential clusterings due to the changes through time; the second pair was selected due to the greatest difference in their qualitative contents. Any clustering relative to this pair would hopefully indicate additional changes in flora to those seen in the first example.

(a) *Reference samples 1 and 39 — Figure 8*

Three major groups of assemblages occur, stratigraphically and numerically divided by two major breaks:

(1) Between samples 24 and 25 (sample 27 is regarded as a weak link), and
(2) between samples 28 and 29.

Therefore three distinct biostratigraphic units are shown which coincide with the visual assemblage zones as follows:

(1) Visual assemblage zone I (Dwyka flora) is seen in group 29–40, with one minor group differentiated (37–40) which corresponds to sub-zone B;
(2) Visual assemblage zone II Transition flora is represented by samples 25–28 (bar sample 27); and
(3) Visual assemblage zones III and IV are incorporated in the youngest numerical group of assemblages 1–24. Three minor groups may be noted, of which only the group bearing samples 20–24 corresponds to a visual assemblage zone (zone III).

It may therefore be concluded that through this

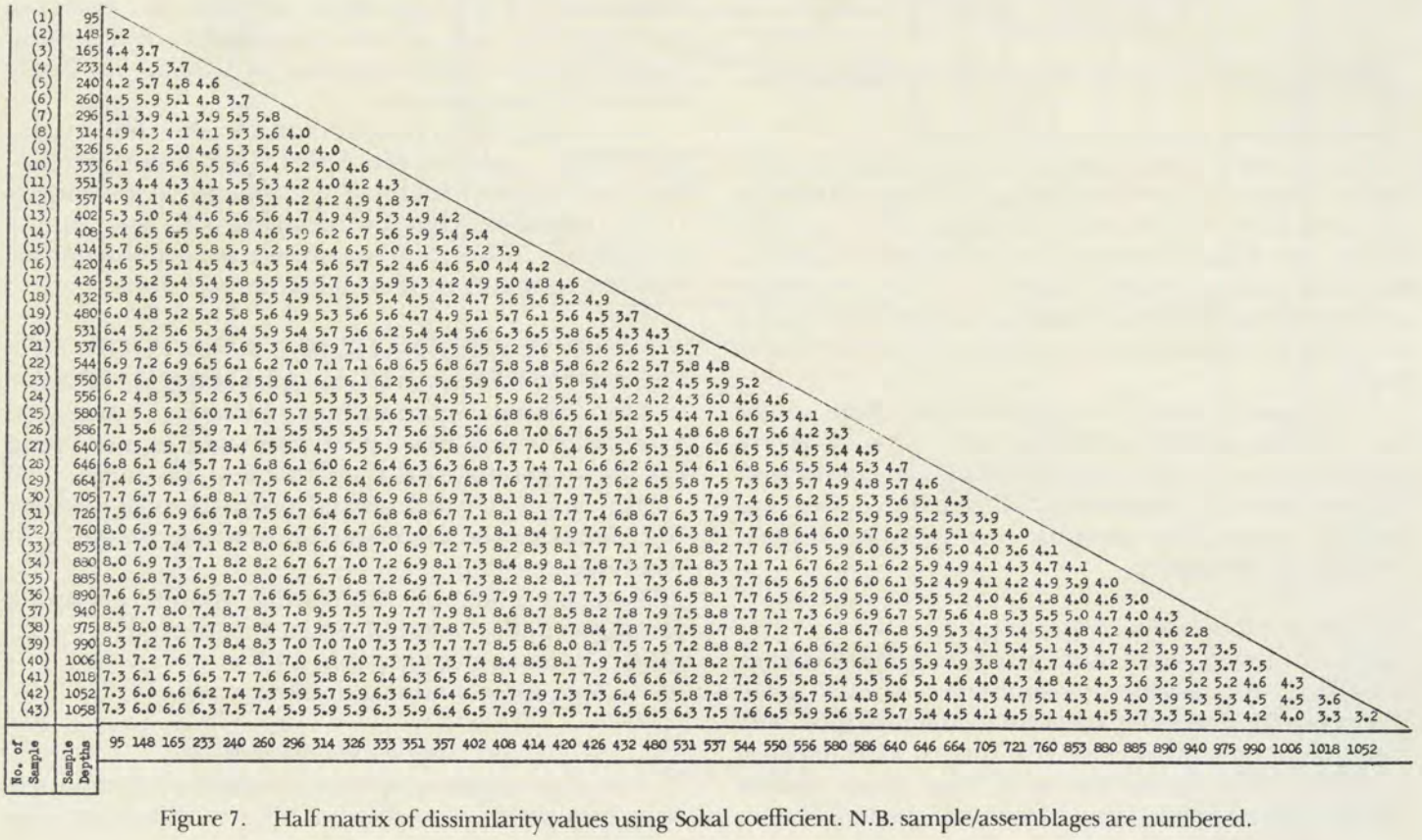
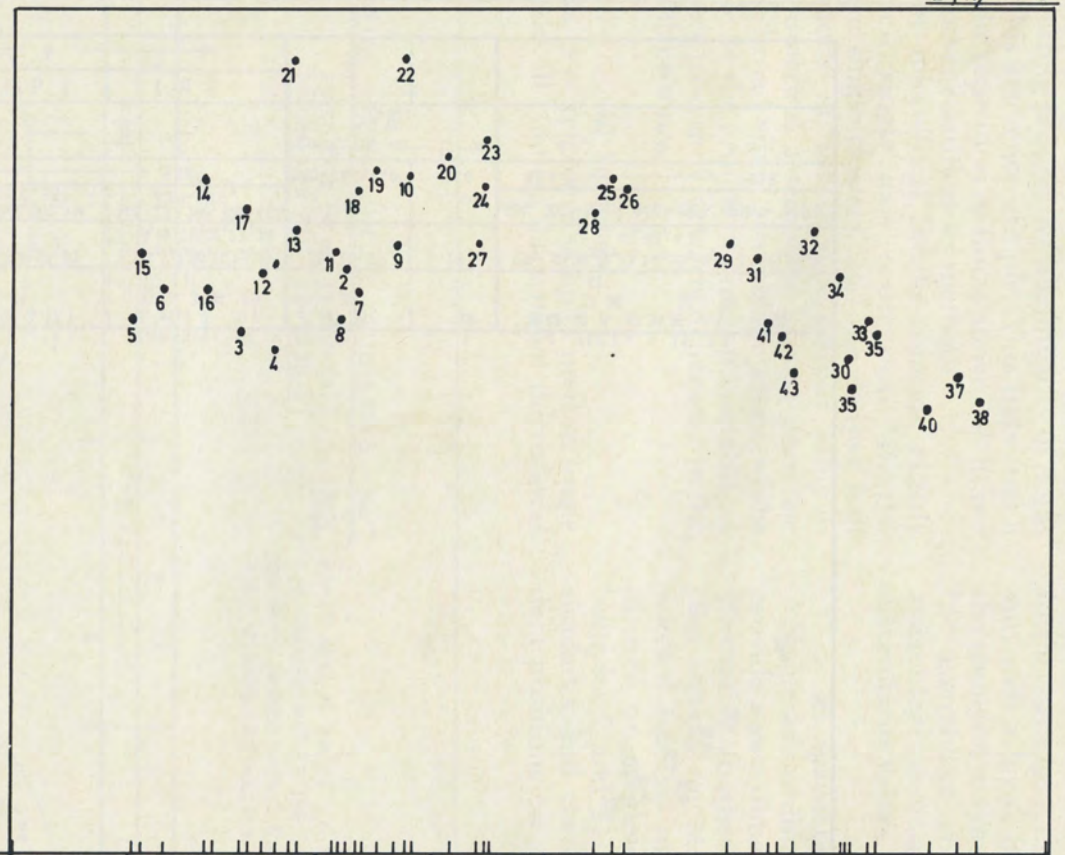


