

Merlewood Research and Development Paper
Number 15

PRINCIPAL COMPONENT ANALYSIS OF PHYSICAL AND
CHEMICAL DATA FROM A PILOT SURVEY OF MORECAMBE BAY

J. N. R. Jeffers, FES

R. & D.69/15

Merlewood Research and Development Papers are produced for the dissemination of information within the Nature Conservancy. They should be regarded as internal documents only, and must not be quoted or referred to in formal publications. All opinions expressed in Merlewood Research and Development Papers are those of the authors and must not be taken as the official opinion of the Nature Conservancy.

Introduction

During 1968, a pilot survey of the physical properties of the sands and muds in Morecambe Bay was undertaken as part of a study of the possible effects of the construction of a barrage across the Bay. Sample cores, from two different zones, i.e. 0-5 cms. and 0-10 cms., were taken at 44 sampling points within the Bay. The material from these cores was analysed by the Chemical Service of the Merlewood Research Station, and the results of the analyses were made available to the research workers concerned with the study of the ecology of Morecambe Bay.

Further statistical analysis of these results was not originally planned. The results for the many variables were plotted on maps of Morecambe Bay, and subsequent interpretation was largely subjective, aided by "contour" lines drawn by hand. The difficulties inherent in this type of interpretation quickly became apparent. First, the interpretation of general trends from variable data is relatively difficult by purely subjective methods, not only because trends of practical significance may be missed by such procedures, but, even more important, because trends may be ascribed to situations where they are not justified. A trend surface analysis of the data from the pilot survey, using statistical techniques of trend-surface fitting, was, therefore, carried out, and is reported in a separate paper (Jeffers, 1969).

In common with many pilot studies, where the relationships between physical and chemical properties are not known in advance, a large number of variables had been determined on the samples of mud and sand from the 44 sampling points. 19 variables were determined on the samples from cores in the 0-5 cms. zone, and 20 variables determined on the samples from the cores in the 0-10 cms. zone. It was expected that many of these variables would be highly correlated, but it was not known in advance which of the variables would be useful in describing the variability of the physical and chemical properties in the sampled area of Morecambe Bay. Again, statistical analysis was not originally planned, but it quickly became apparent that selection of the important variables for further study would be difficult, if not impossible, by subjective methods of interpretation. Principal component analyses of the two sets of data, i.e. from the 0-5 cms. and 0-10 cms. zones, obtained during the pilot survey were therefore undertaken, and the results of these analyses are reported in this paper.

The objects of the principal component analyses were as follows:-

1. The investigation of the correlations between the variables determined on samples from the two sets of cores
2. The determination of the number of essential dimensions of the physical and chemical variability which had been measured by the survey
3. Selection of the variables which should be measured in any future surveys of the physical and chemical properties of the muds and sands of the Bay, assuming that fewer variables could be determined in such surveys without the loss of essential information
4. The classification of the individual sampling points with reference to the essential dimensions of the physical and chemical variability, and the plotting of the resulting groups on a map of the Bay
5. Investigation of the trends of the essential dimensions over the sampled area of the Bay

Method of Analysis

Principal component analysis is described in detail by Kendall, 1957, and by Jeffers, 1965. In essence, it requires the calculation of the latent roots and vectors of the matrix of the coefficients of the correlations between every pair of the original variables. The latent roots define the proportions of the total variability measured by the survey which are accounted for by the linear combinations of the original variables described by the latent vectors. These vectors are calculated in such a way that each vector accounts for the maximum amount of the total variability while remaining orthogonal to all the other vectors. The values of the principal components can be calculated for each of the original sampling points and represent the plotting of the variability in n-dimensional space with orthogonal axes. These points may then be used for further analysis, for example, cluster analysis within the n-dimensional vector space, or for the calculation of trend surfaces.

Further details of the methods of computation will not be given in this paper, which will concentrate on the results from the analyses. The computations were carried out on the I.C.L. Sirius computers at the Blackburn College of Technology, and the PDP-8/I computer at the Merlewood Research Station.

Variables Included in the Analyses

The variables included in the analysis of the data from the cores in the 0-5 cms. zone are given in Table 1. Each of these variables was determined at 44 sampling points within the sampled area of Morecambe Bay.

Table 1. List of Variables Determined on Cores from 0-5 cms.

Variable	Description
1	H ₂ O per cent
2	Loss on ignition per cent dry
3	Density
4	Na (ppm) in interstitial water
5	P (ppm) in interstitial water
6	NO ₃ +NO ₂ -N (ppm) in interstitial water
7	NH ₄ -N (ppm) in interstitial water
8	Na (mg/1) in wet sand
9	P (mg/1) in wet sand
10	NO ₃ +NO ₂ -N (mg/1) in wet sand
11	NH ₄ -N (mg/1) in wet sand
12	Na per cent in dry sand
13	P per cent in dry sand
14	NO ₃ +NO ₂ -N per cent in dry sand
15	NH ₄ -N per cent in dry sand
16	K per cent total
17	Ca per cent total
18	P per cent total
19	N per cent total

The variables included in the analysis of the data from the cores in the 0-10 cms. zone are given in Table 2. For some of the variables, the results of the determinations were clearly erroneous, and the data from three of the 44 sampling points were, therefore, eliminated to simplify the subsequent analysis. Data are, therefore, only available for 41 points for these variables.

Table 2. List of Variables Determined on Cores from 0-10 cms.

Variable	Description
1	H ₂ O per cent
2	Loss on ignition
3	Density
4	Na (ppm) in interstitial water
5	P (ppm) in interstitial water
6	NH ₄ -N (ppm) in interstitial water
7	Na (mg/l) in wet sand
8	P (mg/l) in wet sand
9	NH ₄ -N (mg/l) in wet sand
10	Na per cent in dry sand
11	P per cent in dry sand
12	NH ₄ -N per cent in dry sand
13	K per cent total
14	Ca per cent total
15	P per cent total
16	N per cent total
17	Coarse sand per cent
18	Fine sand per cent
19	Silt per cent
20	Clay per cent

Results: 0-5 cms.

