Segregation: the interaction between spatial pattern and gene flow

Tremlett, M.

How to cite:

Tremlett, M. (1981) *Segregation: the interaction between spatial pattern and gene flow*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/7599/

SEGREGATION: THE INTERACTION BETWEEN SPATIAL PATTERN AND GENE FLOW.

Dissertation for the Degree of MSc. in ECOLOGY

October 1981.

Mark TREMLETT,
Zoology Department, Science Site,
The University of Durham
South Road, DURHAM.

CONTENTS.

LIST OF FIGURES IN THE TEXT.

LIST OF FIGURES IN THE TEXT. Contd.

LIST OF TABLES IN THE TEXT.

LIST OF GRAPHS IN THE TEXT.

LIST OF GRAPHS IN THE TEXT. Contd.

LIST OF GRAPHS IN THE TEXT. Contd.

1 INTRODUCTION.

Much of the ecological research done to date has been descriptive in nature, trying to quantify those processes at work in the environment. Predictive ecology is one of the newest and most exciting fields in this rapidly expanding science. This project is set in the latter vein, being concerned with modelling the evolution of spatial pattern in plant communities, based on the theoretical principles of population genetics, and comparing the predictions of that model with observed spatial patterns in field populations. The model used was a computer program called SPEV, short for spatial evolution, written in AlgolW, by Dr. J.T. Gleaves. It was intended to study segregation, that is the separation of plant populations into discrete single species groups. It was hoped to test the hypothesis that as a result of genetic interactions between individuals of two species, where their hybrids are infertile or inviable, plant populations containing both species will aggregate into groups of one species or the other. This hypothesis was tested using computer simulation results and also comparing these results with field data collected on spatial pattern in populations of _Senecio squalidus L._ and _Senecio viscosus L._, two species of contrasting breeding system. _S.squalidus_ being an outbreeder, and _S.viscosus_, an inbreeder. A variety of field techniques were used, as were a range of statistical treatments of the data collected and some discussion of each is given.

2 METHODS.

2.1 INTRODUCTION

Two very different approaches were used in the study of spatial patterns in plant populations.

2.1.1 Computer Simulation

Populations of plants were simulated using SPEV on the NUMAC computer. The effect of selection, competition, gene flow and seed dispersal were modelled for a simulated population. The impact of these parameters on spatial pattern in the model population, was studied. This will receive thorough discussion in the next chapter.

2.1.2 Field data.

Spatial pattern in field populations of Senecio viscosus L., and Senecio squalidus L. was studied. These two species were chosen because of (a) their relative commonness, and (b) their similar niches. Both plants are common plants of derelict land, especially on old railway lines, and the sites of derelict buildings. (Spatial pattern data was also collected for Senecio jacobaea L., a common plant of derelict agricultural land, this data is included in Appendix A).

Senecio squalidus L. is described by Clapham, Tutin and Warburg, (1959) as:

"Overwintering annual, rarely longer lived with more or less glabrous, tough, diffusely branched ascending stems, 20 to 40

> cm high. Lower leaves narrowed into a winged stalk, upper half clasping; all usually deeply pinnatifid with oblong entire or toothed lobes,...., Heads 1.5 to 2.5 cm diameter in an irregular loose corymb. Involucre bell shaped,...., inner bracts 21; outer 5 to 13; very short; all conspicuously blacked tipped. Spreading rapidly on old walls, waste ground, railway banks, bombed sites etc., throughout England to Yorkshire and Lancashire, being rare in Scotland."

S.squalidus was first introduced to the Botanic Gardens at Oxford in 1699, being recorded in the environs of Oxford in 1794, (loc cit in Lousely 1943,1947). Since that time, especially with the advent of the railways in the last century, S.squalidus has been progressively extending its range throughout the country, such that by now it is common in most vice counties except in the North and West of Scotland, where it is rare. It is a large conspicuous plant whose life cycle is certainly overwintering annual, and arguably longer lived. Gibbs, Milne and Carillo (1975) in their study of the breeding systems of members of the genus Senecio, record S.squalidus as a cross pollinating species, with large showy flowers, high pollen counts, and low recombination index, all characteristics of a cross pollinating species. Individuals in the Durham sites studied were found in flower in late May, and were still in flower in mid-September.

Senecio viscosus L. is described by Clapham, Tutin and Warburg (1959) as:

> "Annual, foetid, with erect very viscid glandular hairy stems, 10 to 60 cm high. Leaves dark green, glandular and very viscid,

deeply pinnatifid with nearly equal toothed or pinnatifid lobes; lower obovate in outline, short stalked, upper oblong and sessile. Heads 8 mm in diameter, long stalked, in a large irregular rounded corymb. Involucre ovoid-conical, densely glandular, its outer bracts almost half as long as the inner ray florets Ray florets c.13, short revolute,..., Probably native. Waste ground, railway embankments and tracks, sea shores etc. Locally common throughout lowland England."

S.viscosus is probably a native of cliffs and costal regions in the British Isles, (Lousley, 1943), having expanded onto road and railway works in the last century. S.viscosus is a self pollinating species whose characteristics are very unlike those of S.squalidus, having less conspicuous flowers, low pollen counts and high recombination index, (Gibbs, Milne and Carillo, 1975).

S.squalidus is native of the southern Mediterranean, its classical habitat being on the slopes of Mount Etna, (Lousley, 1943), thus in its natural environment it is very unlikely to encounter S.viscosus, a species restricted to wetter climes in western Europe. However, since the escape of S.squalidus from the Botanic Gardens at Oxford, and particularly during the last war, which left many areas of derelict land in Britain's major cities, both species encountered one another, on a large scale, for the first time. As a result, a hybrid Senecio xlondenensis Lousley., intermediate in characters between the two parent species, was found, occurring at low frequencies, where populations of the parents met.

S.squalidus, S.viscosus and S.xlondenensis provided useful subject material for the study of segregation and spatial pattern, since both parents are relatively common plants, (although the hybrid is less common, unless there are large populations of both parents), and the environments in which they are found are relatively homogeneous waste ground. Spatial pattern was studied in single and mixed species populations, and a brief site description is given for each of the sites at which field measurements were collected.

Table 1. Site description.

		Smplng	Species		Species		Smplng	
		frame	present		sampled		method	
No	Site Name	No.	sq	vis	sq	vis		Environment
1	Gilegate							Builders rubble
	Builders	1,2	+		+		T.sq	tiles, slates,
	yard Durham							waste ground.
2	Stockton &							Disused railway
	Darlington	3-6	+	+	+	+	T.sq	line, and
	Old Railway	21,22	+	+	+	+	T.sq	embankment.
3	Abandoned							Over grazed
	Field	7-9		S.jacobaea			T.sq	pasture
	Pittington	18		S.jacobaea			T.sq	
4	Old Quarry							
	Entrance	10-11		+		+	T.sq	Rubble,
	Coxhoe	16		+		+	T.sq	road aggregate.
5	Old Field							Derelict
	Quarrington	12-15	+		+		T.sq	cultivated
	Hill							field.

Abbreviations: Smplng = sampling, sq = S.squalidus, vis = S.viscosus
T.sq = T-square, Qrdrt = grid of quadrats.

Table 1. (Contd.) Site description

		Smplng frame	Species present		Species sampled		Smplng method	
No	Site Name	No.	sq	vis	sq	vis		Environment
6	Belmont Viaduct Field	17	S.jacobaea				T.sq	Overgrazed field.
7	Derelict land, Darlington	19, 20		+		+	T.sq	Derelict factory land. Rubble loose top soil.
8	Disused sidings Croft.	23-27	+	+	+	+	T.sq Qrdrt	Railway aggregate.

The grid of contiguous quadrats laid out at site 8, covered four of the T-squares sampling frames (numbers 24 to 27), and the positions of the sampling frames on the grid of quadrats are given so that a comparison of the data from T-square sampling, and the layout of the plants on the ground may be made*.

Spatial pattern was studied using both distance and quadrat methods, a description of these, and the statistical approach to the data collected, is given in the following text:

---

* See section 6.3, Figure 22.

a Distance methods. Besag and Gleaves (1973) propose T-square sampling, a density independent distance method, as a quick 'pilot' technique for assessing spatial pattern in plant populations. T-square sampling is described at length in section 2.2.

b Quadrat methods. These are discussed by Greig-Smith (1964). Two uses of quadrats are described, (1) using random quadrat throws, and (2) using a grid of contiguous quadrats. The amount of information to be obtained from random quadrats is limited, since this only provides information about spatial pattern at the scale of one quadrat size or less. A grid of contiguous quadrats will provide information about spatial pattern at more than one scale. Both techniques are described in section 2.6.

2.2 T-SQUARE SAMPLING.

This method, proposed by Besag and Gleaves (1973), involves measuring the distances between a random point and its nearest plant, and that nearest plant and its nearest neighbour within a given arc.

2.2.1 T-square sampling theory.

Besag and Gleaves (1973), Diggle, Besag and Gleaves (1976).

"Let S denote a sampling frame, containing many events chosen to lie slightly within the region of interest, A, so as to

> eliminate edge effects. Let P denote a randomly selected point in S (see fig. 1), and define U to be the squared distance from P to the nearest event Q. Let PQT denote the perpendicular to PQ passing through Q, and define VT to be the squared distance from Q to the nearest event, excluding those which lie on the same side of PQT as does P itself. Choosing m random points, the sample u and vt of observations on the variate U and VT are obtained."

The Test statistics are defined below, together with their distributions under the null hypothesis of spatial randomness. All the tests are summations over the range $i=1,\ldots,m$.

2.2.2 The Normal Test.

(Besag and Gleaves, 1973).

$$t_N = m^{-1}\sum\left(u_i \,/\, \left\{u_i + 0.5\, vt_i\right\}\right)$$

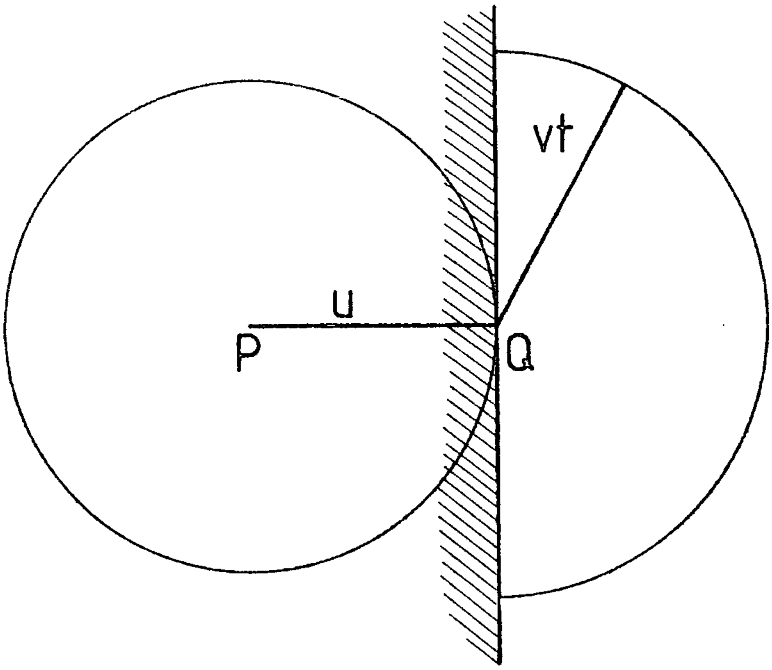
t_N has an approximately normal distribution with mean, 1/2, and variance, 1/12m.

2.2.3 Further Tests.

Diggle, Besag and Gleaves (1976) describe a further test:

$$t_B = \sum u_i \,/ \sum(u_i + 0.5\, vt_i\,)$$

Figure 1 T-square sampling



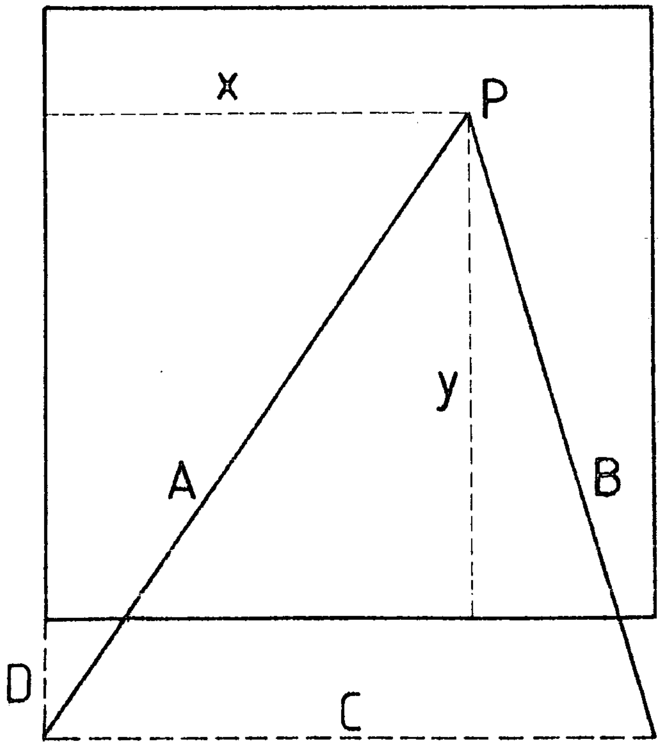


Figure 2 Two tape point location method

which has a B(m,m) distribution under the null hypothesis. (This may easily be transformed to an F(2m,2m) statistic, since F tables are more readily available). Cormack proposes another test (Cormack, 1979), however, this test was not used in this project. Each of these test statistics was evaluated for each of the field populations for which T-square data was collected.

2.3 THE LOCATION OF RANDOM POINTS IN THE SAMPLING FRAME

The location of random points in the sampling frame, S, was very much simplified by the use of "two-tape" methods (Gleaves, personal communication). Two 30m. tape measures were fixed to the ground a known distance, C, apart, and set back a distance, D, (D = C/5), from the edge of the sampling frame. Charts had previously been prepared for scoring the data, and on these the x and y co-ordinates of random points within the sampling frame were printed. In addition, two distances, A and B, were calculated and printed on the data charts where:

$$A = \sqrt{x^2 + (D + y)^2}$$

and

$$B = \sqrt{(C - x)^2 + (D + y)^2}$$

These two distances were measured out, one from each tape measure, and used to locate each random point. This is best explained by fig. 2. This technique reduced the amount of time taken to locate random points, and also the inaccuracies inherent in setting up a grid, of strings or tapes, over the sampling frame, as had previously been used.

2.4 COLLECTION OF THE FIELD DATA.

Most of the sampling frames in the field were of 5m. x 5m. dimensions, and in most cases, except where the population was sufficiently small to count the number of individuals present, an estimation of the population size was made by placing 13, 0.2m. random quadrats. (Thirteen random quadrats were used since this meant that approximately 1/10th of the sampling frame was enumerated. This was considered to be sufficient to obtain an estimate of the population size in the sampling frame). An estimate of the population in the sampling frame was obtained since, if the sampling intensity, m, exceeds n/6, where n = the population size, then the independence of each pair of T-square measurements may be lost. (Diggle, Besag and Gleaves suggest that a sampling intensity of n/10 should be both reasonable and practicable). T-square sampling is density independent so a fixed number of random points per sampling frame could have been used. It was decided to sample at n/6 regardless of the population size, thus allowing for the data to be reduced after collection. The distances measured were from a random point to its nearest plant, and from that plant to its T-square neighbour. In all data was collected from 27 sampling frames, for 7 of these T-square data was collected for both S.squalidus and S.viscosus.

2.5 SEGREGATION DATA.

As well as collecting data relevant to the spatial pattern of plant populations, data was also collected relating to the degree of association between Senecio viscosus and S.squalidus in field populations. (A model of two interacting species was to be the eventual aim with SPEV, modelling populations in which there was hybrid infertility, and equal fitness of the two homozygotes (one of which was a self pollinating species, the other a cross pollinating species)). At the seven sites where T-square data could be collected for both species from the same sampling frame, data about the nearest neighbour of a random plant was collected. In practice, those plants located using the two-tapes technique, and from which the distance to nearest neighbour would be measured, were also used as "base" plants for association data. The species of the base plant, and that of its nearest neighbour were recorded, in this case, the nearest neighbour could be either species, whereas with T-square sampling, the nearest neighbour was the nearest plant, excluding those behind the 'T', of the same species.

2.5.1 Statistical analysis of Segregation field data.

Firstly, a measure of segregation, S, where:

$$S = 1 - \frac{m(b+c)}{(a+b).(b+d) + (c+d).(c+a)}$$

if S=0, the species are not segregated; if S=+1, the two species are completely segregated; and, if S=-1, the two species are associated, that is the nearest neighbour is always the other species. (After Pielou, 1961).

Secondly, Krebs (1978) proposes a chi-squared test, testing the null hypothesis that two species are independently dispersed. This test is designed for use with quadrat data, where the presence of two species in the quadrats has been scored.

Species Two	Species One		
	+	-	Total
+	a	b	a + b
-	c	d	c + d
Total	a + c	b + d	m

$$\chi^2 = \frac{m(ad - bc)}{(a+b).(c+d).(a+c).(b+d)}$$

Figure 3. Two by Two Contingency Table for Segregation Data

As with T-square statistics, it is important that over-sampling should be avoided, this was the reason why the random plants located for T-square sampling were used as base plants in this method.

2.6 QUADRAT METHODS

2.6.1 Random quadrats

Randomly placed quadrats were used, as described in section 2.4, to obtain population estimates for those sampling frames whose populations were too large to enumerate by actual counts. Random quadrats may be used in the detection of spatial pattern, however, if the distribution of individuals in the population is non random, then the amount of information to be obtained from them is limited (Greig-Smith, 1952). Several statistical approaches are available for testing data collected from randomly placed quadrats to determine the intensity of spatial pattern, and these are detailed below. Firstly, the observed quadrat data may be compared with a Poisson distribution of the same mean. By scoring the numbers of quadrats containing 0,1,2,3... individuals, and calculating the mean number of individuals per quadrat, M, the expected numbers of individuals may be derived using the following expansion (Greig-Smith, 1952, Hopkins, 1954):

$$e^{-M}\ ,\ Me^{-M}\ ,\ M^2e^{-M}/2!\ ,\ M^3e^{-M}/3!,\ldots..$$

Observed and calculated values of these terms may be compared, and the significance of any deviation of the observed from expected values measured using a chi-squared test, with n-1 degrees of freedom. Two problems arise with this technique, firstly, that any indication of spatial pattern applies only to one scale (that is to the size of one quadrat alone), and secondly, if the mean M is low, then only the first two or three terms of the series will have an expected value of greater than 5 (which is generally accepted as the lowest expectation permissible for a chi-squared test).

2.6.2 Grid quadrat methods.

Greig-Smith (1952, 1964) proposed that a grid of contiguous quadrats could provide information about pattern in plant populations, at scales other than just the quadrat size. Kershaw (1957) proposed that a transect of quadrats could be used, instead of a grid, requiring less labour. This latter method is very appropriate for the detection of linear pattern, however, the amount of information it will yield about the area of patches in a plant population over the whole study area is limited. When a contiguous grid of quadrats is used, each side having a number of quadrats that is a power of 2, that is 16 x 16, 32 x 64, etc., neighbouring quadrats may be added together to form blocks of 2,4,8,16... grid units, (the even powers of 2 being square blocks, the odd powers of 2 being oblong blocks), in order to detect spatial pattern at more than one scale in the environment. Consequently, in the field, a grid of 32 x 64 contiguous quadrats, each $0.18m^2$ in area, was laid in an area containing large populations of both Senecio viscosus, and S. squalidus. The numbers of both species (and their hybrid, S.xlondenensis), were scored in each quadrat. From this data maps of plant populations were produced. Statistics relating to the amount of spatial pattern in each species, and the amount of association between them, were derived.

2.6.3 Statistical treatment of the contiguous quadrat data.

As was described in the section above, grids of contiguous quadrats are useful in detecting the scale of spatial pattern in plant populations. Greig-Smith (1952) presupposes an area divided by means of a grid of quadrats or plots of identical size and orientation. The whole grid is divided into half and the sum of squares of the halves calculated. Each half is then divided into quarters, and the sum of squares of quarters is calculated. This process of halving blocks within the grid is repeated until the original quadrats are reached. Let $B_m(i)$ be the total of the ith block of size m, where m is the number of plots in the block, and the blocks are arbitrarily ordered. The sum of squares corresponding to the blocks of m, nested within blocks of 2m is:

$$SS_m = \sum_{i=1}^{n/m} B_m(i)^2 / m - \sum_{i=1}^{n/2m} B_{2m}(i)^2/2m$$

for each size, until the original plots are reached, that is m=1. The mean square, MS, of each block is:

$$MS = SS_m / (n/m)$$

where n = total number of quadrats

m = current block size.

If the distribution of plants over the whole area is perfectly random, the mean square of all the block sizes should be the same, that is the variance/mean ratio, should be unity. If the distribution is aggregated then the variance to mean ratio will rise, up to that block size equivalent to the area of the patches. If the patches are themselves random, the variance/mean ratio will maintain these levels, if the patches are regular, variance/mean ratio will fall. Kershaw (1957) demonstrated the reliability of a subjective estimate of the position of peaks in the variance/mean ratio against block size curve to define different scales of pattern, because the statistical proof of different scales of pattern in a grid is complicated because even though the variance/mean ratio should have an F-distribution under the null hypothesis, once the existence of non-randomness has been proved, the F-test is no longer valid. This is not the only problem with this type of analysis. Pattern at one scale, using this method, is not wholly independent of pattern at another smaller one. Zahl (1974) points to a further problem of this method of grid sampling. In Greig-Smith's description of Sums of Squares, blocks are restricted to being non-overlapping. Thus, any cluster that is astride two or more blocks of a given size has less effect on the sum of squares than the same cluster wholly contained within one block.

The lack of statistical verification of the tests discriminating between pattern at different scales has led to a series of papers proposing so-called randomisation techniques.

2.6.4 Mead's Analysis. Randomisation methods for contiguous quadrat data

Mead (1974) proposed a 2-within-4 randomisation test, for use with contiguous quadrat data, which avoids many of the difficulties inherent in the original method proposed by Greig-Smith, and allows independent testing for spatial pattern at different scales. Mead's analysis was originally designed for use with a transect of quadrats as proposed by Kershaw (1957). Besag (1977) proposes a Mead's analysis technique for grids of quadrats, using a "4's-within-16's" randomisation test, and also using a Monte Carlo randomisation (which uses a random sample of the total randomisation of 4s-within-16s, Hope (1968)). This version of Mead's analysis requires that contiguous quadrat data are successively partitioned into 1,4,16...groups, each consisting of 4 x 4 blocks. For each partitioning of the grid, the null hypothesis is that within each group, the observed counts on the four associated 2 x 2 sub-groups form a random sample within the total possible number of randomisations. A test statistic is calculated for the observed data, and 99 (say) values obtained from Monte Carlo randomisations, which are then ranked, the rank of the observed statistic being its exact probability, (e.g. an observed statistic ranked first of 99 randomisations would have a probability of 0.01). The use of different partitions of the grid provides an independent test of spatial pattern at different scales. An analogous procedure may be adopted for the detection of spatial association between two sets of contiguous quadrat counts over a single region (Besag, 1977), i.e. data collected for two species over the same quadrat grid. For each partitioning of the data a statistic measuring the association between corresponding counts within pairs of blocks is calculated, and ranked with Monte Carlo randomisations of the data as before. The mean Spearmans Rank Correlation Coefficient between counts within pairs of blocks, may also be used to give an estimate of the association between two species for which data has been collected over the same grid. The contiguous quadrat data was treated in this way, and the results and usefulness of this technique are fully discussed in chapter 6.

2.6.5 Mead's Analysis Technique.

The method used in Mead's analysis is described, to clarify the usefulness of the technique.

Firstly, testing for spatial pattern in one species only. The data is divided into contiguous groups of 16, and each group of 16 into four sub-groups of 4. Four sub-group total were calculated from the sum of the blocks in that sub-group. Consider the numerical example:

16 Observed Block Scores				Sub group totals		Test statistic
42	14	9	29			
38	21	44	12	115	94	940
62	44	92	29	339	320	
126	107	139	60			

Figure 4. Meads analysis of Observed Quadrat Scores.

The Meads test statistic calculated from these is the sum of the six absolute values of the pairwise differences between sub group totals within blocks. In this case, this is:

$$(|115-94|)+(|115-320|)+(|115-339|)+(|94-320|)+(|94-339|)+(|320-339|) = 940$$

This value is then summed over all groups. The positions of the 16 observed data blocks are then randomized (see below), and a new value for the test statistic is calculated.

16 Randomised Block Scores				Sub block totals		Test statistic
42	29	38	29			
107	92	44	12	270	123	556
9	44	21	139	193	282	
126	14	62	60			

Figure 5. Meads analysis of Randomised Quadrat Scores.

If the rank of the observed test statistic is less than fifth, (out of the 100 randomised+observed statistics), then the null hypothesis is rejected. Sub-group totals are then treated as individuals and the analysis is repeated, at a scale four times that preceding it, until the group size equals the grid size.

Secondly, testing for scales of interaction between plants in a grid of contiguous quadrats. Given the counts for each species over a 2 x 2 squared grid, the data are viewed successively as 1,4,16,....etc. blocks, each comprising four (aggregate) counts in a 2 x 2 arrangement (see Besag, 1977). Using Monte Carlo methods, we may apply the null hypothesis of no association of counts within corresponding pairs of blocks. In this case the test statistic is the summed product of the observed counts at each position within each 2 x 2 block. Once again using a numerical example:

Species One				Species Two			
42	14	9	29	37	19	4	13
38	21	44	12	49	40	11	4
62	44	92	29	54	72	14	10
126	107	139	60	100	58	35	8
174	126	162	120	78	47	23	4
44	123	78	148	26	60	17	9
36	50	50	91	60	30	12	12
30	9	35	100	60	4	24	14

Figure 6. Testing for association between two species (1).

Using Observed Quadrat Data.

For the first 2 x 2 block the Meads test statistic is evaluated as:

$$(42 \times 37) + (14 \times 19) + (38 \times 49) + (21 \times 40) = 4522$$

summed over all eight 2 x 2 blocks, the observed test statistic may be calculated as = 82022.

If the observed values for one species only are then randomised within 2 x 2 blocks, a new test statistic may be calculated as = 73011.