Figure 7. Half matrix of dissimilarity values using Sokal coefficient. N.B. sample/assemblages are numbered.

FIG. 8. — SINGLE AXIS ORDINATION – REFERENCE SAMPLES 1:39 (95 :990') (DISSIMILARITY VALUE Djk) = 8·3



SAMPLES PROJECTED ON TO REFER. AXIS.	① 5 6 14 3 12 13 11 7 19 9 20 27 / 15 16 17 4 21 8 18 10 24 / 2 22 23		28 26 / 25	29 31 43 32 34 33 40 37 38 ③⑨ / 42 30 35 / 43 36
SAMPLES RE-GROUPED.	3, 4, 5, 6, 12, 13, 14, 15, 16, 17, 21. 2, 7, 8, 9, 10, 11, 18, 19, 22. 20, 23, 24, 27.		25, 26, 28.	29, 30, 31, 32, 33, 34, 35, 36, 41, 42, 43. 37, 38, 39, 40.
STRATIGRAPHICAL DISTRIBUTION	165'– 260'; 357 – 426; 537, 148, 296–351, 432–480, 544, 531, 550, 556, 640		580–586, 646'	664' – 890; 1018' – 1058' 940' – 1006'
	95 → 556', 640			664' → 1058
VISUAL ASS. ZONES.	IV. —— III.		II.	I.
NUMERICAL ASS. ZONES.	(c) 3. (b) (a)		2.	1.

Figure 8. Single axis ordination — reference samples 1:39 (95:990') (dissimilarity value Djk) = 8,3.

period of time at least three distinct floras were seen to have occurred with marked breaks between them.

(b) *Reference samples 14 and 38 — Figure 9*

These reference samples, bearing abundant species variation in both, show most clearly six demarkated assemblage groups. The major breaks occur between samples 24 and 25, between 27 and 28, and between 28 and 29, and below 646 ft with minor breaks occurring between three assemblage groups above sample 24. The latter three, unlike the first three, are adjacent.