The data for the 19 variables listed in Table 1 are summarized in Table 3, and the coefficients of the correlations between these variables are given in Table 4.

Moisture content was significantly correlated with loss on ignition, all determinations of P (including total P per cent), all determinations of NH₄-N total percentages of K, P, and N, and was negatively correlated with Na in the interstitial water. Loss on ignition was significantly correlated with moisture content, all determinations of P except total P, all determinations of NH₄-N, and with total K, Ca, and N. Density was not significantly correlated with any other variables.

Table 4. Coefficients of Correlations between Physical and Chemical Variables of Corps at 0-5 kms.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
0.562																			
-0.110	0.275																		
-0.494	0.006	0.094																	
0.382	0.414	0.038	-0.230																
0.178	-0.042	-0.119	-0.361	0.326															
0.607	0.591	0.067	-0.122	0.618	-0.107														
-0.184	0.243	0.131	0.944	-0.119	-0.354	0.085													
0.545	0.530	0.079	-0.312	0.948	0.324	0.645	-0.149												
0.240	0.000	-0.103	-0.399	0.348	0.995	-0.083	-0.371	0.368											
0.672	0.625	0.082	-0.133	0.578	-0.125	0.989	0.100	0.630	-0.096										
-0.045	0.280	0.110	0.844	-0.031	-0.382	0.163	0.940	-0.049	-0.388	0.190									
0.585	0.541	0.040	-0.324	0.939	0.316	0.670	-0.100	0.998	0.362	0.659	-0.047								
0.258	0.004	-0.129	-0.408	0.348	0.993	-0.077	-0.377	0.371	0.999	-0.090	-0.391	0.366							
0.677	0.595	0.072	-0.160	0.532	-0.168	0.963	0.057	0.594	-0.140	0.982	0.156	0.625	-0.132						
0.794	0.555	0.024	-0.463	0.289	0.169	0.437	-0.207	0.449	0.224	0.493	-0.105	0.474	0.234	0.494					
0.200	0.532	0.256	-0.093	0.204	0.139	0.175	-0.009	0.240	0.152	0.174	-0.004	0.232	0.149	0.163	0.316				
0.504	0.284	0.189	-0.480	0.182	0.134	0.349	-0.330	0.292	0.168	0.387	-0.266	0.305	0.168	0.381	0.693	0.159			
0.735	0.603	0.167	-0.242	0.389	0.043	0.655	0.011	0.554	0.088	0.729	0.117	0.580	0.093	0.744	0.690	0.379	0.562		

(Correlations of ± 0.297 and ± 0.384 significant at 0.05 and 0.01 levels of probability respectively)

Table 3. Summary of data from 0-5 cms.

Variable	Minimum	Mean	Maximum	Standard Deviation
1	18.00	21.57	28.70	1.823
2	0.840	1.418	2.570	0.323
3	1.670	1.752	1.800	0.034
4	2370	7389.0	10280	1959.9
5	0.140	0.338	0.770	0.134
6	0.110	0.754	2.270	0.506
7	0.530	2.940	16.25	2.950
8	960	2700.7	3720	654.4
9	0.049	0.129	0.340	0.064
10	0.048	0.280	0.868	0.196
11	0.200	1.143	7.353	1.356
12	101.0	200.4	289.0	43.49
13	0.004	0.009	0.025	0.005
14	0.003	0.021	0.067	0.015
15	0.014	0.083	0.586	0.107
16	0.150	0.279	0.470	0.060
17	1.600	2.475	5.700	0.623
18	0.014	0.028	0.046	0.006
19	0.001	0.008	0.053	0.009

All determinations of Na were significantly intercorrelated, and were correlated with moisture content. They were also negatively correlated with all determinations of P, except for P in the interstitial water, all determinations of $\text{NO}_3 + \text{NO}_2 - \text{N}$, and with total K and P. The determinations of P were significantly intercorrelated, with the exception of total P, and were also significantly correlated with moisture content, loss on ignition, all determinations of $\text{NO}_3 + \text{NO}_2 - \text{N}$, all determinations of $\text{NH}_4 - \text{N}$, and with total N. The determinations of $\text{NO}_3 + \text{NO}_2 - \text{N}$ were significantly intercorrelated, and were also correlated with all determinations of P, except total P. They were also negatively correlated with all determinations of Na.

The determinations of $\text{NH}_4 - \text{N}$ were significantly intercorrelated, and were also correlated with moisture content, loss on ignition, all determinations of P, including total P, and with total K and N.

The total percentage of K was significantly correlated with moisture content, loss on ignition and all determinations of $\text{NH}_4 - \text{N}$. It was also negatively correlated with Na in interstitial water, but not with the other determinations of Na. The total percentage of Ca was significantly correlated with loss on ignition and total K. The total percentage of P was significantly correlated with moisture content, all determinations of $\text{NH}_4 - \text{N}$, and with total K, and negatively correlated with Na in interstitial water and wet sand only.

The total percentage of N was significantly correlated with moisture content, loss on ignition, all determinations of P, all determinations of $\text{NH}_4\text{-N}$, and with total K, Ca, and P.

There was, therefore, a high degree of intercorrelation between the variables determined on cores from 0-5 cms. The correlations were generally consistent, except for Na in interstitial water, and the apparently contradictory behaviour of total P by comparison with the other determinations of P.

The proportions of the total variability accounted for by the first five principal components of the correlation matrix of Table 4 are given in Table 5 and the weighting of the original variables in these components are given in Table 6.