Species One

21	42	44	29
38	14	9	12
126	107	92	60
62	44	139	29
126	123	78	162
174	44	120	148
36	9	50	91
50	30	35	100

(Randomised)

Species Two

37	19	4	13
49	40	11	4
54	72	14	10
100	58	35	8
78	47	23	4
26	60	17	9
60	30	12	12
60	4	24	14

(As before)

Figure 7. Testing for association between two species (2) Using Randomised Quadrat Data.

Again the test statistics are ranked.

Thirdly, using Spearmans rank correlation coefficient to test for association between two species at various block sizes. The mean rank correlation coefficient between counts within corresponding pairs of blocks within fours is calculated over the whole block, tied ranks being resolved by Monte Carlo randomisation methods, (in the example below tied ranks are halved).

Species One

42(1)	14(4)	9(4)	29(2)
38(2)	21(3)	44(1)	12(3)
62(3)	44(4)	92(2)	29(4)
126(1)	107(2)	139(1)	60(3)
174(1)	126(2)	162(1)	120(3)
44(4)	123(3)	78(4)	148(2)
36(2)	50(1)	50(3)	91(2)
30(3)	9(4)	35(4)	100(1)

Figure 8. Ranking Observed data within 2 x 2 blocks (1).

Species Two

37(3)	19(4)	4(3.5)	13(1)
49(1)	40(2)	11(2)	4(3.5)
54(4)	72(2)	14(2)	10(3)
100(1)	58(3)	35(1)	8(4)
78(1)	47(3)	23(1)	4(4)
26(4)	60(2)	17(2)	9(3)
60(1.5)	30(3)	12(3.5)	12(3.5)
60(1.5)	4(4)	24(1)	14(2)

Figure 9. Ranking Observed data within 2 x 2 blocks (2).

Ranks for corresponding fours are multiplied together, and then summed together within each four:

for block 1, $R=(1 \times 3)+(4 \times 4)+(2 \times 1)+(3 \times 2) = 27$

The expected value of R for any one block is

$$R\ (exp) = (1+2+3+4) \times (1+2+3+4)/4 = 25$$

The maximum value of R for any one block, i.e. one in which the two species are perfectly correlated is:

$$R\ (max) = (1 \times 1)+(2 \times 2)+(3 \times 3)+(4 \times 4) = 30$$

The minimum value for R in any block, i.e., the two species are disassociated is:

$$R\ (min) = (1 \times 4)+(2 \times 3)+(3 \times 2)+(4 \times 1) = 20$$

In order to obtain a value of Spearmans rank correlation coefficient between -1 and +1,

$$rho = \frac{(R - R(exp))}{((R(max) - R(min)) / 2)}$$

which in this case is evaluated as:

$$\begin{aligned} rho &= (27 - 25)/((30 - 20) / 2) \\ &= +0.4 \end{aligned}$$

The mean value of rho is calculated by summing over all the blocks of fours, and dividing by their number, N. A value of the standard deviate is calculated, and its value looked up in tables of the normal distribution, to test the significance of the rho.

$$D = \text{rho} \times \sqrt{(3 \times N)}$$

where N = The number of blocks of fours.

In the example given, the mean Spearmans rank correlation coefficient, rho, is +0.4375, with a standard deviate of 2.1433, (N = 8), which is significant at P 0.017. Meads analysis is discussed thoroughly in Mead, (1974), Besag, (1977), and Besag and Diggle, (1977).

3 MODELLING THE EVOLUTION OF SPATIAL PATTERN

3.1. SPEV.

SPEV is a complicated suite of procedures, written by Dr. J.T. Gleaves in AlgolW, to model the evolution of spatial pattern in plant populations. The original version of the program had been written to model the evolution of spatial pattern in communities of lead tolerant plants, where there was strong selection for lead tolerance within the community, and strong gene flow of non-tolerant genes from populations outside the area of high soil lead concentrations. SPEV not only modelled the impact of selection and gene flow on spatial pattern, but also had parameters for seed dispersal and competition. Each of these processes will be discussed directly.

3.1.1 Summary of the program.

SPEV simulates the development of spatial pattern in a plant population, (within a mapped area), based solely on genetic effects.

Starting from a randomly dispersed population of plants, usually with a gene frequency of 0.5, the pollination of each plant is simulated.
The plants in the model are considered to be wind pollinated (although it is thought that the model would hold for insect pollination as well), the amount of pollen received by each plant from every other being a reciprocal power function of the distance between them. As well as pollen received from plants within the population, a parameter, RAIN, is added to the total pollen of each plant representing gene flow from recessive homozygous plants outside the mapped population. For each plant the amounts of 'A' and total (+RAIN)

pollen were accumulated, the gene frequency in the pollen being later used to determine the genotypes of the offspring of that parent plant.
The distances between each plant and every other are also used to determine the intensity of clustering in the population and also the number of competitors for each plant. If the distance between two plants is less than the expected neighbour distance (assuming a random distribution), then a variable counting the number of neighbours in the population is incremented by 1. This is used to calculate the mean number of neighbours per individual (MNNHBR), a statistic reflecting the intensity of clustering in the population. (Large values of MNNHBR indicate a highly aggregated population). Also, if the distance between two plants is less than the competition distance, DCOMP (which is initialised at the start of the program so that the intensity of local competition can be varied), then the number of competitors of both plants is incremented by 1. The number of competitors of a given plant is used to determine the fecundity of that plant.

From the gene frequencies in the pollen of each plant, the frequencies of the offspring genotypes of that plant are calculated. The fitnesses of the offspring genotypes are 1,1-D*S and 1-S respectively for 'AA', 'Aa' and 'aa', where D = Dominance of 'A' over 'a', and S = Selection coefficient, both of which are initialised at the start of the program by the user. Selection acts in the program on differential offspring fitness, not on the parent plants themselves. Offspring survivorships for each parent plant are calculated by multiplying the frequency of a given offspring genotype (for a given parent plant) by the fitness of that genotype. The fecundity (that is the number of offspring that a given parent plant produces) of a plant is regulated by the number of plants with which it is competing, such that, as the number of competitors increases, so fecundity decreases. The fecundity of each plant was modified by a density dependent factor which maintained the population size at approximately that of the initial population.

Using the fecundity of each parent plant and the survivorships of its offspring, appropriate offspring are generated and dispersed from each parent plant. The direction that an offspring plant is dispersed from its parent is completely random, and the dispersal distance from its parent has a normal distribution, centred on that parent plant, the standard deviation of which is the seed dispersal distance, SIGMA, (which is initialised by the user at the start of the program). The offspring plants then become the parent plants of the next generation from which statistics relating to the spatial pattern of the modelled population are obtained, and the cycle is restarted.

3.2 SIMULATING PLANT POPULATIONS USING SPEV.

Having briefly described the processes by which the program generates and maintains a model population, the remainder of this chapter will be devoted to describing the results of simulations, using the initial version of SPEV. The results will, where possible, be augmented with the use of graphs. MNNHBR was used in every case as a statistic of clustering. A value of expected number of neighbours = 0.7853, was derived assuming a random distribution of plants, and is drawn for comparison on each graph where MNNHBR is plotted.

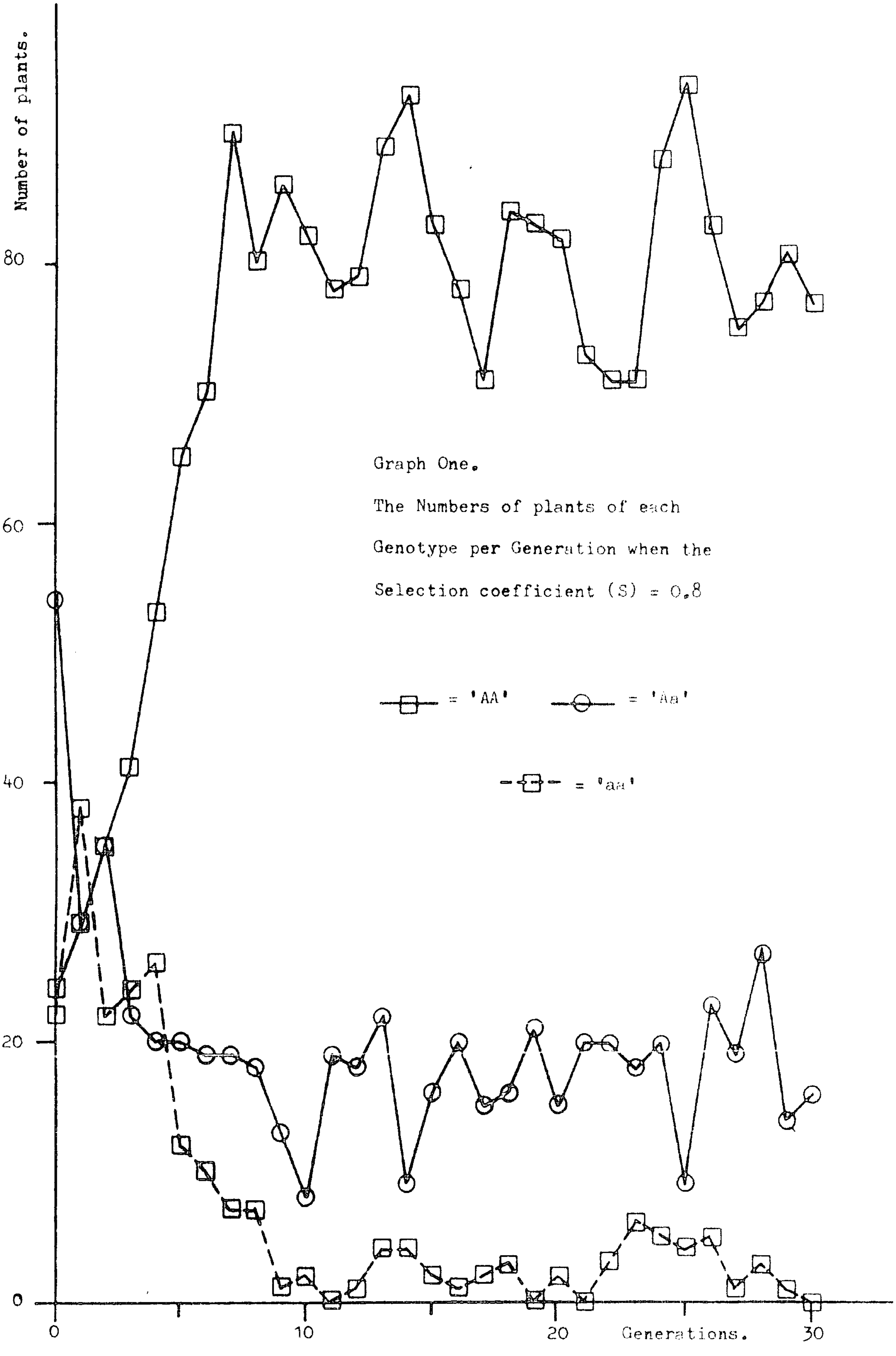
3.2.1 Selection (S).

When the value of selection coefficient was high, say S=0.8, the 'a' genes were rapidly eliminated from the gene pool (see Graph 1.). However, the speed at which 'a' was eliminated from the gene pool was dependent on two factors:

Dominance (D). If D had a low value, say D=0.3, then the fitness of the heterozygote would be nearly equivalent to that of the 'AA' homozygote. As a result of this, 'a' genes would remain in the gene pool for a considerable time, in the heterozygote. Graph 2. plots results for D=1.0, 0.5, 0.3 at S=0.8. The number of heterozygotes persistently increases as D decreases, and the recessive genes in the gene pool are thus protected in the population, as would be expected.

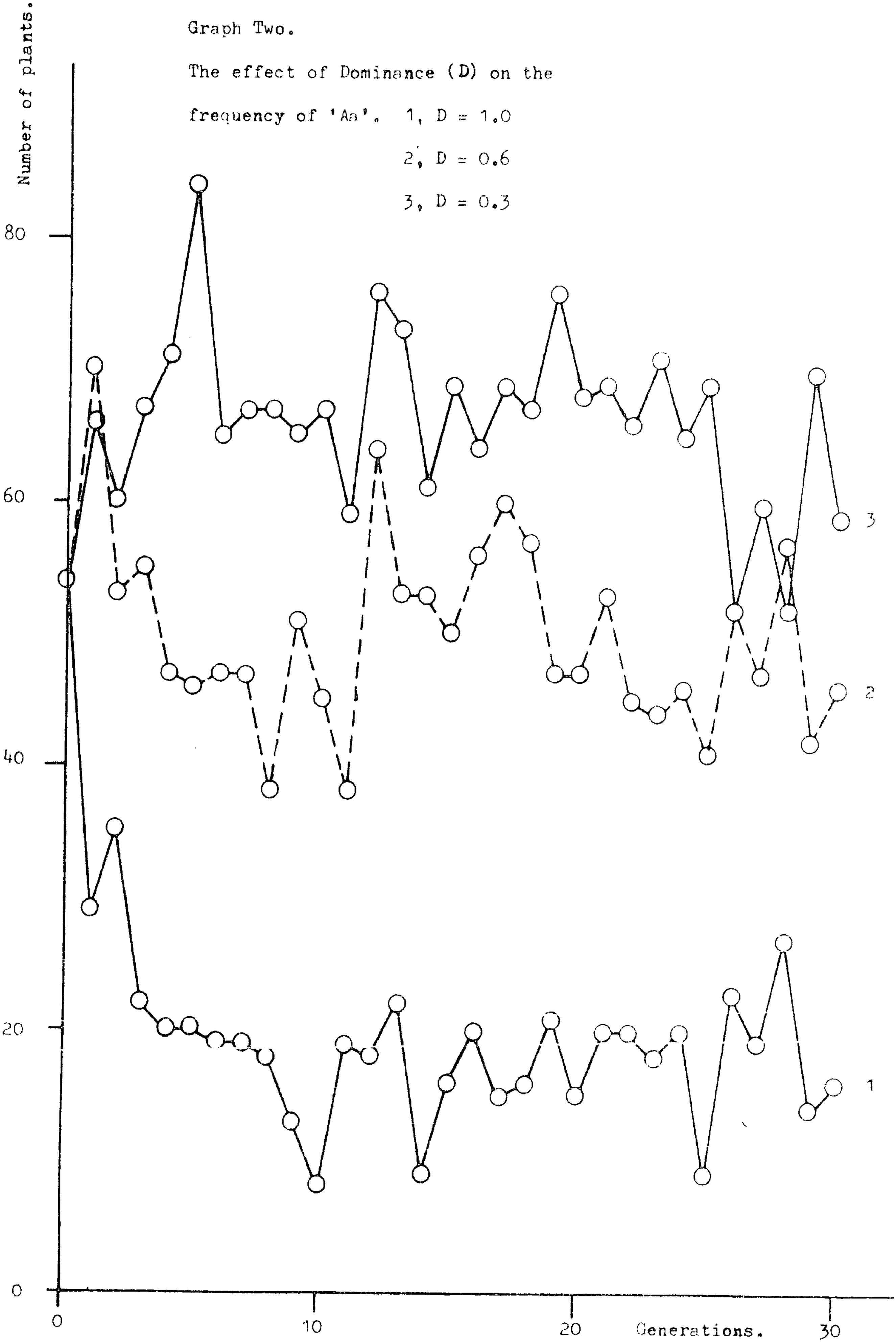
Pollen rain (RAIN). If RAIN was intense, then the flow of 'a' genes from outside the gene pool would be sufficient to maintain 'a' genes in the population. The effect of RAIN on the spatial pattern of populations was profound, since those individuals which were sufficiently isolated to receive most of their pollen from the background rain, would, if selection were intense, fail to produce viable offspring. This point is thoroughly discussed in section 3.2.3. Graphs 3(a) and 3(b) plot the frequencies of heterozygotes and recessive homozygotes, when RAIN = 10.0 and 5.0. In general, the higher the levels of gene flow from outside the population, the longer heterozygotes and recessive homozygotes will persist in the population.

As Selection decreases, so the gene frequencies in the population become stable, at S=0.5 the effects of RAIN and selection hold one another in



Graph One.

The Numbers of plants of each Genotype per Generation when the Selection coefficient (S) = 0.8

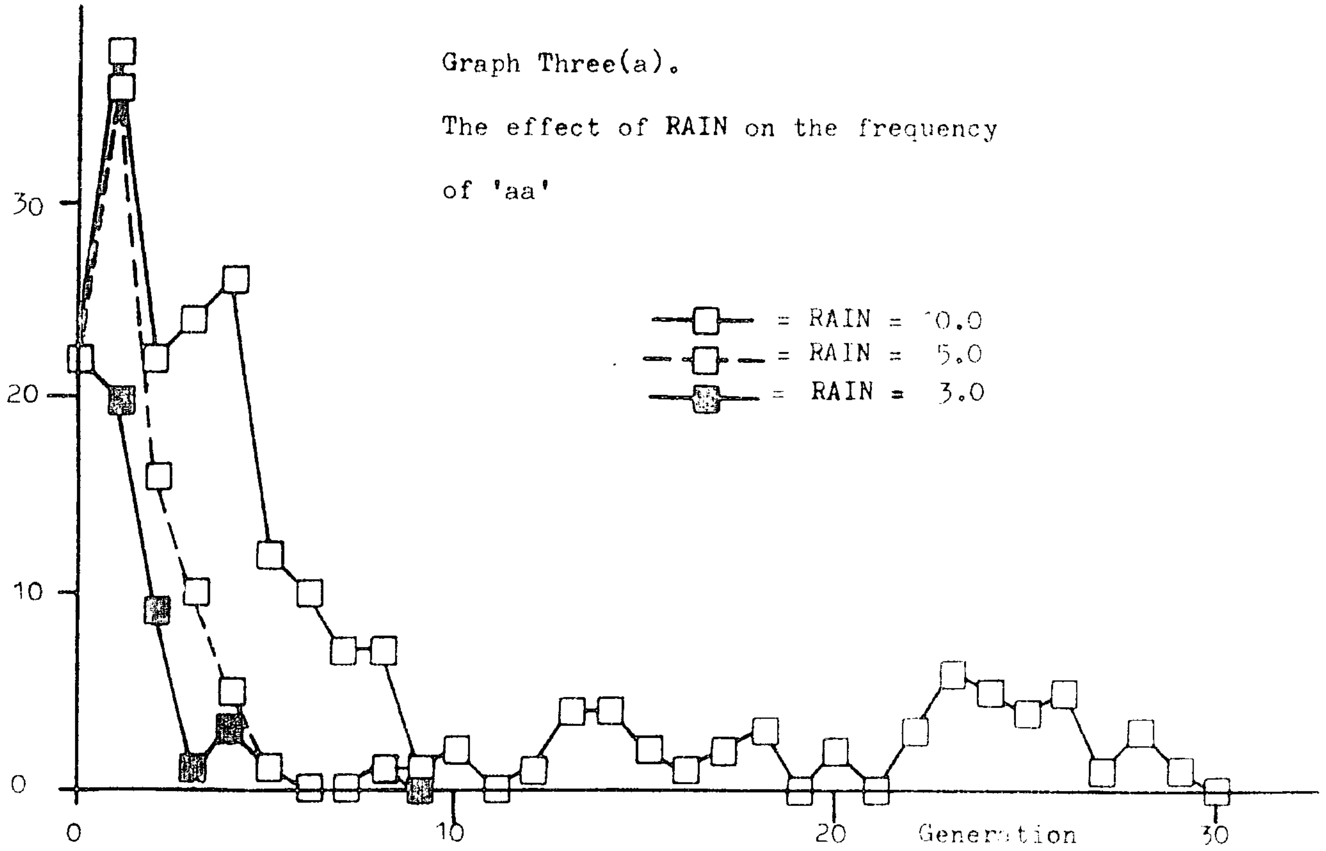


Graph Two.

The effect of Dominance (D) on the frequency of 'Aa'. 1, D = 1.0

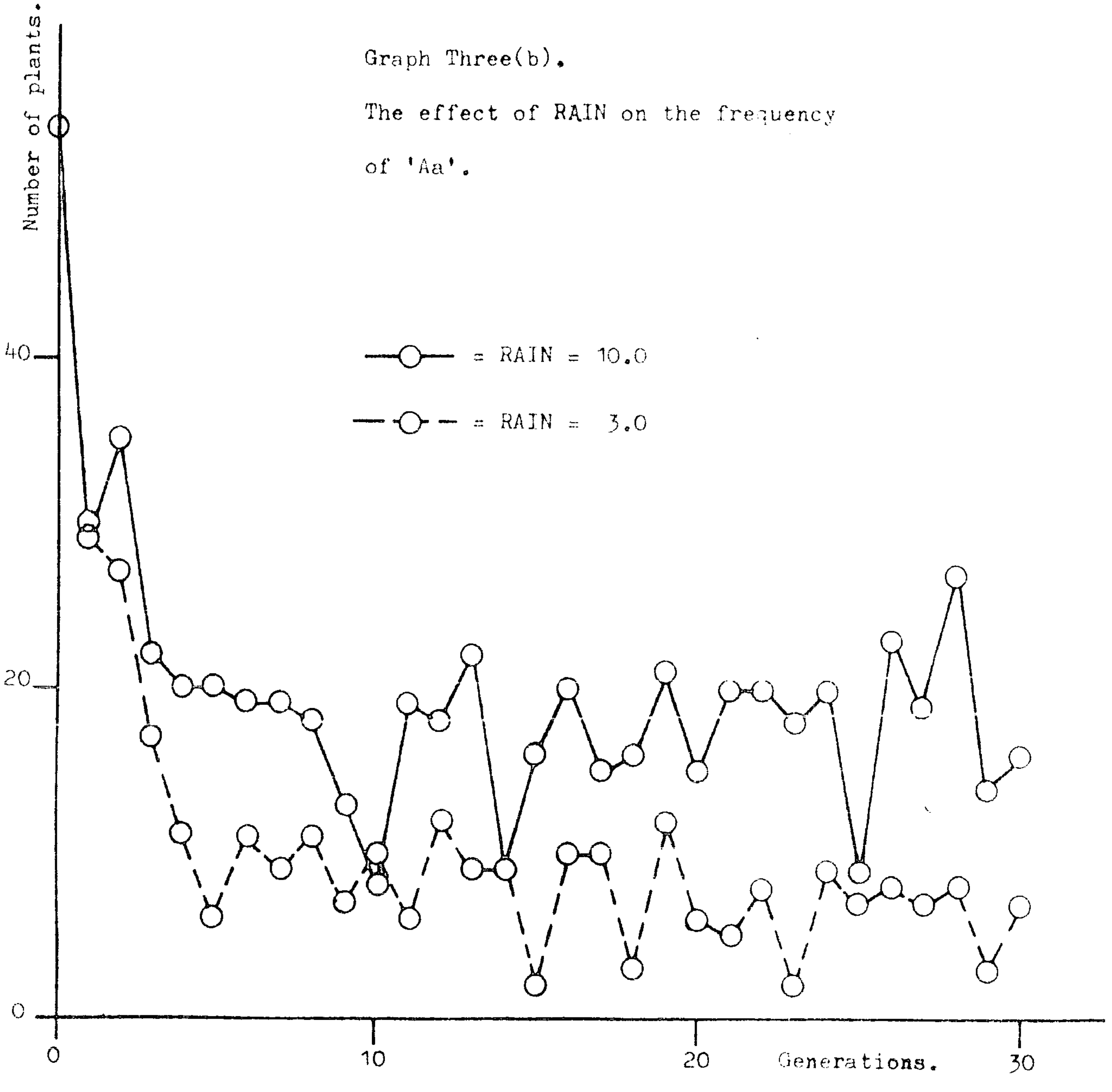
2, D = 0.6

3, D = 0.3



Graph Three(a).

The effect of RAIN on the frequency of 'aa'



Graph Three(b).

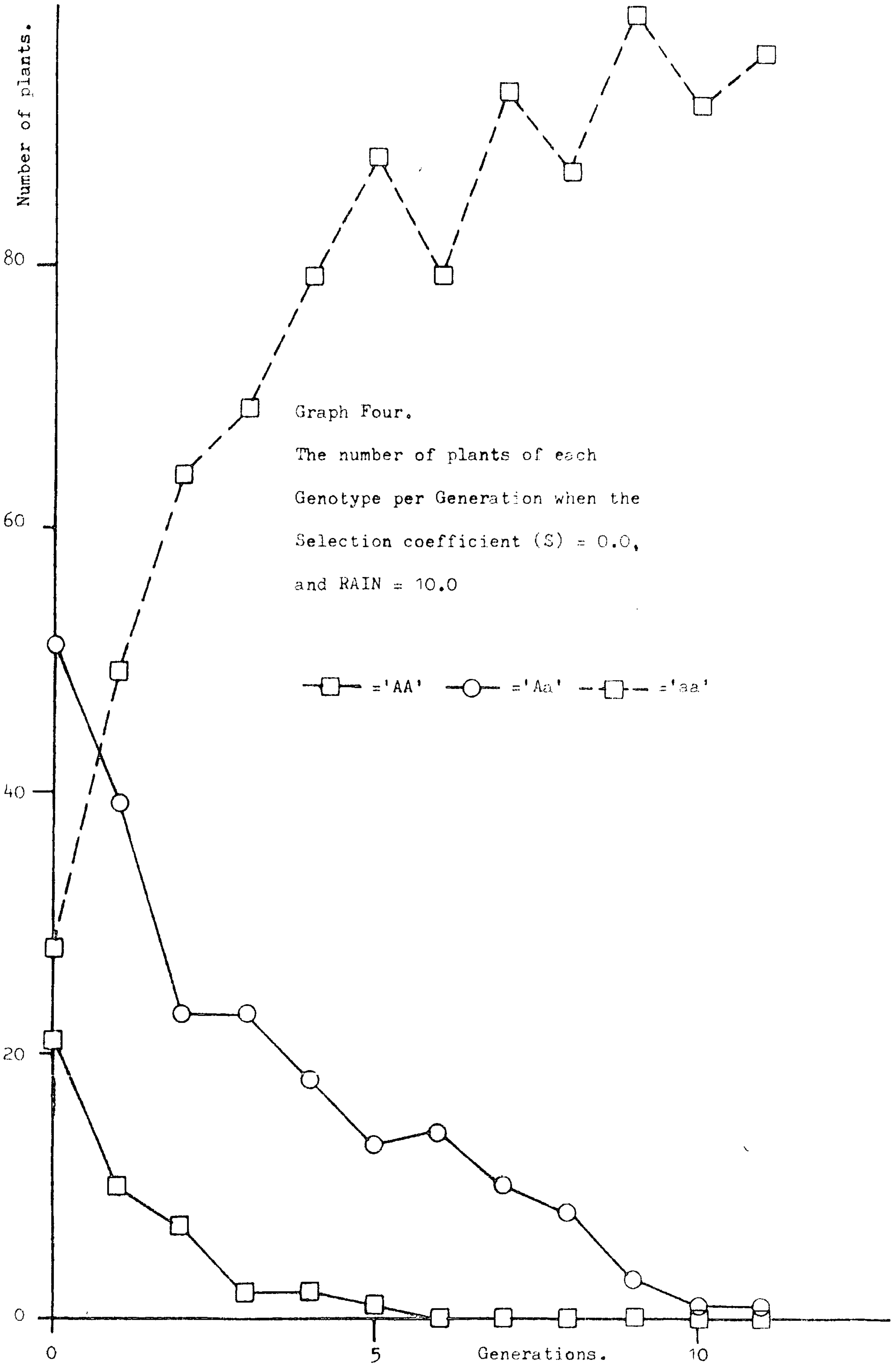
The effect of RAIN on the frequency of 'Aa'.

balance, eventually becoming unstable again as S=0.0, see Graph 4. The elimination of 'A' from the gene pool, is a result of the effect of RAIN effectively increasing the total pollen available to a given plant, whilst thus reducing the relative availability of 'A' pollen. This fits well with the situation which the model was originally designed to study, in that in an environment where lead tolerance was of no advantage, migration would lead to the loss of lead tolerance genes, in much the same way as selection for them had led to the elimination of 'a' genes in highly polluted soils. A very stable population, whose gene frequency, P, remained near its original value, was obtained by setting S=RAIN=0, (results of this simulation are plotted on Graph 5.). The effect of RAIN may be seen by comparing Graphs 4 and 5, in the former, the population goes to fixation, whilst in the latter, stable gene frequencies are maintained.