The numerical groups 1, 2 and 3 are stratigraphically clearly divided and represent the Visual Assemblage Sub-zones A, B and C (=NAZ I); VAZ D (=NAZ II) and VAZ E (=NAZ III). Above

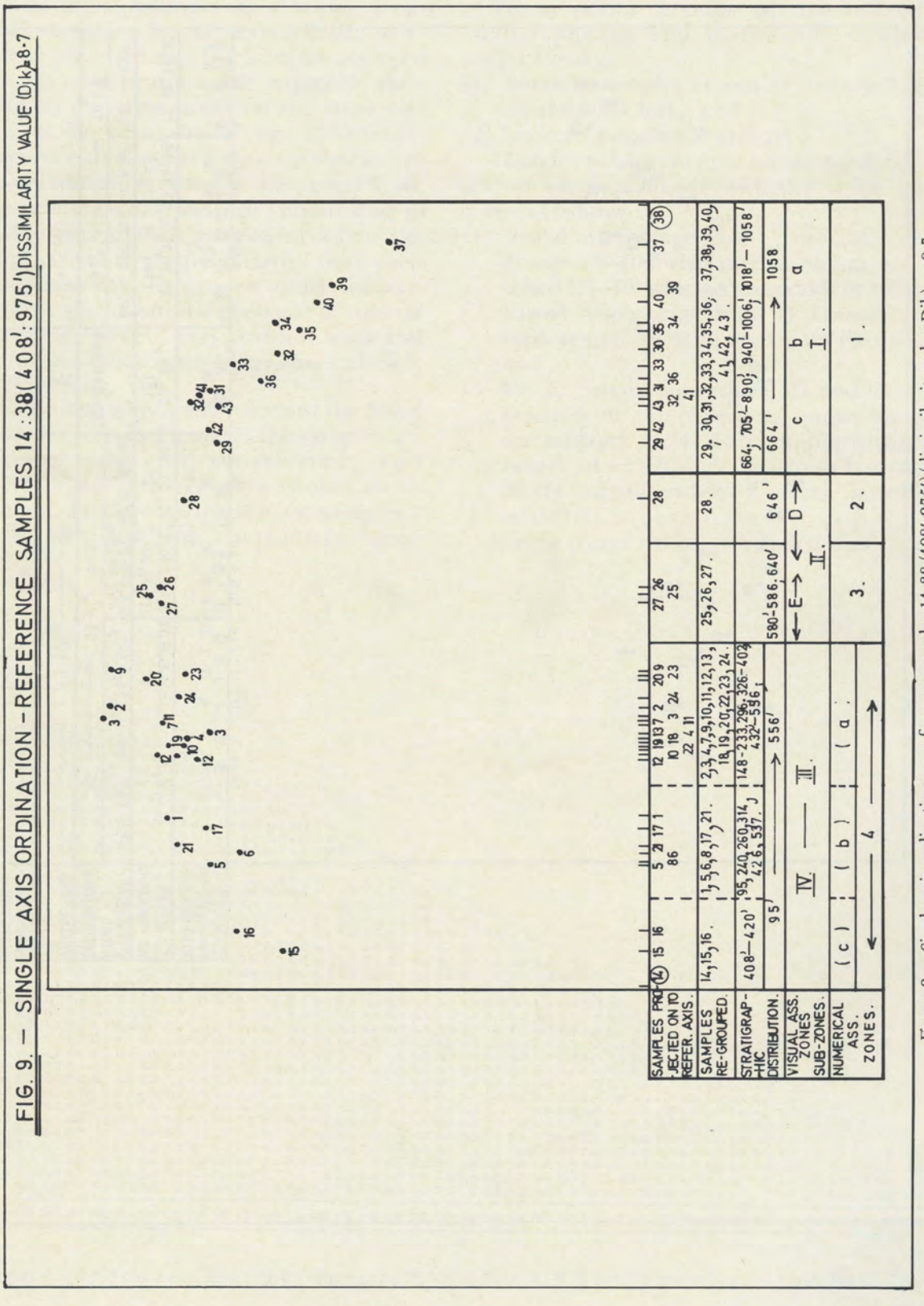


Figure 9. Single axis ordination — reference samples 14:38 (408':975') (dissimilarity value Djk) = 8,7.

sample 24 (556 ft) the assemblages are randomly arranged in three groups of proportional dissimilarity relative to the reference assemblages. NAZ (C) with samples 14–16 are stratigraphically adjacent and very similar, whilst groups (a) and (b) possess assemblages older and younger than those in (c) falling into two marked groups of dissimilarity. When stratigraphically divided into concentric bands of dissimilarity above and below assemblage NAZ (c), the random assemblages fall into similar associated assemblage groupings to those of the single linkage clusters, indicating progressive dissimilarity of floras away from this reference sample. This indicated a potential cyclic repetition of floras above 586 ft, not indicated in other classificatory methods.