Table 5. Proportions of Total Variability of Samples from 0-5 cms. accounted for by the first five principal components

Component	Proportion	Cumulative Proportion
1	0.3895	0.3895
2	0.2295	0.6190
3	0.1121	0.7311
4	0.0805	0.8116
5	0.0616	0.8732

The first five components accounted for 87.3 per cent of the total variability measured by the 19 variables. The first of these components, accounting for about 39 per cent of the variability, is an index of the amounts of P and $\text{NH}_4\text{-N}$ in the interstitial water, wet sand, and dry sand, together with the moisture content, loss on ignition, and total percentages of K and N. The second component, accounting for a further 23 per cent of the total variability, is a contrast between the amounts of Na and $\text{NO}_3\text{+NO}_2\text{-N}$. The third component, accounting for a further 11 per cent of the variability, is a contrast of the amounts of Na and $\text{NO}_3\text{+NO}_2\text{-N}$ with the total percentage of P. The fourth and fifth components, accounting for 8 per cent and 6 per cent, are measures of the percentage of calcium and the density, respectively.

The analysis, therefore, suggests that, because of the intercorrelations between the variables, there is a high degree of redundancy in the information contained by the data, and only five essentially independent dimensions of the variability in physical and chemical properties have been measured. Two of these dimensions are measured by single variables, i.e. the calcium per cent and the density, but the other three represent joint indices or contrasts of the remaining variables, and the interpretation of these components invites clearer definition of the objectives of the survey, if further survey work is to be done in Morecambe Bay.

For all essential purposes, almost all of the information contained by the 19 variables could be obtained by measuring only five of the variables. The weighting of the variables in Table 6 suggest some degree of choice, particularly as to whether determinations are made in the interstitial water, wet sand, or dry sand. Certainly, there is little value in determining any variable in all three, and the choice could be made according to the ease of laboratory tests, without any loss of information. All other things being equal, the choice of variables for future surveys should probably include the amount of P or $\text{NH}_4\text{-N}$ in dry sand, the amount of Na and $\text{NO}_3\text{+NO}_2\text{-N}$ in the interstitial water, and the density and calcium content of the total sample.

The values of the five components were computed for the 44 sample points included in the pilot survey. Trend surfaces were fitted to the first three components, using the method described by Jeffers, 1969, and the resulting trend surfaces are plotted in Figures 1, 2, and 3. The fourth and fifth components are measures of single variables, and no trend surface was fitted to these, as the trend surface for the calcium content of the pilot survey is given in the paper already mentioned, and no significant trend was found for the variable of density.

A cluster analysis, using a slightly modified version of the minimum spanning tree method of Gower, 1969, was carried out on the component values of the 44 sample points. Twelve primary clusters were obtained, and these were grouped into five secondary clusters. The mean values of the first five components for the clusters are given in Table 7, the clusters being listed so that the cluster with the highest value of the first component occurs at the top of the classification and the cluster with the lowest value of the first component at the bottom. The distribution of the secondary clusters is plotted in Figure 4.

Table 7. Cluster Analysis of Sample Points: 0-2 cms.

Group	Mean Value of Component:-				
	1	2	3	4	5
A	8.751	-4.868	-1.268	-1.491	-0.303
B 1	2.291	0.819	1.671	0.239	-0.062
B 2	1.691	3.386	0.833	-0.204	-0.337
C 1	1.242	1.814	-1.987	-1.227	1.473
C 2	-0.037	0.082	-2.728	1.463	0.233
C 3	-0.650	1.545	-1.570	0.304	-0.417
D 1	-1.059	-1.261	1.052	0.673	0.874
D 2	-1.303	-1.905	1.443	-0.239	-0.662
D 3	-1.409	-1.766	0.104	0.754	-0.711
E 1	-2.004	-0.395	-0.523	-0.206	-0.662
E 2	-2.704	-1.521	0.428	-0.793	-0.939
E 3	-3.019	-0.781	0.078	-0.835	-0.062

Cluster A, which consists of only two sampling points, is characterised by very high values of the first component, i.e. high values of P, $\text{NH}_4\text{-N}$, moisture content, loss on ignition, and total percentages of K and N, and low values of the second component. Both of these sampling points were fairly close to the shore on the Morecambe side of the Bay.

Cluster B is characterised by relatively high values of both the first and the second component, so that the sampling points within this group would have high amounts of phosphate, $\text{NH}_4\text{-N}$, and $\text{NO}_2\text{+NO}_3\text{-N}$, and low amounts of sodium. The two primary groups within this cluster differ mainly in their percentages of calcium. The sampling points of this cluster occur mainly close to the Morecambe shore and on the shallow flats between the Kent and Leven estuaries.

Table 6. Weighting given to Original Variables by Components

Variable	Component				
	1	2	3	4	5
1	0.94	0.00	-0.39	0.12	-0.59
2	0.78	-0.49	0.25	0.59	0.11
3	0.10	-0.28	0.04	0.81	1.00
4	-0.53	-0.83	1.00	0.15	-0.19
5	0.86	0.07	0.73	-0.59	0.45
6	0.32	1.00	0.85	0.21	-0.24
7	0.92	-0.53	-0.07	-0.42	0.01
8	-0.24	-0.95	0.96	0.30	-0.37
9	0.98	0.07	0.58	-0.40	0.35
10	0.38	1.00	0.81	0.24	-0.24
11	0.94	-0.56	-0.14	-0.33	-0.07
12	-0.11	-0.98	0.87	0.23	-0.42
13	1.00	0.06	0.53	-0.42	0.27
14	0.39	1.00	0.80	0.22	-0.27
15	0.92	-0.57	-0.27	-0.34	-0.08
16	0.85	0.07	-0.54	0.58	-0.45
17	0.42	-0.03	0.24	1.00	0.49
18	0.66	0.17	-0.85	0.50	-0.13
19	0.09	-0.29	-0.25	0.41	-0.22

Cluster C has average values of the first component, and moderately large values of the second component, but is also characterised by low values of the third component, the contrast of the amounts of Na and $\text{NO}_3 + \text{NO}_2 - \text{N}$ with the amounts of total phosphate. The first primary group of this cluster is more similar to the primary groups of cluster B, apart from the characteristically low value of the third component; the other two primary groups differ mainly in the amounts of calcium and the density. The sampling points from this cluster occur mainly along the Morecambe and Grange shore, and at the mouth of the Leven estuary.

