STUDIA GEOBOTANICA
5: 143-165, 1985

THE USE OF ELLIPSES OF EQUAL CONCENTRATION TO ANALYSE ORDINATION VEGETATION PATTERNS

Mario LAGONEGRO and Enrico FEOLI

Keywords: Ellipses, Forest, Ordination, Pattern, Structure.

Abstract: The use of ellipses of equal concentration is suggested for the analysis of pattern in ordination scattergrams produced directly by environmental variables, or with axes obtained by numerical methods. The potentials of the method are illustrated by some examples using structural characteristics of mixed forest type of NE Italy and environmental variables estimated by ecological indicator values.

Introduction

The analysis of vegetation patterns detected by ordination scattergrams can be done by the method suggested by Feoli & Ganis (1986). It consists in the use of a revised version of the Moran formula (Cliff & Ord, 1981) of autocorrelation for testing if the values of a variable, simple or composite, have a non-random pattern in the spaces defined by other ecological variables. The method does not give any information about the barycentre of the variable or about the correlation between the variables which define the space. We think that the method of ellipses of equal concentration (see Daboni, 1967; Mardia, Kent & Bibby, 1979) should constitute a further extension of the autocorrelation method in describing vegetation pattern related to ordination scattergrams.

The method

Given N points with coordinates (x_i, y_i) for i = 1, ..., N, with which are associated N scores (p_i) of a variable i, the coordinates of the barycentre of the variable i in the space defined by the axes x and y are:

$$X_G = \frac{1}{T} \sum_{1}^{N}{}_i \, p_i x_i \qquad Y_G = \frac{1}{T} \sum_{1}^{N}{}_i \, p_i y_i$$

where $T = \sum_{1}^{N}{}_i \, p_i$. The parameters of dispersion of x and y around the barycentre are:

$$s_x^2 = \sum_{1}^{N}{}_i \left(\frac{p_i}{T}\right)(x_i - X_G)^2$$

$$s_y^2 = \sum_{1}^{N}{}_i \left(\frac{p_i}{T}\right)(y_i - Y_G)^2$$

$$s_{xy} = \sum_{1}^{N}{}_i \left(\frac{p_i}{T}\right)(x_i - X_G)(y_i - Y_G)$$

$$R = \frac{s_{xy}}{s_x \cdot s_y}$$

R is the correlation coefficient. One can define the ellipse of equal concentration k as:

$$\frac{1}{1 - R^2}\left[\left(\frac{x - X_G}{s_x}\right)^2 - 2R\frac{(x - X_G)(y - Y_G)}{s_x \cdot s_y} + \left(\frac{y - Y_G}{s_y}\right)^2\right] = k$$

$$|R| \neq 1$$

which is centered at (X_G, Y_G) with axes and inclination defined by s_x^2, s_y^2, R. It corresponds to the curve which, on the (x, y) plane, given a binormal distribution having the maximum at the barycentre, leaves outside itself a residual equal to exp (—k) (tail probability). It must be remembered that the total amounts to unity. It follows that if one wants an ellipse excluding 5% of the total binormal distribution, K must be estimated as:

$$e^{-k} = \frac{5}{100} = 0.05$$

$$k = -\ln 0.05$$

In general if a percentage P is to be excluded:

$$k_P = -\ln\left(\frac{P}{100}\right)$$

The corresponding ellipse will satisfy the following relation:

$$\left(\frac{x - X_G}{s_x}\right)^2 - 2R\left(\frac{x - X_G}{s_x} \cdot \frac{y - Y_G}{s_y}\right) + \left(\frac{y - Y_G}{s_y}\right)^2 = k_P(1 - R^2)$$