The impact of selection on spatial pattern is indirect, having its effect via gene flow, this is thoroughly explained in section 3.2.3.

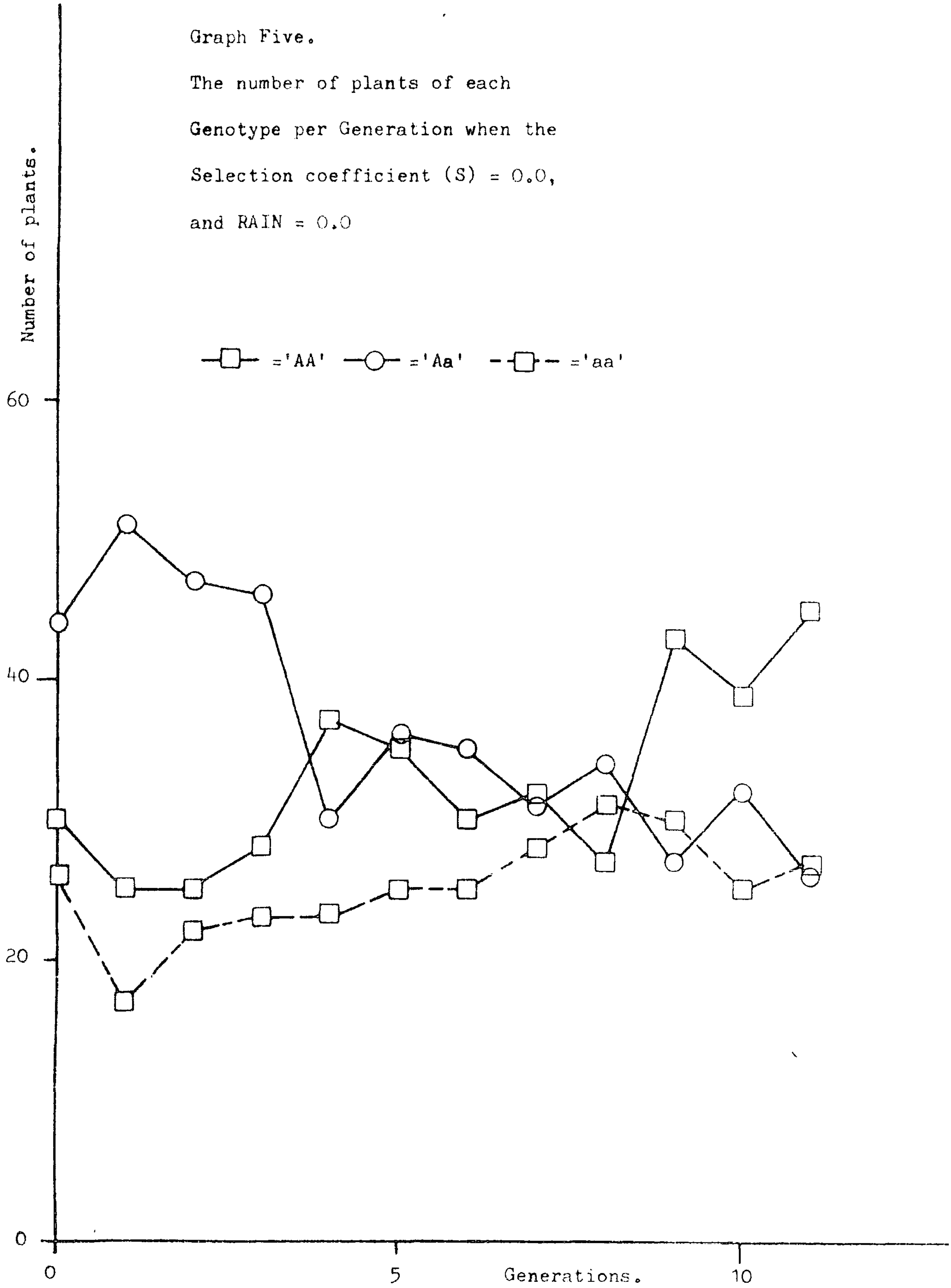
3.2.2 Seed dispersal (SIGMA).

Seed dispersal has a direct effect on the spatial pattern of a given population, since a low root mean square seed dispersal distance, SIGMA, will result in the offspring of a given plant always being located close to it. If SIGMA is low, this can lead to the formation of point clusters, that is many small, highly intense clusters, each containing only a few individuals, (see Figure 10). If SIGMA is high, this will lead to more random distribution in the population, depending on the values of S and DCOMP. The results of simulations of SIGMA= 0.05., 0.25, and 0.5 are given on Graph 6.



Graph Four.

The number of plants of each Genotype per Generation when the Selection coefficient (S) = 0.0, and RAIN = 10.0



Graph Five.

The number of plants of each Genotype per Generation when the Selection coefficient (S) = 0.0, and RAIN = 0.0

Key to the symbols used in Figures 10, 14, 15, 16, 18 and 19.

+ = 'aa' individual

= 'Aa' individual

✳ = 'AA' individual

2, 3, 4... = more than one individual at the same location.

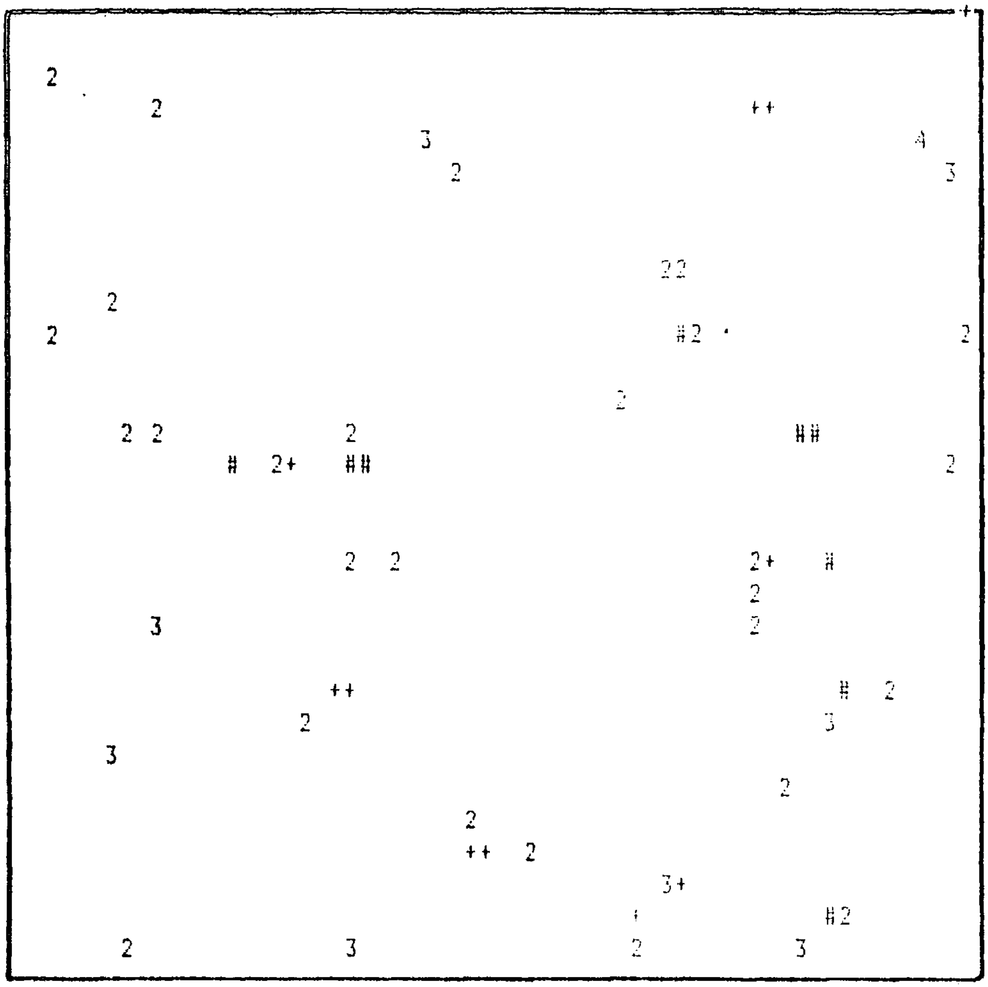
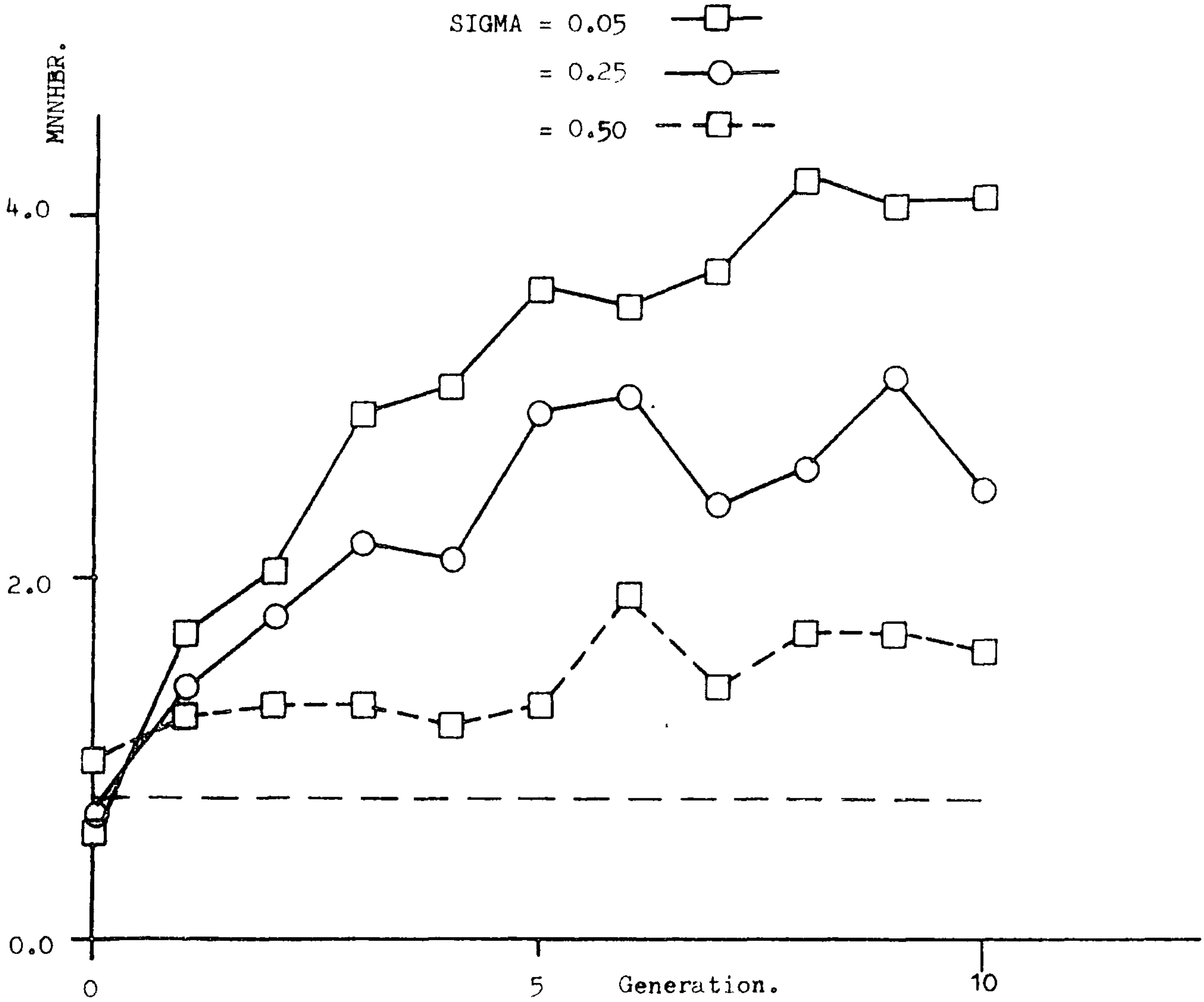


Figure 10. Point clusters of model plants, simulated when r.m.s. seed dispersal distance (SIGMA) is low. (N = 104, MNNHB = 3.1150)

Graph Six.

The effect of r.m.s Seed Dispersal Distance on the intensity of clustering (MNNHB) in a model population.



3.2.3 Pollen dispersal.

Since the amount of pollen received by each plant, from every other, is a distance related phenomena (see section 3.1), (the exception to this rule being the case where K=0[1]), plants will tend to receive more pollen from those plants nearest to themselves. Therefore the genotype of the offspring of any plant will reflect firstly the genotype of its parent, and secondly that of those plants that are nearest to it. Any isolated individual in the population is certain to find a large proportion of its total pollen being composed of background RAIN, the pollen of other plants in the population playing a lesser part. It is at this stage that selection plays its part in determining spatial pattern.

If the selection coefficient, S, is low, say 0, then the offspring of each genotype have an equal chance of survival, each having the same fitness. Two situations may now arise:

> If RAIN was zero also, then the population would tend to maintain a stable gene pool, as in Graph 4.
>
> If RAIN is large, then, as has been previously explained, 'A' will be lost from the population, as in Graph 5.

---

[1] This models a random mating population, since the amount of pollen received by any plant will be a constant

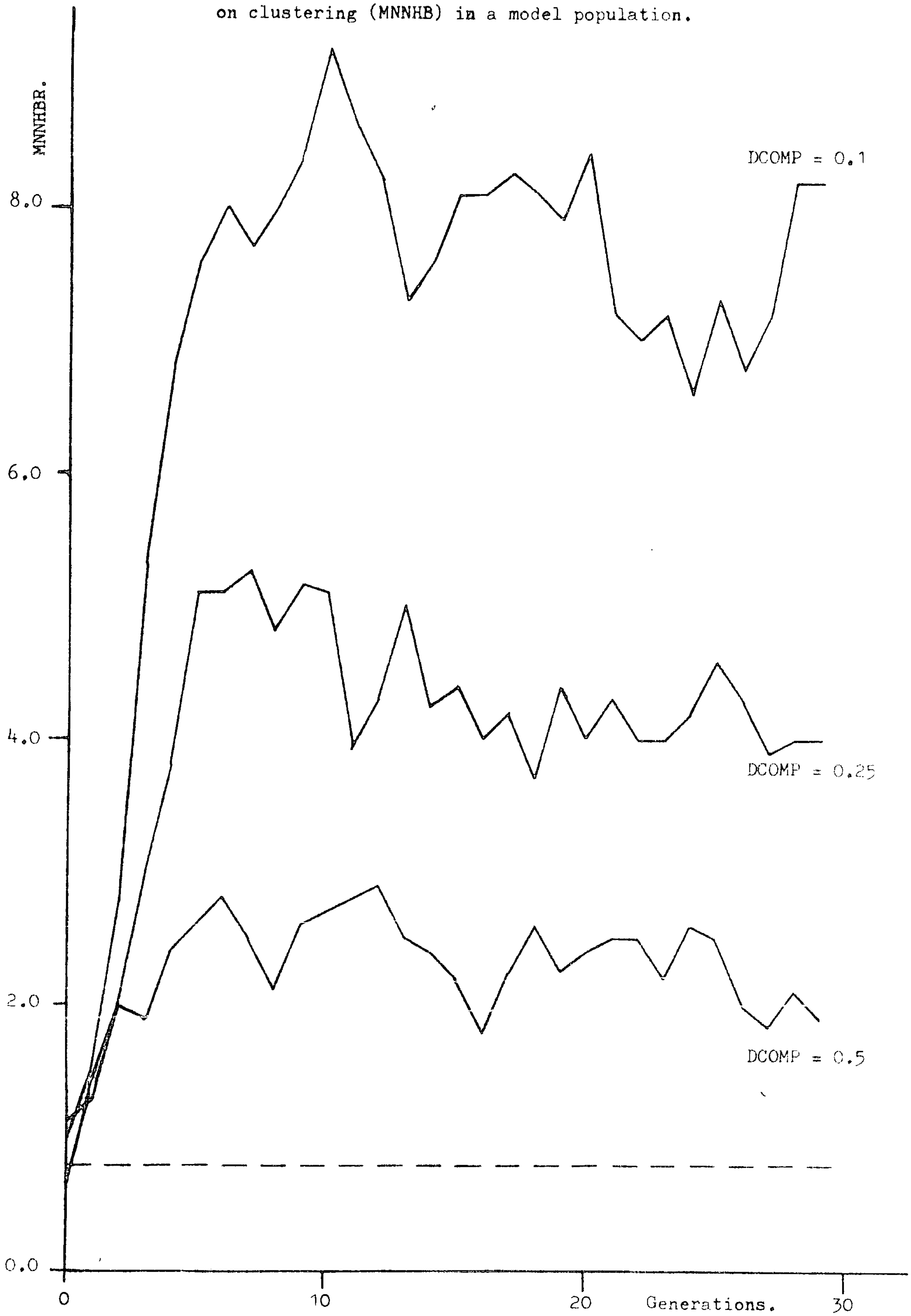
If the selection coefficient, S, is high, (say 0.8), then the survivorship of 'Aa' and 'aa' offspring would be greatly reduced. The importance of this relates not only to heterozygous and recessive homozygous plants in the population, but also to isolated dominant homozygous individuals which receive more of their pollen from RAIN, 'Aa' and 'aa' individuals than they do from other 'AA' individuals. In their case the chances of their offspring surviving are reduced, resulting in the loss of those isolated homozygous plants from subsequent generations. In order that 'AA' plants should survive in the model environment, spatial patterns should develop in such a way as to maximise the quantity of 'A' pollen received by 'AA' plants, and yet not reduce their fecundity (by being so close together as to compete with other plants). The result of these processes is that, (according to the limits defined by the competition distance, DCOMP), where selection is intense, there will be a strong tendency towards aggregated spatial patterns in the model populations, brought about by the elimination of 'Aa' and 'aa' plants and also those 'AA' individuals isolated from other 'AA' plants, because of the low survivorship of their offspring.

3.2.4 Competition (DCOMP).

DCOMP represents the degree of local competition in the population. If DCOMP is large local competition is increased, and if DCOMP is small local competition is decreased, (allowing for more intense clustering in the population). The results of a series of simulations of DCOMP=0.1, 0.25, 0.5 are given on Graph 7. The effect of increasing DCOMP is to increase the number of competitors that a given plant may have (for the

Graph Seven.

The effect of Competition distance (DCOMP) on clustering (MNNHB) in a model population.



same population size). When DCOMP is high plants that aggregate into clusters would tend to be lost since, although this would secure an abundance of pollen of the same genotype, the number of offspring that each plant in the cluster would produce would be reduced. Conversely, if DCOMP is low, then plants may aggregate, thus gaining a pollen advantage by being closer together, without any loss of fecundity. One problem with representing competition in this way is that, when local competition (DCOMP) was low, plants would tend to aggregate into one very large cluster, a model not wholly representative of natural populations. For this reason the model was developed to include a tesselation, which represented competition in terms of the resources of each plant. (see section 4.1).

3.3 SUMMARY.

Each of the processes outlined in the preceding sections act independently of one another, yet the populations that arise as a result of the simulations are a complicated mixture of the effects of each. It is hoped that this chapter has clarified the original form and operation of the model, without having clouded its methods with a surfeit of detail.

4 DEVELOPMENT OF SPEV: THE DIRICHLET TESSELATION.

4.1 A NEW VERSION OF SPEV.

Although DCOMP had been a useful and quick tool for describing plant to plant interactions in SPEV, it was rather clumsy, and a more sensitive analysis of these interactions was required. An alternative method of assessing plant fecundity is to allocate fecundity according to the resources of a model plant. Each plant was allocated a tile, that is that portion of the mapped area nearer to that point than to any other (see section 4.1.1), and the area of this tile was taken as a measure of its fecundity, relative to the sum of the areas of all the tiles of the plants in the population. These tiles were produced using a very efficient algorithm for producing a Dirichlet tesselation written by Green and Sibson (1977).

4.1.1 The Dirichlet Tesselation.

The Dirichlet Tesselation is defined most simply as

> "a subdivision of the plane determined by a finite set of points, each point has associated with it that region of the plane that is nearer to that point than to any other. Each tile, T_N, is the intersection of the half open planes bounded by the perpendicular bisectors of the lines joining P_N, with each other P_M."

(see Figure 11). Green and Sibson (op cit) have produced a suite of programs in ANSI FORTRAN, called TILE, which will produce for an array of m points a Dirichlet tesselation of T_m tiles, each corresponding to point P_m. Using this tesselation, various statistics about each tile can be derived in SPEV, the area of each tile was to be the crucial statistic.

4.1.2 Disadvantages of the Dirichlet Tesselation.

The major assumptions of the Dirichlet tesselation, particularly in the way in which it was to be implemented in SPEV, are that (1) each individual has the same ability to use the resource that is describing its fecundity, and (2) that this resource, in this case tile area, is the sole determinant of the fecundity of a given individual. Selection is operating on offspring survivorship, and not on parent plants, once these are established, therefore in the model established parent individuals are equally fecund. The latter point, that the area of the tile should be the sole determinant of the 'strength' of an individual is discussed by Cormack (1979), who stresses that this is biologically misleading, since a number of other factors, such as relative growth rate, or germination time might have an equally important effect on the success of an individual. However, for the purposes of a spatial pattern model, where the growth rates of individuals are not considered, and all germination events are assumed as happening instantaneously, the use of the Dirichlet tesselation would represent a considerable increase in the flexibility of the model. Cormack (op cit) discusses alternative forms of tesselation, but disregards each on the premise that none is as efficient or reliable as the Dirichlet tesselation proposed. (One final assumption is that the borders of a tile are clear cut, i.e. there is no interaction between plants in neighbouring tiles once they have encountered

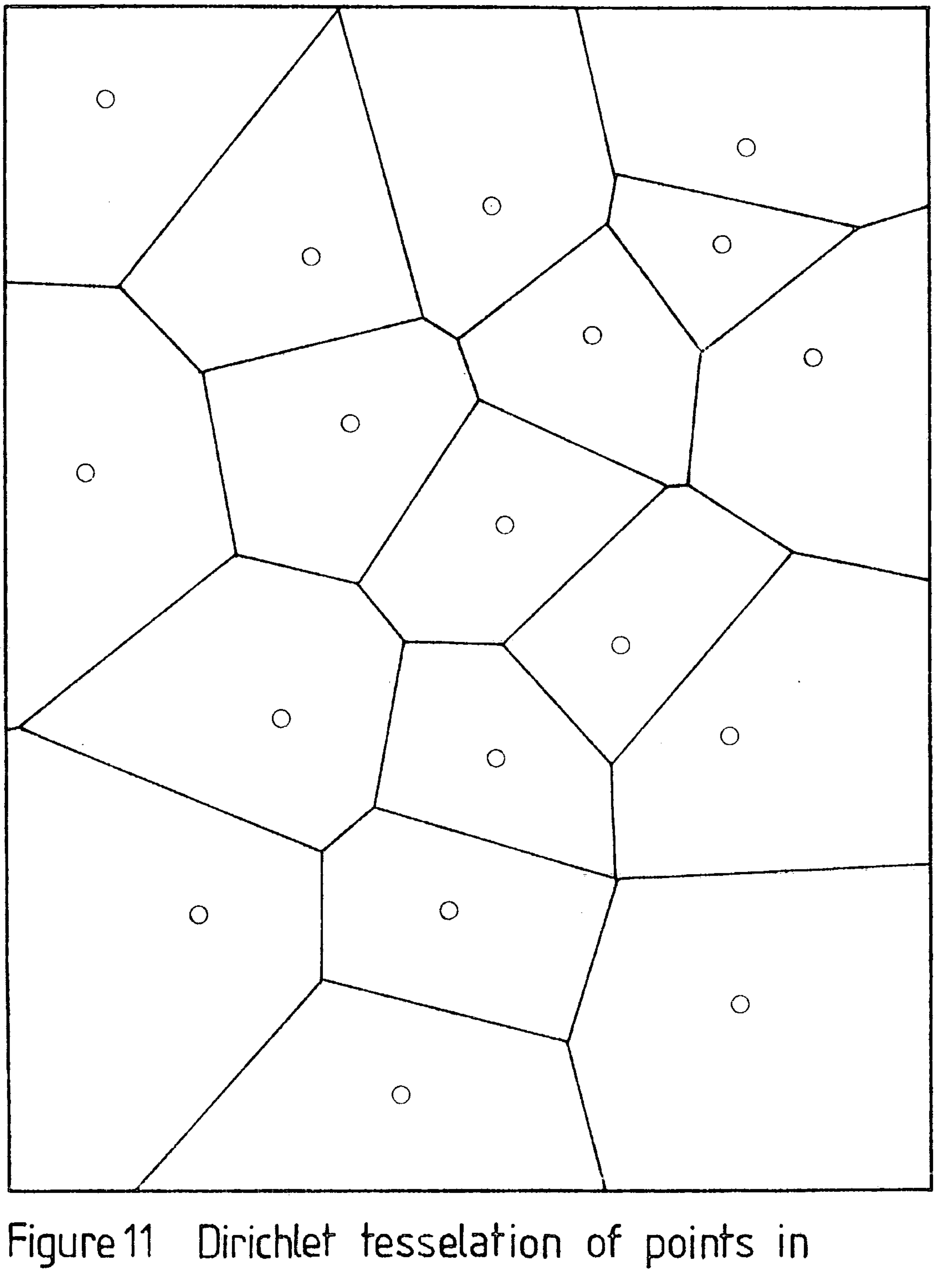


Figure 11 Dirichlet tesselation of points in a plane.

one another at their borders. This, of course, is incorrect biologically, since, considering for example the root systems of neighbouring plants, these are intertwined to a very large extent. However, this objection, although noted, was not considered in the model).