Cluster D is characterised by relatively low values of both the first and second component, the three primary groups within the cluster differing only in the values of the third, fourth and fifth components. All of the sampling points from this cluster occur in the mid-Bay section of the sampled area, or off Rampside Sands and Point of Comfort Scar.

Cluster E has the lowest values of the first component, with values of the second component very similar to those of cluster D. Most of the sampling points of the cluster fall in the centre of the Bay, but a few occur in the Leven estuary, and off Point of Comfort Scar.

Results: 0-10 cms.

The data for the 20 variables listed in Table 2 are summarised in Table 8, and the coefficients of the correlations between these variables are given in Table 9.

Table 8. Summary of Data from 0-10 cms.

Variable	Minimum	Mean	Maximum	Standard Deviation
1	18.00	21.53	26.80	1.721
2	0.670	1.362	2.930	0.398
3	1.630	1.743	1.800	0.036
3	2880	7683	10470	2220
5	0.030	0.403	1.620	0.320
6	0.750	2.353	6.760	1.442
7	104.0	2869	3920	772.3
8	0.041	0.154	0.657	0.121
9	0.239	0.891	2.502	0.570
10	79.00	208.3	285.0	55.21
11	0.003	0.011	0.048	0.009
12	0.016	0.065	0.187	0.043
13	0.170	0.275	0.450	0.052
14	1.900	2.461	3.600	0.335
15	0.021	0.029	0.042	0.005
16	0.001	0.006	0.040	0.008
17	0.500	2.463	43.00	6.905
18	53.00	90.93	99.00	6.875
19	0.500	0.866	6.000	1.078
20	0.000	5.439	9.000	2.180

Moisture content was significantly correlated with loss on ignition, NH_4-N in wet and dry sand, the percentages of K, Ca, P, and N, and the clay per cent, and negatively correlated with Na in the interstitial water and with the coarse sand per cent. Loss on ignition was significantly correlated with moisture content, the percentages of K, P, and N, and the clay per cent. Density was correlated with all determinations of Na and with silt per cent, and negatively correlated with the fine sand per cent.

The determinations of Na were significantly intercorrelated, and were also correlated with density. These determinations were negatively correlated with moisture content, and the percentages of Ca, P, and K. The determinations of P were significantly intercorrelated, and were also correlated with the determinations of Na in wet and dry sand, the percentage of Ca, and the silt per cent. The determinations of NH_4-N were significantly intercorrelated and were also correlated with the percentage of N.

The total percentage of K was significantly correlated with moisture content, loss on ignition, the determinations of NH_4-N , the percentages of Ca, P, and N, and the clay per cent. It was negatively correlated with the determinations of Na, and the coarse sand per cent. The total percentage of Ca was significantly correlated with moisture content, all determinations of P, the percentages of K and N, and the silt and clay per cents. It was negatively correlated with the determinations of Na. The total percentage of P was significantly correlated with moisture content, loss on ignition, NH_4-N in dry sand, and the percentages of K and N, and negatively correlated with the Na determinations. The total percentage of N was significantly correlated with moisture content, loss on ignition, NH_4-N , and the percentages of K, Ca, and P.

The coarse sand per cent was negatively correlated with moisture content, the percentage of K, and the fine sand per cent. The fine sand per cent was negatively correlated with density, and with the coarse sand per cent. The silt per cent was significantly correlated with density, and the percentages of P and Ca. The clay per cent was significantly correlated with moisture content, the loss on ignition, and the percentages of Ca, and K.

The proportions of the total variability accounted for by the first five principal components of the correlation matrix of Table 9 are given in Table 10, and the weighting of the original variables in these components are given in Table 11.

Table 10. Proportions of Total Variability of Samples from 0-10 cms. accounted for by the first five principal components

Component	Proportion	Cumulative Proportion
1	0.2609	0.2609
2	0.2175	0.4784
3	0.1263	0.6047
4	0.1192	0.7239
5	0.0898	0.8137

Table 11. Weighting Given to Original Variables by Components

Variable	Component				
	1	2	3	4	5
1	0.95	0.06	-0.04	0.37	0.32
2	0.67	0.17	-0.12	0.22	0.98
3	-0.20	0.59	0.22	-0.21	1.00
4	-0.85	0.60	0.63	0.28	0.19
5	-0.11	1.00	-0.73	-0.00	-0.37
6	0.56	0.47	1.00	-0.39	-0.67
7	-0.67	0.68	0.66	0.42	0.36
8	-0.03	0.99	-0.78	0.04	-0.36
9	0.69	0.48	0.97	-0.35	-0.49
10	-0.62	0.68	0.69	0.47	0.32
11	-0.00	0.98	-0.77	0.06	-0.38
12	0.72	0.46	0.96	-0.33	-0.49
13	1.00	-0.05	-0.03	0.28	0.29
14	0.71	0.45	-0.81	-0.19	-0.18
15	0.85	-0.16	0.13	-0.01	0.50
16	0.69	0.46	0.36	0.18	0.70
17	-0.43	0.01	0.01	-0.97	0.54
18	0.23	-0.12	0.18	1.00	-0.64
19	0.12	0.59	-0.39	-0.32	0.45
20	0.59	0.14	-0.37	0.20	0.19