If

$$X_S = \frac{x - X_G}{s_x} \qquad Y_S = \frac{y - Y_G}{s_y}$$

then:

$$X_S^2 - 2\,R\,X_S\,Y_S + Y_S^2 - k_p\,(1 - R^2) = 0$$

by reordering and taking X_S as an independent variable

$$Y_S^2 - (2\,R\,X_S)\,Y_S + [\,(X_S^2 - k_p\,(1 - R^2)\,] = 0$$

$$a\,Y_S^2 + b\,Y_S + c = 0$$

where:

$$a = 1 \qquad b = -\,2\,R\,X_S \qquad c = X_S^2 - k_p\,(1 - R^2)$$

the solutions are real if:

$$b^2 - 4\,a\,c \geq 0$$

it follows that values of X_S may be derived from:

$$(X_S^2 - k_p)\,(1 - R^2) = 0\ ,$$

since $-\,1 < R < 1$ and consequently $(1 - R^2) \neq 0$

it follows that:

$$-\sqrt{k_p} \leq X_S \leq \sqrt{k_p}$$

an in consequence

$$X_G - s_x\sqrt{k_p} \leq x \leq X_G + s_x\sqrt{k_p}$$

The parameters computed in this way are used to compute the y values and to draw the ellipses.

When more than one group of objects is considered, or when a group is considered under different weighting approaches, the ellipses can be computed for each group. The differences between the groups may be tested by Student's estimated as follows:

$$t = \frac{|D|}{S_{12}} \left[\frac{M N}{M + N} \right]^{1/2}$$

where:

$$D = \left[(X_{G1} - X_{G2})^2 + (Y_{G1} - Y_{G2})^2 \right]^{1/2}$$

$$S_{12} = \left[\frac{(N-1)\, s_1^2 + (M-1)\, s_2^2}{M + N - 2} \right]^{1/2}$$

$$s_1^2 = \left(\frac{s_x^2 + s_y^2}{2} \right)_1 \qquad s_2^2 = \left(\frac{s_x^2 + s_y^2}{2} \right)_2$$

N being the number of objects in group 1, M that of group 2.

The program also computes the regression lines for each group and tests for differences between the population slopes and intercepts, two by two, (Zar, 1974). The formulae are:
for slopes

$$t = \frac{|a_1 - a_2|}{S a_1 a_2}$$

$$S a_1 a_2 = \left[\frac{S^2_{xy}}{\left(\sum_1^N {}_i\, x_i^2 \right)} + \frac{S^2_{xy}}{\left(\sum_1^M {}_i\, x_i^2 \right)} \right]^{1/2}$$

$$S^2_{xy} = \frac{\left[\sum_1^N {}_i y_i - \left(\sum_1^N {}_i\, x_i y_i \right) / \sum_1^N {}_i\, x_i^2 \right] + \left[\sum_1^M {}_i y_i - \left(\sum_1^M {}_i\, x_i y_i \right) / \sum_1^M {}_i\, x_i^2 \right]}{N + M - 4}$$

for elevations

$$t = \frac{(b_1 - b_2) - a_c (X_1 - X_2)}{\left\{ S^2_{xy} \left[\frac{1}{N} + \frac{1}{M} + \frac{(X_1 - X_2)^2}{\left(\sum_1^N {}_i\, x_i^2 \right) + \left(\sum_1^M {}_i\, x_i^2 \right)} \right] \right\}^{1/2}}$$

$$a_c = \frac{\left(\sum_{1}^{N}{}_i\, x_i y_i\right) + \left(\sum_{1}^{M}{}_i\, x_i y_i\right)}{\left(\sum_{1}^{N}{}_i\, x_i^2\right) + \left(\sum_{1}^{M}{}_i\, x_i^2\right)}$$

The program ELLIPLOT (1), written both in GW-BASIC and FORTRAN 77 for personal computers operating under MS-DOS, is listed in the Appendix.

Examples

The method has been applied to data of mixed forest types obtained by Lagonegro and Feoli (1985) from Poldini (1982). They consist of two matrices describing 20 vegetation forest types. In one matrix the description of the types is based on the combination of life forms and growth forms (life-growth forms); in the other it is based on the average indicator values, Landolt (1977). The method of ranking characters proposed by Orlóci (1973) has been applied to the matrices, in order to select variables for the examples.

In the first place, the environmental variables as indicated by the average indicator values of the species were used as axes.

Among the environmental variables, humidity, temperature and pH have been chosen because they explain more than 90% of the total sum of squares. These variables, two at a time, served as axes to test for the pattern of all the life-growth forms. Among those with significant patterns (according to the autocorrelation method), we used those accounting for more than 90% of total sum of squares. They are: scapose hemicryptophytes, rhizome geophytes and suffrutescent chamaephytes.