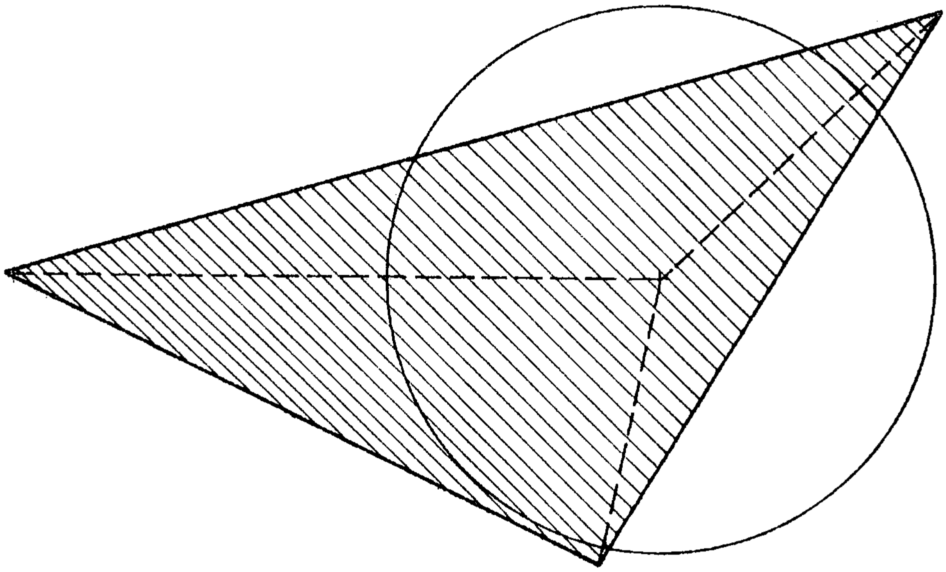
4.1.3 Implementing the Dirichlet Tesselation in SPEV.

The fecundity of parent plants that had originally been determined by the number of plants with which it competed, was now evaluated according to the area of the tile allocated to that plant. Initially the total area of the tile was used as a plant resource, however (as is explained in the next section) this was replaced with a more subtle tesselation in which only part of the tile is used. The area of each tile, and the tesselation itself, were computed by using a set of subroutines called TILE written in ANSI FORTRAN by Dr. P.J. Green.

4.1.4 Modification of the Dirichlet Tesselation.

In the initial implementation of the Dirichlet tesselation in SPEV the entire area of the tile was used as the resource available to that plant. However, since the area of a tile may be very irregular, or specifically 'needle' shaped, it is conceivable that, were this the area available to a real plant, the plant would be unable to utilise a considerable part of the tile. To explain this, if a plant's maximum growth radius is R, then it is possible that the distance from the centre of the plant to the border of the tile in one direction may be less than R, whilst in another it may be considerably greater than R (see Figure 12). Thus a plant may be unable to

Figure 12 Original tile area.



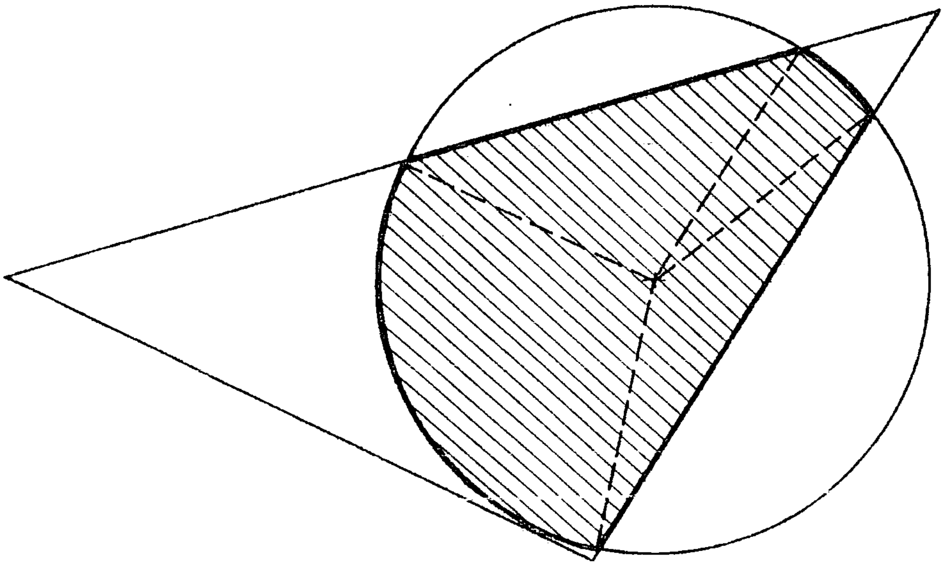


Figure 13 Modified tile area.

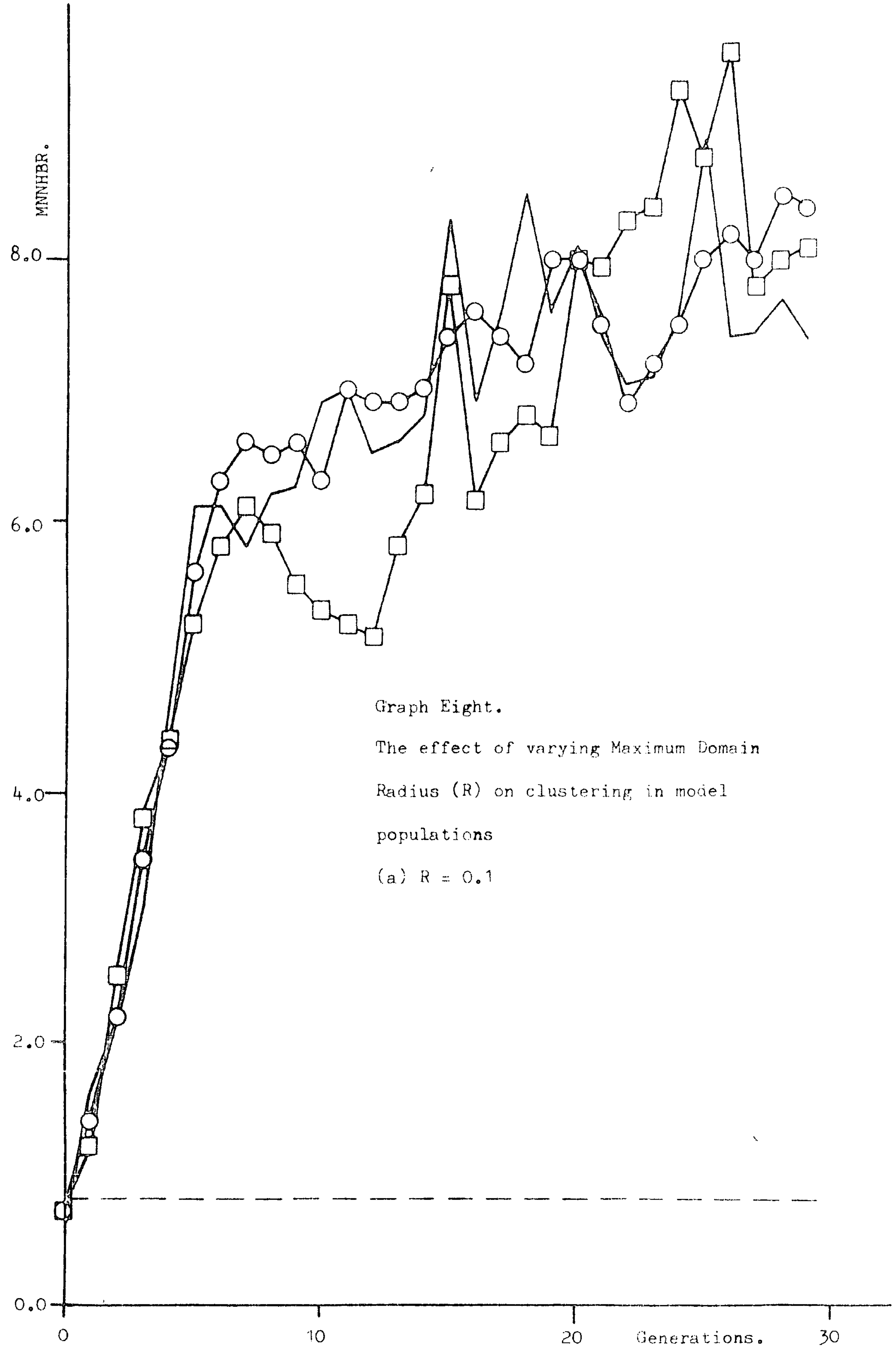
utilise its resources in one direction, whilst being restricted by the presence of another plant in another. As a result of this problem the subroutine that calculates the area of the tile was altered, such that, only that area which lay within both the maximum domain radius and the tile, was treated as the resources available to the plant (see Figure 13).

This modification is again based on assumptions that need to be pointed out. Firstly, it is assumed that a maximum growth radius can be defined for a given plant, in a given set of environmental conditions, spanning a set time period. Secondly, it is assumed that plants do not direct their growing points into those areas with available, unexploited resources, whether they be light, soil, space, or any other. This is a difficult assumption, since it is well demonstrated that plants may direct the growth of their organs.

This new tesselation, based on the old, was considered to be a further improvement on the flexibility and realism of the model, since it allowed for variable maximum domain radii to be used, allowing for (1) variation in the degree of local competition and, (2) partial or complete use of the tesselation.

4.2 RESULTS OF SIMULATIONS USING THE DIRICHLET TESSELATION.

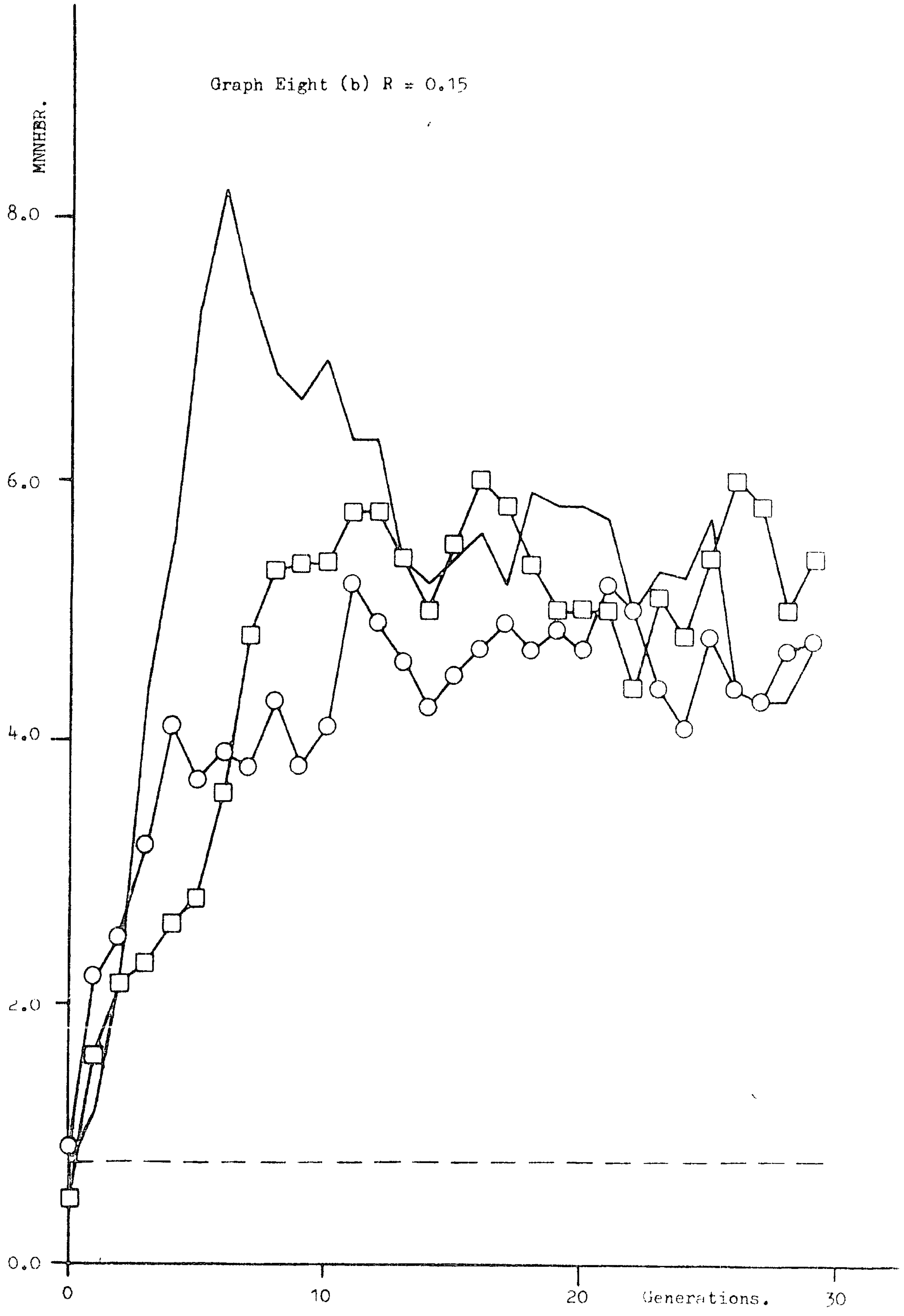
SFEV was now re-run many times using a range of values of maximum domain radius, R. These results are plotted out on graphs 8a to 8h. The mean number of neighbours is plotted out for three runs of thirty generations, for each R value. Simulations were done at R = 0.1, 0.15, 0.2, 0.25, 0.375, 0.5, 1.0, 10.0. On each of the graphs the selection coefficient, S, was fixed at 0.8, in general as the maximum domain radius, R, increased so the intensity of clustering in the population decreased. It is interesting to note that, once the maximum domain radius, R, increased to about the same value as the seed dispersal distance, SIGMA, (usually given the value of 0.25), there was no further decrease in clustering in the population (Graphs 8d to 8h), although the mean levels of clustering observed were above the expected mean number of neighbours. Maps of model populations after 30 generations when R = 0.1, 1.0 and 10.0 are given on Figures 14, 15 and 16 respectively. The domain radius, R, is related to the degree of local competition. When R is larger than the average tile size (in general when R is greater than 1.0), then the intensity of local competition is increased, since plants are interacting at the borders of their respective tiles. Thus the formations of clustered spatial patterns is inhibited because, any reduction in tile area for a given plant, will result in a loss of fecundity for that plant. This situation can be averted by the formation of many small clusters containing only a few plants (see Figure 16 and 17). Small clusters may enable plants to maintain their fecundity (when R is large), and receive most of their pollen from plants of the same genotype in the cluster. This is discussed rather more thoroughly in chapter Five. When R is small, local competition is diminished, and the tesselation is ignored (because in general R is less than the average tile size). In this case, plants will tend to become aggregated into a single large cluster, as had happened when DCOMP was small in the original version of SPEV.



Graph Eight.

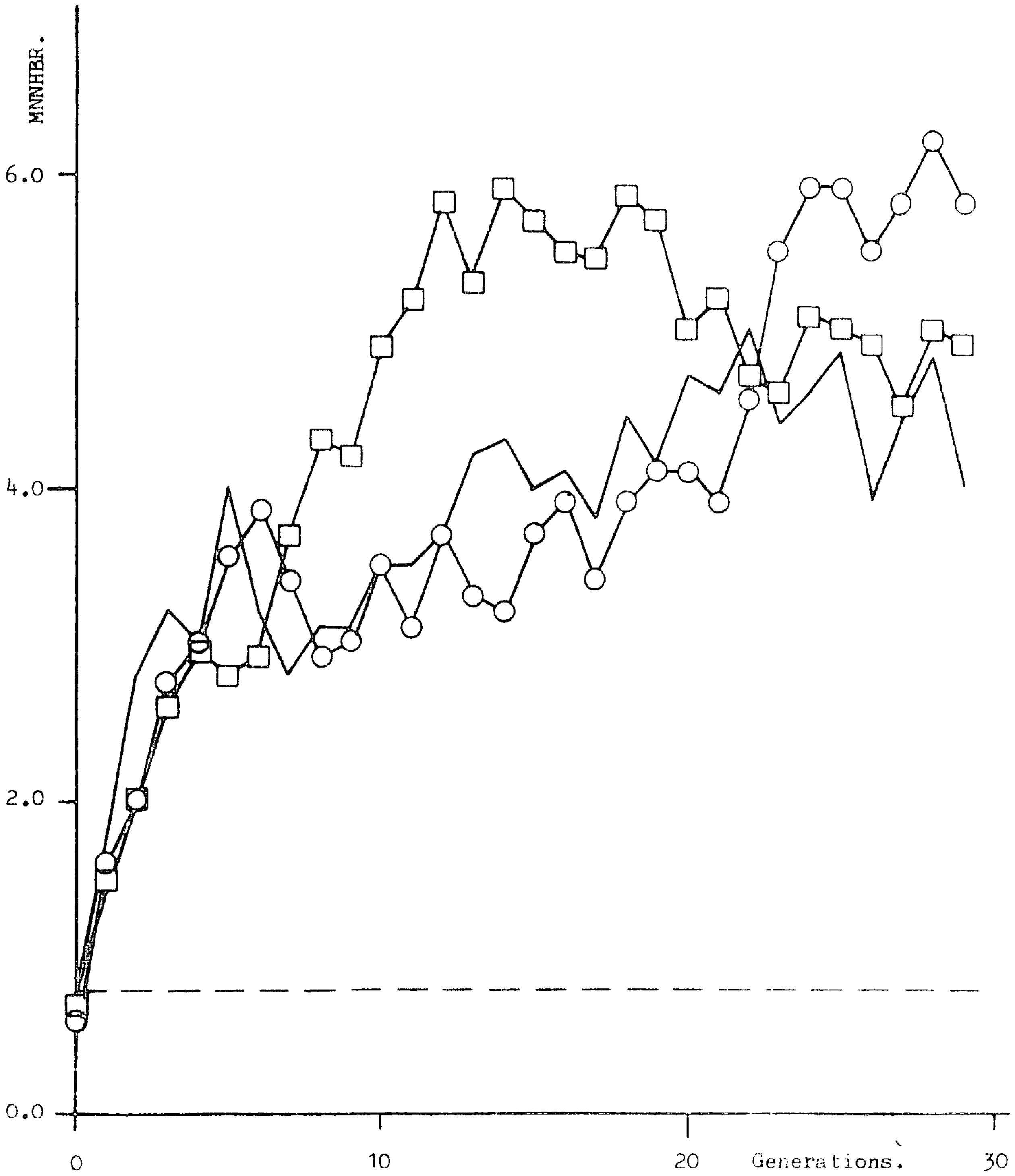
The effect of varying Maximum Domain Radius (R) on clustering in model populations