Two major microfloral breaks coincide exactly with the first single axis ordination (a), the agglomerative single clusters, and the divisive half matrix methods of sub-division. From samples 25–40 the numerical assemblage groups correlate well with the broad visual assemblage zones, whilst above sample 24 the groups of random assemblages form a distinct pattern not clearly delineated in the visual assemblage zones.

(b) Summary

The two pairs of reference samples used in single axis ordination all confirm three major microfloral zones with large-scale breaks between samples 24 and 25 (i.e. 556 ft and 580 ft) and between samples 28 and 29 (i.e. 646 ft and 664 ft).

Three of the major visual assemblage units are thus well defined with visual assemblage zones III and IV apparently absorbed into the youngest numerical zone. Below the oldest microfloral break an assemblage occurs which has the possibility of further sub-division with different reference samples.

In the middle numerical assemblage zone, two further stratigraphically separated assemblages occur, with a break between 640 ft and 646 ft, whilst the youngest major zone (NAZ III) shows four possible assemblage groupings composed of random assemblages associated together by means of similar dissimilarity values.

(c) Spatial or double-axis ordination

Single axis ordination may be further extended to provide a two-dimensional arrangement of samples illustrating another method of similarity grouping. (See section on Methods.) The reference samples selected for this method are those previously ordinated on single axes, as well as a further pair (samples 7 and 11) selected due to high degree of similarity and therefore occupying almost coincident positions on a previous axis. This single axis

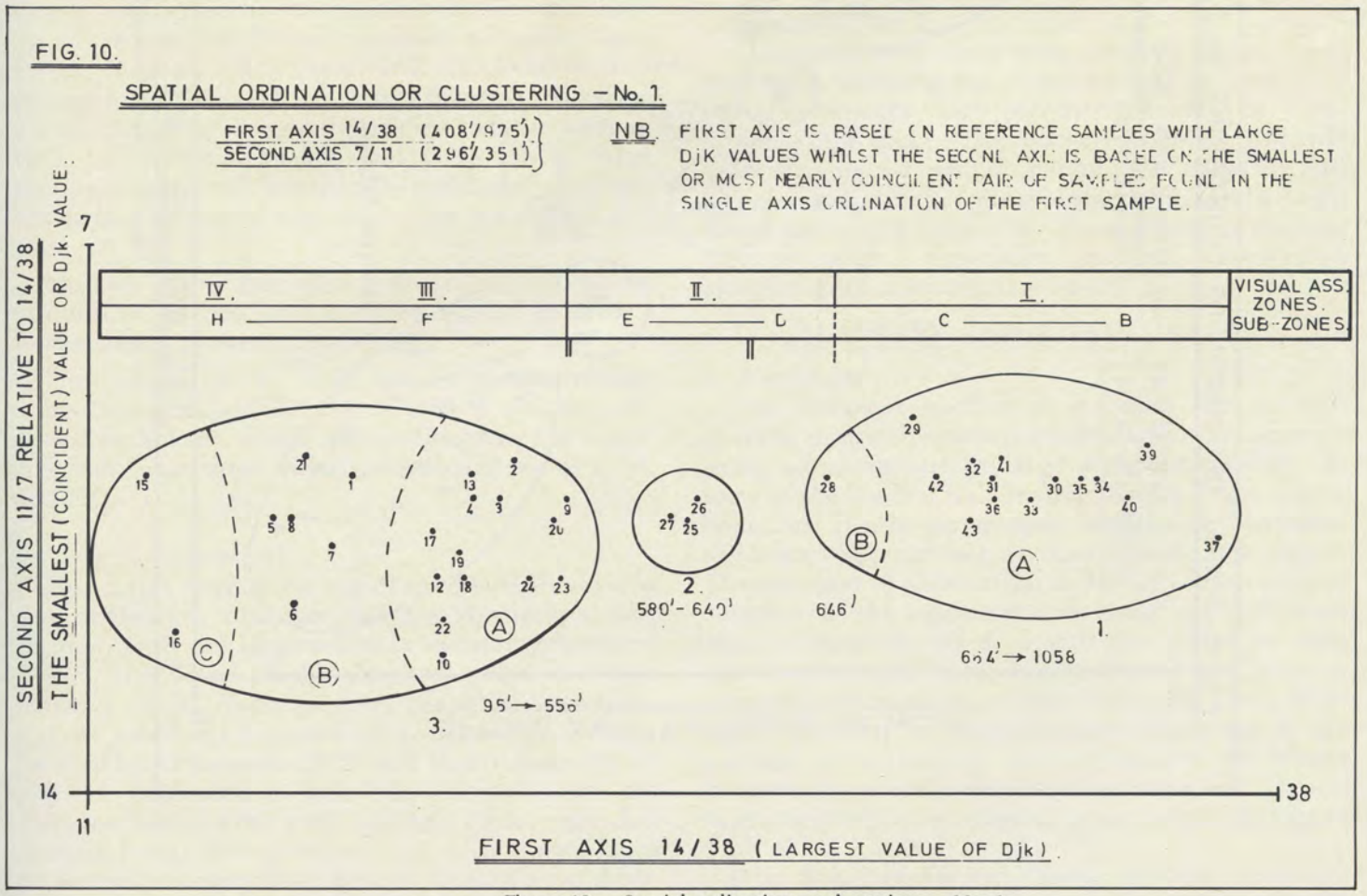


Figure 10. Spatial ordination or clustering — No. 1.

