

Stichting voor Bodemkartering
Postbus 98
6700 AB Wageningen
Tel. 08370-19100

Rapport nr. 1607
Project nr. 176.34

THE SPATIAL VARIABILITY OF THE SOIL OF THE
TRIALFIELD PAGVI AT LELYSTAD

Dr. P.A. Burrough

CONTENTS

		Page
	ABSTRACT	5
	INTRODUCTION	7
1	QUANTITATIVE DETERMINATION OF THE NATURE OF THE SOIL SPATIAL VARIABILITY UNDER PAGV 1	9
1.1	General data	9
1.2	Soil Data	9
1.3	Statistical analysis	11
1.4	Data exploration	11
1.5	One-way analysis of variance of block differences	11
1.6	Nested analysis of variance to determine the components of variance at different spatial scales	14
1.7	Ultra-short range studies of variability	15
1.8	Principal component analysis of the soil data	17
1.8.1	Correlation matrix	17
1.8.2	Eigenvectors and eigenvalues	19
1.8.3	Spatial analysis of the component scores	20
1.9	Comparison of the soil patterns in Figures 11 and 12 and the patterns of slaked soil (verslemping)	20
1.10	Conclusions about the nature of soil variability on PAGV 1	21
2	EXAMINATION OF SPATIAL VARIATIONS OF YIELD ON PAGV 1	23
2.1	Data	23
2.2	Problems of analysis	23
2.3	Data analysis via standardization to zero mean and unit variance	24
2.4	Data analysis via summed annual yields	24
2.4.1	Stratification of soil profiles according to the four quartal annual yield classes.	25
2.4.2	Correlations between yields and soil parameters for the 2-year rotation blocks	27
2.5	Comparison of yield patterns with slaked ground patterns,	27
2.6	Conclusions from the yield data	28
3	OVERALL CONCLUSIONS	29
	REFERENCES	31

Tables

1	Soil data used for variability analyses	10
2	Distribution of variance components with respect to major cropping blocks in PAGV1 (180 profiles)	12
3	Components of variance at different spatial scales for 144 soil profiles in the 1- and 2-year rotation blocks on PAGV1	16
4	Estimates of variance over 1 m for selected attributes	17
5	Eigenvectors of the first 10 principal components (180 profiles, 32 variates)	19
6	Contributions of original variates to the components	18
7	Spatial variabilities of the first 6 principal components	20
8	Analysis of variance of standardized potato yields for 1-year replicates	24
9	Means and standard deviations of summed yields	25
10	Analysis of variance of soil parameters according to 4 yield classes	26

Figures

1	Layout of the PAGV1 trialfield
2	Profile schematics of the soil under PAGV1
3	Soil mechanical resistance (cone penetrometer) of top soil (0-20 cm)
4	End depth of Almere clay (upper) deposit/depth at which Almere sand (upper) begins
5	Blocks for oneway anovar
6	Nested anovar for PAGV1
7	Blocks for nested anovar
8	Sources of variance with respect to sample spacing/block size
9	Contributions of variates to factors, PC1 v PC2
10	Contributions of variates to Components PC3 x PC4
11	Isoline map of first principal component scores
12	Isoline map of second principal component scores
13	Variability vs distance for principal component scores (144 sites)
13a	Pattern of slaked top soil (1979)
14	Standardized yields 1976 - 1979
15	Standardised average yields for the 4 years 1976 - 1979

ABSTRACT

A soil survey carried out in 1976-77 of the trial field PAGV1 at the Proefstation voor Akkerbouw en Groenteteelt in het Vollegrond, Lelystad, Flevoland suggested that the variability of some critical soil properties was higher than had been thought. However, statistical spatial analysis of the soil data showed that for most soil properties measured, more than 80% of the total variability within the field occurred within 17.5 m, that is within distances much smaller than the basic plot replicates used in the trials. Further sampling suggested that more than 60% of the total variability was present within distances of 1 m. Thus, in spite of high variances of certain critical soil properties, the PAGV1 field must be regarded as homogenous because most of the variability occurs well within a trial plot. Analysis of the yields from the trials for 1976-79 inclusive also failed to show any marked relation to site or soil. The conclusions are that soil and site differences have not markedly influenced the results of the yield trials on PAGV1.

INTRODUCTION

It is well known that soil varies considerably from place to place, and that generally speaking, the greater the distance between two observations of the soil, the greater the differences are likely to be. Over relatively large distances (more than a few 100's of metres), soil changes often go hand in hand with landscape patterns. Within shorter distances, however, soil changes are less easy to predict. If the variations in soil properties are small, the soil is said to be homogenous, but if the variations are large the soil is thought of as being complex. This degree of homogeneity or complexity and its spatial scale are of considerable importance when working in trial fields because they may have considerable effect on the numbers of samples necessary to characterise the soil, or on the interpretation of the experiments carried out. It is therefore important to know how the soil varies over the distance scales encountered in trial fields.

During 1976, 1977 and 1978, the Stichting voor Bodemkartering was asked to map the soil differences and study the soil conditions of the trial PAGV1 at the Proefstation voor Akkerbouw en Groenteteelt in het volle grond, at Lelystad in Oost Flevoland (Ovaa 1980). The trial field PAGV1 has for the last 8 years been used for detailed yield studies for potatoes and sugar beet (Lamers 1980). The trials have attempted to relate yields to rotation period, cultivation practice, nitrogen levels, organic fertilizers, and soil fumigation. When the experiment was originally conceived, the soil under PAGV1 was considered to be homogenous - that is that soil differences would be relatively unimportant and would have a negligible effect on yields. The results of the soil studies (Ovaa, op cit) suggested otherwise; large differences in soil profile appearance were found that could well influence plant performance, and hence the results and interpretations of the whole yield experiment.

The variations in the soil under PAGV1 result from variations in the thickness and texture of various marine and lacustrine deposits making up the first 120 cm of the soil. These deposits (from the surface downwards, the IJsselmeer, the Zuiderzee and the Almere - Pons and Wiggers 1959, 1960) vary greatly in thickness over distances of a few metres. The texture and thickness differences of these deposits have a visible effect on root development (Ovaa op cit). Because patterns of these soil variations appeared to coincide with various trial plot replicates it was by no means clear whether the yield differences resulted from different treatments or from different soil conditions. Other research (v.d. Graaff 1979) on slaking suggested also that there was a connection between soil conditions and the tendency for slaking in various parts of the field.

The aims of this study were initially threefold:

1. To determine quantitatively the nature of the soil variability under PAGV1
2. To attempt to determine the role of observed soil variability with respect to crop yield differences
3. To determine if the soil variability in other parts of the PAGV station was similar to that observed under PAGV1.

As the work progressed, however, two further aims became apparent. These were:

4. The need to define methods for determining the soil variability of trial fields before soil mapping or crop trials are attempted
5. Theoretical studies into the stochastic processes of soil variability,

This report handles aims 1 and 2. Aim 3 was undertaken by J.B. Kool, Doctoral Student at the Landbouwhogeschool, Wageningen (Kool, 1981). Aims 4 and 5 will be reported shortly.

1 QUANTITATIVE DETERMINATION OF THE NATURE OF THE SOIL SPATIAL VARIABILITY UNDER PAGV1

1.1 General data

Figure 1 is a plan of PAGV1 showing the locations of the trial plots and their subdivisions. The Figure also shows the plot dimensions, plot numbers and soil profile locations. The area is split into three main plots, one for one-year rotations (mono culture), one for two-year rotations and one for three-year rotations. Potatoes and sugar beet are the two main crops, with winter wheat as the third crop in the three-year rotation. Plots are divided according to tillage practice (plough, cultivator, and controlled-traffic 3 m-wide cultivation techniques), the use of champost (mushroom compost), and nitrogen levels (see Ovaa, op. cit. Figure 1).

PAGV1 has been accurately surveyed (Kuiper 1978). The field lies between -4.18 and -4.52 m NAP (sd = -3.5 cm). It appears to have no perceptible slope or trend within the limits of the survey.

1.2 Soil Data

Soil profile data were collected at 180 sites (Ovaa 1980). The sites were not spaced on a regular grid, but were located to minimise interference with the crop trials. The sample spacing was regular, however, and was 30 m in the N-S, and 17.5 m in an E-W direction (see Figure 1), except for the 3-year area where the E-W spacing was 20 m. Soil profiles were sampled with a Guts auger to a depth of 120 cm. Data were recorded on a horizon basis. For each horizon, the depths, thicknesses, clay content (lutumgehalte), M50 (median sand size larger than 50 μm), degree of layering of the horizon, clay content of the most sandy and most clayey layers in the horizon, and ripening class were recorded. Penetrometer data were recorded using a hand-held penetrometer to a depth of 80 cm.

Because of the intense and variable layering of the different deposits, especially the Almere, the number of horizons recorded per profile was not constant and varied between 5 and 9. This complicated the statistical analysis because ideally one should have equal numbers of observations on strictly comparable variables. In order to make statistical analysis possible the profile data were reworked so that all profiles had six layers that corresponded as far as possible with the various deposits (Figure 2). The procedure has undoubtedly introduced extra error into the data but it must be stressed that the resolution into the 6 horizons has been done as rationally as possible following on-site work at PAGV together with Ovaa and members of the PAGV staff. Table 1 lists the soil data that were actually used in the following analyses.

Table 1 Soil data used for variability analyses

Variate code	Description
NH	Number of horizons described at survey time
DEn	Depth at which the nth layer ended (all horizons)
Dn	Thickness of the nth layer (layers 2, 3, 4 and 5)
LUn	Average texture of the nth layer (all)
LAn	Degree of layering of the nth layer (layers 2-6)
LUn1	Texture (lutum content) of the most clayey sub-layer in the layer (layers 2-6)
LUn2	Texture (lutum content) of the least clayey sub-layer in the layer (layers 2-6)
LDn	Texture differences (lutum content) between most clayey and least clayey layers in a layer (layers 2-6)
PEn	Penetrograph data for the nth layer. Data are average of 3 values at the middle of the layer (layers 1-4).

Depths were measured in centimetres to an estimate accuracy of +/- 1 cm with respect to the soil in the Guts bore. Textures are measured in percent lutum (hand estimates). Penetrograph data in M. Pascals. Other data are dimensionless.

Two further complications are worth mentioning here. One is that not all data were recorded during the initial survey if it was not thought necessary to record them. This has led to the somewhat unfortunate situation where the data base contains a relatively large number of missing values, especially for the textures of the deeper horizons. Unlike the field scientist, the computer cannot judge what data are relevant, so data sets with as few missing values as possible are desired. The second is that the data were collected on a semi-regular grid. If either grid spacing corresponded to a semi-regular pattern in the landscape one could obtain biased data that could give misleading results. In the event, however, semi-regular grid spacing was probably of little consequence.

1.3 Statistical analyses

The analyses were used to determine the nature of the differences for each soil property over the PAGV1, both independently, and together. Four kinds of analysis were used: data exploration (frequency analysis, histograms, etc.) one-way analysis of variance for major blocks, one-way nested analysis of variance to determine the components of variance at different sample spacings, and principal component analysis for property interactions. Details of all methods are given in Webster (1977). Frequency analysis, one-way analysis of variance and principal component analysis were all done on the SPSS package on the IWIS-TNO Cyber in Den Haag; nested analysis of variance was done using a FORTRAN program specially written by the author and run on the same computer.

1.4 Data exploration

Means, standard deviations, variances, ranges and histograms were computed for all variates. Most showed normal, unimodal distributions (c.f Figures 3, 4) but the depths and thicknesses of the subsoil horizons showed non-normal, elongated distributions. These results suggested that the depths and thicknesses of the various subsoil horizons would be responsible for the soil differences in the PAGV1.