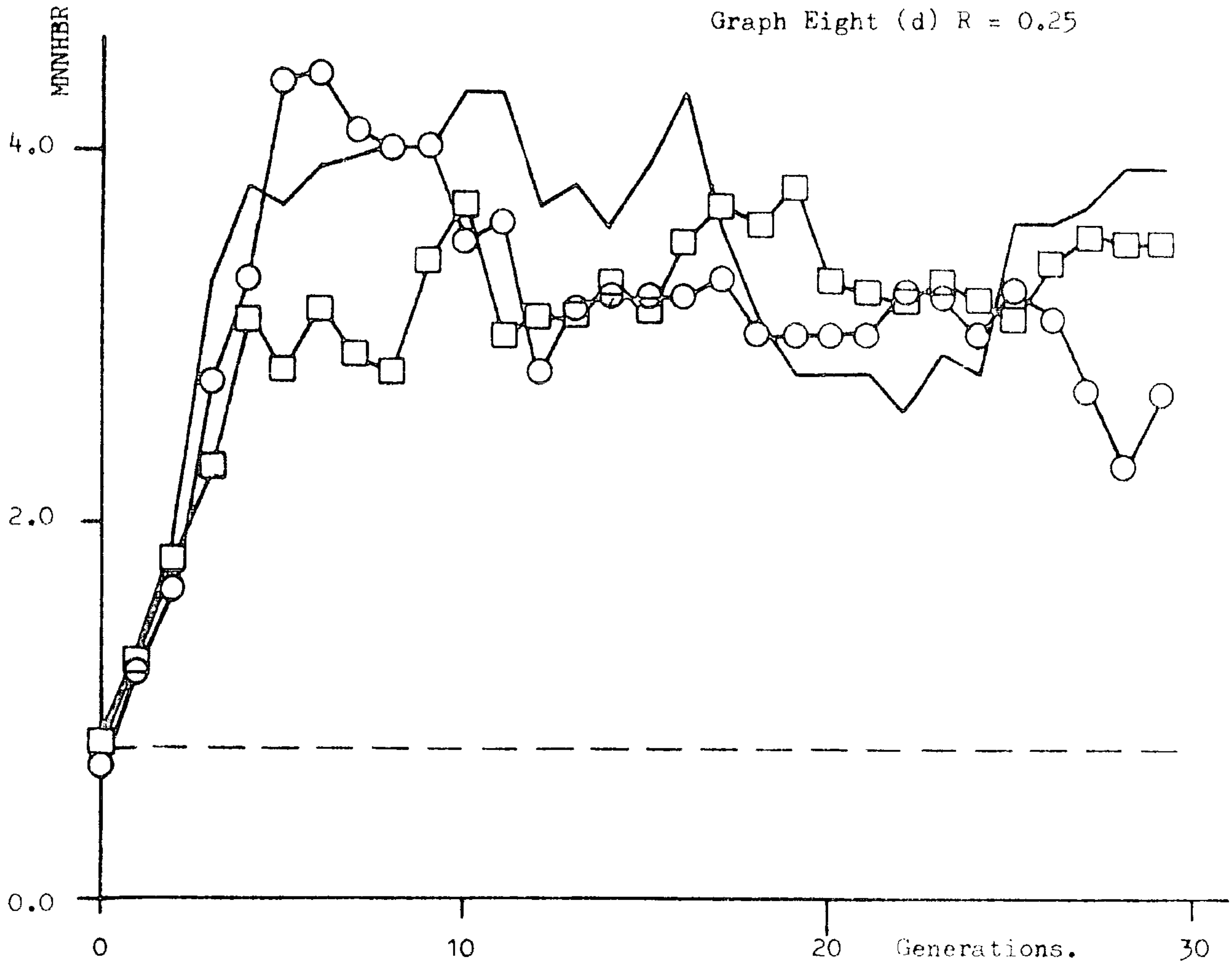
(a) R = 0.1



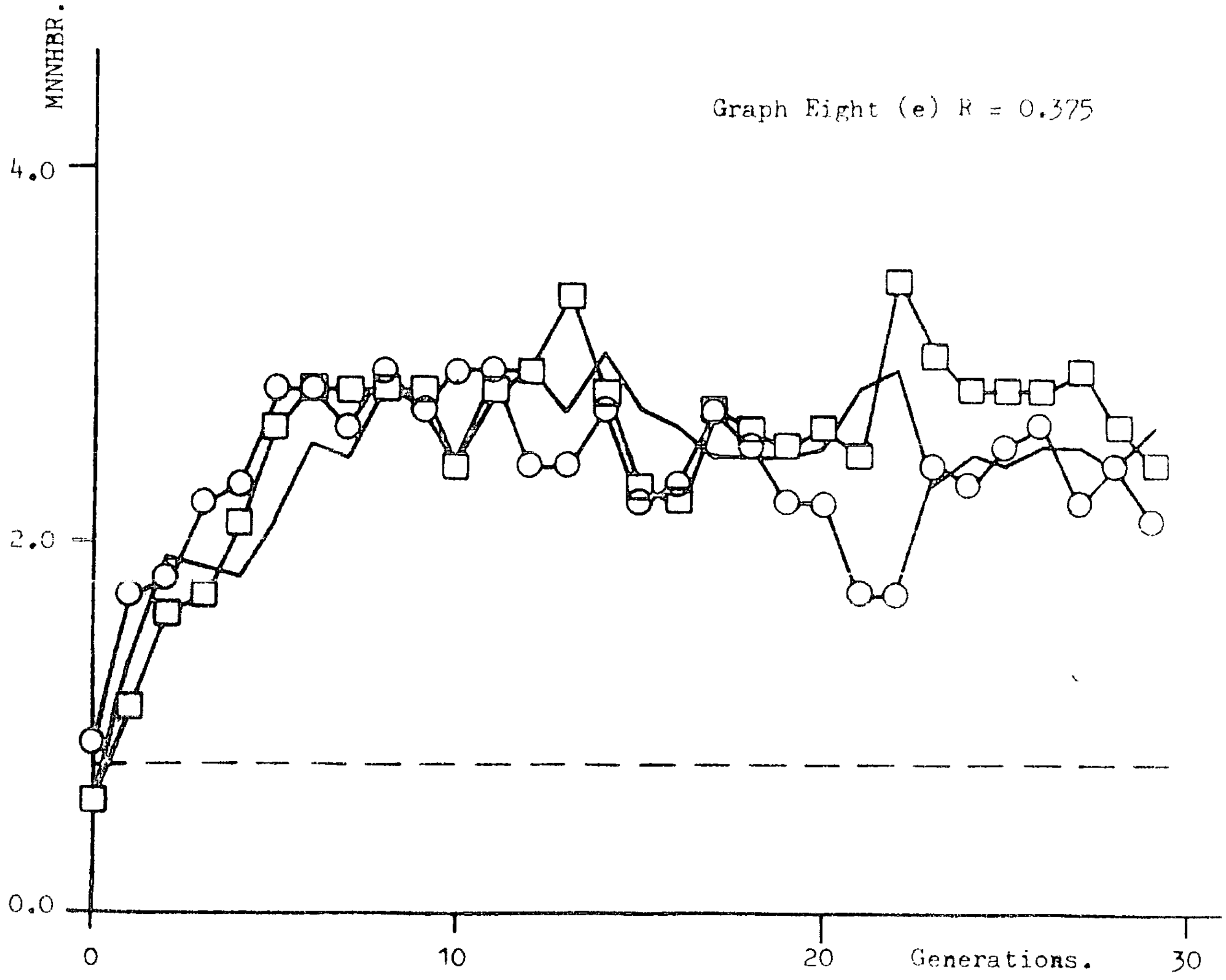
Graph Eight (b) R = 0.15

Graph Eight (c) R = 0.2

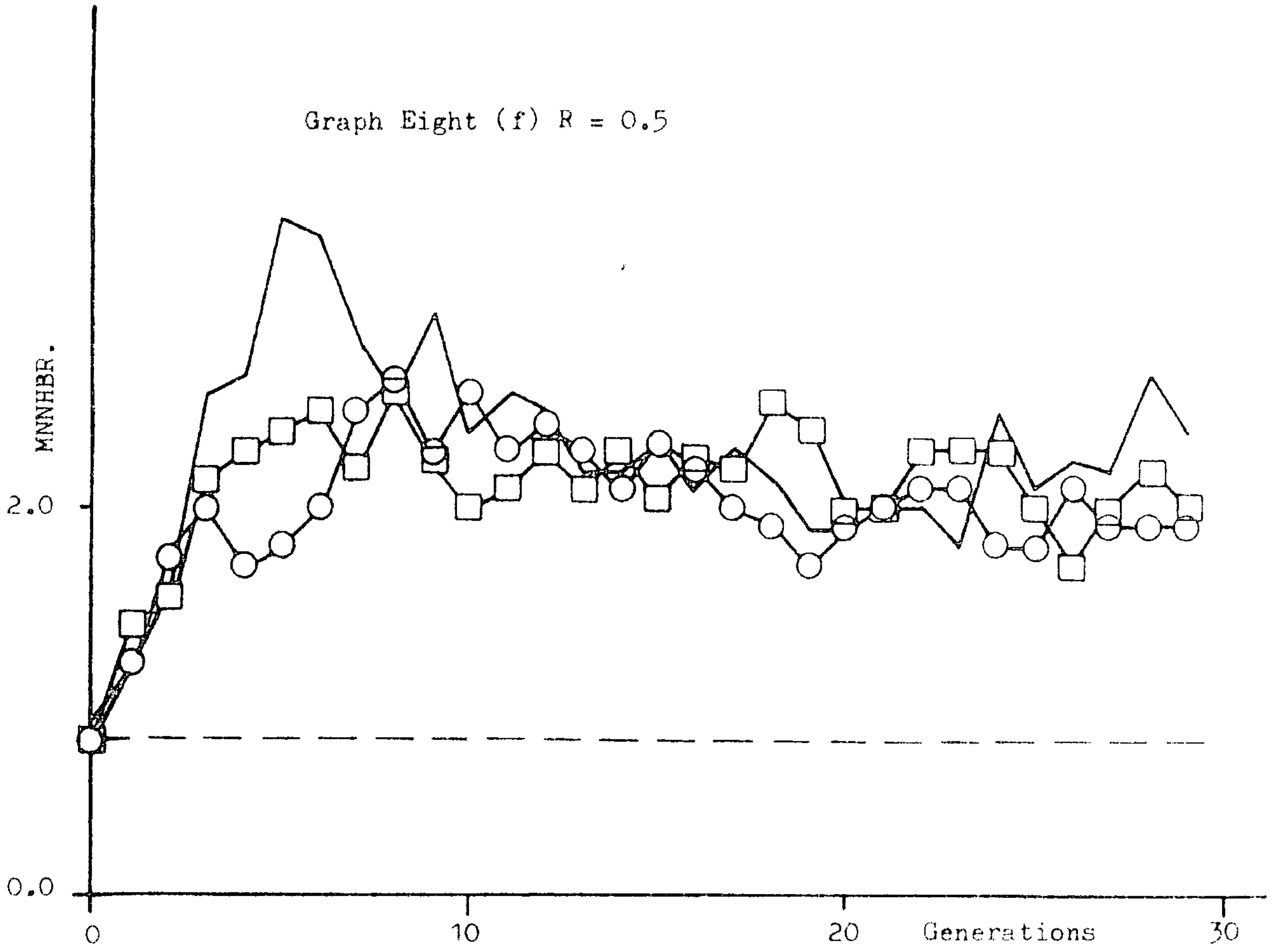


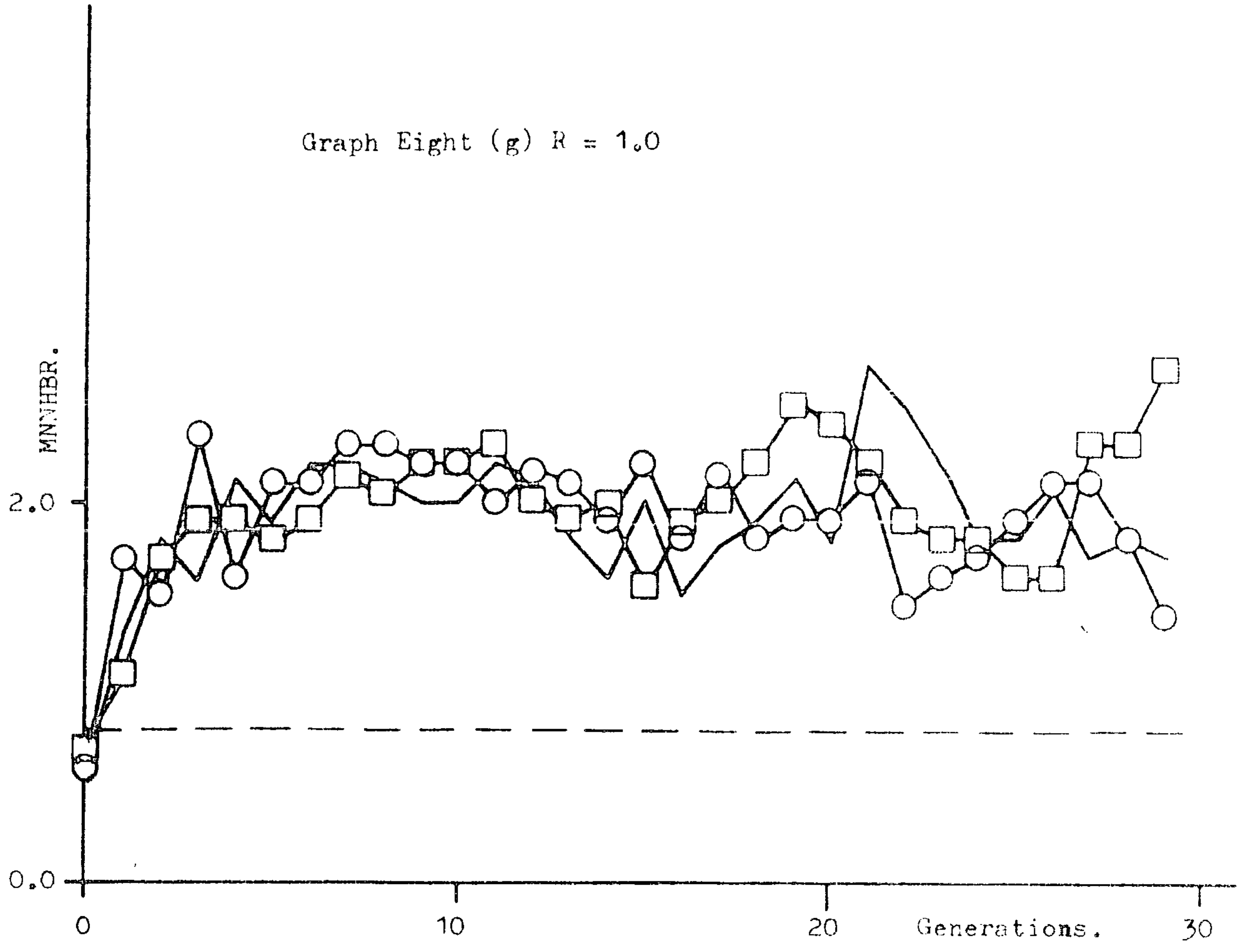


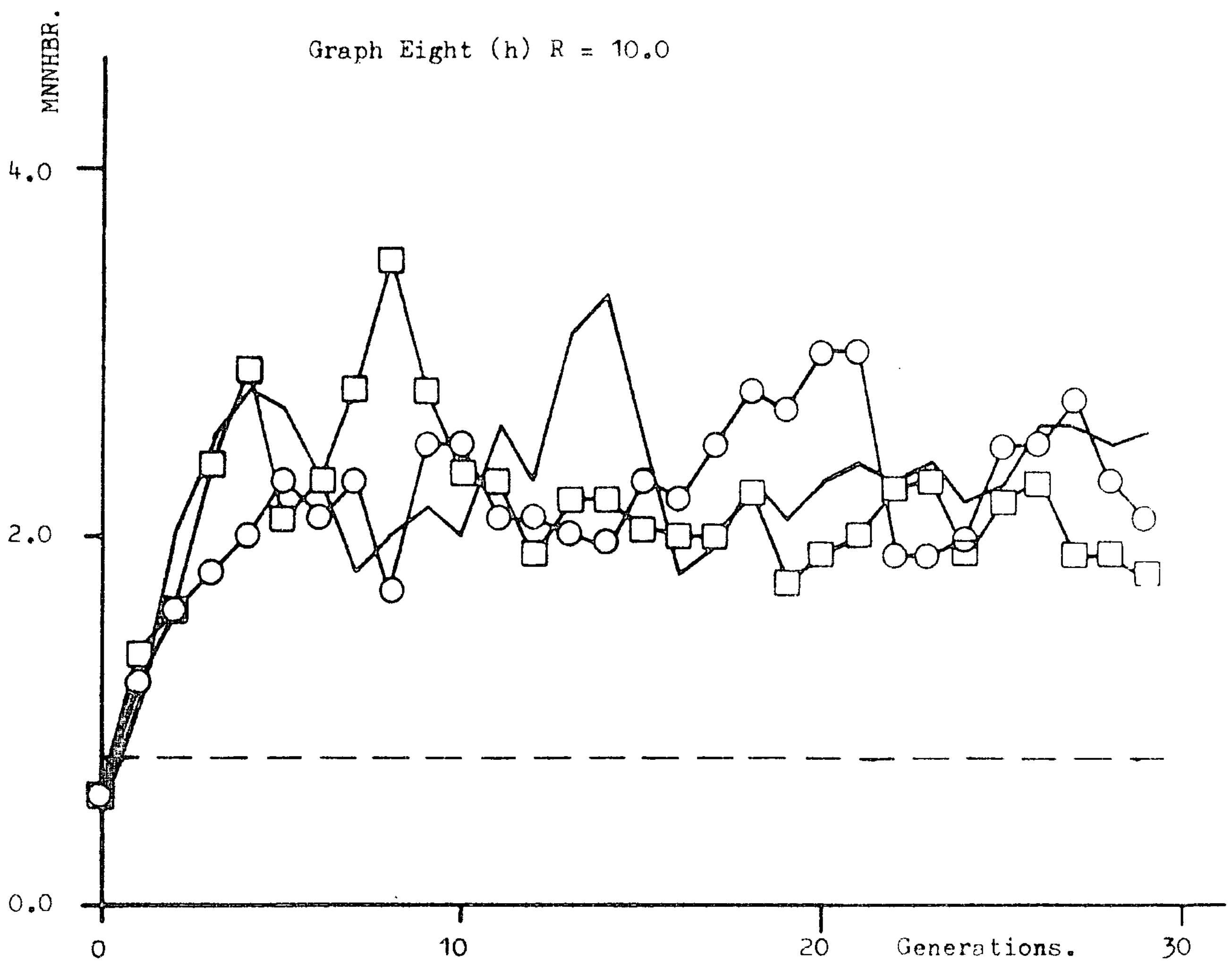
Graph Eight (d) R = 0.25



Graph Eight (e) R = 0.375


Graph Eight (f) R = 0.5
MNNHBR.
2.0
0.0
0
10
20
Generations
30


Graph Eight (g) R = 1.0
MNNHBR.
2.0
0.0
0
10
20
Generations.
30


Graph Eight (h) R = 10.0
MNNHBR.
4.0
2.0
0.0
0
10
20
Generations.
30

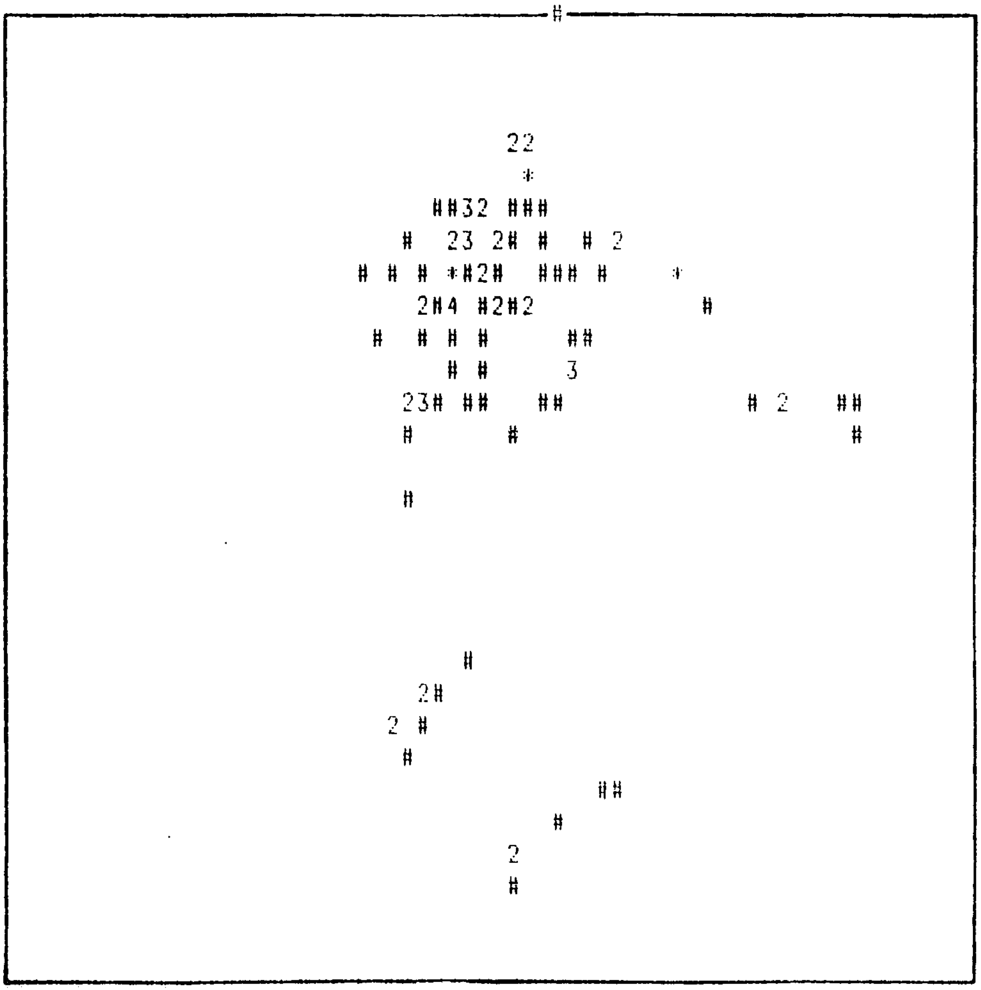


Figure 14. Population of model plants at Generation 30 when the maximum domain radius (R) = 0.1 (MNNHB = 8.0400)

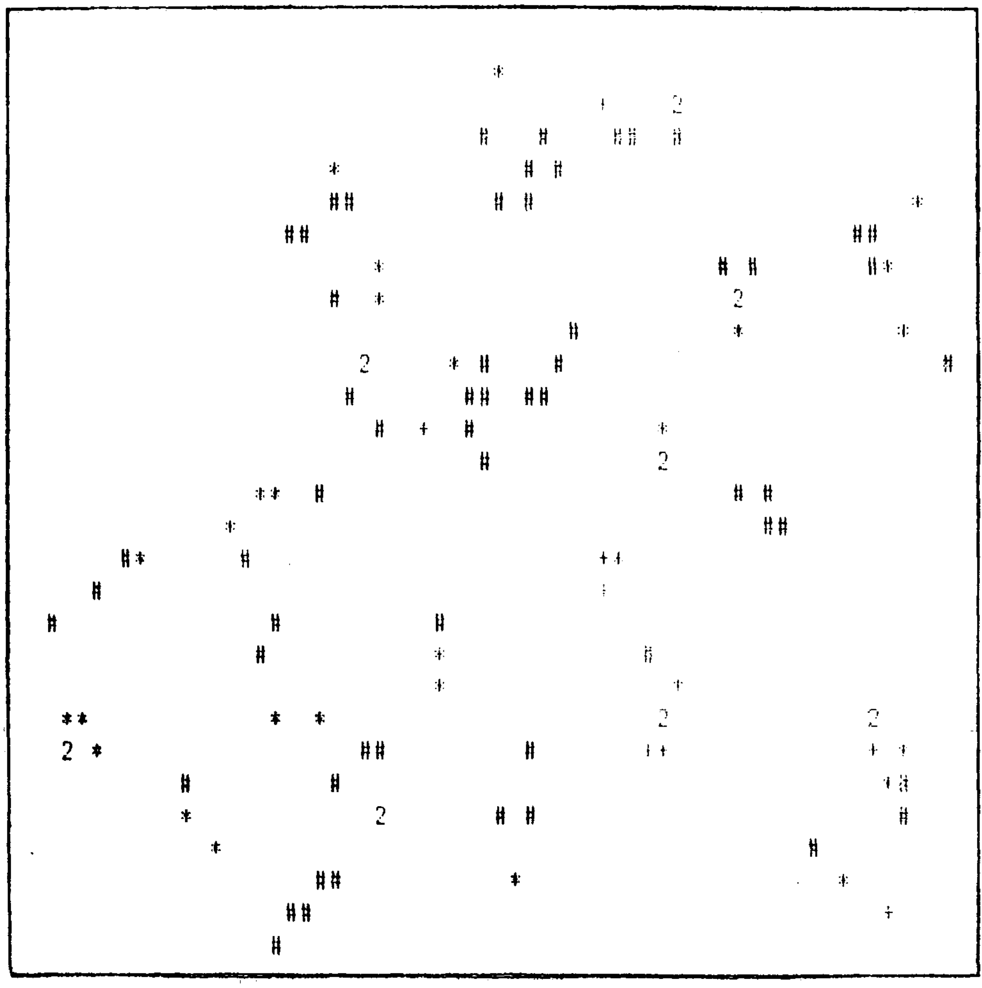


Figure 15. Population of model plants at Generation 30 when the maximum domain radius (R) = 1.0 (MNNHB = 1.9464)

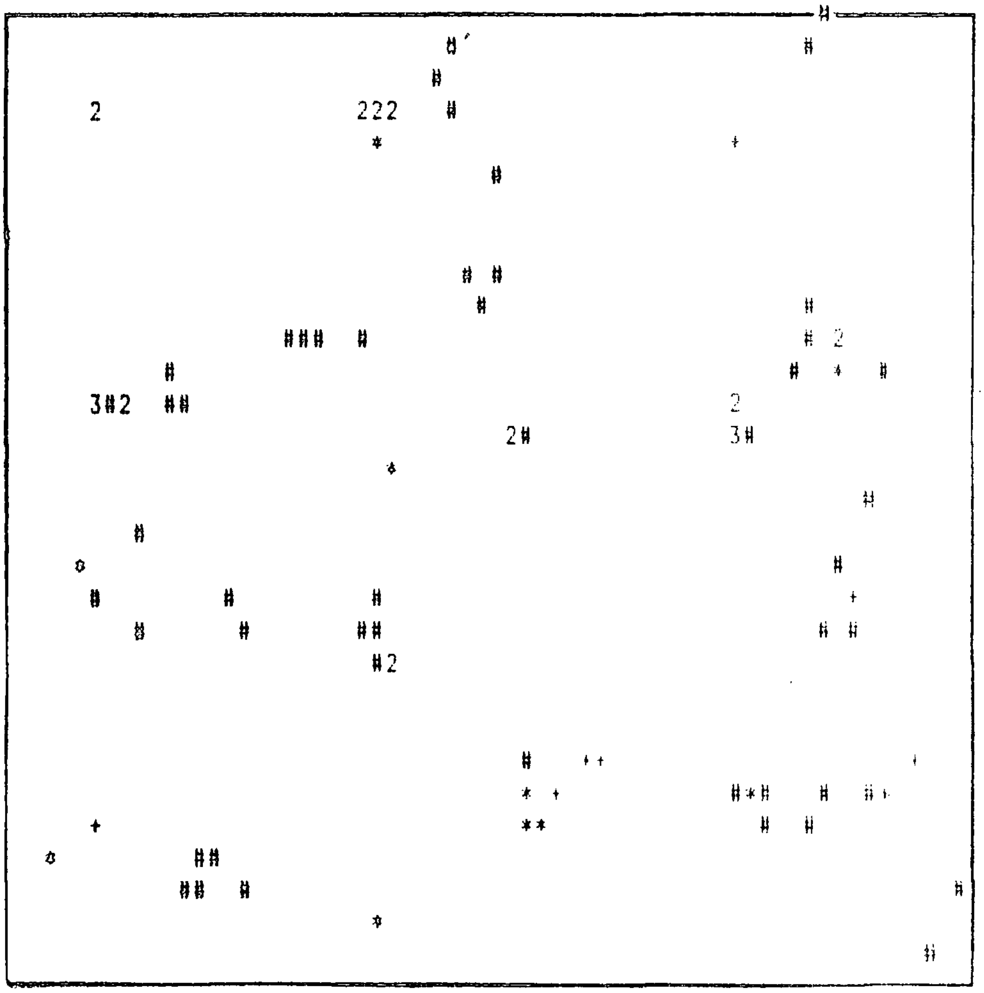


Figure 16. Population of model plants at Generation 30 when the maximum domain radius (R) = 10.0 (MNNHB = 2.5576) (n.b. Local clustering effect, plants aggregating into small clusters in order to maximise resource usage, see figure 17.)

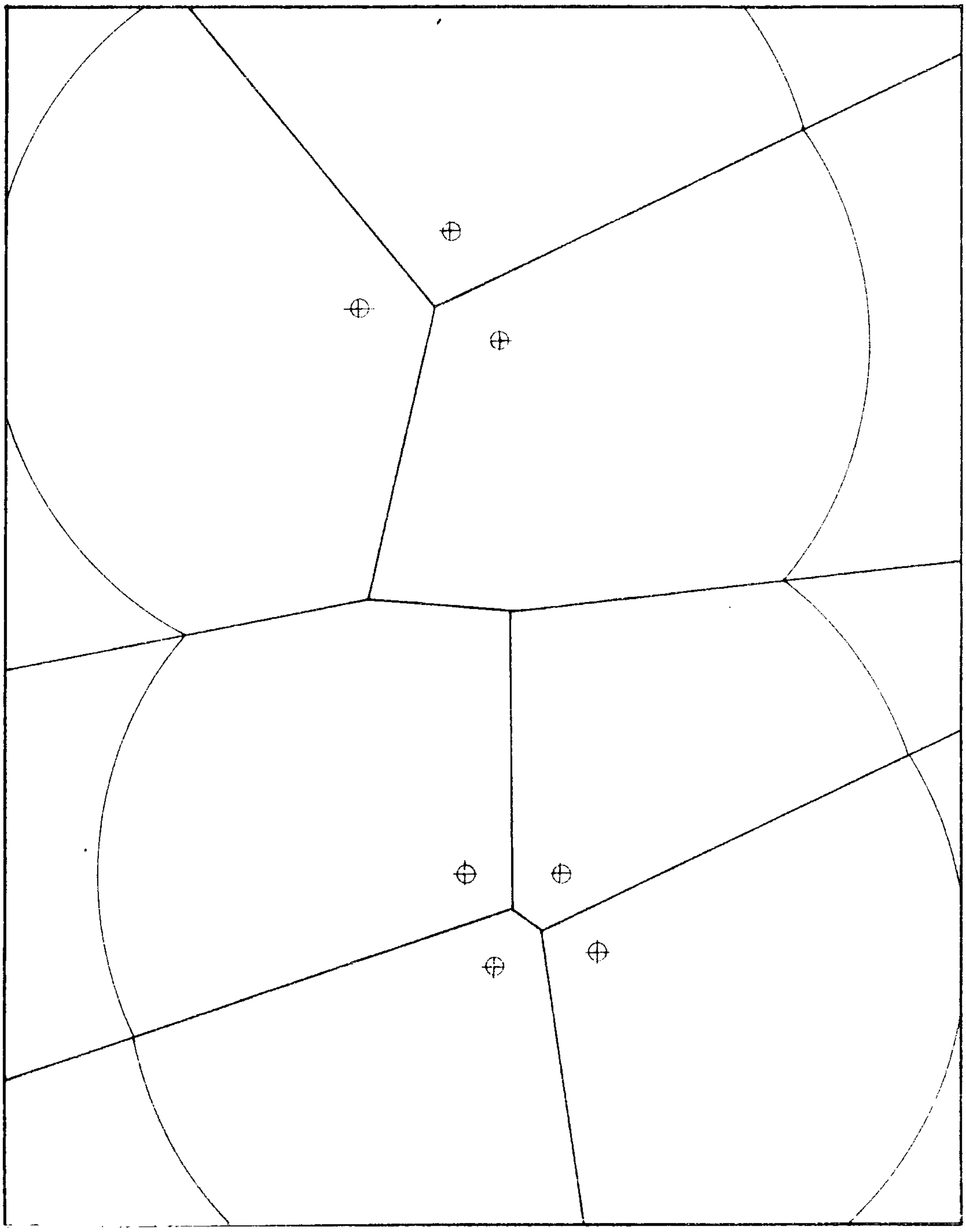


Figure 17 Point clusters when R is large.

4.3 SUMMARY.

The application of the Dirichlet tesselation to SPEV increased the amount of central processor time required per generation of the program by 50%, yet it is justified as a useful, efficient method of modelling spatial interaction between plants, as well as providing a versatile tool for varying the intensity of spatial pattern in the population, by varying the maximum domain available to each plant. The general conclusion from the model is that spatial pattern is a resource related phenomena.
If selection is intense, and demand for 'resources' low, then spatial aggregation may arise. If model plants have high 'resource' demands, that is R is large, then the degree to which spatial pattern may develop is reduced.