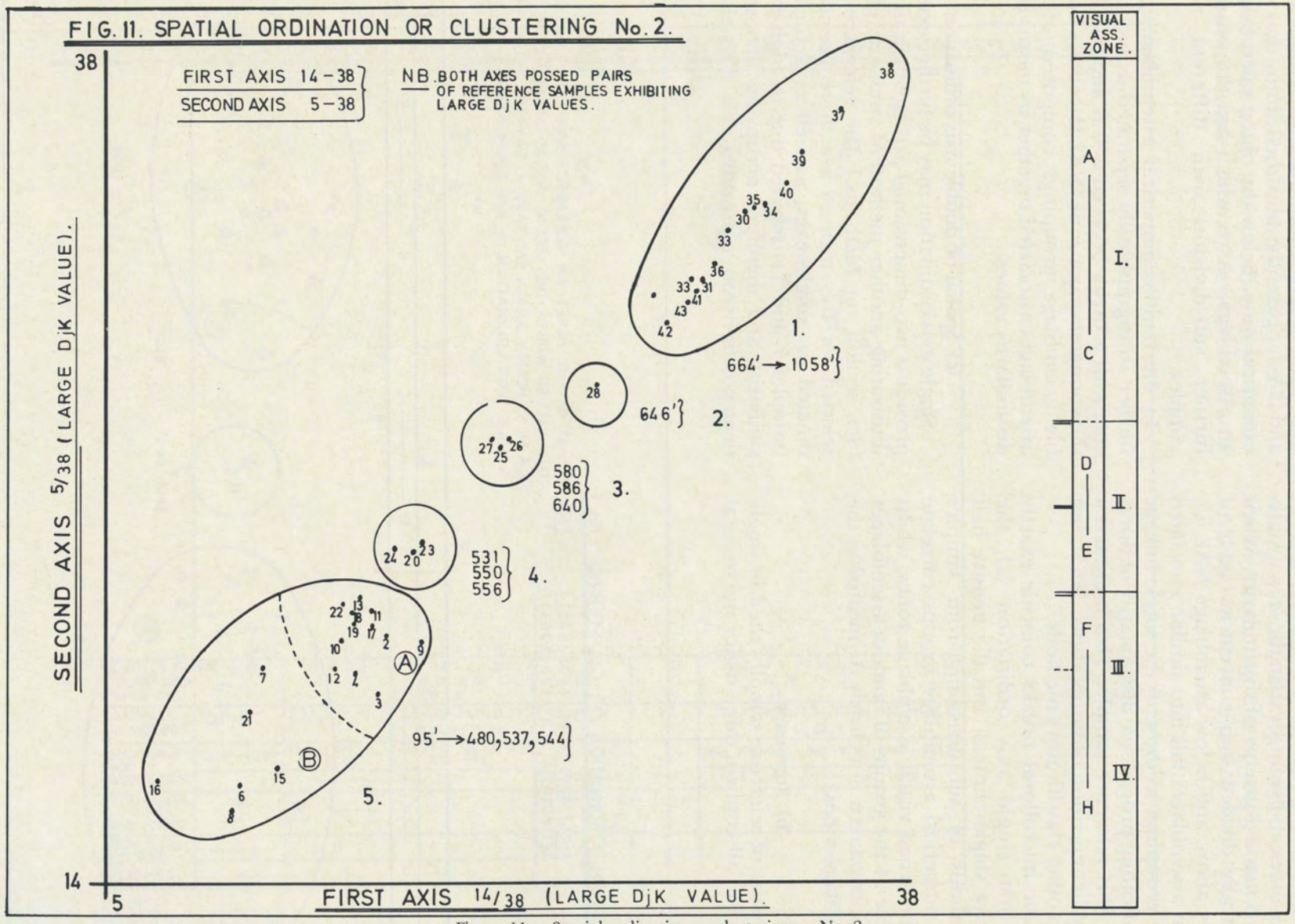


Figure 11. Spatial ordination or clustering — No. 2.

ordination is not illustrated here. The purpose in choosing such wide ranging reference samples is to illustrate the maximum possible variations relative to these.

In both examples of spatial ordination the results of two single axes ordinations are arranged as co-ordinates, with each sample assemblage then located relative to both axes, and perpendicular to points on those axes.

(1) *Spatial ordination number 1* (Figure 10) shows the result of clustering or ordination of samples based on reference samples 7 and 11, and 14 and 38. Reference pair 7 and 11 were chosen as the closest coincident pair of samples, and therefore the most similar relative to reference samples 14 and 38 on a single axis ordination not shown here.