1.5 One-way analysis of variance of block differences

PAGV1 was divided into 11 major landuse blocks (Figure 5). One-way analysis of variance was used to partition the variance of each soil property into a between-block and a

Table 2 Distribution of variance components with respect to major cropping blocks in PAGV1 (180 profiles)

Attribute	Grand Mean	Total variance (TV)	Within'block Var (% TV)	F'value*
NH	6.9	0.69	93	2.347
DE1	26.1	10.29	90	3.040
LU1	22.0	1.38	66	10.326
PE1	7.1	6.17	74	7.235
DE2	45.1	40.60	78	6.073
LU2	9.6	9.26	93	2.449
LA2	1.0	0.16	100	0.428
LU21	15.7	13.86	100	0.979
LU22	3.5	10.53	83	4.638
PE2	18.1	15.08	87	3.690
DE3	62.0	70.54	93	2.376
LU3	15.4	7.29	89	3.254
LA3	1.8	0.48	87	3.606
LU31	19.5	9.50	97	1.570
LU32	7.5	12.51	90	2.911
PE3	12.8	12.99	96	1.810
DE4	79.7	74.78	94	2.204
LU4	11.5	11.22	96	1.752
LA4	2.2	0.35	98	1.435
LU41	16.3	15.80	97	1.612
LU42	4.8	7.57	100	0.769
PE4	14.3	13.38	85	3.870
DE5	96.4	89.44	97	1.605
LU5	6.8	13.30	99	1.259
LA5	2.0	0.25	89	3.196
LU51	10.4	24.26	100	0.838
LU52	3.3	5.9	99	1.219
LU6	4.1	0.58	97	1.453
LA6	1.9	0.16	100	1.494
LU61	6.8	8.95	99	1.217
LU62	2.2	0.33	88	3.033
D2	19.0	39.56	87	3.784
D3	16.9	39.34	86	3.830
D4	17.6	39.16	95	1.930
LD2	12.2	28.62	97	1.464
LD3	11.9	19.44	99	1.250
LD4	11.0	20.88	99	1.103
LD5	6.0	20.14	94	2.062

* for 5% level at DF = 11, 169 F = 1.85
 " 1% " " " " " F = 2.37

within-block term. Table 2 presents the means, total estimated variance, within block variances as percentage of the total, and F-values for each variate, computed for all 180 profiles. Although many variates show statistically significant values, it is clear that the only important reductions in overall variance at the block level occur for the properties DE1, LU1, PE1, DE2, LU22, PE2, LU3, PE4, D2, D3 and LU62. For all other properties more than 89% of all variance is present within a landuse block. Of the above properties, only 3 show within-block variance levels of less than 80%. The nature of these differences, and their overall importance will now be examined.

- DE1 - The overall variability is small. It arises from the effects of deep ploughing during the reclamation. From block means and standard deviations it appears that blocks 1, 2, 3 and 9 have slightly shallower than average first horizons, but the differences are negligible.
- LU1 - The overall variability is small, ranging from 19-25% lutum. Variations due to estimation error are not known. Blocks 5, 6, 7 and 8 appear to have consistently higher lutum levels than the more northerly blocks. This is consistent with earlier soil mapping of the whole area (RIJP 1962), but within PAGV1, the differences are small.
- PE1 - Variations are small but appear to be related to differences in land use because blocks 5, 6, 7, 8 and 9 have higher values than the other blocks. In 1977, the year of the survey, blocks 5, 6, 7, 8 and 10 were all under potatoes, and 9 was under wheat. The remaining blocks were all under sugar beet.
- DE2 - The variability is relatively high ($\bar{x} = 45$ cm \pm sd 6.4 cm) with 95% confidence limits of 32-58 cm over the whole area. The differences do not follow any major N-S division: higher values occur in blocks 1, 4, 9 and 10, and lower than average in 3, 5, 6, 7 and 8. Such large differences in the depth at which this sandy deposit ends could have an effect on plant performance.
- D2 - This also varies considerably, which is not unexpected as it is correlated with DE2. The Zuiderzee deposit appears to be thicker than average in blocks 1, 4, 6 and 10, and thinner than average in blocks 3, 5, 6, 7 and 8.
- LU22 - The differences are small, but the lutum levels in blocks 5, 6, 7 and 8 are roughly 2% lutum higher than elsewhere.
- PE2 - The variation is greater than for PE1, but the block differences shown by that variable do not reoccur. Blocks 1, 4, 5, 7 and 10 have lower than average values, and blocks 3 and 11 higher, but the differences are vague.
- LU3 - The variability is low, but again blocks 5, 6, 7 and 8 are different, this time having slightly lower than average values.

- LA3 - The differences are small and are related almost entirely to differences in blocks 3 and 9. These could have arisen from estimation differences.
- PE4 - Blocks 1, 2, 10 and 11 have lower than average values; 6 and 7 higher. These differences are difficult to interpret and probably mean very little.
- D3 - Blocks 5, 6 7 and 8 have clearly more variable values of D3 than other areas. This again suggests N-S differences.
- LU62 - Blocks 5, 6, 7, 8 and 11 have on average higher values, but the overall variability is small. Differences could be due to estimation differences, but the data reinforce the N-S trends.

To sum up. For all soil properties, the variability within the major land use blocks is a very large proportion of the total variance over the whole 12 ha field. A few top-soil properties appear to have management controlled differences, while other textural properties show small differences between the northern and southern parts of the field. Were it not for the data from the RIJP survey, which also indicate such a N-S textural trend, these textural differences could be explained by estimation error because they are so small. It is possible that variations in the thickness of the second and third horizons have a regional component of variability at a scale that could affect the performance of crops in different blocks.

1.6 Nested analysis of variance to determine the components of variance at different spatial scales

Because the within-land use block variances formed such great proportions of the variance, the components of variance at various spatial scales were examined by a nested one-way analysis of variance. The data from the 144 profiles located in the one- and two-year trial areas were used as these formed a balanced and symmetrically laid-out sample set.

Nested analysis of variance seeks to partition the variance of an attribute within each level of the nesting. If each nested level has a different spatial scale (sampling distance) the variance components per nested level are thus variance components corresponding to the sample scale at that level. The 144 sites were grouped into 4 main levels each containing 36 sites; these were further divided into 4 sublevels of 9 sites each, and each sublevel was further divided into 3 sub-sublevels (Figure 6). Each sub-sublevel corresponded to a major trial plot consisting of four 7.5 m x 55 m strips having identical crops and tillage, but comprising 4 nitrogen levels. Figure 7 shows how this was achieved. The resulting grouping of sites meant that the variance components, going from the highest to the lowest levels, corresponded to

spatial variations over 140 m, 55 m, 30m and 17.5 m. It was somewhat unfortunate that the 17.5 m spacings were aligned E-W, and the 30 m N-S, as anisotropic variation could well affect the interpretation of the results. However, studies on a nearby field - D1 - showed that strong directional anisotropy was unlikely (Kool, 1981). Strictly speaking, this analysis of variance can be represented as a Mixed Model I (Model II system in which at the highest levels the blocks are fixed, and in which the lowest levels are pseudo-random. The actual estimation of variances with respect to conditions outside the field is difficult to judge. However, comparisons of the contributions to the mean squares from the different levels gives a good idea of where the greatest levels of variation lie.

Table 3 presents the results of the nested analysis of variance for all soil variables excluding those with a high proportion of missing values. The data are also presented graphically in Figure 8. The results confirm the results obtained from the one way analysis based on the major land use blocks, but show that the variability is even more short-range than that analysis showed. With the exception of LU1 and PE1 (not shown on the figure), every variate has more than 62% of the total variance present within 17.5 m, and most have more than 75% present. Only 5 variates, LU1, PE1, DE2, LU22 and LU3, have less than 75% of the variance present within 30 m, and only 2, LU1 and PE1 have more than 20% present at distances greater than 55 m.

Excluding the first horizon, horizon depths and thicknesses show the greatest variability. 95% Confidence limits for the distribution of values of these variables within 17.5 m range from +/- 10 cm for D2 to +/- 17 for DE5. Texture differences for the second and fourth horizons, and the clayiest texture of a fourth horizon layer also have high variability. This last is caused by the frequent absence of a clay sub-layer in the fourth horizon. All other variables have a low, and almost constant level of variability, indicating "homogeneity" over the whole field.

1.7 Ultra-short range studies of variability

Because of the high levels of variance shown above, a small sampling survey was undertaken with the help of Messrs Ovaa and De Smet to attempt to determine how much of the variance present within 17.5 m was present within much shorter distances. Six sites were randomly chosen, and sampled in duplicate 1 m apart. The 1 m replicates were also randomly oriented. The studies were confined to the top three deposits and the variance analysis was performed with respect to their thickness and depth at which they ended. Table 4 presents the results:

Table 3 Components of variance at different spatial scales for 144 soil profiles in the -1 and 2-year rotation blocks on PAGV1

Soil property	Overall mean	Variance	Components expressed as % of variance present within			
			17.5 m*	30 m	55 m	140 m
DE1	25.7	11.08	77	23	0	0
LU1	22.2	1.19	51	3	22	24
PE1	7.2	6.25	51	16	13	20
DE2	44.6	37.76	70	0	23	7
D2	18.9	37.77	68	8	14	10
LU2	9.8	11.42	77	14	0	9
LA2	1.0	.22	90	9	0	1
LU21	15.6	15.14	78	18	0	4
LU22	3.9	13.12	64	8	13	16
LD2	11.8	32.60	70	16	13	1
PE2	17.9	16.48	77	9	13	0
DE3	62.3	75.35	83	9	8	0
D3	17.7	43.13	68	16	10	6
LU3	15.1	8.21	62	11	20	7
LA3	1.9	.51	81	0	18	1
LU31	19.2	15.88	86	4	8	2
LU32	7.2	15.69	64	19	0	17
LD3	12.0	18.98	78	9	8	5
PE3	12.7	14.14	87	1	12	0
DE4	79.9	76.61	91	0	8	1
LU4	11.5	12.02	72	22	6	0
D4	17.7	42.69	78	15	1	6
LA4	2.2	.36	80	18	0	2
LU41	15.2	46.80	77	17	3	3
LU42	4.1	16.07	85	12	2	1
LD4	11.1	22.60	80	20	0	0
DE5	96.2	90.15	80	18	0	2
D5	16.2	37.60	87	12	0	1
LU5	6.6	16.64	74	20	0	6
LA5	1.8	1.95	93	0	0	7
LD5	5.8	17.73	83	7	0	10
LU6	3.4	9.19	82	0	13	5

*residual error included (see text)

Table 4 Estimates of variance over 1 m for selected attributes

Attribute	Total mean square	Mean square within 1 m	% of total
D2	28.85	29.66	100
DE2	12.28	8.25	67
D3	54.56	40.46	74
DE3	56.88	32.30	57

Because of the small sample (12 points) the estimates of total variance differ somewhat from the variance estimates obtained from all 144 sites. However, it is noteworthy that the *proportions* of variance occurring within 1 m remain high.

These results by themselves are not completely convincing as during the sampling some difficulties were experienced with the soil smearing in the Guts auger. However, linked with the studies on a 1 m transect on field D1 (Kool, 1981), which also showed very marked variations within one metre, one may reasonably conclude that a very large proportion of the variance of the attributes measured is not only present within 17.5 m, but is also present within distances of less than a few metres. These variations can be readily seen on the photographs of the profiles shown in Figures 5 and 6 in Ovaa's report (Ovaa op. cit). Estimates of measurement errors suggest variances of c. 6 cm² for depths, 12 cm² for thicknesses and 2.5-3% clay² for textures.

1.8 Principal component analysis of the soil data

Although the soil attributes showed different levels of variability, the common feature of short-range variation suggested that some of them might similarly covary -in other words that they might be correlated. The data from all 180 sites were analysed by principal component analysis to see if any important linear correlations occurred, and how great their role might be in explaining the total variability of the soil. Because of many missing values in the deeper layers, the number of variates analysed was limited to 32.