The first five components accounted for 81.37 per cent of the total variability measured by the 20 variables. The first of these components, accounting for 26.09 per cent of the variability, is an index of the moisture content, the determination of Na in the interstitial water, and the percentages of K and P. The second component, accounting for a further 21.75 per cent of the variability, is an index of the determinations of P in the interstitial water, wet sand, and dry sand. The third component, accounting for a further 12.63 per cent, is a contrast of the amounts of $\text{NH}_4\text{-N}$ in the interstitial water, wet sand, and dry sand with the percentage of calcium. The fourth component, accounting for 11.92 per cent, is a contrast of the coarse sand and fine sand per cents. The fifth component, accounting for 8.98 per cent, is an index of the loss on ignition and density. The sixth and seventh components, which account for 4.16 and 3.84 per cent respectively, are direct measures of the clay per cent and silt per cent.

The analysis again suggests that there is a high degree of redundancy in the information contained by the data. Only three essential dimensions of the chemical variability have been measured by the survey. The remaining four dimensions are derived from the physical properties, and it is interesting to see that four independent dimensions have been derived from five physical measurements. It is also of some interest that the components derived from the 0-10 cms. zone are different from those derived from the 0-5 cms. zone. To some extent, this will have been due to the lack of measurements of the $\text{NO}_2 + \text{NO}_3\text{-N}$ in the 0-10 cms. zone, but the negative association of Ca with the amounts of $\text{NH}_4\text{-N}$ is of interest, as is the emphasis on the amount of Na in the interstitial water rather than in the wet sand or dry sand. There are also suggestions of greater stability in the relationships between the variables measured in the 0-10 cms. zone than in the 0-5 cms. zone.

For all essential purposes, almost all the information contained by the 20 variables measured in the 0-10 cms. zone could be obtained by measuring only seven of the variables. As for the variables of the 0-5 cms. zone, there is some degree of choice, but the choice of variables for future surveys should probably include the total percentage of K, the percentage of P in the interstitial water, the percentage of $\text{NH}_4\text{-N}$ in the interstitial water, the density, and the coarse sand, fine sand, clay, and silt per cents.

The values of the first five components were computed for the 41 sample points included in the pilot survey. Trend surfaces were fitted to these components, and the resulting surfaces are plotted in Figures 5, 6, 7, 8, and 9. The trend surfaces for the clay per cent and silt per cent have already been given in a previous paper (Jeffers, 1969).

Accluster analysis was carried out on the component values of the 41 sample points. Eleven primary clusters were obtained, and these were grouped into three secondary clusters. The mean values of the first five components for the clusters are given in Table 12, and the distribution of the secondary clusters is plotted in Figure 10.

Table 12. Cluster Analysis of Sample Points: 0-10 cms.

Group	Mean value of components:-				
	1	2	3	4	5
A 1	3.777	-0.270	-0.243	-0.574	0.760
A 2	3.225	0.979	1.360	-1.021	-1.887
A 3	1.260	-3.313	-2.309	-0.304	-1.830
A 4	0.909	-1.647	-0.674	0.505	0.608
A 5	0.802	2.079	3.446	-1.141	-2.855
A 6	0.286	-2.016	-0.330	-0.564	-0.951
B 1	0.353	0.568	0.568	0.958	0.807
B 2	-1.690	-0.140	1.110	0.962	-1.127
B 3	-3.237	-0.774	0.478	-1.755	0.535
C 1	-1.121	2.040	-1.233	0.311	-0.310
C 2	-1.364	1.745	-0.346	0.601	-0.923

Cluster A is characterised by relatively high values of the first component, i.e. high values of moisture content, Na in the interstitial water, and total percentages of K and P. The primary groups differ widely in the values of the second and remaining components, but nearly all of the sample points in this cluster occur along the eastern side of the Bay and on the shallow flats between the Kent and Leven estuaries.

Cluster B has moderate to low values of the first component combined with generally low values of the second component, i.e. the amounts of P in the interstitial water, wet sand and dry sand. The primary groups within this cluster differ mainly in their values of the fourth and fifth components, i.e. in their physical properties. The sample points of this cluster fall mainly in the mid-Bay sections.

Cluster C has moderately low values of the first component and relatively high percentages of P in the interstitial water, dry sand, and wet sand. The sample points of this cluster are mainly from the western parts of the Bay.

Discussion

It is of interest at this point to relate the results of the analyses to the objectives stated in the introduction to this paper. The first of these objectives was the investigation of the correlations between the variables determined on the samples from the two sets of cores, and the significant correlations have been described in detail. The variables of both sets were heavily intercorrelated, and it may be of interest to study these correlations, or sometimes the lack of correlations, in close relation to what is already known about the chemical and physical properties of sands and muds from the same or similar localities. Such a study can clearly not be attempted by the author of this paper; but the importance of the correlations themselves should not be glossed over by the later stages of the analyses.

The marked reductions in the numbers of essential dimensions indicated for both sets of data emphasises the redundancy of much of the information obtained from the survey. The variables measured on the 0-5 cms. cores had only five essentially independent dimensions, while those from the 0-10 cms. cores had seven independent dimensions. Most of the redundancy of information occurred in the chemical variables, and it was notable that, where the physical variables were included in the determinations on the 0-10 cms. cores, the components based on physical variables dominated the analysis. There were also very few correlations between the physical and chemical variables, and none of the components gave high weighting to both chemical and physical variables. The analysis suggests that the relationships within the variables determined on the 0-10 cms. cores were more stable than those determined on the 0-5 cms. cores, presumably because the surface layers of the sands and muds are more likely to reflect local changes. The reduced number of essential dimensions provides the minimum set of dimensions that have to be considered in characterising the variability of the chemical and physical characteristics of Morecambe Bay.