(2) *Spatial ordination number 2* (Figure 11) was constructed using reference samples 14 and 38, and 5 and 38. Both axes possess pairs of reference samples exhibiting large dissimilarity values. This variation from no. 1 was attempted in order to explore the possibility of any further microfloral patterns.

Results

(a) *Spatial ordination no. 1*

Three broad microfloral assemblage groups are clearly visible, each stratigraphically divided from the other two. NAZ I is comprised of assemblages 28–43, NAZ II, 25–27, and NAZ III, 1–24. Within groups I and III two further sub-divisions based on a transitional or marginal position may be seen — NAZ I—28, from 29–43; and NAZ III — three groups of randomly assorted assemblages which are not stratigraphically aligned, except for 14, 15 and 16.

The two major microfloral breaks occur between samples 24 and 25, and between 27 and 28, with a further break between 28 and 29.

The major breaks and numerical assemblage zones thus grouped are closely similar to those obtained in Sokal's single axis ordination, and coincide well with three visual assemblage zones I, II, and III + IV.

(b) *Spatial clustering No. 2*

Five major microfloral zones are distinctly clear in this ordination diagram; all are stratigraphically divided, with the largest breaks occurring between 24–25, 27–28 and 28–29. One minor break occurs between 19–20. In comparison to spatial cluster no. 1 more emphasis is placed on the more prominent form of break between 28/29 and assemblages 20, 23 and 24 represent a transition assembly to the youngest major flora 1–19, (21 and 22 are regarded as weak links). The groups of random assemblages in numerical assemblage Zone 5 are similar to those already shown by other methods.

(d) Summary
(Figure 12)

The clustered groups of assemblages in these two sets of spatial ordination graphs shows a convincing degree of clustering separated by significant breaks. Three major numerical assemblage zones are obvious (with a possible fourth) which are comprised of stratigraphically adjacent samples in the borehole core. They may therefore be termed biostratigraphic units. Major breaks occur between assemblages 24/25, 27/28 and 28/29.

Minor variations occur, for example, in:

(1) assemblages 14, 15 and 16. These are stratigraphically grouped together in single ordination No. 1, and not so evidently in No. 2. This small biostratigraphic unit occurs at the base of visual assemblage zone G.

(2) assemblages 1–24 are divided into three sub-groups showing repetitive floras as seen already in single ordinations No. 1 and 2, but with slightly different magnitudes. Assemblages 25–27 coincide in both. The older assemblages 29 to 40 are potentially further sub-divided into two, varying again in degree of magnitude.

(3) one sample assemblage, 28, is spatially equidistant from both adjacent assemblage groups showing a separate, transitional microflora. In spatial ordination No. 1, this assemblage is a sub-group (B) in NAZ I, but it is sufficiently distant to warrant its own sub-group in spatial ordination No. 2.

This method of illustrating numerical division is successful, showing the positions and groupings of the assemblages into distinct clusters in two-dimensional space. The numerical zones coincide with the broad visual assemblage zones in all respects with the exception of the boundary between visual sub-zones G and H. These methods warrant further study, especially in the "associated assemblages" of assemblages 1–24.

SUMMARY AND CONCLUSIONS

1. *Coefficients*

The Simpson coefficient provides for the upgrading of poorly stocked assemblages. This appears most advantageous in large scale subdivision. As seen in the shaded half-matrix (Figure 2) the major floras are clearly demarcated, whilst the *relationships of all assemblages* to one another is seen at a glance. This method of illustration is, for this same reason, superior to the agglomerative clustering technique seen in Figure 4. Here, similar groupings are seen but transitional floras and assemblages bearing complex relationships are not indicated. The advantage, however, in single linkage clustering is the method of portraying the *magnitude* of the breaks between the numerical assemblage zones at a glance. (See summary of numerical zone breaks in Figure 12.)

The Sokal coefficient, being a linear measure, allows for a one and two dimensional spatial or-