1.8.1 Correlation matrix

The correlation matrix proved to be disappointing in that most variates showed low correlations with each other, except for horizon thicknesses with horizon depths, and texture differences with texture extremes. These were not unexpected as horizon thicknesses and texture differences

Table 6 Contributions of original variates to the components

FILE MONAME (CREATION DATE = 02/10/80)

FACTOR MATRIX USING PRINCIPAL FACTOR. NO ITERATIONS

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
/	-.07909	-.26782	-.04990	.03681	.02358	-.19443
YH	-.35704	.55134	-.30136	-.01890	.00289	.08445
DE1	.12289	.09304	-.11515	-.22232	-.01916	-.01755
LU1	.11869	-.14766	-.03849	.19069	.29948	.26210
PE1	.12445	-.10412	-.13533	.34041	-.12056	.10756
DE2	.55032	.33046	.39558	-.32098	-.31634	.37336
LU2	-.40169	-.49997	-.04850	.19812	-.16907	.17159
LAP	-.00068	.20331	.28832	.32868	-.07796	-.21006
LU21	-.25536	.01590	.33459	.64232	-.42298	-.02086
LU22	-.17870	-.66951	-.40948	-.28363	.10325	.18396
PE2	.11525	.35901	-.08188	.02999	.28803	-.02439
DE	.49523	.28731	.45936	-.21197	-.31112	.38666
LD2	-.04903	.41732	.48110	.61901	-.35717	-.12688
DE3	.90317	.08814	-.00218	-.14958	-.15298	.11154
LU3	.08116	.73339	.06836	-.05957	.31931	-.10064
LA3	.62694	-.04978	-.06389	.23513	.34705	.26359
LU31	.16551	.52237	-.07484	.30160	.45121	-.15177
LU32	-.39527	.39445	.10564	-.45918	-.10423	-.42241
PE3	-.41201	-.17057	-.09402	.08459	.01047	-.27745
DE	.65570	-.21685	-.40761	.12575	.11306	-.24139
LD3	.45255	.08718	-.15579	.55832	.40504	.26808
DE4	.76522	-.30177	.26846	-.05075	-.14720	-.16922
LU4	-.58140	-.09883	.43950	.00210	.11231	.41593
LA4	.03628	.05242	.58183	-.21084	.45577	-.16287
LU41	-.56771	-.21304	.47161	-.03671	.36590	.31073
LU42	-.44976	-.07859	-.23012	.06855	-.22718	.41738
DE	-.15537	-.54869	.37260	.13818	-.00587	-.37684
LO4	-.28644	-.11975	.68259	-.14718	.43671	.03648
DE5	.62288	-.43357	.32684	-.06747	-.04844	-.09910
LU5	-.32305	.54539	-.30114	-.13563	-.15330	.16683
LA5	-.15038	.25660	-.00009	-.12736	-.13284	-.07333

were computed from horizon depths and texture extremes respectively.

Correlations between different layers was low. This is also not unexpected. The different horizons/layers represent geomorphologically distinct depositional phases. The soils are too young to show profiles development, apart from mottling and ripening, and real soil horizons have yet to form.

1.8.2 Eigenvectors and eigenvalues

Table 5 presents data for the first 10 eigenvectors, which together express only 74% of the total variability of the 32 variates used. This is not uncommon for soil data (c.f. Nortcliff 1978, Webster and Burrough 1971, Burrough and Kool, in preparation), but can be regarded as somewhat low.

Table 5 Eigenvectors of the first 10 principal components (180 sites, 32 variates)

Eigenvectors	Eigenvalue	% of variance	Cum. %
1	5.365	16.8	16.8
2	3.824	12.0	28.7
3	3.035	9.5	38.2
4	2.256	7.0	45.3
5	2.043	6.4	51.6
6	1.727	5.4	57.0
7	1.473	4.6	61.6
8	1.398	4.4	66.0
9	1.286	4.0	70.0
10	1.191	3.7	73.7

The eigenvectors are not easy to interpret in terms of the contributions of the original attributes. Component 1 is dominated by contributions from the thickness and depth attributes of layers (Table 6), though texture extremes of the 4th layer also play a part. Component 2 has largest contributions from the number of horizons, average texture, and most sandy texture of the second and third horizons, and by the thickness of the fourth horizon. Average texture of the fifth layer also contributes.

Component 3 is dominantly the layering and texture of the fourth layer with contributions from thickness and texture of the second. Component 4 has contributions mainly from texture and texture differences in layer 2 and 3.

Penetrograph data make insignificant contributions to all of the first six components.

Figures 9 and 10 display the contributions of the original variates to the first 4 eigenvectors in graphical form. Figures 11 and 12 are hand-contoured of these components. Figure 11 has a distinctly "spotty" appearance but Figure 12 appears to show consist differences between the northern and southern parts of the area.

1.8.3 Spatial analysis of the component scores

Component scores were computed for each of the 180 sites, for each of the first six components. The first 144 sites were analysed to determine the levels of spatial variation present in the same manner as used for the raw data (Section 1.6). Table 7 presents the results and Figure 13 displays them.

Table 7 Spatial variabilities of the first 6 principal components

Components	Proportion of variance over			
Number	17.5 m*	30 m	55 m	140 m
1	80	10	10	0
2	54	20	6	20
3	7	20	3	0
4	42	14	10	34
5	84	12	0	4
6	66	1	33	0

*residual error includes measurement error

These results display the not unexpected pattern of high residual error, all components excluding 2 and 4 having all variance present within 55 m, and components 1, 3 and 5 having more than 90% present within 30 m. Only component 4 has less than 50% present within 17.5 m. These results suggest that there is a part of the variation of the attributes contributing to components 2 and 4 that has a spatial variation distances of 30-100 m and this is supported by Figure 12. Further interpretation has little meaning for practical land use because such spatial variations would only account for something like 6% of the total variation of the measured attributes.

1.9 Comparison of the soil patterns in Figures 11 and 12 and the patterns of slaked soil (verslemping)

Comparison of the map of slaked areas made by the PAGV staff (Figure 13a) with the plots of the principal components fails to reveal any correspondence in pattern. One must conclude that the slaking pattern is unrelated to the soil properties represented by these components.

1.10 Conclusions about the nature of soil variability on PAGV1

From the foregoing, it can be concluded, in spite of the problems of rearranging the data for statistical analyses, that all measured soil attributes have the largest proportion of their variance present within a few metres. Estimates of measurement errors suggest that these often represent no more than 10-20% of the total variance, and thus do not affect these conclusions. Soil differences within trial blocks used for the yield experiments are thus highly likely to be greater than those between them.

In spite of low correlations and the high absolute variances of some soil attributes thought critical to plant growth, namely the thickness of the sand or clay layers in or directly under the root zone, the variations take place over such short distances that they are unlikely to have affected the outcome of the yield experiments. This hypothesis will be examined in the light of the yield data in the following section.

2 EXAMINATION OF SPATIAL VARIATIONS OF YIELD ON PAGV1

2.1 Data

Yield data are available for all three crops for the period from 1976 to 1979 inclusive, (excluding the 3-year rotation for 1976). The data used were net potato (>35 mm), net sugar yields and net wheat yields. The data were available for each 7.5 m x 55 m trial plot separately. However, each set of four of these plots formed a replicate unit when different nitrogen levels were discounted. Since there were only 3 soil profiles per set of 4 nitrogen levels (a block of 30 m x 55 m), all yield figures were converted to average yields/hectare within a 30 m x 55 m block. In this way, it was hoped to be able to determine if site and yield differences were related.

2.2 Problems of analysis

The analysis of yield differences with respect to soil differences is made very complex here because of the changing factors of the trial. Besides possible soil differences, the following contribute to yield differences:

1. Annual climatic differences
2. Rotation and crop type
3. Cultivation techniques (2-year rotations only)
4. Cultivation technique plot replicates within years
5. Cultivation technique plot replicates between years (2 and 3 year rotation areas)
6. Nitrogen levels and composting
7. Other factors such as disease, slaking, etc. that are difficult to quantify.

Of these, 7 must be regarded as residual error, and 6 has been removed by averaging yields over all nitrogen replicates (see 2.1). Rotation differences and plot replicate differences between years (2 and 5) can to some extent be compensated for by reducing all annual yield data per rotation period and crop type to zero mean and unit variance. Because the number of replicates and yield levels differ in the 1-, 2- and 3-year rotation, each are should be treated separately. Standardizing to zero mean and unit variance on an annual basis should also remove much of the yearly differences. Year differences can also be removed by calculating cumulative yields for each plot. Providing each has had the same history of cropping, comparisons will be valid.

Accordingly, two approaches have been taken. The first attempts to standardize the data and look for anomalous areas whose yield might be explained by site differences. The second sums yields over the 4 years and examines how 4-year yields relate to site and soil.

2.3 Data analysis via standardization to zero mean and unit variance

For each of the following areas, the 1-year potato plots, the 1-year sugar beet-plots, and the 2-year potato/sugar beet plots, standardized yields are computed for each year. The 1-year plots had 6 replicates each, the 2-2 year plots 18 replicates each. A map was drawn (Figure 14) showing for each year whether a particular plot had yields greater or less than 1 and 2 standard deviations from the mean.

Examination of the map shows that no plot was "good" or "bad" for all four years. Some plots were better some years than others, few were dominantly "good" or "bad". The most extreme picture is found for plots 177-180 (1-year potato rotation) which was "bad" for three of the four years. However, analysis of variance of the yields of the 6 1-year potato replicates grouped by years fails to reject the null hypothesis (Table 7)

Similar analyses were repeated for the one-year sugar beet yields, the sugar beet yields for the 2-year rotations in 1977 and 1979 from the more northerly 2-year block, and for all standardised 2-year yields over both 2-year replicate blocks. Table 8 presents the results. In no case was there a possibility of a conclusion that plot location had any influence on yields.

Table 8 Analysis of variance of standardized potato yields for 1-year replicates

Source	Degrees of freedom	Sum of squares	M. square	F
Sites	5	8.106	1.6212	
Sites within years	18	17.364	0.965	1.68
Total	23	25.470		

2.4 Data analysis via summed annual yields

For the 1- and 2-year plots, yields were summed over all four years. Each block (1-year potatoes, 1-year sugar beet, and each 2-year block) was handled separately thereafter because of the differences of crop and annual yield. For each of these 4 blocks means and standard deviations were computed. A map was drawn (Figure 15) showing for each trial plot per block which quartal the summed yields fell in. Table 9 presents the means, standard deviations and coefficients for variation per cent for these four blocks.

Table 9 Means and standard deviations of summed yields

Trial block and crop	4-year average yield(ton/ha)	Standard deviation	Coefficient of variation %
1-year potatoes	108.63	2.58	2.38
1-year sugar beet	40.68	1.03	2.54
2-year south block	87.18	1.56	1.79
2-year north block	101.02	2.80	2.77

The coefficients of variation are all low, indicating considerable homogeneity in the areas. From both Figure 13 and Table 9 it can be seen that the southerly 2-year block is more uniform than the others: in fact, all plots return similar yields except for a single plot (rows 121-124). This area returned an exceptionally low potato yield in 1979, but has otherwise performed normally. There is thus no reason to think that the soil here is responsible.

The apparent contiguity of the "good" and "bad" plots in the northerly 2-year block suggests that soil factors might be playing a role in affecting yields. This is extremely difficult to pin down however, because of the obscuring effects of annual variations in climate and crop, plus the differences in management practised in the 2-year blocks. Correspondences between yield differences and soil differences can be sought through 2 methods:

1. On the basis of Figure 15, stratify the soil profiles into 4 classes and see if there are statistically significant differences in soil parameters for the classes
2. Within a given 2-year rotation block, look for correlations between the 4-year yield and soil parameters.

2.4.1 Stratification of soil profiles according to the four quartal annual yield classes

The profiles from the 144 sample points under the 1- and 2-year rotation blocks were grouped according to Figure 13.