In Figure 1 (a, b) is given an example of output of the program ELLIPLOT applied to the ordination scattergram with humidity and pH as axes, using the values of suffrutescent chamaephytes. Figure 1a shows the barycentre of the life-growth form in the space defined by humidity and pH, and the ellipses corresponding to 0.1, 1, 5, 31.7 percent tail probability. The regression line of the two variables is also printed.

Figure 1b shows the centroid of the two variables. The program also provides for the two figures to be drawn both on a graph as in Figure 2. In this case only two ellipses are drawn (those at 5% probability) to avoid excessive complexity.

The deviation of the centroid from the barycentre may be used as a measure of asymmetry of the pattern of the life-growth form with respect to the two variables used as ordination axes. The t test, estimated as described above, can serve as such a measure. In the case of Figure 2 the deviation of the barycentre from the centroid

(1) The program is included in a library (47 programs) written in GW-BASIC for personal computers running under MS-DOS operating system (Lagonegro & Feoli, 1985).

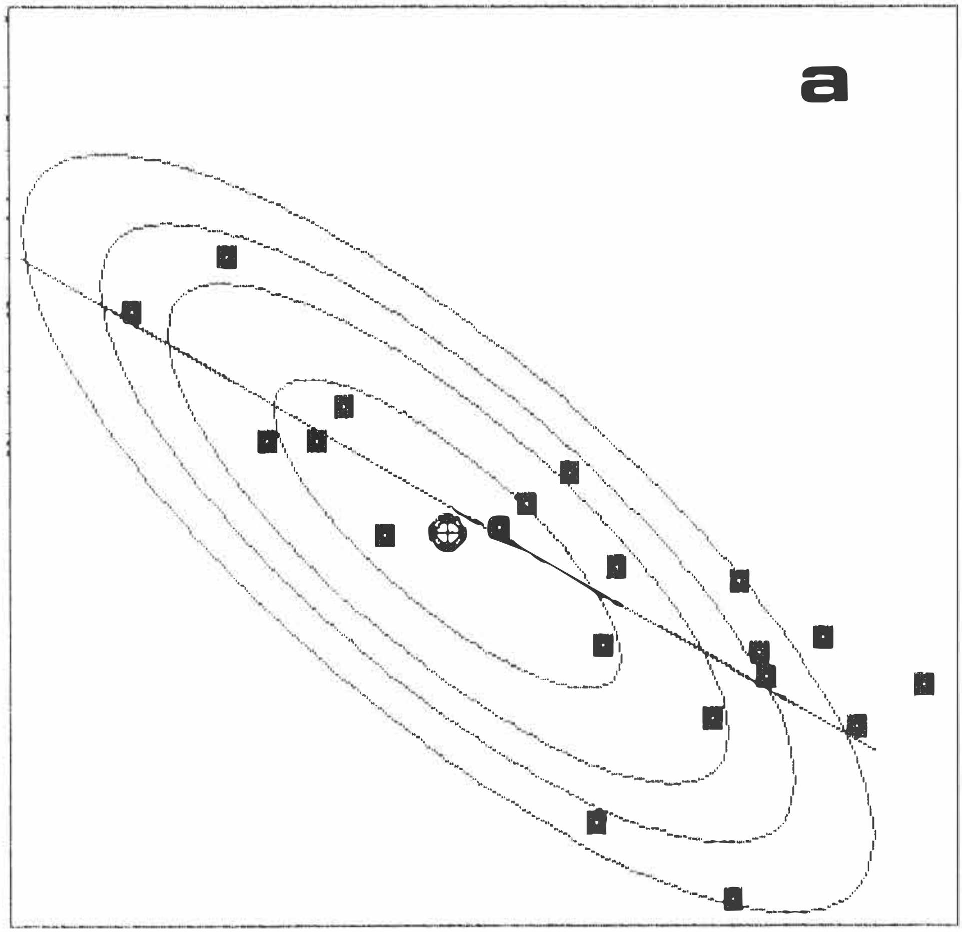

Fig. 1(a) — Ellipses of equal concentration drawn at probability levels of 0.317, 0.05, 0.01 and 0.001 with reference to the barycentre of suffrutescent chamaephytes.

is significant $t - p < 2\%$. This proves that the suffrutescent chamaephytes have a pattern that is asymmetric with respect to the centroid of the two variables. This does not happen for the other two life-growth forms.