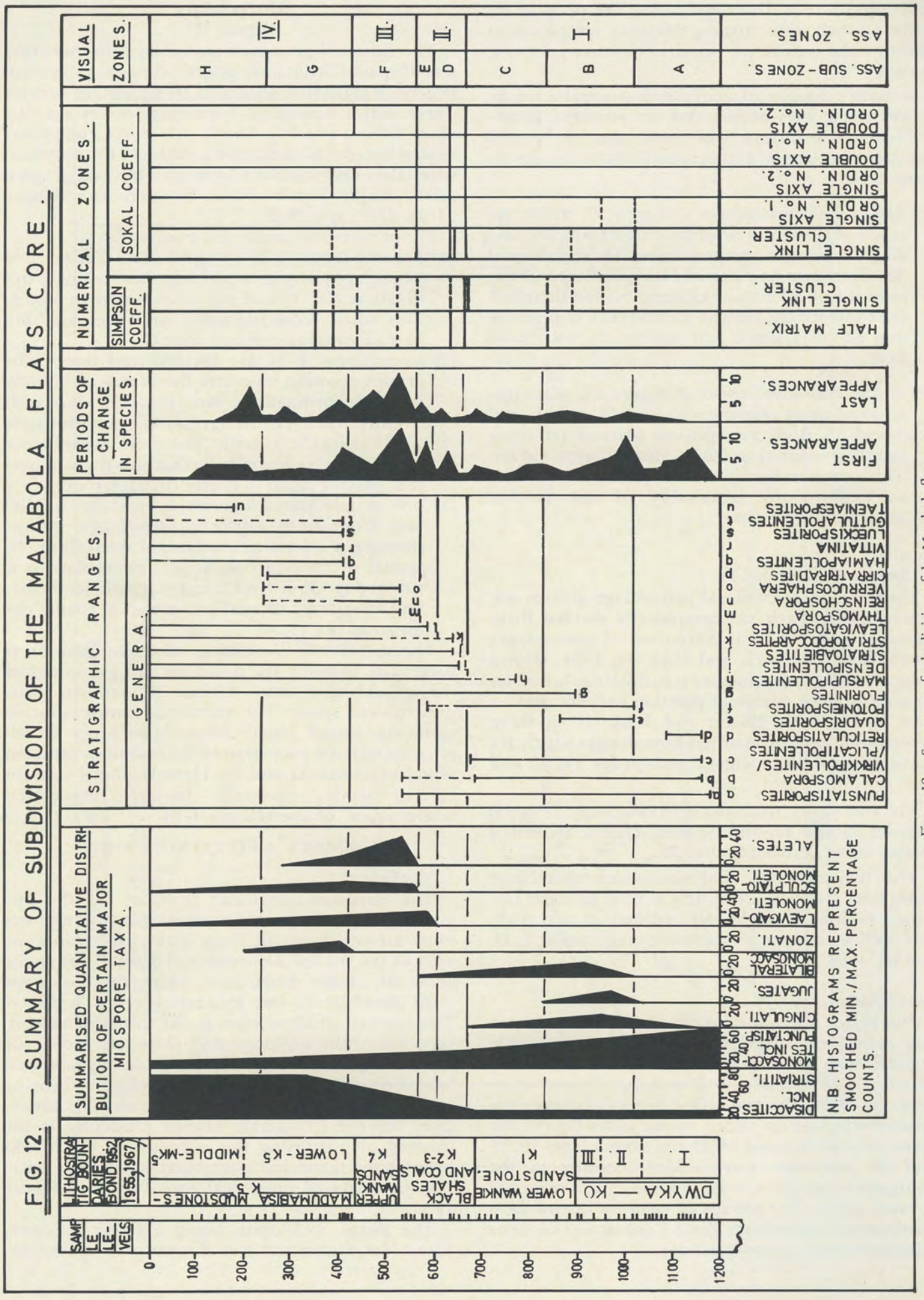


Figure 12. Summary of sub-division of the Matabola flats core.

dination of similar assemblages and ease in clustering of randomly distributed assemblages. Furthermore it shows more clearly the existence of recurrent or repetitive floras within the assemblage group 1–24 not clearly seen elsewhere.

2. *Parameters*

Despite all possible combinations of the presence/absence data (i.e. positive and negative matches, common or common to only one sample; linear and non-linear coefficients; agglomerative, divisive and non-hierarchical techniques of illustration; and spatial ordinations employing various parameters such as maximum time or maximum quantitative values), it appears that the major floral zones always occur distinctly and with considerable definition.

The conclusion drawn is that the floral changes are of sufficient magnitude to mask almost all variations that occur in numerical techniques such as these.

3. *Correlations*

In comparison with the (visual) palynological biostratigraphic zones and sub-zones of Falcon (1975) (see Figure 12), the major *numerical* zones appear to conform to the major *visual* zones, with the major numerical breaks occurring during deposition of the Black Shales and coals (K^{2-3}). Minor breaks in the Dwyka (K^{0}) Lower Wankie Sandstone (K^{1}) Flora are consistent with floral changes during and immediately after a period of glaciation, whilst a minimum of changes are noted during the deposition of the youngest sediments, the Madumabisa Mudstones (LK^{5}, MK^{5}).

The major points at which first-appearances and last-appearances of species occur, together with selected quantitative and qualitative data obtained visually (Falcon, 1975), are illustrated in conjunction with the boundaries of the numerical assemblages in Figure 12. In comparing the biozones erected by Hart and Fiehler (1971) in Tanzania and those illustrated here (based on species ranges and frequency data), it becomes obvious that the major change seen by Hart in his transition zone (based on points of inflection for the regression equations of *Cordaitina* and *Protohaploxypinus*) appears to be equivalent to that seen as the major numerical break between assemblage zones I and II in the basal Black Shales and Coals.

4. *Sub-division*

In terms of sub-dividing a single stratigraphic sequence into practical delineated assemblage zones, it is concluded here that visual methods (analysing qualitative and quantitative date for various ranks of miospore taxa) are equally reliable to numerical methods (processing only presence/absence data of all species) when both forms of analyses are taken as a whole. This bears out the conclusions previously reached by Gordon and Birks (1972) in Quaternary palaeoecology.