The grouping was as follows:

Group 1-	lowest yield quartal	27	sites
" 2 -	next lowest "	39	"
" 3 -	next highest "	15	"
" 4 -	highest yield quartal	63	"

No significant differences were found for any of the 35 soil attributes analysed (Table 10), nor for the principal components.

Table 10 Analysis of variance of soil parameters according to 4 yieldclasses

Attribute	Within MS (DF 140)	Between MS (DF 3)	Fratio
NH	.760	.455	.598
DE1	11.209	5.067	.452
LU1	1.108	1.532	1.383
PE1	5.986	1.692	.283
DE2	36.737	13.021	.354
D2	36.834	20.380	.553
LU2	10.327	4.454	.431
LU21	14.932	9.585	.642
LU22	12.550	10.066	.802
PE2	15.284	32.198	2.107
DE3	73.502	29.698	.404
D3	41.693	64.875	1.556
LU3	8.0831	2.527	.313
LU31	10.333	5.511	.533
LU32	12.247	7.361	.601
PE3	14.229	1.736	.122
DE4	7.277	16.008	.222
D4	42.594	1.432	.034
LU4	11.518	23.107	2.006
LU41	17.279	11.433	.662
LU42	7.476	7.716	1.032
PE4	14.751	0.115	0.008
DE5	90.332	1.544	0.017
LU5	13.972	2.215	0.159
LU51	23.088	22.373	0.969
LU52	6.497	6.884	1.059
LU6	.534	.254	.478

2.4.2 Correlations between yields and soil parameters for the 2-year rotation blocks

Each 2-year rotation block was treated separately because although each had had two potato and two sugar beet crops, these had been in different years. Yearly differences were much greater than other differences. For each soil profile within a block (54 profiles) the 4-year average yield data were added, and the whole set submitted for principal components analysis in the same manner as described in Section 1.8. The results were as follows:

Block 1 (north). Yield shows only weak correlations with the following soil properties - PE1 (.207), LA2 (-2.09), LU21 (-.251), LU22 (.254), LD2 (-.261), PE3 (-.230), LU4 (.258), LA4 (-.234), LD4 (-.292*). Of these, only the last is weakly correlated at the 5% level.

These correlations could be interpreted as showing a weak increase in yields for less layered, less texture-contrasted dominantly sandy horizons. This interpretation can be examined in the light of the map of the distribution of the second Principal Component calculated from all 180 sites (Section 1.8, Figure 12). This map shows that the second component appears to have a more regional distribution than the first, and that the more northerly parts of the area tend to have higher values of that component than occur in the south. Higher values of component 2 reflect less textured contrasted and sandier Zuiderzee deposits in combination with sandier Almere 1 deposits. In other words, profiles that have lower clay contents, and thus possibly lower moisture and mineral supply just in and under the root zone than average. Comparison of Figure 15 with Figure 12, however, shows that the link between yields and values of the second component are little more than vague.

Block 2 (south). Here the 4-year yields showed weak correlations with PE1 (-.303) and LA4 (-.26). All other correlations were less than 0.2.

2.5 Comparison of yield patterns with slaked ground patterns

Comparison of the pattern of slaked ground (Figure 13a) with the yield patterns (Figures 14 and 15) fails to show any correspondence. Areas showing the most extreme slaking also return above average yields, while areas having little slaking have shown lower than average yields.

2.6 Conclusions from the yield data

1. It is difficult to isolate site yield differences from the effects of climate/year, crop type, cultivation technique and other factors, When this is attempted, it appears that yield differences due to site are insignificant. Yields show no correlations with soil profile differences.
2. The variation in yields over the whole of PAGV1 is small. Yield variation is greatest in the most northerly 2-year rotation block where there is a suggestion of a very weak link between yield and the properties of the Zuiderzee and first Almere deposits. The links are so weak, however, as to be statistically insignificant when studied using the existing data. There are no suggestions of a link between patterns of slaking and long term yields.

3 Overall conclusions

1. The thickness and depth of the Zuiderzee and Almere deposits in the area of PAGV1 vary considerably, whereas texture, and texture differences within deposits vary much less. The bearing strength of the ground, as measured by a penetrometer is also relatively uniform over the whole field. The thickness and depth of the deposits vary not only absolutely, but also over very short distances (of the order of a 1 to 10 metres). Such differences are not mappable without a ridiculously high sampling density.
2. Yield does not appear to vary in a manner directly related to soil or site. Since soil differences that could affect yields are mainly present over very short distances, large site-related yield differences should not be expected. There is a weak suggestion that the Zuiderzee and Almere 1 deposits in the northern 2-year rotation block are sandier, and may have depressed yields slightly, but the effect is so slight as to be statistically insignificant, given the present data.
3. Considering the large differences in yields resulting from annual and rotation differences, any yield differences of the fields resulting from a soil or site component are so small that they appear to be negligible). For the purposes of the trials on PAGV1, the soil is homogenous.

REFERENCES

- Burrough, P.A. (in preparation). 1/f noises in soil distributions.
- Burrough, P.A, and Kool, J.B. A comparison of statistical techniques for estimating the spatial variability of soil properties in trial fields. (Submitted to ISSS International Colloquim on Handling of Soil Data 14-17 Sept. Paris, France 1981.)
- Graaf, J. v.d., 1979. Verslemping en Remote Sensing, PAGV Lelystad.
- Kool, J.B., 1981. De bodem variabiliteit van het proefveld D1 te PAGV, Lelystad. PAGV Lelystad.
- Lamers, J., 1980. Continuïteit en nauwe rotataties van aardappelen en suikerbiet: resultaten van vruchtwisselingsonderzoek op PAGV1 van 1973 t/m 1978. PAGV, Lelystad. Publicatie nr. 12.
- Nortcliff, S., 1978. Soil variability and reconnaissance soil mapping: a statistical study in Norfolk, J. Soil Sci. 29(3) 404-419.
- Ovaa, I., 1980. De bodemgesteldheid van het proefveld PAGV1 te Lelystad. Rap. 1530, Stichting voor Bodemkartering Wageningen, 48 pp. ISBN 90 327 0085 5.
- Pons, L.J. en A.J. Wiggers, 1959. De holocene-wordingsgeschiedenis van Noordholland en het Zuiderzeegebied. Dl. I Tijdschr. Kon.Ned.Aard.Gen. 76, 104-152 1960 De holocene-wordingsgeschiedenis van Noordholland en het Zuiderzeegebied Dl. II. Tijdschr. Kon. Ned. Aard. Gen. 77, 3-77.
- RIJP, 1962. Bodemopname van Oostelijk Flevoland. Rijksinstituut voor IJsselmeerpolders.
- Webster, R., 1977. Quantitative and numerical methods in Soil Survey O.U.P.
- Webster, R. and P.A. Burrough, 1971. Computer based soil mapping of small areas from sample data I. Multivariate classification and ordinations. J. Soil Science 23(2) 210-221.