The ellipses are useful in quantifying the pattern of a variable in ordination scattergrams. The pattern may be described by the cumulative frequencies of the points included in the ellipses corresponding to different probability levels, as in Table 1. The deviations from the expected frequencies, which can be calculated by multiplying the expected cumulative probability values by the total number of objects in the scattergram, may be used in a chi square test. From Table 1 one can conclude that the pattern deviating most from that expected is that of suffrutescent

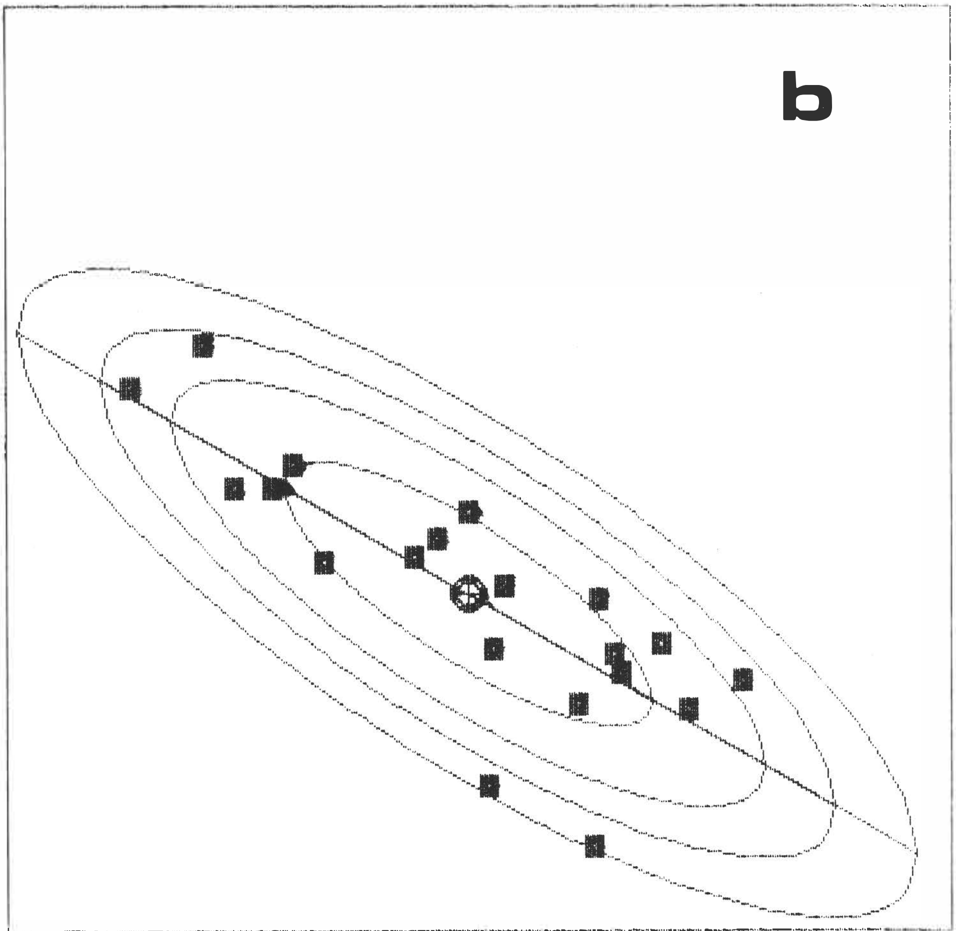

Fig. 1 (b) — Ellipses of equal concentration drawn at probability levels of 0.317, 0.05, 0.01 and 0.001 with reference to the centroid of the two variables defining the space: humidity and pH.