5 SEGREGATION.

5.1 DEVELOPMENT OF SPEV TO MODEL SEGREGATION.

Once the model had been developed this far, the final model was within sight. The aim of the final version of SPEV, in this project, was to model the outcome of a system where two homozygous species, 'AA' and 'aa', one self-pollinating ('AA'), the other cross-pollinating ('aa'), coexist in the same spatial area, but where hybrids are inviable.

The aim was to model the spatial patterns that were anticipated in the two species of Senecio for which the field data had been collected. In order to model this situation it was once again necessary to make some alterations to SPEV, and these will be described first, followed by the results of simulations using the program as it now stands.

5.2 FINAL ALTERATIONS TO SPEV TO MODEL SEGREGATION.

The model was altered to represent two species, 'AA' and 'aa', the hybrid 'Aa' being inviable. The self-pollinator, in this case 'AA', had SELFPOL added to its 'A' pollen. (SELFPOL was initialised at the start of the program, representing the amount of pollen the self-pollinator received from its own anthers). Also RAIN was considered to consist of pollen of both species, whereas previously it had consisted only of 'aa' pollen. Selection and dominance were no longer used as parameters in the program, the fitnesses of 'AA', 'Aa' and 'aa' being fixed as 1, 0 and 1 respectively. Finally, the mean number of neighbours, MNNHBR, was calculated for each species independently, so that the intensity of clustering of both species could be compared.

5.3 RESULTS OF SIMULATING SEGREGATION.

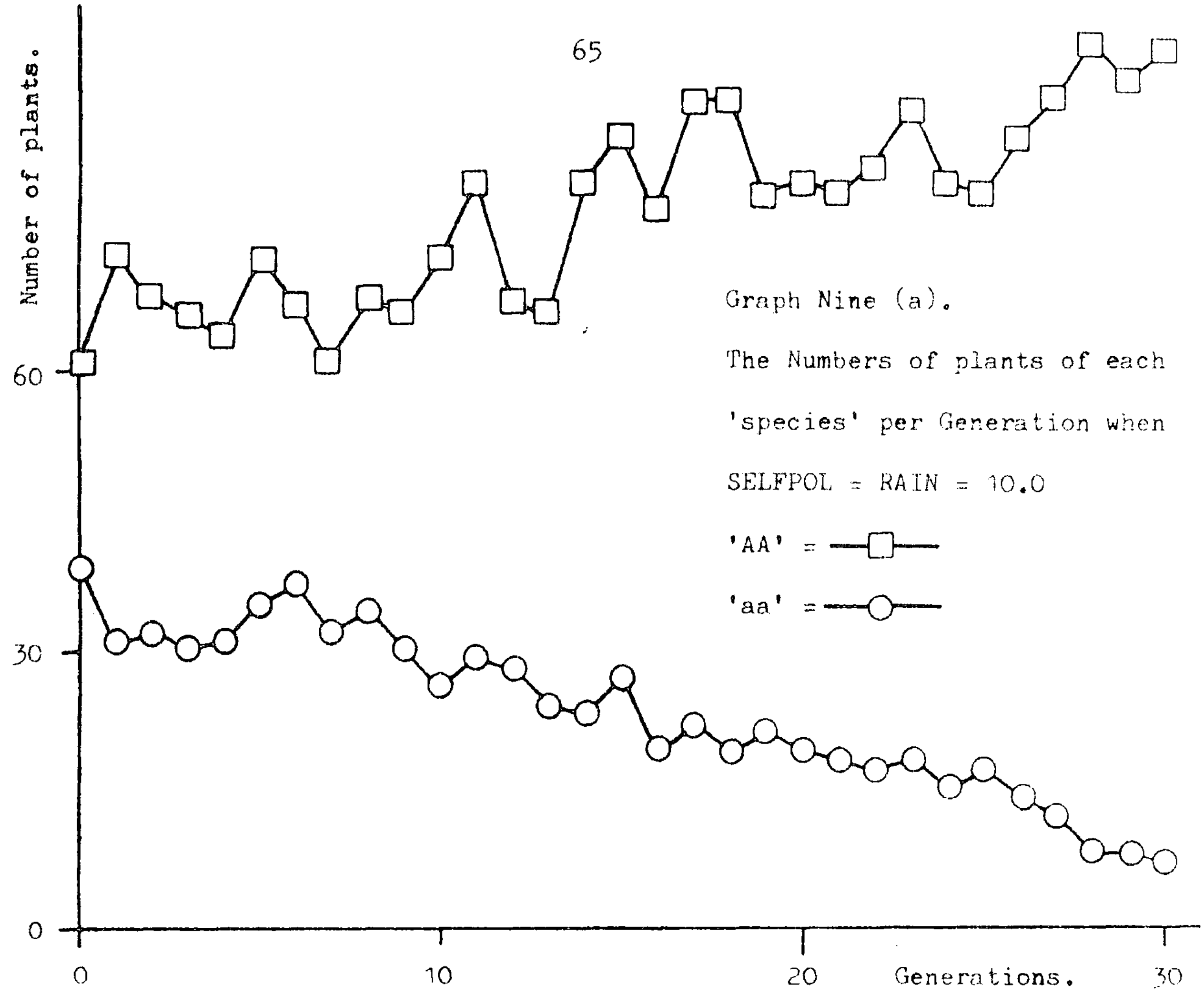
5.3.1 SELFPOL.

If SELFPOL was greater than 0, then the population would generally go to fixation leaving only the self-pollinating species. The reason for this was the same as that for the model in which RAIN had led to the loss of 'A' from the population. When SELFPOL is added to the 'A' pollen counts of 'AA' only, these have a pollen advantage which increases the frequency of 'AA' offspring, which, after the passage of a few generations, led to the elimination of 'aa' plants. This effect is dependent on several factors. If RAIN is increased, whilst R and SIGMA are reduced, then the rate at which cross pollinating 'aa' plants are eliminated, is decreased, and the intensity of spatial pattern of both species is increased. This is demonstrated on Graphs 9 and 10, plotting the frequency of both species, and the intensity of clustering in both species when RAIN is 10.0, and 100.0 respectively. (On both graphs R = 0.15, SIGMA = 0.05, and SELFPOL = 10.0).

5.3.2 SELFPOL = 0.

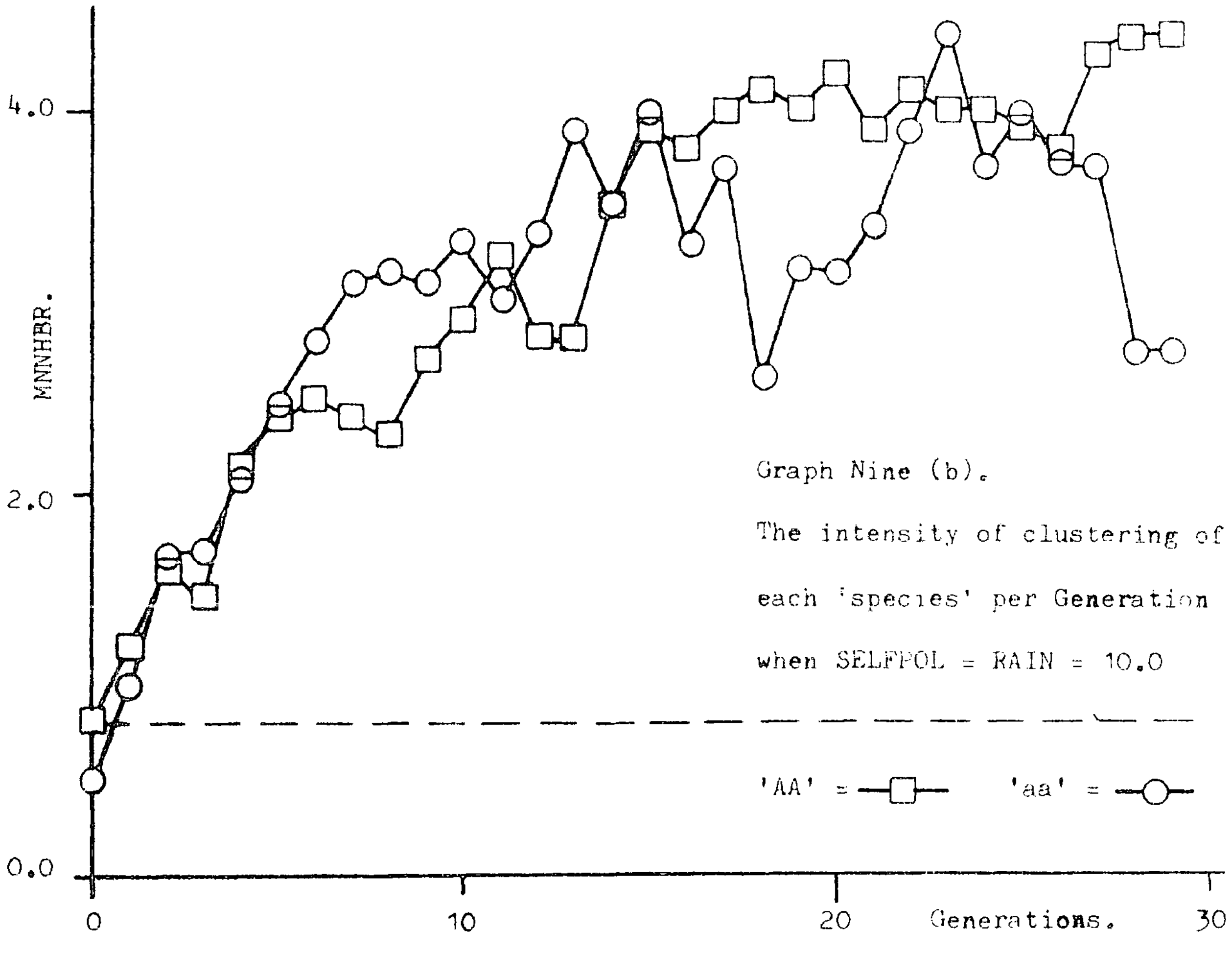
If SELFPOL and RAIN were both zero, then the outcome of the simulation was random. Two general cases were observed:

Firstly, the population would go to fixation, that is either one species or the other.



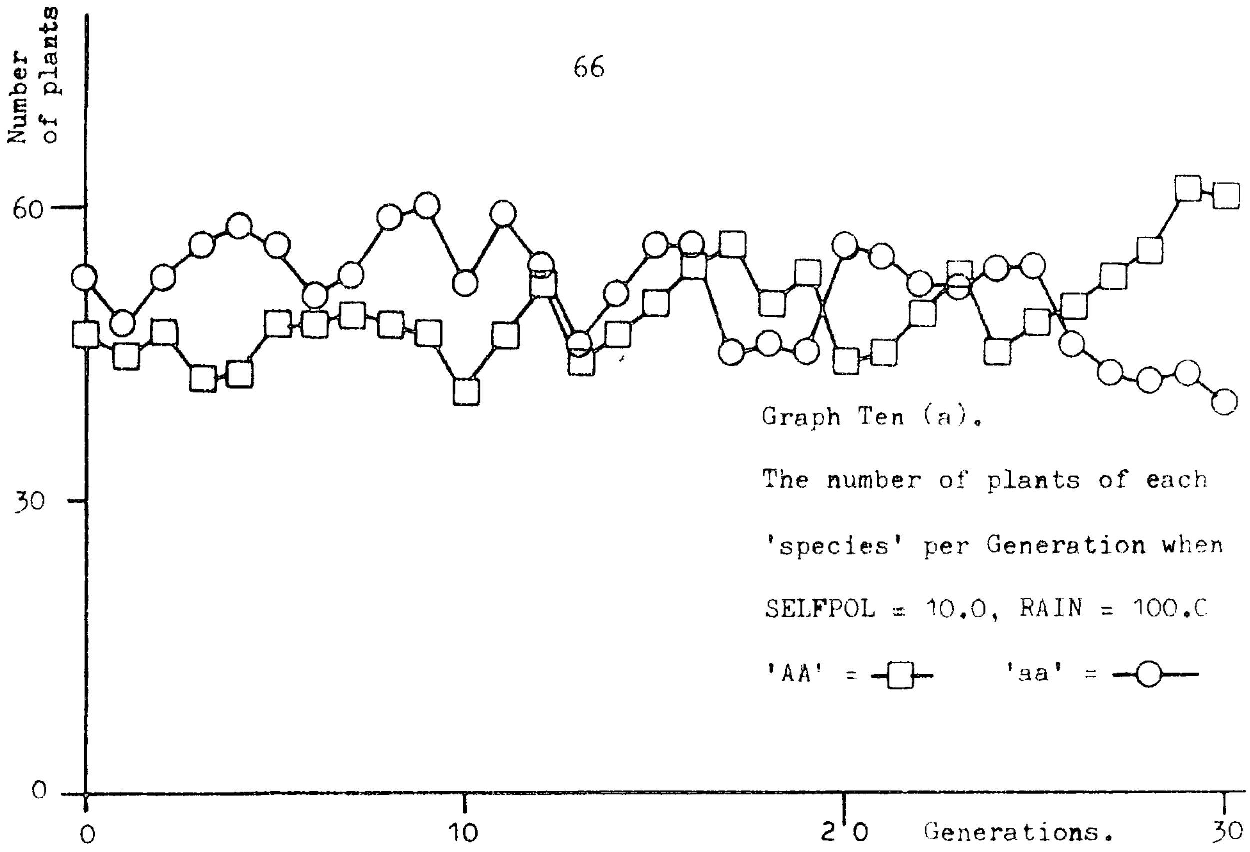
Graph Nine (a).

The Numbers of plants of each 'species' per Generation when SELFPOL = RAIN = 10.0



Graph Nine (b).

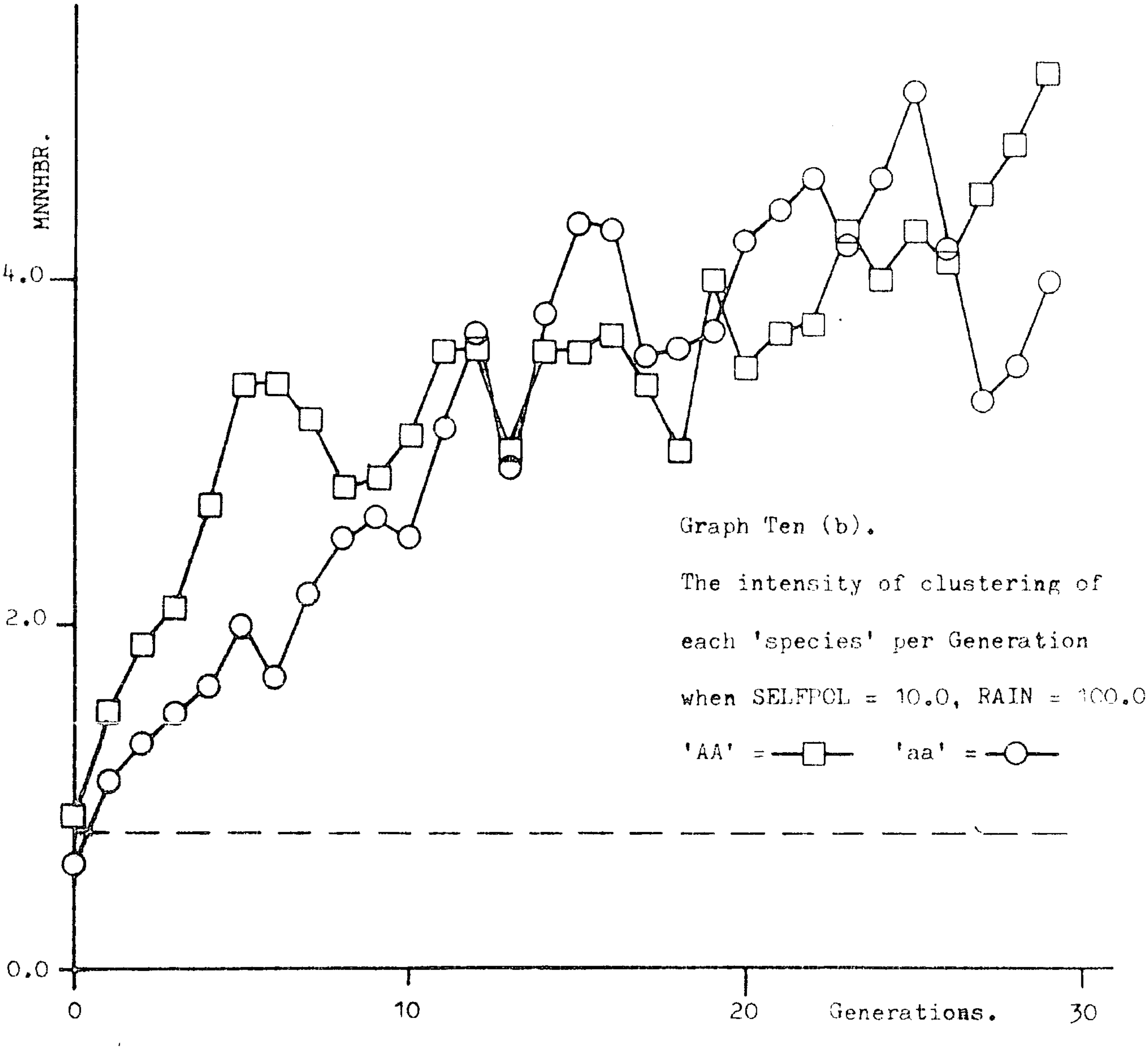
The intensity of clustering of each 'species' per Generation when SELFPOL = RAIN = 10.0



Graph Ten (a).

The number of plants of each 'species' per Generation when SELFPOL = 10.0, RAIN = 100.0

'AA' = —☐—   'aa' = —○—



Graph Ten (b).

The intensity of clustering of each 'species' per Generation when SELFPOL = 10.0, RAIN = 100.0

'AA' = —☐—   'aa' = —○—

Secondly, and more interestingly, the population would segregate into two discretely clustered groups.

The first result was often observed because at the outset of the program one 'species' chanced to be slightly more aggregated than the other, this slight clustering advantage often determined the direction in which fixation would go. The second, and more interesting result was often the case where the seed dispersal distance and maximum domain radius were diminished, so that tight clusters could develop. If an 'AA' plant happened to occur in a cluster of 'aa', it would tend to be lost, because of the amount of 'aa' pollen it received from its close neighbours, resulting in the loss of 'hybrid' offspring. The converse effect is observed with 'aa' in clusters of 'AA', there being strong pressure to form mutually exclusive clusters. Figures 18 and 19 demonstrate this effect, the randomly dispersed population at generation 0, has aggregated into 7 discrete clusters by generation 30.

One final subtle change to SPEV was made rather experimentally. RAIN in the segregating version of SPEV had been pollen of both species, this was reverted to the same situation as the original model where RAIN was the pollen of one genotype only. The reason for this was to model the different amounts of pollen produced by self pollinating and cross pollinating species. RAIN is used in this instance to model the excess of pollen produced by the cross pollinating species. The results of simulations done in this way showed that the cross pollinating species would eliminate the self pollinator, unless SELFPOL was 3 times greater than RAIN, but where the population appeared stable, the intensity of clustering in the self pollinating species was increased, whereas the cross pollinating species

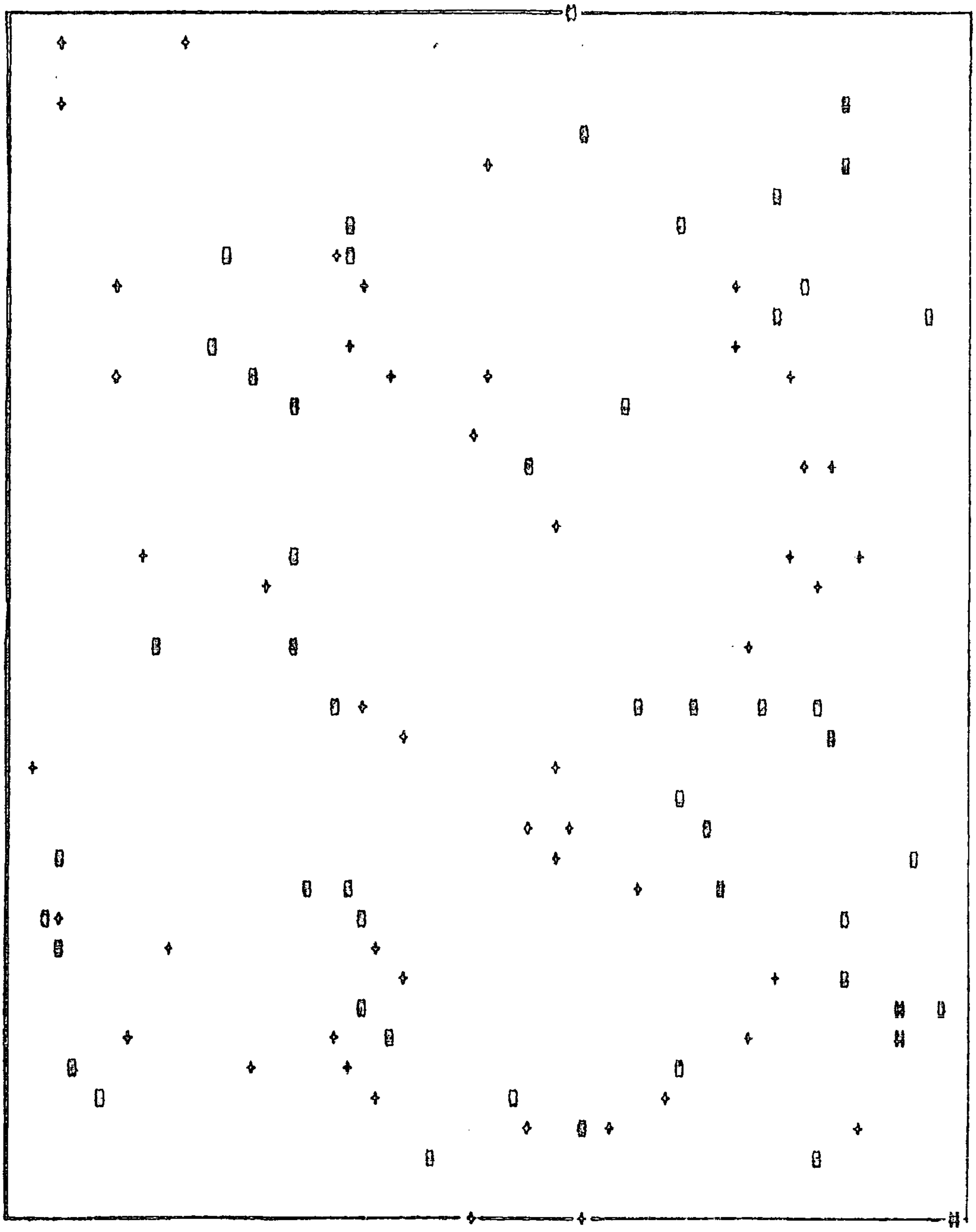


Figure 18. Two species model plant population, randomly dispersed at Generation 0

(N('AA') = 51, MNNHB('AA') = 1.01961)

(N('aa') = 49, MNNHB('aa') = 0.93878)

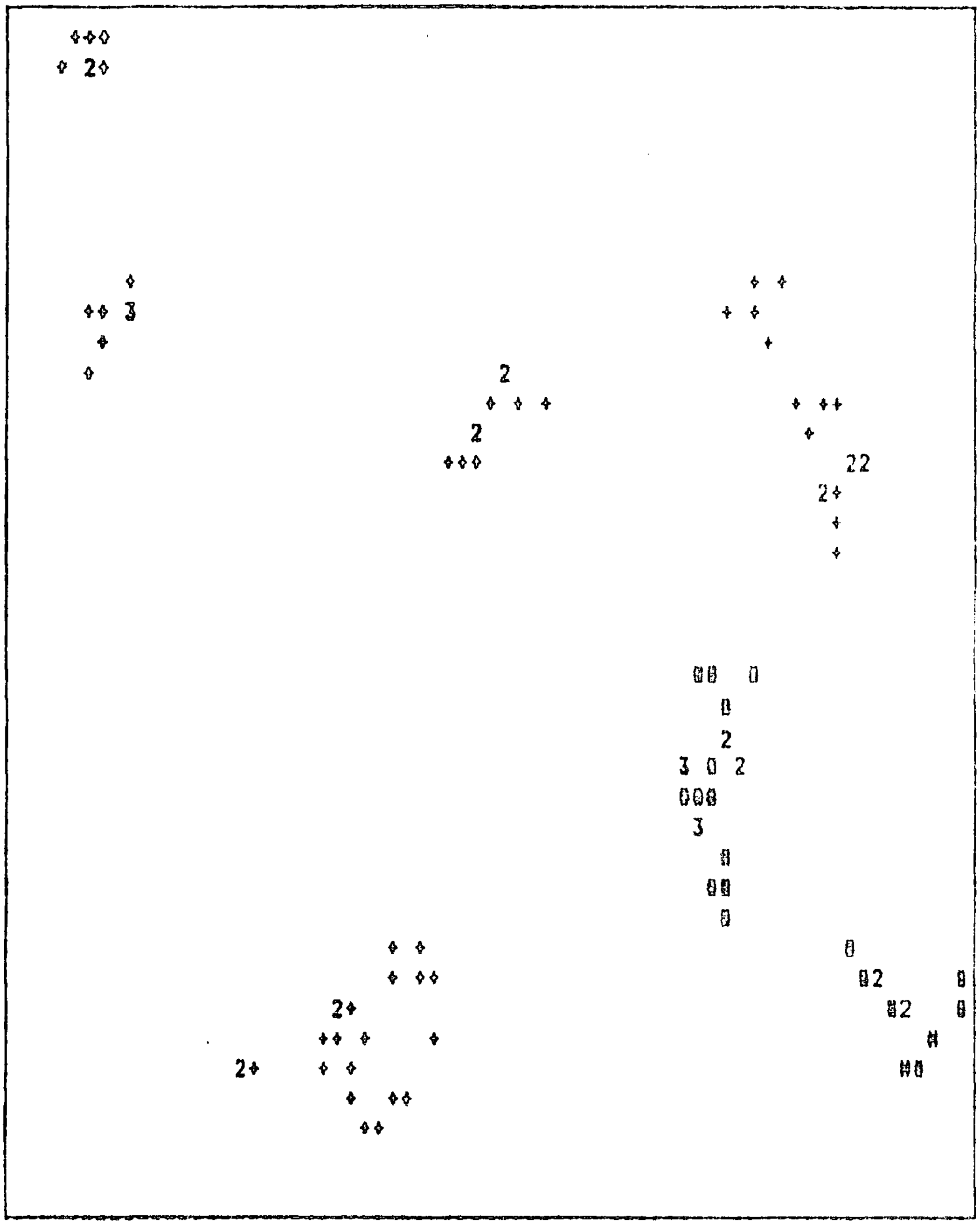


Figure 19. The same model population as Figure 18, at Generation 30 of the simulation. The population has segregated into seven isolated single species groups.