The components also enable the selection to be made of the variables to be included in any future surveys of the chemical and physical properties of the Bay, although it should be stressed that the close correlations between many of the variables indicate that, within the groups of variables defined by the components, there is a fair degree of choice which might, for instance, be influenced by the ease with which some of the determinations can be made in the laboratory. Generally, however, future surveys should probably concentrate on determinations of the total percentage of K, the percentages of P, Na, and $\text{NH}_4\text{-N}$ in the interstitial water, the density, and the coarse sand, fine sand, silt⁺ and clay per cents, all measured on cores from the 0-10 cms. zone.

The analyses also enabled reasonably consistent classifications of the sample points with respect to their chemical and physical properties to be made, and these classifications have been plotted in Figures 4 and 10. Such classifications may well enable research workers to think more deeply about the causes of variation in the Bay, and to interpret the results of the survey in relation to other observations not included in the analysis. The cluster analysis technique has the advantage of being orientated towards the actual sample points included in the survey, and may, for this reason, be more acceptable, although not necessarily more informative, than the plotting of the trend surfaces of the components. To the outside observer, the clusters seem less informative, but this may be because of a lack of appreciation or knowledge of other observations.

The trend surfaces of the components have the merit of summarising the maximum information about the underlying variation in chemical and physical properties in the minimum space, without significant loss of information. These components and their trend surface representation will also be valuable in relating the distribution of the invertebrates in the Bay to chemical and physical properties. It must be admitted, however, that the value of the trend surfaces may be reduced by the relative complexity of their derivation, so that their role in the interpretation of the variability of the properties of the Bay may be diminished.

References

- Jeffers, J. N. R. 1965. Principal component analysis in taxonomic research. Proceedings 2nd Conference Advisory Group of Forest Statisticians, International Union of Forest Research Organisations, Stockholm, 1965.
- Jeffers, J. N. R. 1969. Trend surface analysis of chemical and physical variables from a pilot survey of Morecambe Bay. Merlewood Research and Development Paper No. 14, 1969.
- Kendall, M. G. 1957. A course in multivariate analysis. Griffin, 1957.