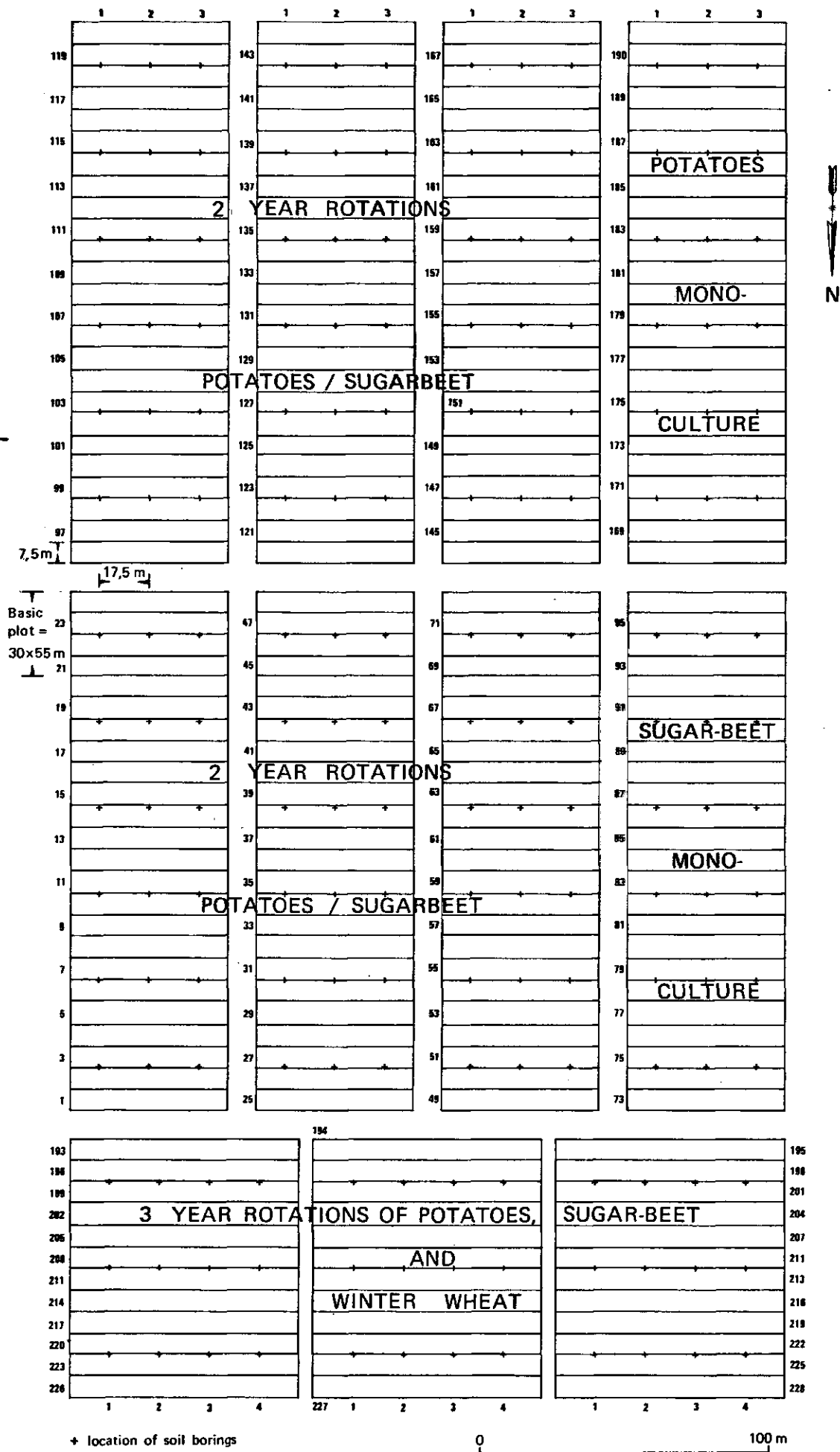


Figure 1 Layout of the PAGVI trialfield

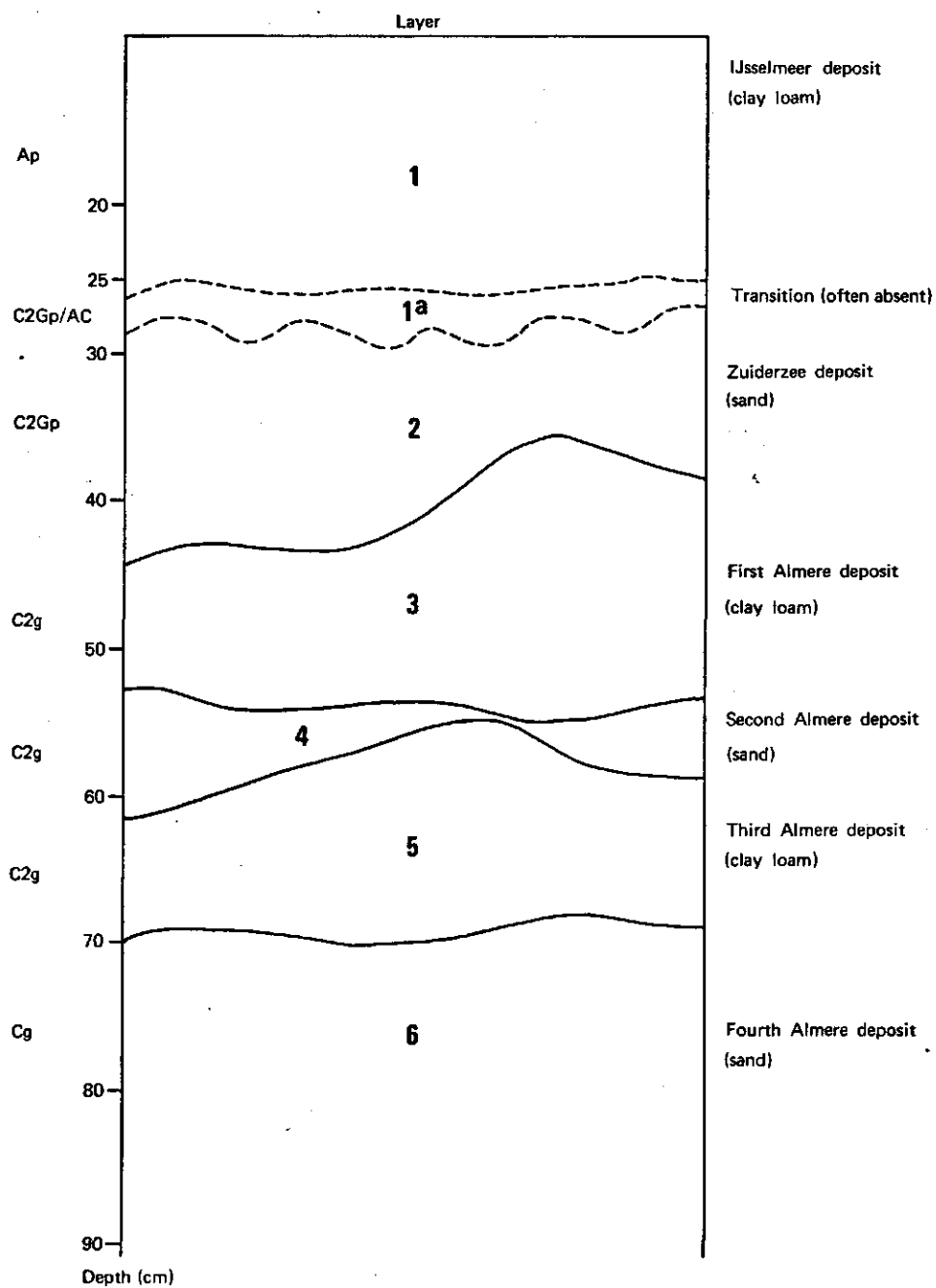


Figure 2 Profile schematics of the soil under PAGVI

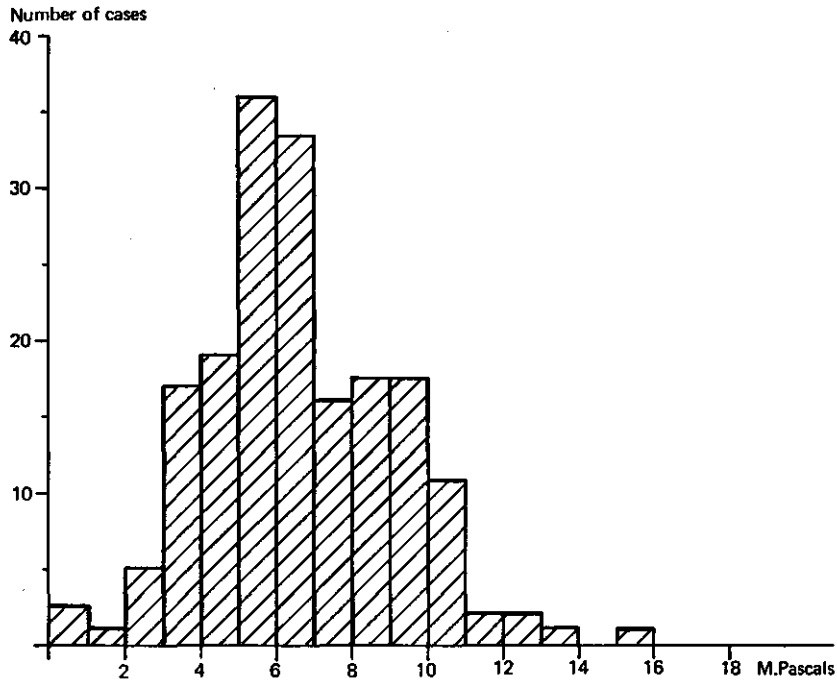


Figure 3 Soil mechanical resistance (cone penetrometer) of top soil (0-20cm)

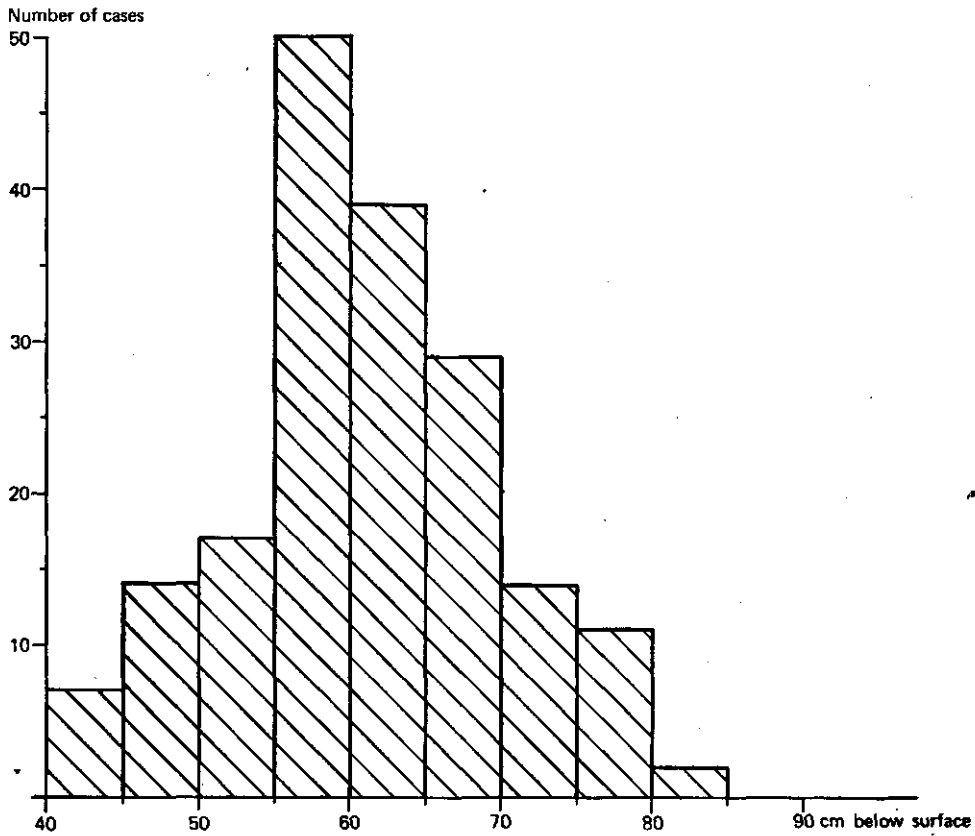


Figure 4 End depth of Almere clay (upper) deposit/depth at which Almere sand (upper) begins

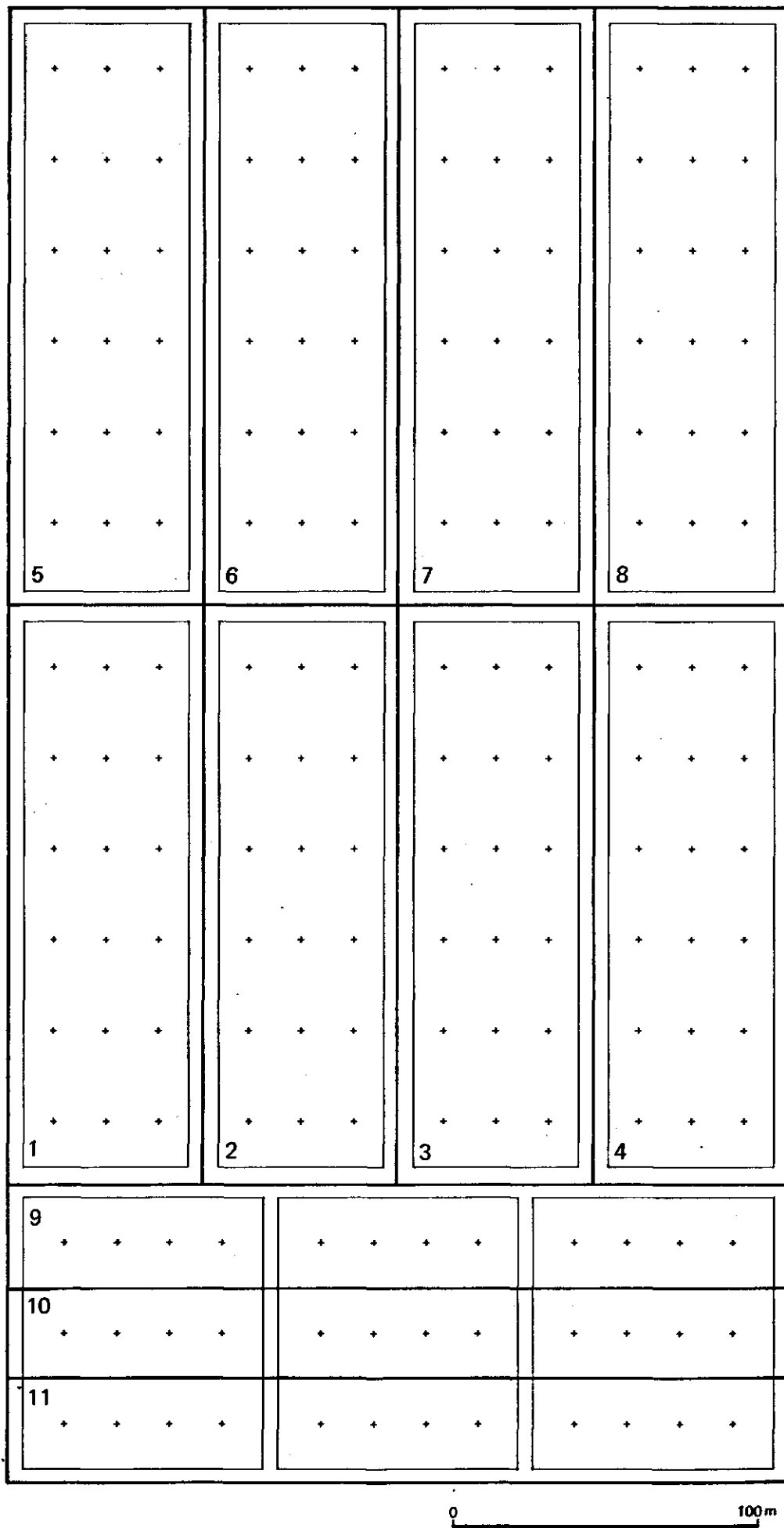
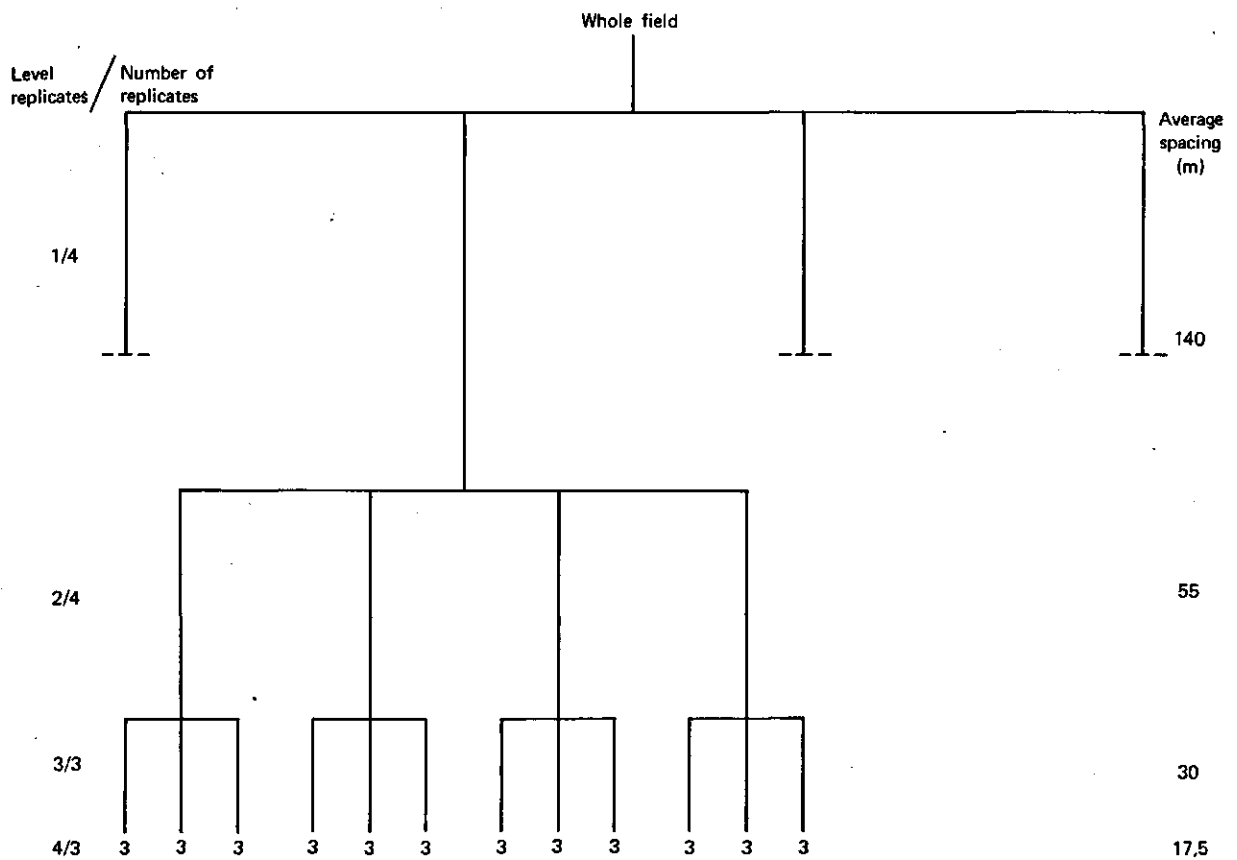