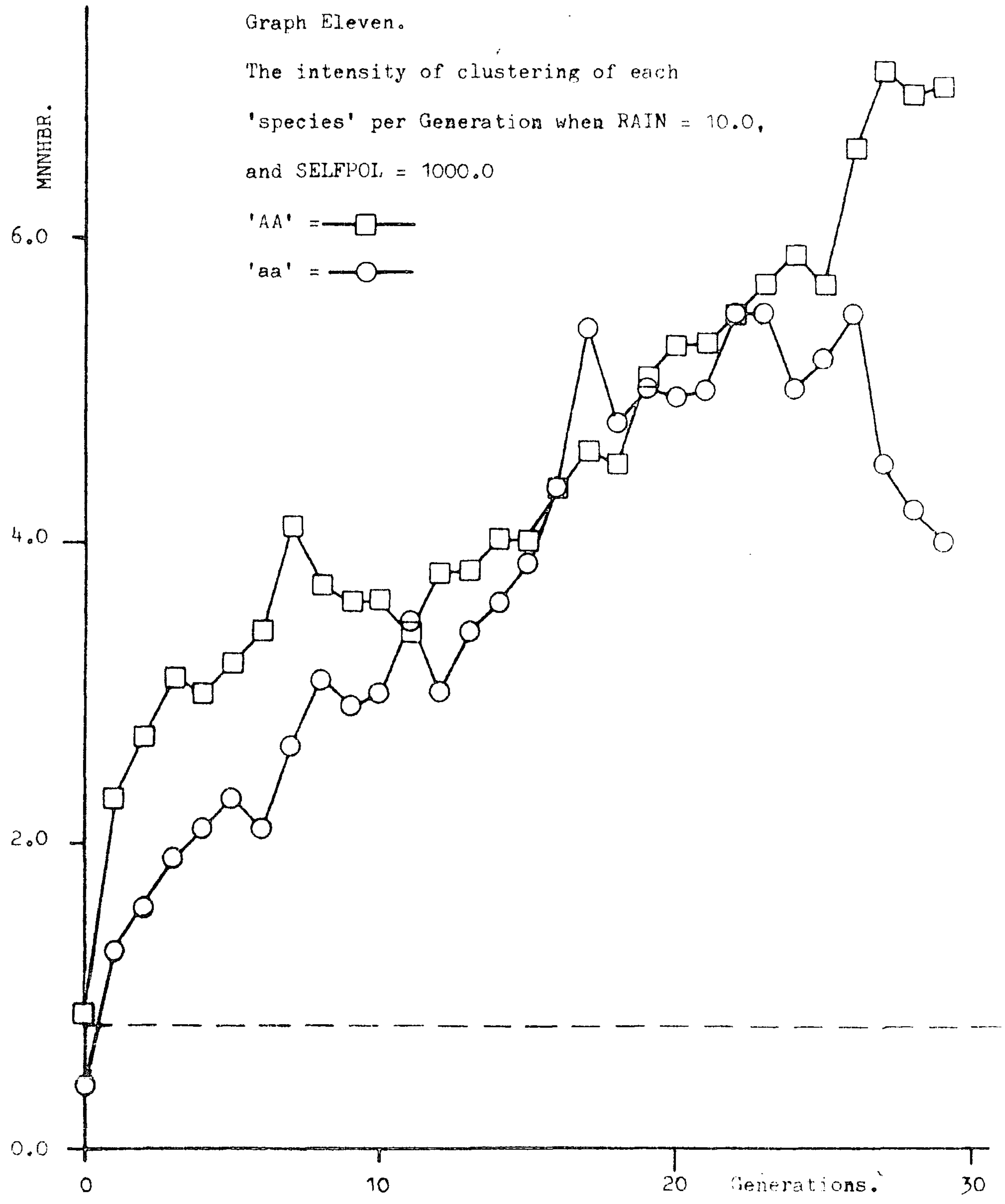
(N('AA') = 34, MNNHB('AA') = 11.17647)

(N('aa') = 65, MNNHB('aa') = 6.41176)

maintained a more random dispersion. Graph 11 plots the mean number of neighbours of both species, where RAIN = 10.0, and SELFPOL = 1000.0.

5.4. SUMMARY.

The results of simulating two differing breeding systems has shown clearly that self-pollinators have an advantage over cross-pollinators, which led to the extinction of the latter. This advantage could be reduced by increasing RAIN, effectively swamping the advantage of self-pollination. Also if RAIN was pollen of one species only then this would counteract the advantage of self pollination, and in both cases discrete single species groups would arise. If SELFPOL and RAIN were 0 then the population would segregate, as had been expected, into single species groups because of the loss of plants of one species within clusters of the other. However, it should be stressed that, in a model in which the hybrids are inviable, the only stable population is one containing one species or the other, i.e. one species will always eventually be eliminated.



Graph Eleven.

The intensity of clustering of each 'species' per Generation when RAIN = 10.0, and SELFPOL = 1000.0

6 RESULTS OF THE FIELD STUDY OF SPATIAL PATTERN IN POPULATIONS OF S.VISCOSUS AND S.SQUALIDUS

6.1 T-SQUARE FIELD DATA.

Data was collected from a total of sixteen sampling frames for S.viscosus, and 13 sampling frames for S.squalidus. At each site the population was estimated, and the number of random sample points adjusted accordingly. (This is in theory unnecessary, since T-square sampling is independent of the density of the population, and a fixed number of sample points per sampling frame could have been used). The data for each sampling frame has been summarised as values for t_N and t_B, as well as their respective probabilities under the null hypothesis of random distribution. Also for each site where the null hypothesis is accepted, M/C (Diggle, 1977) is also calculated. An asterisk marks those sampling frames where data was collected for both species. (Significant probabilities only are given).

Table 2: Senecio squalidus T-square sampling data.

Sampling frame.	N	t_B	Pt_B	t_N	Pt_N	M/C	$P_{M/C}$
1	24	0.8300	0.001	0.6974	0.001		
2	11	0.8164	0.001	0.7603	0.001		
12	6	0.7095		0.6149		8.0566	
13	6	0.3791		0.5960		3.2332	
14	7	0.6135		0.5877		2.5005	
15	8	0.5670		0.4988		6.7245	
21 *	8	0.4950		0.5013		2.9605	
22 *	5	0.5781		0.4552		1.8090	
23 *	20	0.7530	0.001	0.6438	0.012		
24 *	40	0.5036		0.5693		23.4421	
25 *	12	0.9030	0.001	0.6443	0.042		
26 *	17	0.4984		0.6293		45.8047	0.001
27 *	12	0.4161		0.4594		8.7701	

Abbreviations: $\left.\begin{array}{l} Pt_N \\ Pt_B \\ P_{M/C} \end{array}\right\}$ probabilities of t_N, t_B, and M/C under the null hypothesis

Table 3: Senecio viscosus T-square sampling data.

Sampling frame	N	t_B	Pt_B	t_N	Pt_N	M/C	$P_{M/C}$
3 *	16	0.5230		0.5614		7.8670	
4 *	17	0.6825	0.014	0.5001			
5 *	11	0.6970	0.028	0.6837	0.017		
6 *	9	0.8487	0.001	0.8340	0.001		
10	40	0.5133		0.4277		25.0008	
11	48	0.7379	0.001	0.5596	0.05		
16	35	0.7748	0.001	0.6298	0.004		
19	8	0.7344	0.025	0.5919			
20	14	0.3529		0.5653		20.9458	
21 *	8	0.6348		0.5536		5.2432	
22 *	7	0.6920		0.4953		8.7955	
23 *	20	0.4602		0.5174		27.6846	
24 *	27	0.5236		0.5924	0.048		
25 *	20	0.9252	0.001	0.7284	0.001		
26 *	6	0.3423		0.5221		5.7563	
27 *	13	0.6517	0.058	0.6882	0.009		

The T-square data collected for S.jacobaea is given in Appendix A.

Where either of the test statistics yielded a significant result, that is $P<0.05$, the null hypothesis was rejected, and the population was considered to be aggregated. On this basis the data was condensed:

Table 4: Summary of the T-square sampling data.

	Random	Non-random
S. squalidus	9	4
S. viscosus	7	9

This was further broken down into tables summarizing the numbers of sampling frames showing clustering for S. squalidus and S. viscosus when;

1. In the same area as a population of the other species (sympatric populations).

2. In isolated populations of one species or the other (allopatric populations).

Table 5:

	S. squalidus		S. viscosus	
	Random	Non-random	Random	Non-random
+ other spp.	5	2	5	6
- other spp.	4	2	2	3

(The data broken down in this way could have been used as the basis of a 2 x 2 contingency test, however the number of samples was too small to yield any useful information).

As an alternative the mean values of t_N obtained for both species in allopatric and sympatric populations were compared using an analysis of variance;

Table 6 (a). Analysis of Variance of t_N data obtained for S. squalidus

	t_N				N	Total	$\bar{x}$
+ vis	0.5013	0.4552	0.6438	0.5693	7	3.9026	0.5575
	0.6443	0.6293	0.4954				
- vis	0.6974	0.7603	0.6149	0.5960	6	3.7551	0.6258
	0.5877	0.4988					
	Totals........................				13	7.6577	0.5890

	SS	df	MS	F	
Between treatments	0.01518	1	0.01518	1.9548	n.s
Residual	0.08542	11	0.00776		

Table 6 (b). Analysis of Variance of t_N data obtained for S. viscosus

	t_N				N	Total	$\bar{x}$
+ squ	0.5614	0.5001	0.6837	0.8340	11	6.6766	0.6069
	0.5536	0.4953	0.5174	0.5924			
	0.7284	0.5221	0.6882				
- squ	0.4277	0.5596	0.6298	0.5919	5	2.7743	0.5548
	0.5653						
	Totals........................				16	9.4509	0.5906

	SS	df	MS	F	
Between treatments	0.0094	1	0.0094	0.9001	n.s
Residual	0.1462	14	0.01044		

Both of the F values obtained in Table 6 are not significant. This leads to the conclusion that within treatment (allopatric or sympatric population classes) variation is greater than between treatment variation.

6.2 NEAREST NEIGHBOUR FIELD RESULTS.

The data collected at those sites where both species grew together, relating to the species of a random plant, and that of its nearest neighbour, are given, as well as values of chi-squared and the segregation coefficient (Pielou, 1961) for the data.

	Neighbour Species	
Base plant	squalidus	viscosus
squalidus	a	b
viscosus	c	d

Table 7: Segregation data for Mixed Populations of S. squalidus and S. viscosus

Site Name	a	b	c	d	S	
Old Railway No 5	2	2	4	10	0.1818	0.6428
Old Railway No 6	0	2	1	10	0.1142	0.1969
Sidings No 1	18	6	9	5	0.1121	0.4935
Sidings No 2	19	17	15	16	0.4576	0.0728
Sidings No 3	11	7	6	8	0.1811	1.0537
Sidings no 4	13	4	1	5	0.5147	6.6591[2]
Sidings No 5	10	3	6	6	0.3534	2.2435

2 Significant at P 0.05

Chi-squared has one degree of freedom, the null hypothesis being that the two species are not segregated, (see Krebs, 1978 , pp 376, and Pielou, in Poole, 1974, pp 341). The value of S is 0 if the two species are not segregated, +1 if the two species are completely segregated, and -1 if the nearest neighbour is always the other species.

6.3 RESULTS FROM THE CONTIGUOUS QUADRAT GRID.

The aim of the contiguous quadrat grid was to detect spatial pattern, firstly in each species independently, and secondly to assess the degree of association of the two species over the quadrat grid. It was hoped to detect pattern at more than one scale in the population. In each 0.18m quadrat of a 32 x 64 quadrat grid, the numbers of both species in the quadrat were scored. (Maps of the quadrat grid for each species and both species together are given in Figures 20, 21 and 22. Each character represents the number of plants in that quadrat). These counts were firstly compared with a Poisson series, see Section 2.6.1 The numbers of quadrats containing 0,1,2,3... individuals were scored for both species, and the mean number of plants per quadrat for both species was also calculated.

The mean number of plants per quadrat
of <u>S.squalidus</u> = 1.0957

The mean number of plants per quadrat
of <u>S.viscosus</u> = 0.4922

Key to the symbols used in Figures 20,21 and 22.

Symbol.	Number of plants.
+	1
⧺	2
⁎	3
4	4
5	5
.	.
.	.
.	.
9	9
A	10
B	11
.	.
.	.
.	.
J	20

On Figure 22 Green symbols represent quadrats in which only *S. viscosus* was present, Blue symbols represent quadrats in which only *S. squalidus* was present, Black symbols represent the total number of plants present in quadrats containing both species.

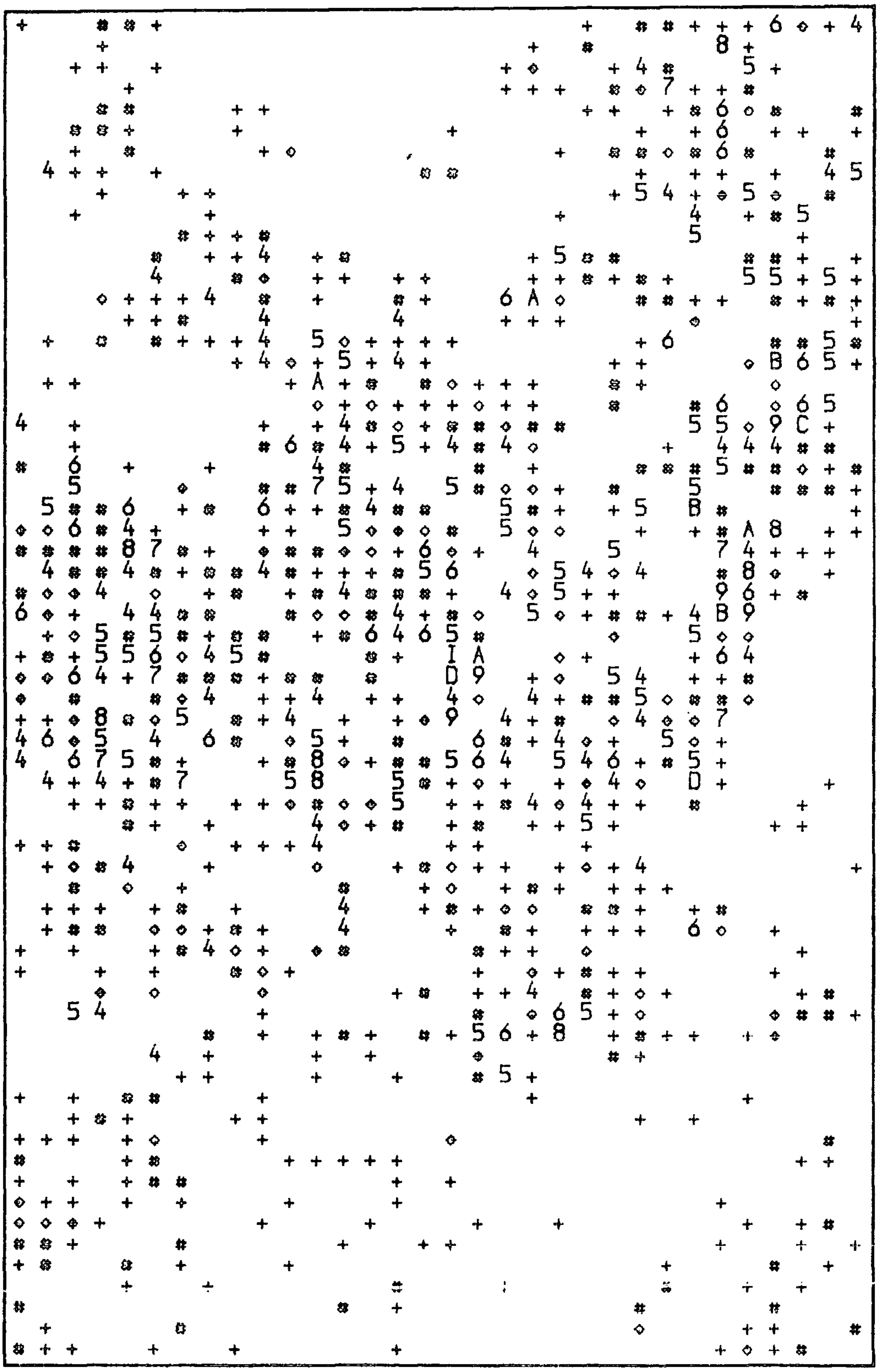


Figure 20. Map of the grid of contiguous quadrats for <u>Senecio squalidus L.</u> (The key to the symbols is given on page 79).

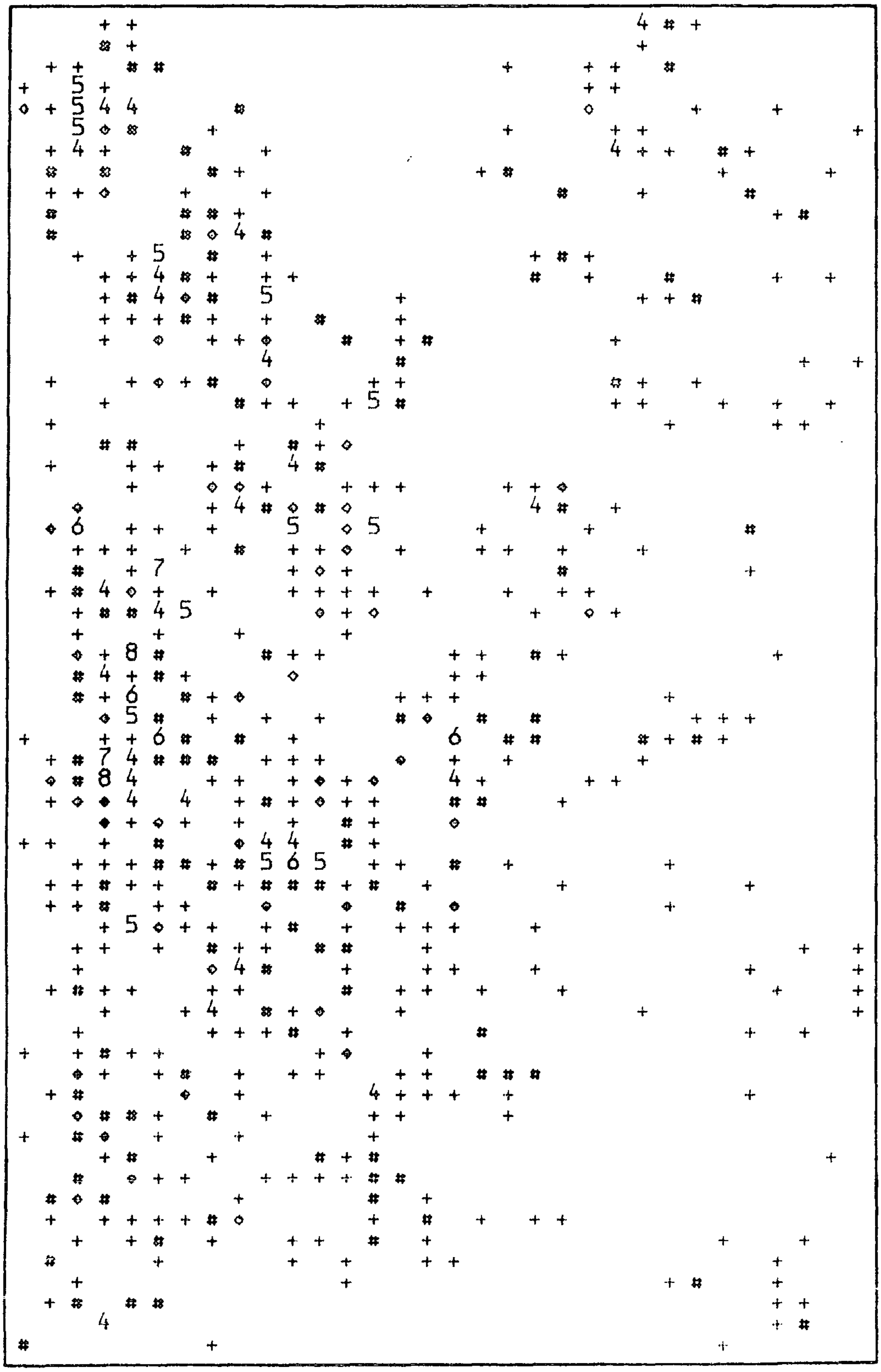


Figure 21. Map of the grid of contiguous quadrats for <u>Senecio viscosus</u> <u>L.</u> (The key to the symbols is given on page 79).

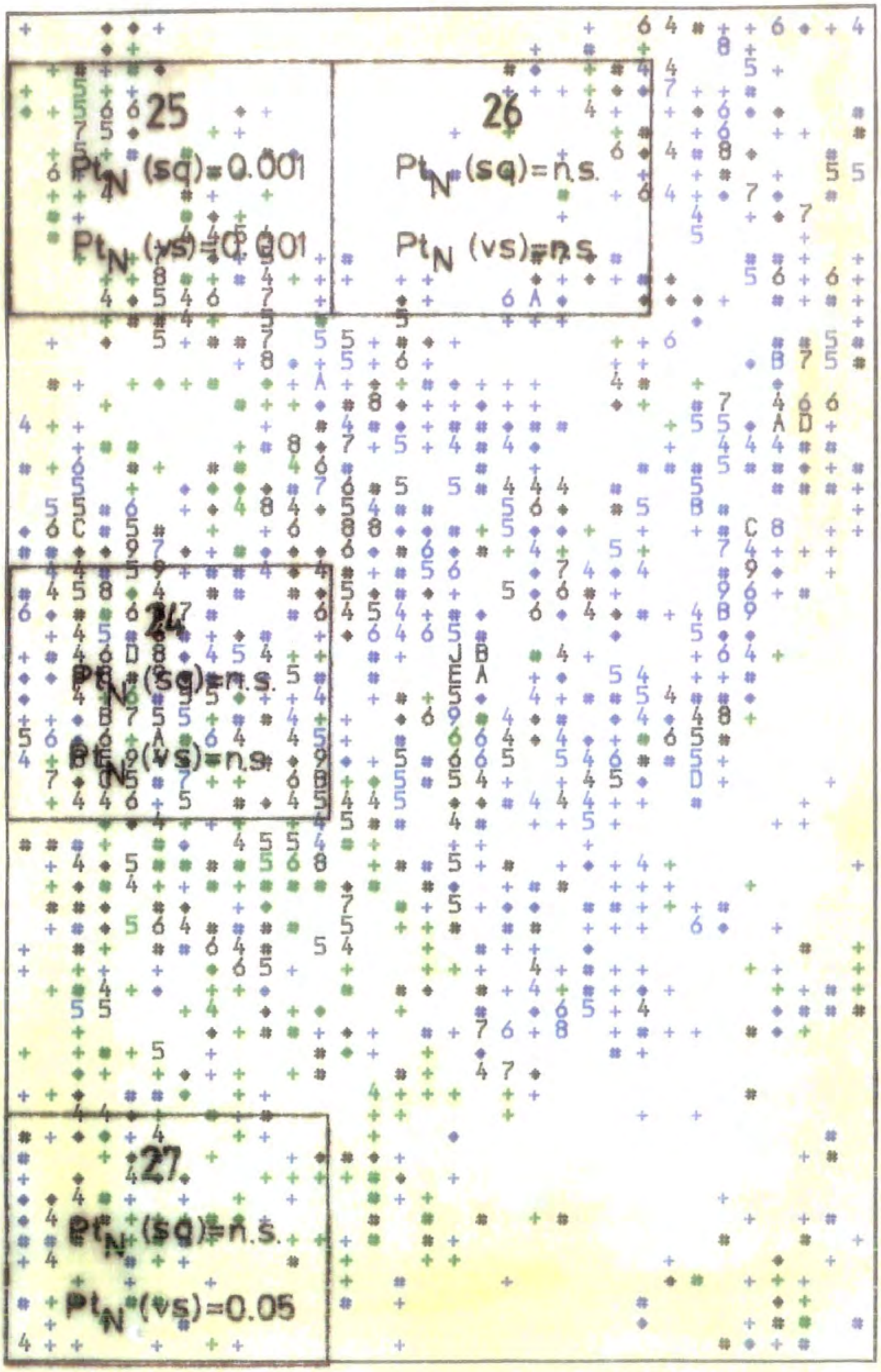


Figure 22. Map of the grid of contiguous quadrats for both species superimposed on one another. (The key to the symbols is given on page 79).

Table 8: Observed Quadrat Counts and Expected Poisson Series for S.squalidus and S.viscosus.