chamaephytes, followed by that of scapose hemicryptophytes. The pattern for rhizome geophytes does not deviate significantly from that expected. If one considers the two main groups of communities defined by the dendrogram of Figure 7 in Lagonegro & Feoli (1985), then one can see from Figure 3 that for cluster a there is no deviation from expectation, while for cluster b there is a significant deviation. This means that the cluster a represents a homogeneous environment for the suffrutescent chamaephytes as far as humidity and pH are concerned, while cluster b represents a heterogeneous environment. The regression lines corresponding to the two clusters suggest that they constitute two parallel series of vegetation with respect to humidity and pH. The distance between the barycentres of suffrutescent

(a)					
probability	0.317	0.05	0.01	0.001	chi square
expected	13.6	19	19.8	19.98	—
suffrutescent chamaephytes	4	10	15	17	12.70*
scapose hemicryptophytes	6	13	14	17	8.28*
rhizome geophyte	7	13	15	17	6.70
(b)					
probability	0.317	0.05	0.01	0.001	chi square
expected	13.6	19	19.8	19.98	—
suffrutescent chamaephytes	4	10	14	17	13.18*
scapose hemicryptophytes	6	14	14	17	7.70
rhizome geophytes	7	14	15	17	6.1

Table 1 — Cumulative frequency distributions of the 20 vegetation types, as shown by the pattern of their inclusion in four ellipses of equal concentration, respectively at probability levels of 0.317, 0.05, 0.01, 0.001. The ellipses are drawn with reference to the barycentre of the life-growth forms. The axes in (a) are humidity and pH and in (b) the first and second principal components of the matrix of seven environmental variables. The asterisk indicates that the frequency distribution deviates significantly from expectation (see the text).

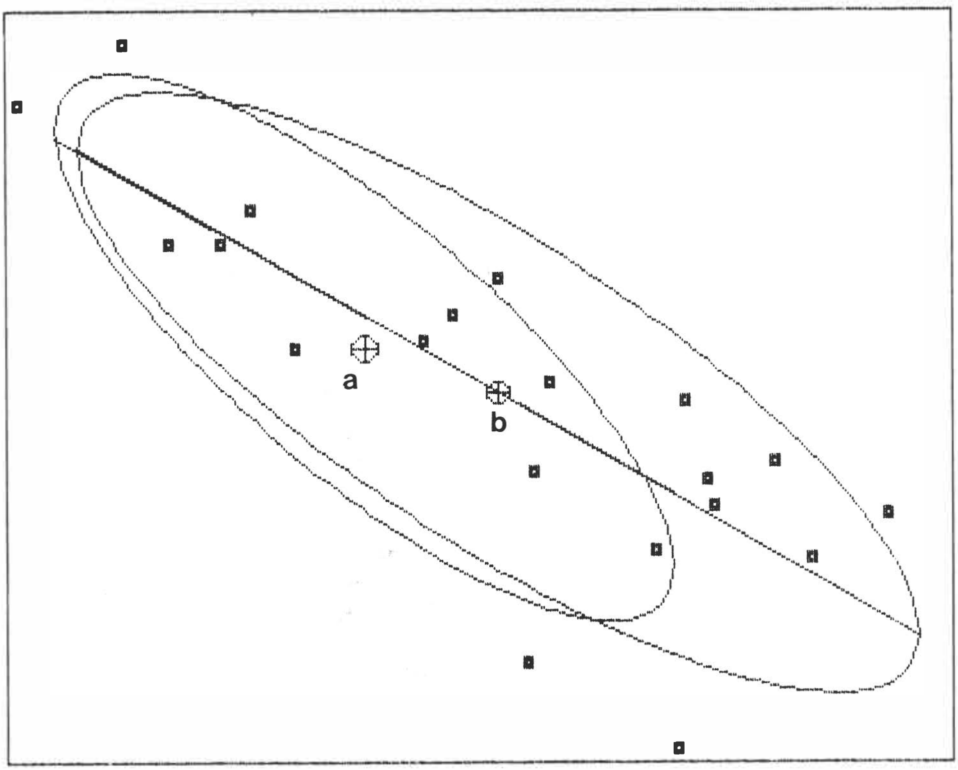

Fig. 2 — Ellipses of equal concentration at 0.05 probability drawn around the centroid (b) of the variables (humidity and pH) and around the barycentre (a) of suffrutescent chamaephytes. The distance between the barycentre and the centroid is significant *($p < 0.02$)*.