Figure 5 Blocks for oneway anovar



Analysis of variance

Source	Degrees of freedom	Sums of squares	Variance components
level 1	3	$\sum_{i=1}^4 36(\bar{x}_i - \bar{x})^2$	$\sigma_4^2 + 3\sigma_3^2 + 9\sigma_2^2 + 36\sigma_1^2$
level 2	12	$\sum_{i=1}^4 \sum_{j=1}^3 9(\bar{x}_{ij} - \bar{x}_i)^2$	$\sigma_4^2 + 3\sigma_3^2 + 9\sigma_2^2$
level 3	32	$\sum_{i=1}^4 \sum_{j=1}^3 \sum_{k=1}^3 3(\bar{x}_{ijk} - \bar{x}_{ij})^2$	$\sigma_4^2 + 3\sigma_3^2$
level 4	96	$\sum_{i=1}^4 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 (x_{ijkl} - \bar{x}_{ijk})^2$	σ_4^2
Total	143	$\sum_{i=1}^4 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 (x_{ijkl} - \bar{x})^2$	

Figure 6 Nested anovar for PAGVI

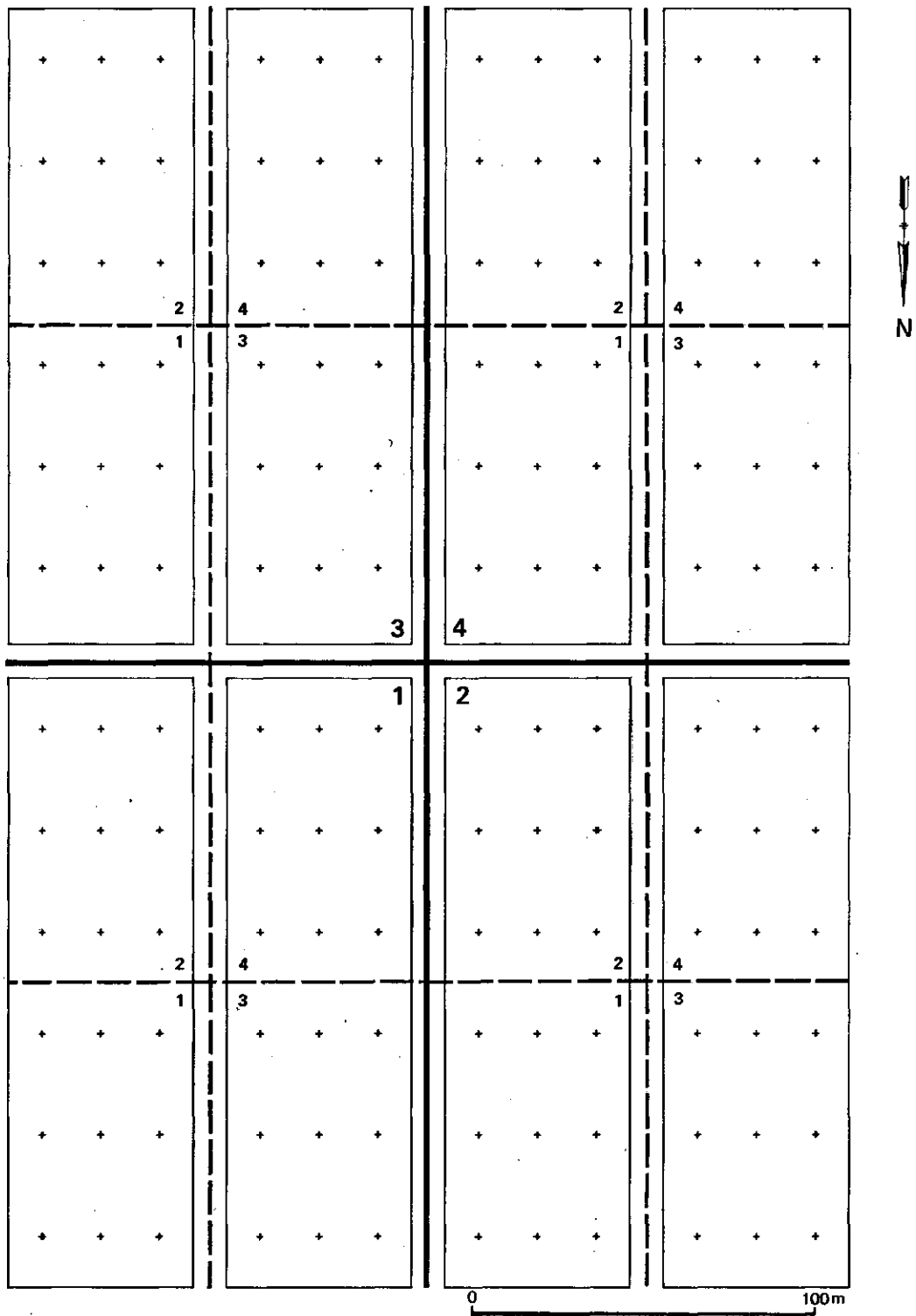


Figure 7 Blocks for nested anovar

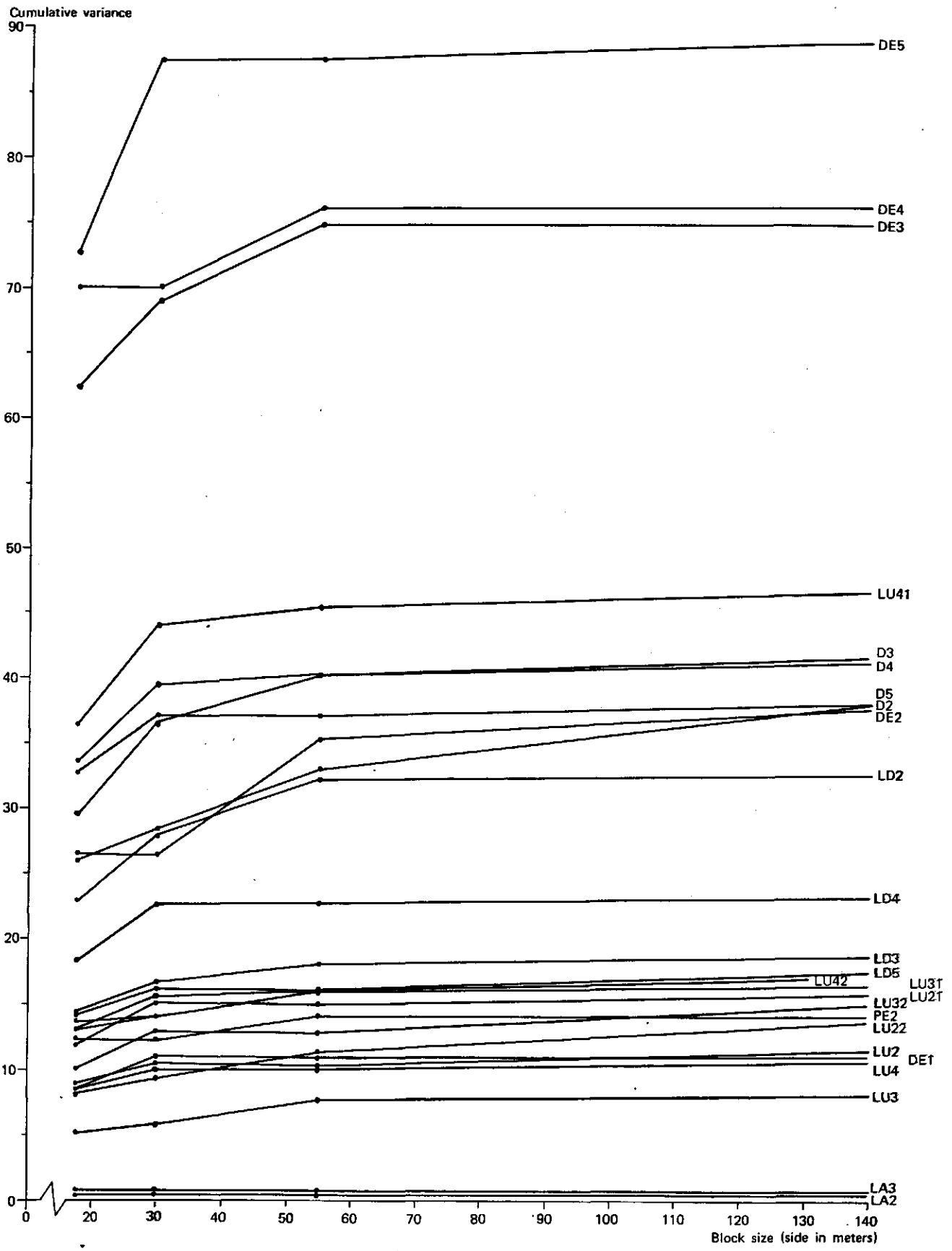


Figure 8 Sources of variance with respect to sample spacing/block size

FILE NONAME (CREATION DATE = 02/10/80)