		0	1	2	3	4	5
S.viscosus	Observed	1465	346	134	103		
	Expected	1252	616	152	28		
S.squalidus	Observed	1133	383	218	119	73	122
	Expected	684	750	411	150	41	12

Sum of Chi-squared (Viscosus) = 357.605 P = < 0.001

Sum of Chi-squared (Squalidus) = 1604.669 P = < 0.001

The Observed and Expected Quadrat counts are plotted on Graphs 12 and 13 for S.squalidus and S.viscosus respectively. Using a 2 x 2 contingency table to test the null hypothesis that two species are independently dispersed (over the grid of quadrats),

	S.squalidus	
S.viscosus	+	-
+	357	226
-	558	907

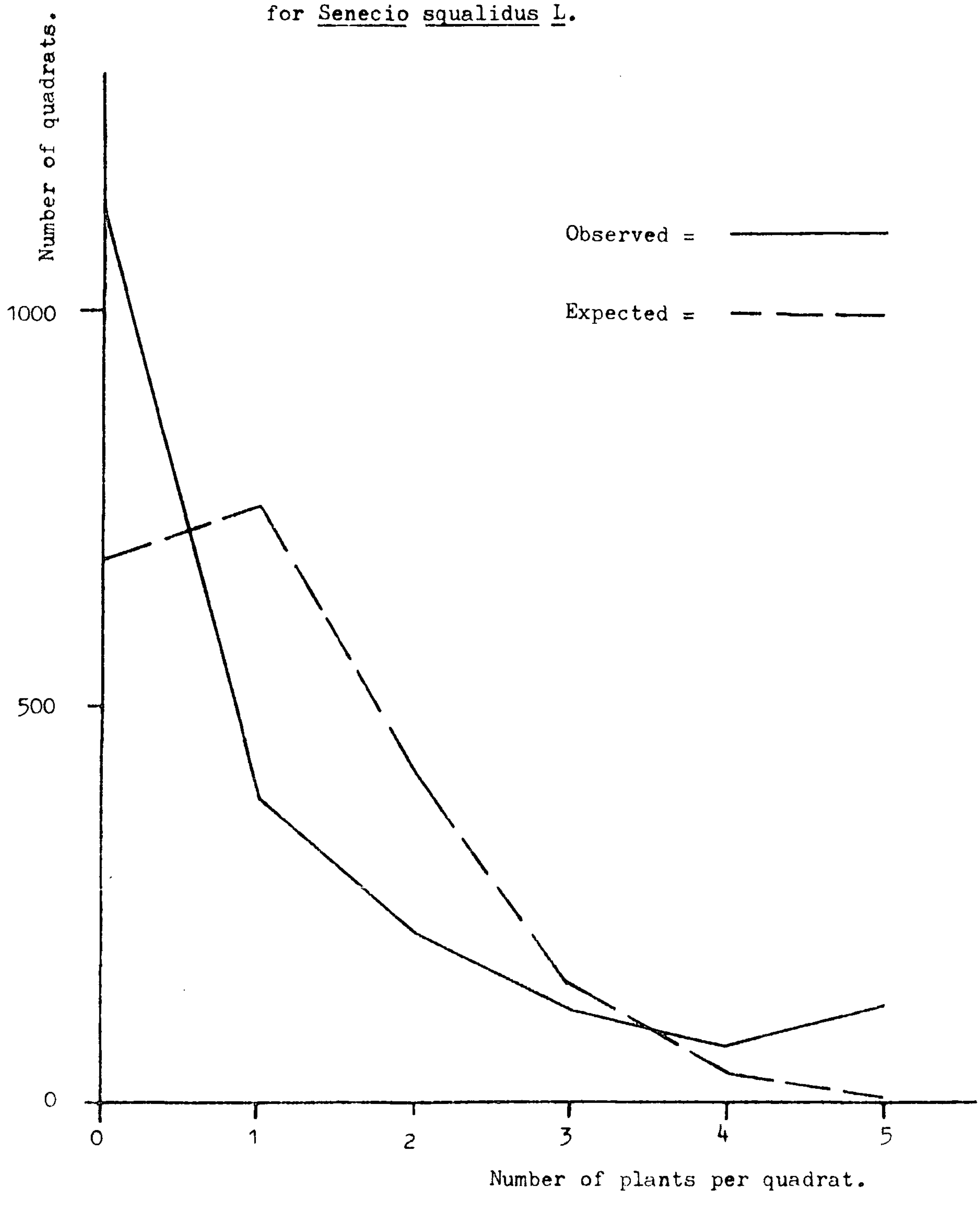
χ^2 = 90.3954 P=0.001

Graph Twelve.

Comparisson of Observed quadrat counts

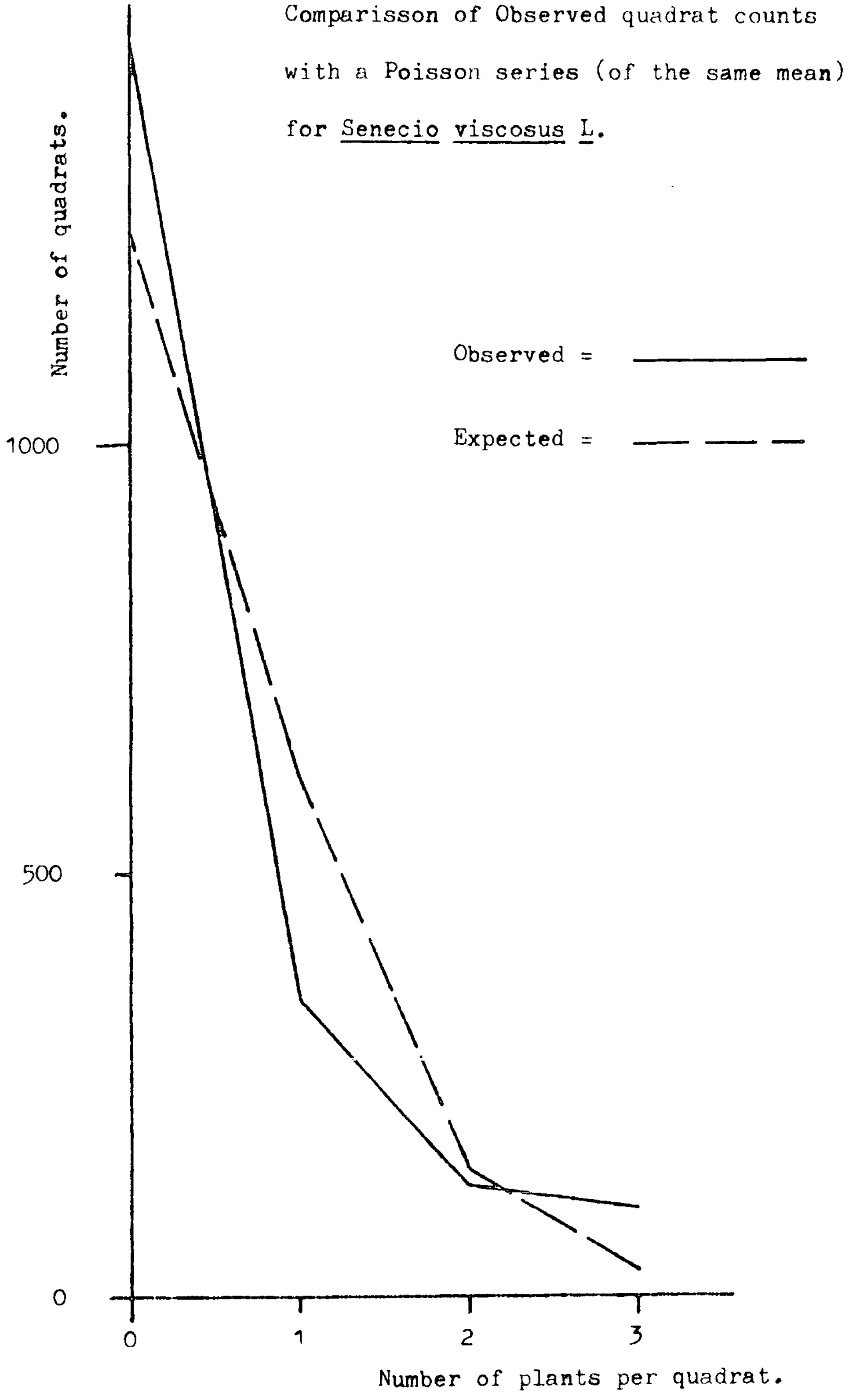
with a Poisson series (of the same mean)

for <u>Senecio</u> <u>squalidus</u> <u>L</u>.



Graph Thirteen.

Comparisson of Observed quadrat counts with a Poisson series (of the same mean) for Senecio viscosus L.



(numbers are numbers of quadrats in the 2048 quadrat grid).

The comparison of observed and expected quadrat counts will detect spatial pattern at one scale only, that is at the size of one quadrat or smaller. Monte Carlo simulations of the observed quadrat data were computed, and Meads test statistics calculated for the observed and simulated results. The results given on Table 9 are the observed Meads test statistic, as well as the top ten scores in the rank. This method was used, firstly in the detection of spatial pattern in S.squalidus and S.viscosus at different scales, and secondly in studying the interactions between the two species. (For the first method (Table 9) the data was divided in two blocks of 32 x 32 for simplicity of computation).

Table 9: Meads Analysis of Contiguous Quadrat Data.

Blocksize	Meads test statistic	Top ten values in Rank
S.squalidus rows 1 to 32.		
4 x 4	1070	1070 1040 1034 1016 1012 1010 1006 1002 1000 994
8 x 8	806	806 800 770 752 752 750 740 726 726 772
16 x 16	876	876 830 786 774 756 736 716 696 686 678
32 x 32	940	988 962 960 942 940 910 894 884 874 868
S.squalidus rows 33 to 64.		
4 x 4	1668	1668 1522 1516 1512 1500 1498 1484 1484 1480 1474
8 x 8	1142	1174 1172 1160 1150 1148 1142 1130 1126 1118 1112
16 x 16	1038	1038 992 938 928 924 922 914 904 888 870
32 x 32	1340	1424 1340 1306 1300 1242 1238 1234 1212 1168 1110

Table 9: Meads Analysis of Contiguous Quadrat Data.

Blocksize	Meads test statistic	Top ten values in Rank
S.viscosus rows 1 to 32.		
4 x 4	696	696 696 682 680 672 664 660 658 654 652
8 x 8	494	516 514 510 502 502 494 492 490 484 480
16 x 16	364	442 404 384 364 356 342 336 336 328 328
32 x 32	834	834 754 740 672 660 654 652 648 600 578
S.viscosus rows 33 to 64.		
4 x 4	792	792 746 736 728 720 716 712 710 708 706
8 x 8	502	560 554 538 522 518 512 512 510 506 506
16 x 16	470	470 446 442 436 428 418 416 412 406 400
32 x 32	566	608 568 566 552 540 528 522 518 512 496

Table 10: Meads interaction analysis for two species.

Blocksize	Meads test statistic	Top ten values in rank
2 x 2	1666	1720 1690 1666 1662 1654 1654 1651 1651 1649 1648
4 x 4	6329	6329 5932 5878 5820 5772 5756 5754 5725 5721 5719
8 x 8	21949	29149 21517 21379 21367 21316 21238 21209 21200 21142 21114
16 x 16	82022	82022 81291 81106 80583 80402 79840 79065 78967 78870 78771
32 x 32	299222	349895 336546 335532 334847 333782 331409 331409 326998 325906 325728

The mean Spearmans Rank Correlation Coefficient, rho, between four sub-blocks within blocks in the two data sets was calculated, and its value and standard deviate are given in Table 11.

Table 11: Interaction analysis for counts of two species using Spearmans Rank Correlation Coefficient.

Blocksize	Rho	Standard deviate	P
2 x 2	0.02	0.81	
4 x 4	0.25	4.87	0.000005
8 x 8	0.38	3.74	0.00001
16 x 16	0.42	2.08	0.019
32 x 32	0.20	0.49	

The significance of the standard deviate may be obtained from tables of the normal distribution.

7 DISCUSSION OF THE FIELD DATA AND SIMULATION RESULTS.

7.1 RESULTS OF THE T-SQUARE DATA.

The summary of the T-square sampling data in Table 4 shows that there is little difference between the proportion of non-random populations of S. squalidus (31%) and S. viscosus (46%). From Table 5 the proportion of populations in which non-randomness was detected was statistically indistinguishable, regardless of whether each species was isolated or in mixed stand with the other species. These results were dissappointing in that they failed to pick out clusters larger than the size of the sampling frame. This raises an interesting problem, the major objection too distance methods (in the analysis of spatial pattern) is that they deal with spatial pattern at one scale only, that is plant to plant distances, Diggle (1977) proposed the M/C test to deal with this problem, aiming to detect gradual changes in density within the sample, and so distinguish between random heterogeneous and random homogeneous patterns. This test too, where the null hypothesis of spatial randomness had been accepted, yielded little extra information, only at Site 32 (Table 2) did it indicate a randon heterogeneous distribution, suggesting that there was spatial non-randomness at a scale different to that being tested.

The analysis of variance of the t_N results yielded non-significant results for both species when comparing allopatric and sympatric populations (Table 6), this was because the amount of variability between sites within treatments was such that any variation between treatments was undetectable. This result is discussed more thoroughly on page 98.

7.2 RESULTS FROM THE NEAREST NEIGHBOUR FIELD DATA.

The results on Table 8 show that at all but one of the seven sampling frames studied the two species could not be shown to be segregated (although there was slight segregation, and never positive association). The sampling intensity was deliberately kept low (as with T-square sampling), and it may have been that the sampling intensity was too low to be representative, Pielou (1961) gives no indication of the intensity of sampling to be used. It is interesting, with reference to the maps[3], that those sampling frames at which S is highest (that is, that the degree of segregation is most intense), are also those with the lowest populations of plants. The implication from this is that when the plants are at low density there is a tendency to have clusters of either S.viscosus or S.squalidus, but that this effect is masked when the plants occur at high densities. This also brings to the fore a problem with the sampling, since there was a tendency to sample areas containing large numbers of either or both species, which may have given misleading results with this kind of data. The reason for this may be that those areas of high plant densities may be optimum habitats for Senecio species, that is both species would grow there preferentially. If this were the case spatial pattern, and in particular segregation, in the populations would be masked by the high population densities. Segregation in the population may only be apparent at low population densities, and this appears to have been the case from the field data.

---

3. Figures 20, 21 and 22.

7.3 RESULTS FROM THE CONTIGUOUS QUADRAT GRID

1. From the raw quadrat counts two conclusions may be drawn, firstly, that the observed quadrat counts deviate significantly, for both species, from the expected quadrat counts (according to a Poisson series). Secondly, that the null hypothesis that the two species are independently dispersed is rejected, the two species being strongly positively associated. Therefore, at the single quadrat level, both species are exhibiting aggregated spatial patterns, and they are associated together over the grid (it should be pointed out that this refers only to presence and absence of either or both species, and not their numerical scores).

2. The results of Meads Analysis, on Tables 8 and 9, indicate that there is independent spatial pattern at most block sizes in each of the four 32 x 32 blocks of quadrats. The only exception to this is in the lower <u>S.squalidus</u> block (rows 33 to 64), and in both <u>S.viscosus</u> blocks, the observed statistic at the 8 x 8 block size is not significant, the rank of the observed statistic is greater than 5. (This is adopting rather a purist interpretation of the data, accepting the null hypothesis at any probability of greater than 0.05). At the 8 x 8 block size spatial pattern is not as intense as could be produced 5 times in a hundred by randomising the observed data. The conclusion from this is that, in both species there are aggregated distributions at the smallest and largest block sizes, and that in the mid range, pattern is less intense, clusters of plants being randomly dispersed. The impression being, for both species, of small clusters, aggregated

into very large groups of clusters, but at block sizes larger than the clusters themselves and smaller than the groups of clusters, clusters are randomly distributed.

3. The Meads analysis of interaction between the two species, (Table 10), suggests that there is strong positive association between both species at every block size tested, except for the largest blocksize. Study of the maps, particularly Figure 22, shows that in the central portion of the grid both species regularly occur within the same quadrat, but there are areas of the grid containing only one or the other species. The mean Spearmans Rank Correlation coefficient (Table 11) indicates a slightly different aspect of the data, in that it demonstrates that there is no association between the two species at the smallest scale tested (2 x 2), but that, as the block size is increased, so does the degree of association between the two species. From the rank correlation coefficient, it can be concluded that, at the smallest block sizes, the two species are independently mingled, but as block size increases so the degree of spatial association increases.

7.4 SUMMARY OF THE SIMULATION RESULTS.

The final version of SPEV in which two species of differing breeding system were modelled posed several questions:

Firstly, in the simulations in which the self pollinating species had a pollen advantage over the cross pollinating species, the final population consisted, invariably, of only one species, the self polliantor. The cross pollinating species was more often than not eliminated from the population. This conclusion could be slowed down by reducing the seed dispersal distance, such that the population would form tight clusters of one or other species, ensuring that cross pollinating plants, in tight clusters would recieve sufficient pollen from their neighbours to counter the advantage of the self pollinator, or by increasing RAIN to such an extent as to reduce the advantage of self pollination.

How well does this fit with the observed field data?. In the field study the self pollinating species was S. viscosus, and the cross pollinating species, S. squalidus. From other sources, it is clear that S. squalidus is an expanding plant in the British Isles, moving into those habitats formerly occupied by S. viscosus. Two of the reasons for this encroachment by S. squalidus are 1) its over wintering habit & 2) the greater quantities of pollen it produces. S. viscosus is an annual, which flowers in mid July to August, by which time S. squalidus has already been in flower for two months, flowering in early May, and remaining in flower till mid September. In the field situations studied, it was S. viscosus, the self pollinator, that was being lost because of the more robust nature of S. squalidus. Clearly the field data do not wholly match the predictions of the model, firstly because the biology of the two species differs from that of the two 'model' species. This is particularly true of S. squalidus where the model takes no account of over wintering. Secondly (as is discussed later) because of the unpredictability of the enviroment which in the model was homogeneous. However, in those simulations in which the cross pollinating species produces a surfeit of pollen, the model predicts that which has been observed in the field, in that the cross pollinator eliminates the self pollinating species by 'swamping' any advantage that the self pollinator previously had.

Secondly, in those simulations in which neither species has a self pollinating advantage, the maintainence of both species in the environment was often the result of the simulation. This depended on the ability of individuals of each species to form clusters of one species or the other, (in order to recieve sufficient pollen from plants of the same species to avoid loss of offspring because of reduced hybrid fertility). Those plants associated with plants of the other species would tend to produce hybrid offspring that were subsequently lost. This fits with the intuitive notion that where there is strong selection against hybridisation, plants will aggregate in such a way as to avoid loss of reproductive output through hybridisation. (It is worth pointing out once again that in a model of hybrid infertility, the only stable population is one in which only one of the hybridising species remains). A population containing both hybridising species will always remaih unstable, even though both species may persist in the population for many generations.

From the field data four conclusions may be drawn;

1. The T-square data showed that there were no detectable differences in the intensity of spatial pattern shown by allopatric and sympatric populations of either species.

2. Comparing the raw quadrat data with a Poisson series showed that there was significant pattern in both species over the 32 x 64 grid at the one quadrat blocksize. Also at this level the two species were strongly associated.

3. The use of Meads analysis to study spatial pattern at more than one scale showed significant spatial pattern at large and small blocksizes for both species, but also that in the mid range the intensity of spatial pattern in the population was not so great.

4. The rank correlatinn coefficient indicated that, at the smallest scale of spatial pattern studied, 2 x 2 blocks, both species were independently dispersed, but that at every other blocksize there was significant association between the two species.

The apparent clash between (2) and (4), the chi-squared showing significant association of the two species at the one quadrat level, and the Spearmans rank correlation coefficient suggesting that the two species are not associated at the same scale, may be explained because the chi-squared deals solely with presence or absence of the two species, whereas the Spearmans rank correlation deals with more information, using the numerical scores of the two species in a given quadrat. This may be interpreted as indicating that where S. squalidus is found, S. viscosus is likely to be found in the same quadrat. However the numerical scores of the two species within that quadrat will not coincide, (where large numbers of one species are found, low numbers of the other are likely to occur in that quadrat).

The results of the field work for this study were dissappointing since little supported the original hypothesis directly. Only one population was demonstrated to be segregated (Table 7), and there was some supportive evidence evidence for segregation at small scales from the tests of association using Spearmans rank correlation coefficient. The analysis of variance of the T-square data (Table 6) though not supporting the hypothesis directly does indicate why the two species were not behaving as predicted. Between site variation within each treatment is so marked that between treatment variation is negligable. It is therefore not possible to predict the intensity of spatial pattern at a site either by species or by treatment. Both S. squalidus and S. viscusus are plants of derelict land, which is inherently disturbed and heterogeneous. It follows that the evolution of segregated spatial patterns may be prevented a) by the frequency of disturbance at a given site, and b) because of the patchy distribution of germination sites, consequently the distribution of seedlings and adult plants will also be patchy without respect to species, increasing the probability of finding the two species together and reducing the detectability of segregation. If these arguements hold then the hypothesis should be tested in hybridising populations with; a) Low disturbance frequency, and b) Even distribution of germination sites. The hypothesis could be tested experimentally by establishing a tilled, weeded, homogeneous site, sown with seed of both species, in which stable populations were allowed to develop, whilst maintaining the site weed free. The intensity of spatial pattern could be tested annually testing the hypothesis, that gene flow interacts with spatial pattern to cause segregation in hybridising species, directly. Alternatively seed could be collected from a mapped natural population of both species, and subsequently germinated, recording germination and hybrid frequencies in seed gathered from specific plants, attempting to obtain a correlation between these and the proximity of plants of the other hybrid parent. This would test the hypothesis indirectly by supporting the major assumption of SPEV that fecundity is inversely related to the proximity of plants of the other hybrid parent.

8. CONCLUSIONS.

The computer simulations using SPEV lead to the conclusion that where two hybridising species were of equal breeding abilities, and also where their hybrids are inviable, the two species will tend to segregate into clusters containing plants of only one species or the other, and co-existing in an apparently stable population. This segregation is caused by the reduction of fecundity of plants of one species that are spatially close to plants of the other. (If one species has a self pollinating advantage then that species will eliminate the other from the model population).

From the field study the hypothesis (generated from the computer simulations), that gene flow interacts with spatial pattern to cause segregation in hybridising populations, was rejected because of the degree of between site variation in intensity of spatial pattern in allopatric and sympatric populations of both species. It was found that it was not possible to predict, either by species or by treatment, the intensity of spatial pattern at a given site, and it was concluded that segregated spatial patterns in sympatric populations may have insufficient time to evolve because of the unpredictability of the environment. Despite this the logic of the original hypothesis seems correct and experimental methods of testing it are proposed.

9 ACKNOWLEDGEMENTS.

I am very grateful to all who have helped, and encouraged me throughout this MSc. course, especially my freinds on the course who made the awkward moments less so. I am greatly indebted to the staff of the Durham University computer unit for their help and guidance, to Peter Green for the use of TILE, and to Julian Besag for his patient assistance with Mead's analysis. My work on this course was funded by a grant from the Natural Environment Research Council. Most special thanks and congratulations are due to my mother who typed the manuscript, my father who checked it, and finally to Tim Gleaves for many hours of discussion without which this project would not have happened.

28th October 1981. M.T.

10 REFERENCES.

Besag, J.E. (1977). Some methods of statistical analysis for spatial data. Bull. Int. Stat. Inst., 47, 77-92.

Besag, J.E., & Diggle, P.J. (1977). Simple Monte Carlo tests for spatial pattern. J. Roy. Stat. Soc. Ser. C., 26, 327-333.

Besag, J.E., & Gleaves, J.T. (1973). On the detection of spatial pattern in plant communities. Bull. Int. Stat. Inst., 45, 153-158.

Clapham, A.R., Tutin, T.G. & Warburg, E.F. (1959). Excursion Flora of the British Isles. Cambridge University Press.

Cormack, R.M. (1979). Spatial aspects of Competition. Statistical Ecology Vol 8: Spatial and Temporal Analysis in Ecology (Ed. by R.M. Cormack & J.K. Ord), pp. 151-211. International Co-operative Publishing House, Fairland, USA.

Diggle, P.J. (1977). The detection of random heterogeneity in plant populations. Biometrics, 33, 390-394.

Diggle, P.J., Besag, J.E. & Gleaves, J.T. (1976). Statistical analysis of spatial point patterns by means of distance methods. Biometrics, 32, 659-667.

Gibbs, P.E., Milne, C. & Vargas Carillo, M. (1975). Correlation between the breeding system and recombination index in 5 species of Senecio. New Phytol., 75, 619-626.

Gleaves, J.T. (1973). Gene flow mediated by wind borne pollen. Heredity, 31, 355-365.

Green, P.J. & Sibson, R. (1977). Computing Dirichlet Tesselations in the plane. The Computer Journal, 21, 168-173.

Greig-Smith, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. Ann. Bot. N.S., 16, 293-316.

Greig-Smith, P. (1964). Quantitative Plant Ecology, 2nd edn. Butterworths, London.

Hope, A.C.A. (1968). A simplified Monte Carlo significance test proceedure. J. Roy. Stat. Soc. Ser. B., 30, 582-598.

Hopkins, B. (1954). A new method for determining the type of distribution of plant individuals. Ann. Bot. N.S., 18, 213-227.

Kershaw, K.A. (1957). The use of cover and frequency in the detection of pattern in plant communities. Ecology, 38, 291-299.

Krebs, C.J. (1978). Ecology: The Experimental Analysis of Distribution and Abundance, 2nd edn. Harper & Row, New York.

Lousley, J.E. (1943). A new hybrid from the London area. Bot. Exch. Rep., 12, 869-874.

Lousley, J.E. (1947). A new hybrid Senecio (S. squalidus L. x S. viscosus L.) from the London area. Proc. Linn. Soc. Lond., 158, 21-22.

Mead, R. (1974). A test for spatial pattern at several scales using data from a grid of contiguous quadrats. Biometrics, 30, 295-307.

Pielou, E.C. (1961). Segregation and symmetry in two species populations as studied by nearest neighbour relations. J. Ecol., 49, 255-269.

Poole, R.W. (1974). An Introduction to Quantitative Ecology. McGraw-Hill Kogakusha Ltd, Tokyo.

Zahl, S. (1974). Application of the S-method to the analysis of spatial pattern. Biometrics, 30, 513-524.

APPENDIX A - SENECIO JACOBAEA T-SQUARE DATA.

T-square sampling data was collected for S. jacobaea at 5 sites, and this is included here for completeness.

Table 1: Senecio jacobaea T-square sampling data

Sampling frame	N	t_B	Pt_B	t_N	Pt_N
7	10	0.6418		0.6731	0.028
8	15	0.3844		0.4060	
9	19	0.5685		0.6432	0.015
17	60	0.5888	0.025	0.5574	0.034
18	22	0.5269		0.5753	

APPENDIX D - FIELD STUDY SITE GRID REFERENCES.

Grid references are taken from the Ordinance Survey 1: 50,000 scale, Second Series, sheets number 88 and 93.

1	Gilesgate Builders Yard, Durham	NZ 282 429 (sheet 88)
2	Stockton and Darlington Old Railway	NZ 315 153 (sheet 93)
3	Abandoned Field, Pittington	NZ 333 445 (sheet 88)
4	Old Quarry Entrance, Coxhoe	NZ 319 363 (sheet 93)
5	Old Field, Quarrington Hill	NZ 333 367 (sheet 93)
6	Belmont Viaduct Field, Belmont	NZ 290 452 (sheet 88)
7	Derelict Land, Darlington	NZ 273 164 (sheet 93)
8	Disused sidings, Croft, Darlington	NZ 294 132 (sheet 93)