THIRD COMPONENT

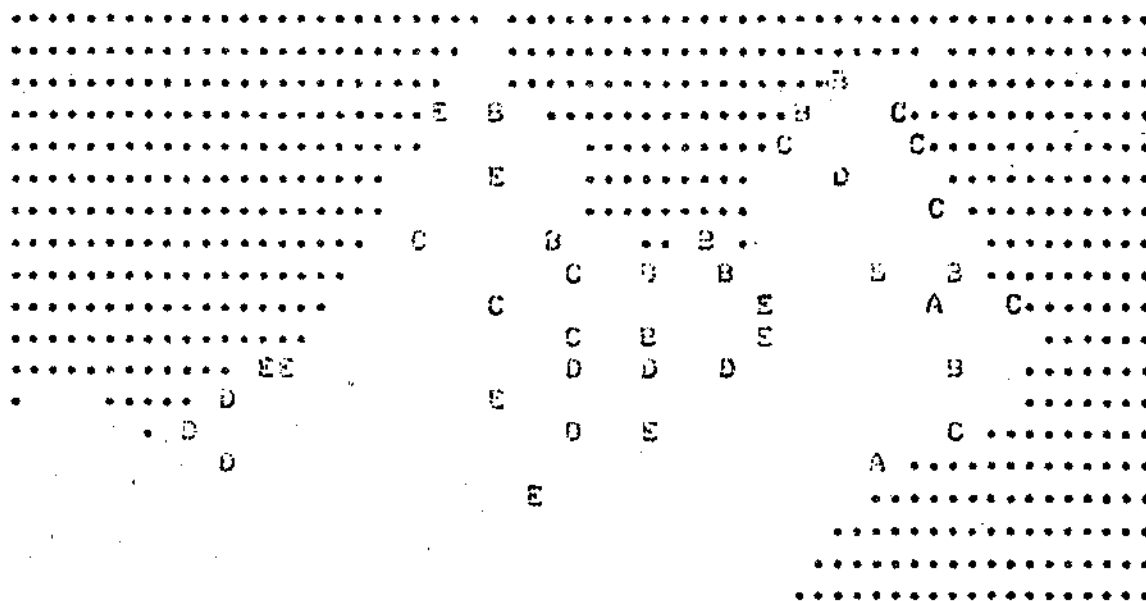
FIGURE 3



Contour step = 0.5

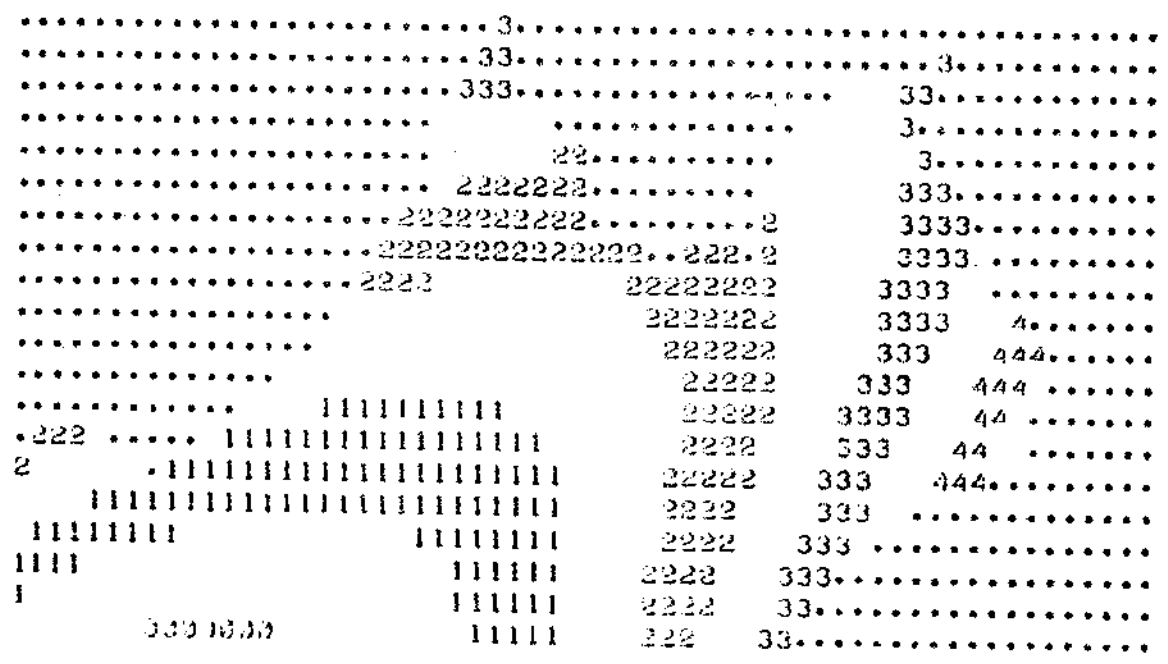
DISTRIBUTION OF CLUSTERS: B-SC4S

FIGURE 4



FIRST COMPONENT: W-180MS

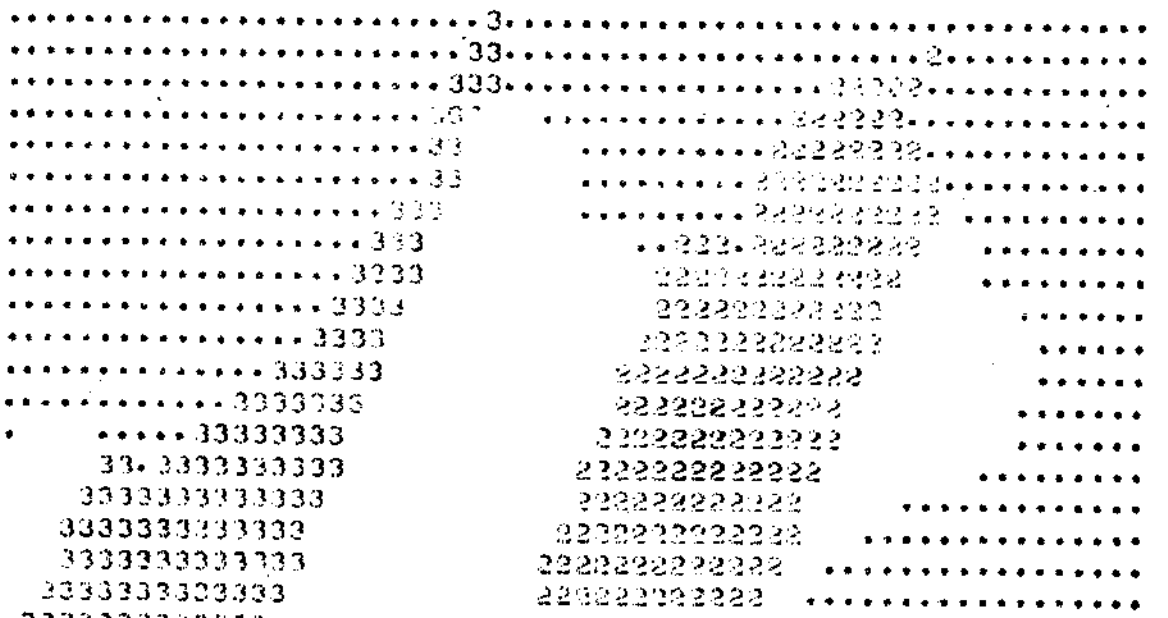
FIGURE 5