- 1 = Z
- 3 = DE1
- 5 = PE1
- 7 = LU2
- 9 = LU21
- 11 = PE2
- 13 = LD2
- 15 = LU3
- 17 = LU31
- 19 = PE3
- 21 = LD3
- 23 = LU4
- 25 = LU41
- 27 = O4
- 29 = DES
- 31 = LAS
- 2 = NH
- 4 = LU1
- 6 = DE2
- 8 = LA2
- 10 = LU22
- 12 = D2
- 14 = DE3
- 16 = LA3
- 18 = LU32
- 20 = D3
- 22 = DE4
- 24 = LA4
- 26 = LU42
- 28 = LD4
- 30 = LU5
- 32 = LU6

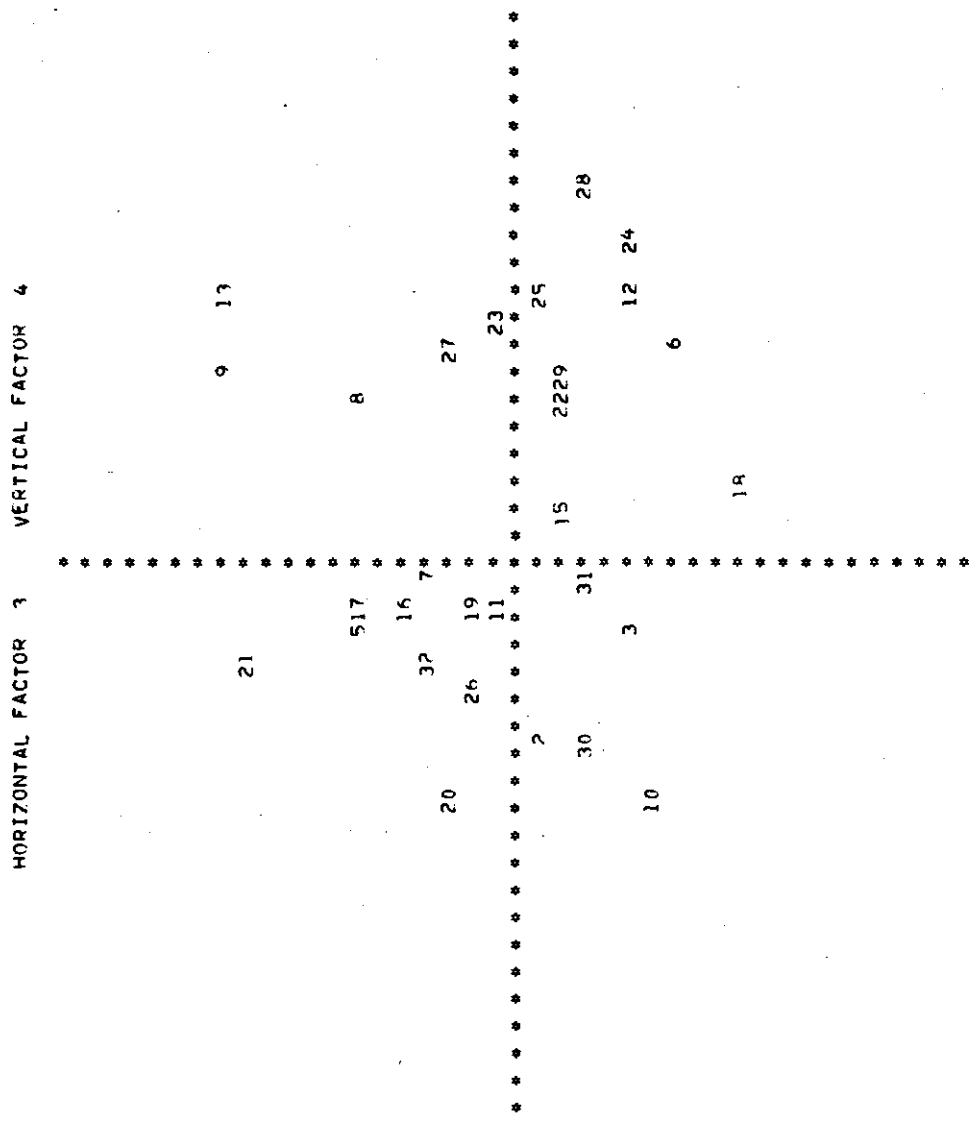


Figure 10 Contributions of variates to Components PC3 x PC4

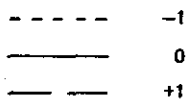
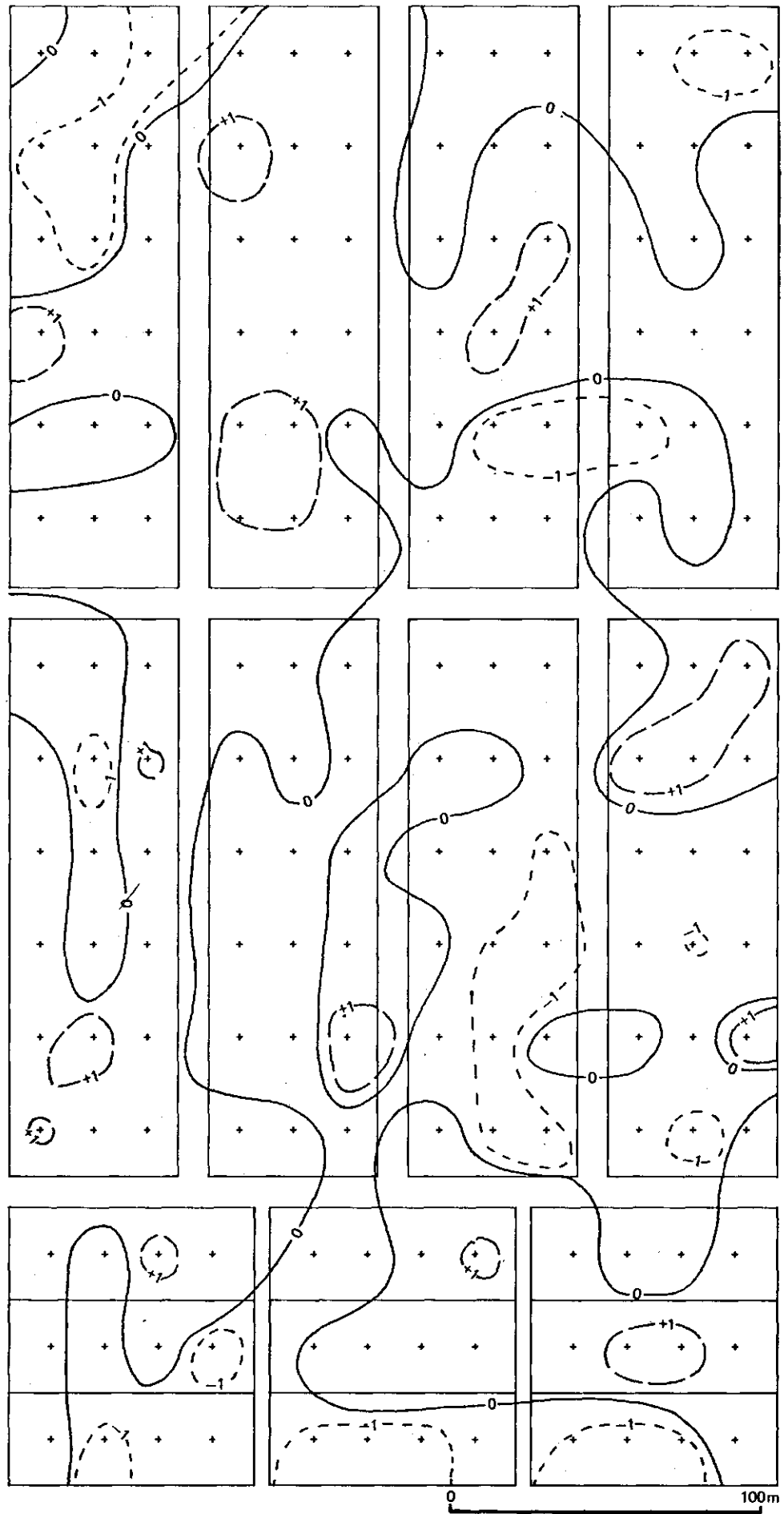