Visual methods are easier, quicker and more direct to process than the numerical methods (when computer facilities are limited), but they possess the disadvantage of subjective bias. Numerical methods are essentially objective, unbiased and well defined, allowing for a further advantage, that of two-dimensional projection of assemblages into groups where magnitude of the floral change is measured. Repetition of similar microfloral assemblages at random stratigraphic levels, i.e. cyclic or repetitive assemblages, is also better illustrated by numerical methods. This may allow for certain palaeoecological, climatic or depositional (reworking) influences to be studied.

At present, visual methods of microfloral analysis are prevalent in Gondwana palynological studies but it is hoped that this paper, along with Hart and Fiehler (1971) and Christopher and Hart (1971), may serve to introduce numerical methods for further development in detailed stratigraphic analysis. This may be particularly so when a broader and more universal systematic scheme has been accepted, and numerical methods based on more reliable and well-circumscribed taxa have been further refined and thereafter projected on to a more sophisticated computer basis.

ACKNOWLEDGEMENTS

I am gratefully indebted to Prof. W. G. Chaloner (Birkbeck College, London) and Dr. B. Goldsmith and Dr. M. Mortimer (both of London University) for their guidance and help in initiating this form of study and this paper. Furthermore, my thanks are due in no small measure to Dr. A. R. I. Cruickshank, Dr. P. Fatti, Dr. R. Pienaar, Dr. B. Turner and Dr. J. T. Brown for reading the manuscript and for much fruitful criticism and discussion. For assistance in presentation my sincere thanks are also due to Mrs. R. Bourhill and Mr. Bob Foster.

REFERENCES

BALME, B. (1964). *The Palynological record on Australian Pre-Tertiary Floras.* In *Ancient Pacific Floras,* University of Hawaii press.

CHRISTOPHER, R. A. and HART, G. F. (1971). A statistical model in Palynology. *Geoscience and Man*, **3**, 49–56.

DALE, M. B. (1971). *Information Analysis of Quantitative Data.* In *Statistical Ecology*, 3, Pennsylvania, State Univ. Press., 133–148.

———— and Walker, D. (1970). Information Analysis of Pollen Diagram I. *Pollen et spores,* **12(1)**, 21–37.

EVANS, E. R. (1967). Upper Carboniferous and Permian palynological stages and their distribution in Eastern Australia. *Bureau Mineral Resources (Australia),* Report No. **1967/99**.

FALCON, R. M. S. (1975). Palynostratigraphy of the Lower Karroo Sequence in the Central Sebungwe District, Mid-Zambezi Basin, Rhodesia. *Palaeont. afr.,* **18**, 1–29.

GORDON, A. D. and BIRKS, H. J. B. (1972). Numerical Methods in Quaternary Palaeoecology — I Zonation of Pollen diagrams. *New Phytologists,* **71**, 961—979.

HART, G. (1967). Micropalaeontology of the Karroo Deposits in South and Central Africa. *In* I.U.G.S. Reviews prepared for the first symposium on Gondwana stratigraphy, 1967. Haarlem, Netherlands, I.U.G.S. Secretariat; 161–172.

———— (1970). The stratigraphic subdivision and equivalents of the Karroo sequence as suggested by Palynology. *In* Proc. 2nd I.U.G.S. Symposium on Gondwana Stratigraphy and Palaeontology. Cape Town and Johannesburg. S. H. Haughton, Ed. Pretoria, C.S.I.R.

———— and FIEHLER, J. (1971). A computer-aided storage and retrieval system for Lower Karroo Stratigraphic palynology. *Geo-science and Man,* **3**, 61–64.

HUGHES, N. F. and MOODY-STEWART, J. C. (1969). A Method of Stratigraphic Correlation using early Cretaceous spores. *Palaeontology,* **12(1)**, 347–358.

MORTIMER, M. G. (1969). *Devonian Spores of Southern Britain.* Ph.D. thesis, University of London, Birbeck College.

SEGROVES, K. (1970). The Sequence of Palynological Assemblages in the Permian of the Perth Basin, Western Australia. *In* Proc. 2nd I.U.G.S. Symposium on Gondwana Stratigraphy and Palaeontology. Cape Town and Johannesburg. S. H. Haughton, Ed. Pretoria, C.S.I.R.

SOKAL, R. R. and SNEATH, P. H. A. (1963). *Principles of Numerical Taxonomy.* Freeman and Co., London.

WILLIAMS, W. T. and DALE, M. B. (1965). Fundamental Problems in Numerical Taxonomy, Advances in Botanical Research. Academic Press, London, New York. **2**, 35–67.

———, LAMBERT, J. M. and LANCE, G. N. (1966). Multivariate Methods in Plant Ecology. Similarity Analyses and Information Analysis. *Journal of Ecology,* **54(2)**, 427–444.