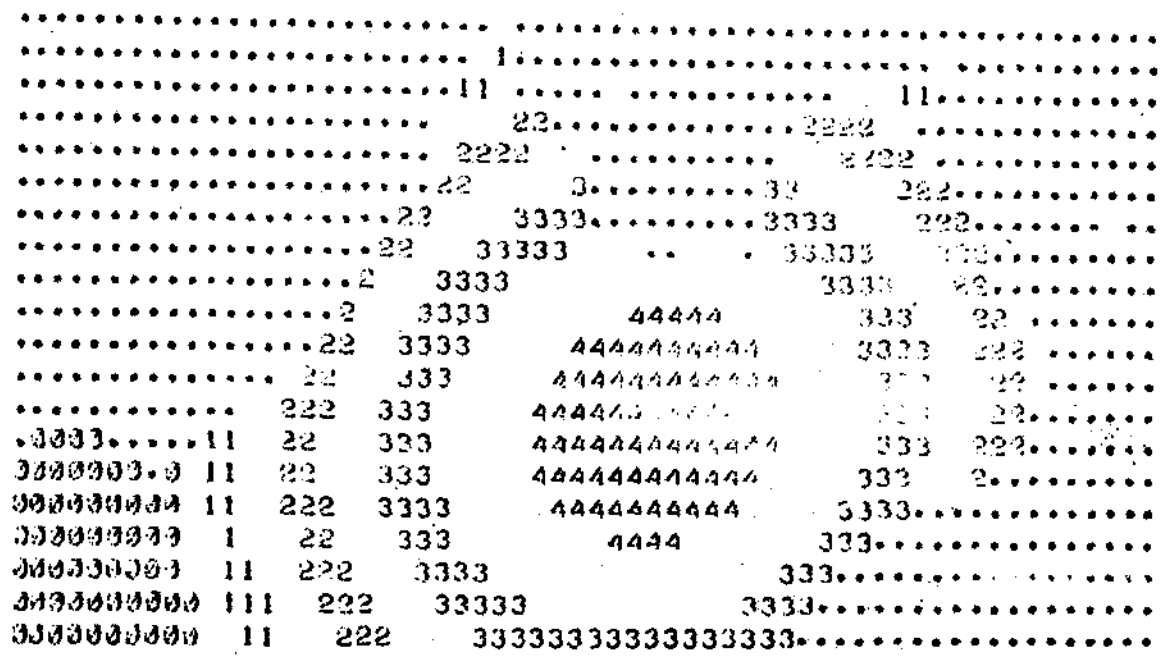


Contour step = 1.0

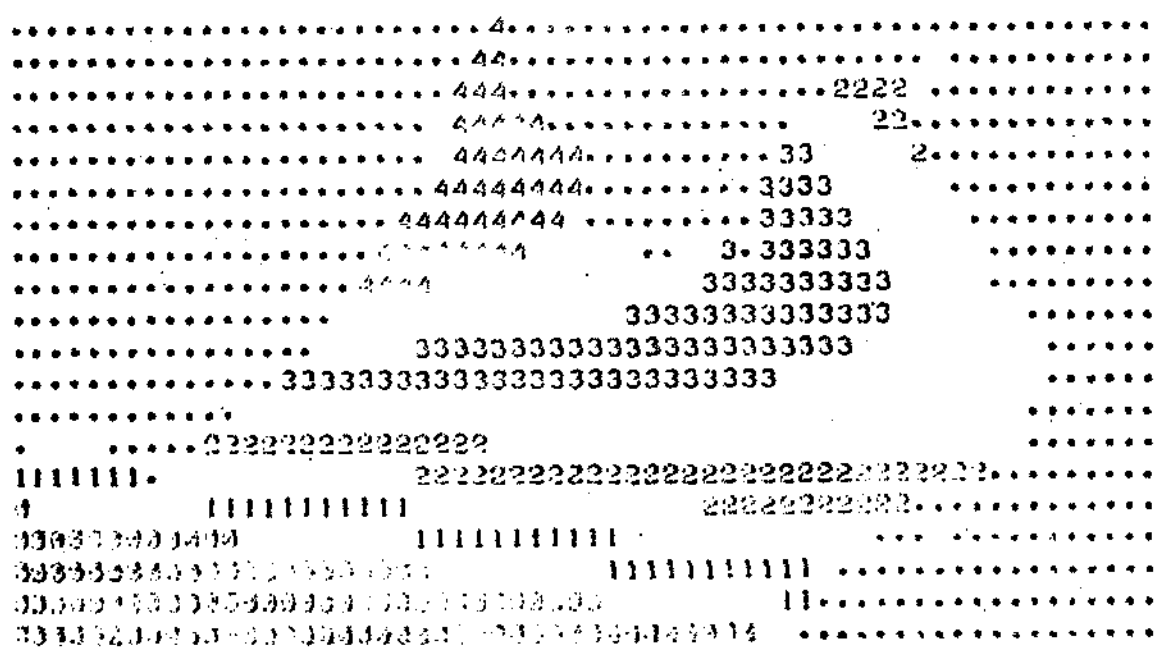
SECOND COMPONENT: W-180MS

FIGURE 6

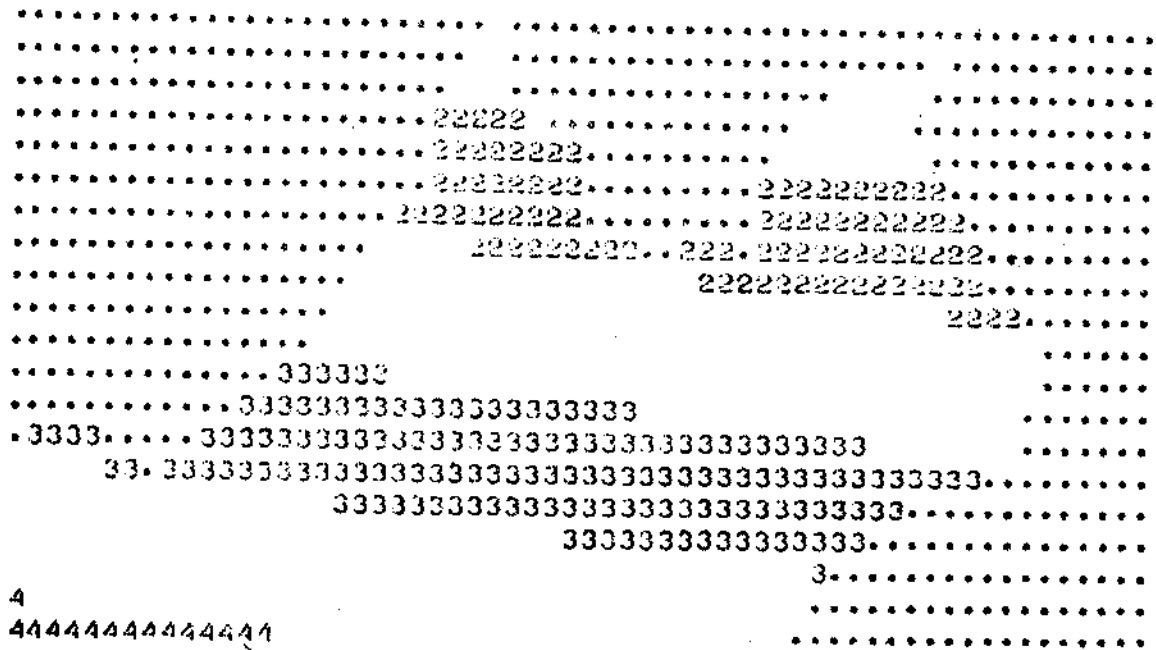




Contour step = 0.5



Contour step = 0.5



Contour step = 0.5