Figure 11 Isoline map of first principal component scores

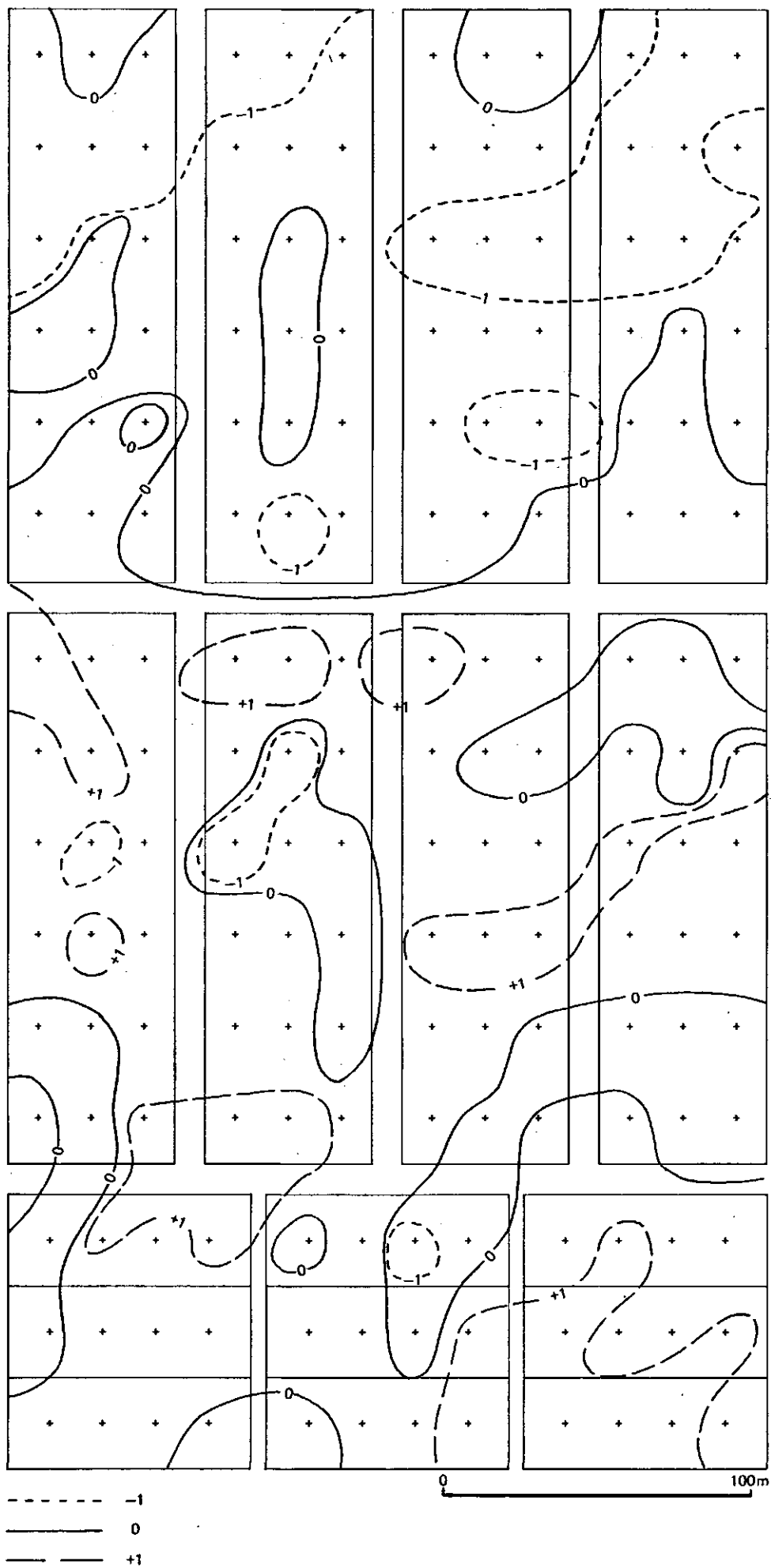


Figure 12 Isoline map of second principal component scores

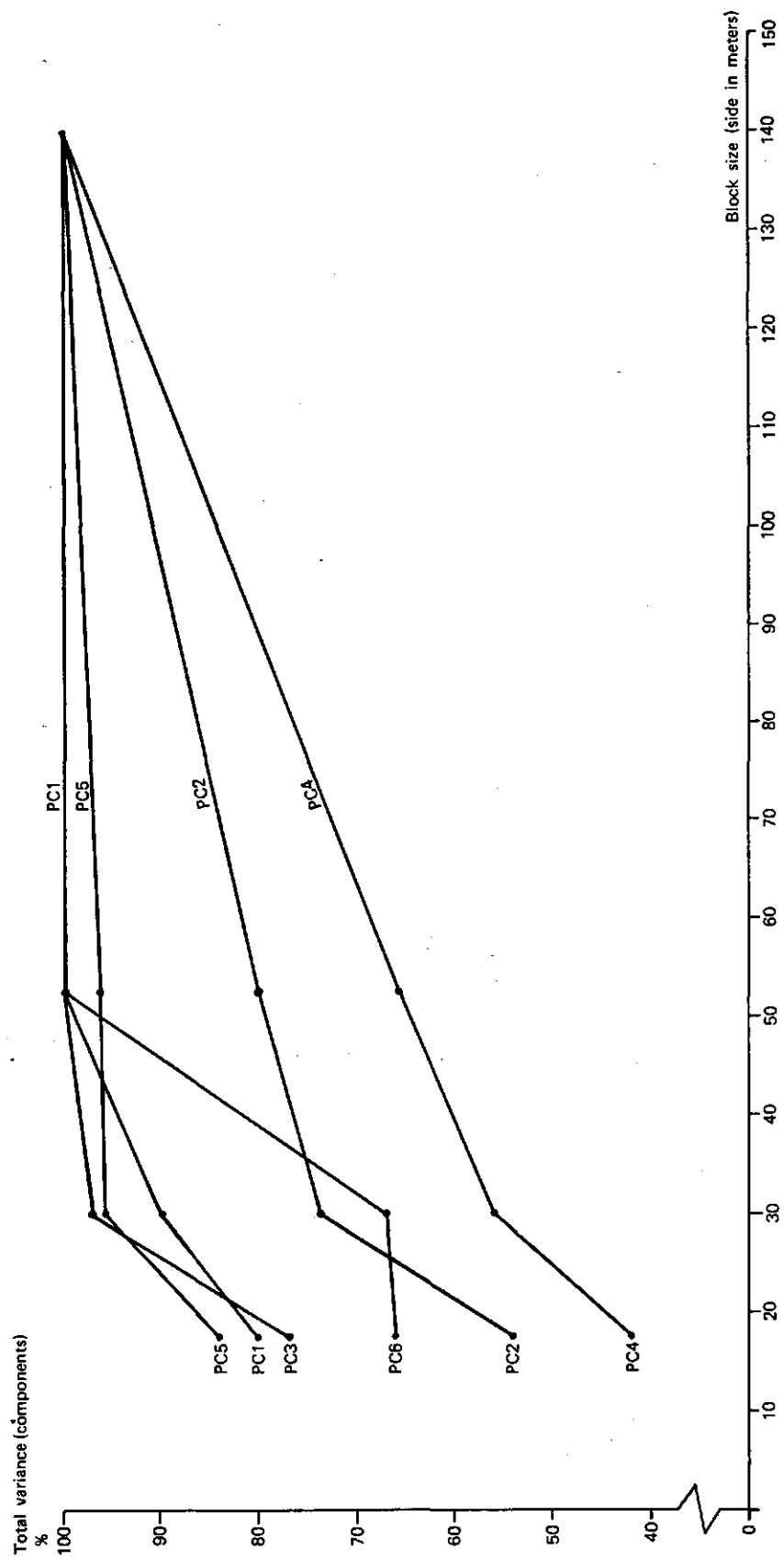


Figure 13 Variability vs distance for principal component scores (144 sites)

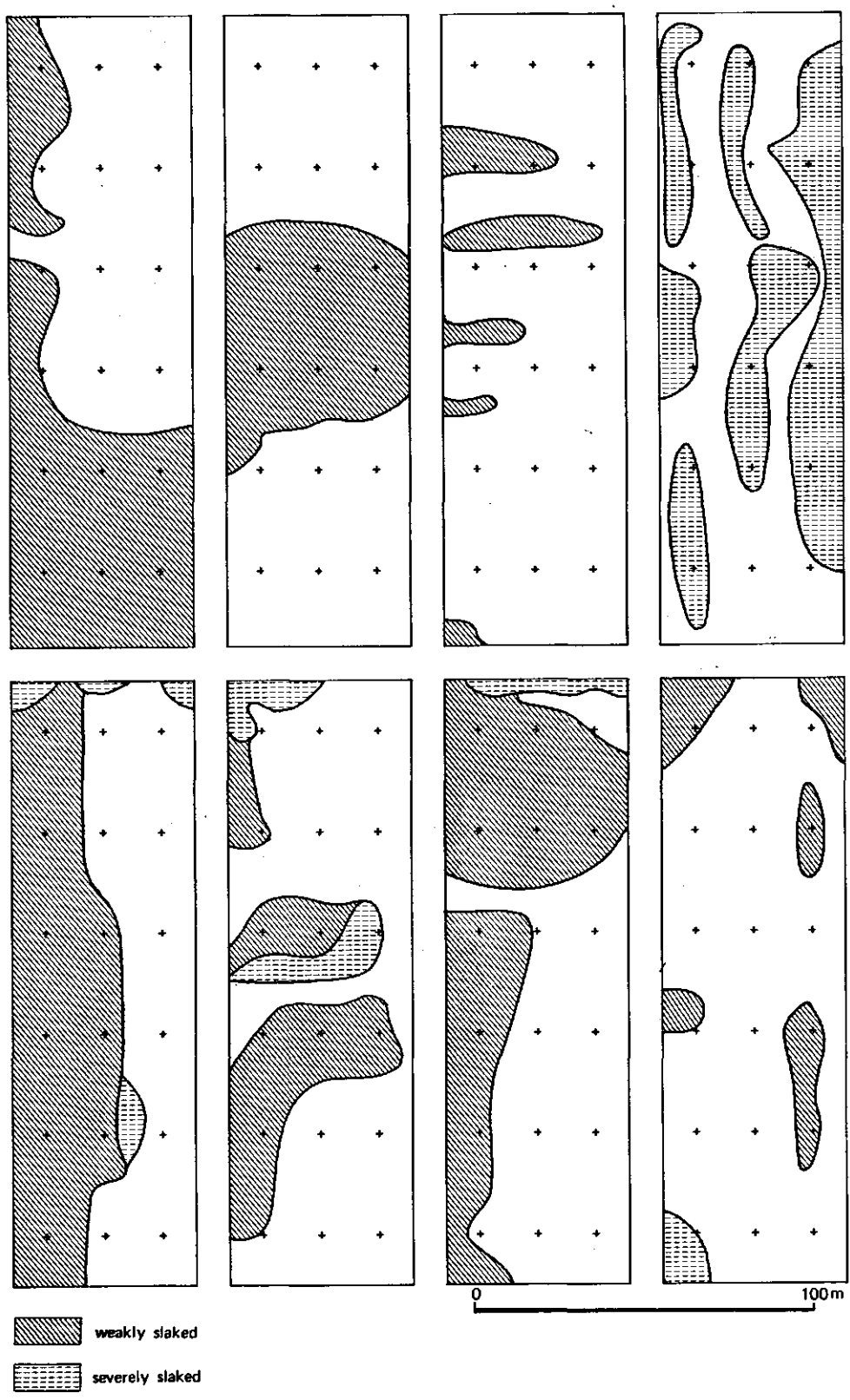


Figure 13a Pattern of slaked top soil (1979)

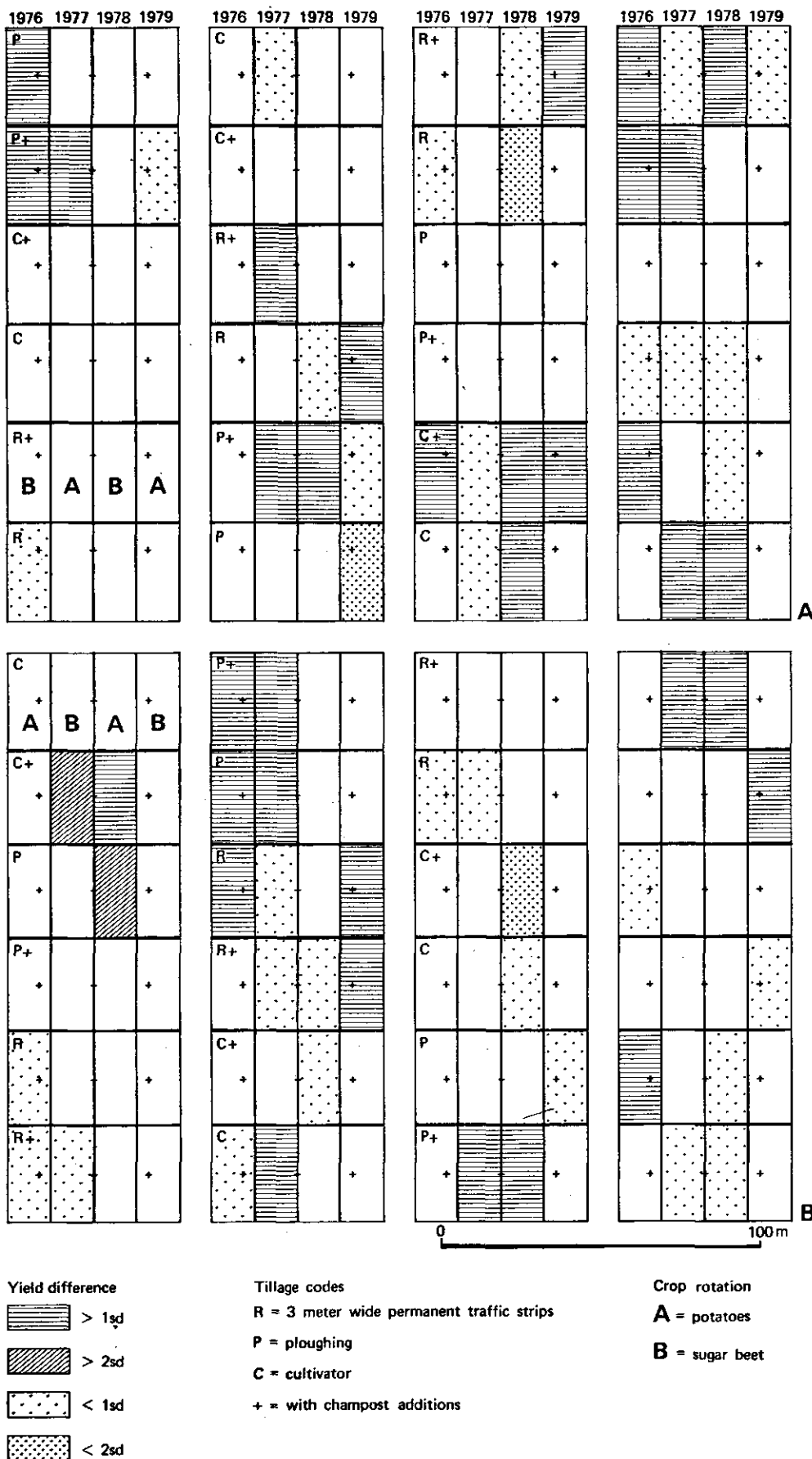


Figure 14 Standardized yields 1976 - 1979 Yields are standardized in terms of $\frac{x - \bar{x}_{ij}}{sd_{ij}}$
 - for 2 year and 1 year rotations where \bar{x}_{ij} = mean yield year i for rotation j
 sd_{ij} = std.dev. yield year i for rotation j