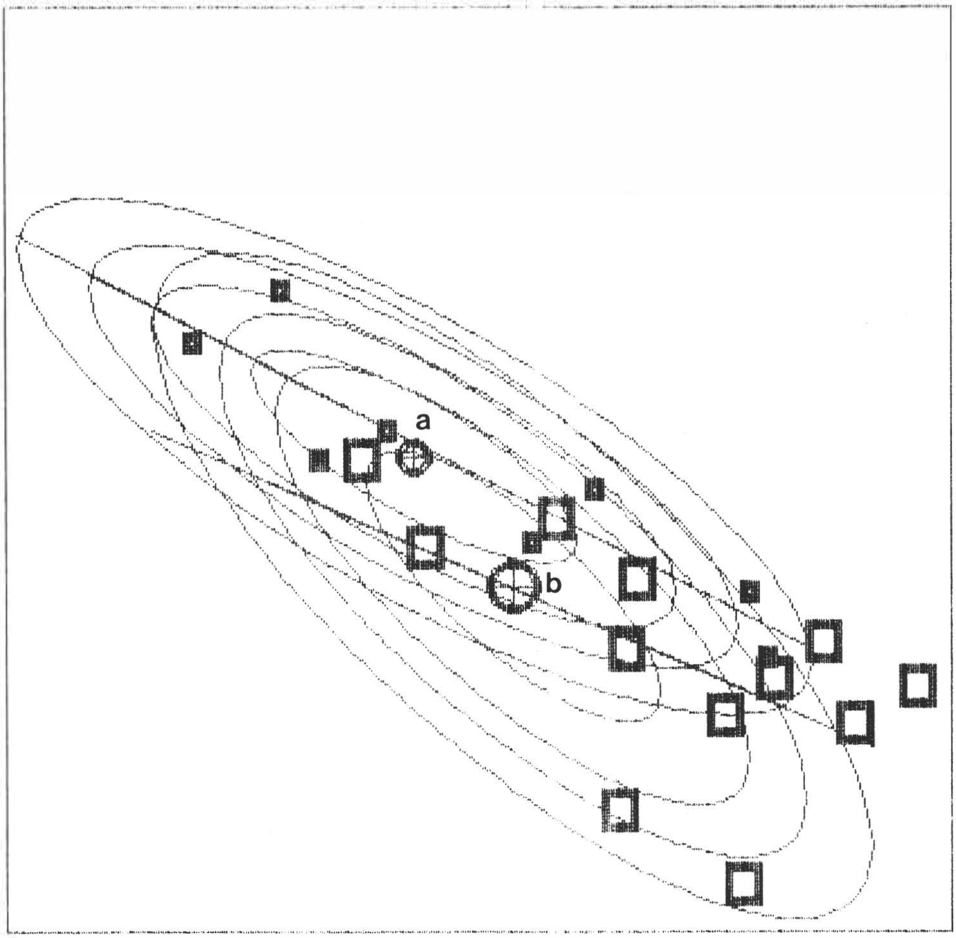

Fig. 3 — Ellipses of equal concentration at 0.317, 0.05, 0.01 and 0.001 probability around the barycentres of suffrutescent chamaephytes, computed for two clusters of forest types. Only the cumulative frequency distribution of the types of cluster a) fits expectation. The distance between the two barycentres is not significant.

chamaephytes is not significant; however the distance between the centroids of the two clusters is significant (Figure 4). The significant shift of the barycentre of cluster b from the centroid (Figure 5) proves further that the environment defined by such cluster is not homogeneous for suffrutescent chamaephytes. Figure 6 shows the ellipses of equal concentration for the cluster a and b at 5% probability in the space defined by temperature and humidity. The barycentres calculated for suffrutescent chamaephytes and the centroids of the clusters are indicated. The deviation of the barycentre from the centroid is significant only for cluster b, while the distance between the centroids and the distance between the barycentres of the

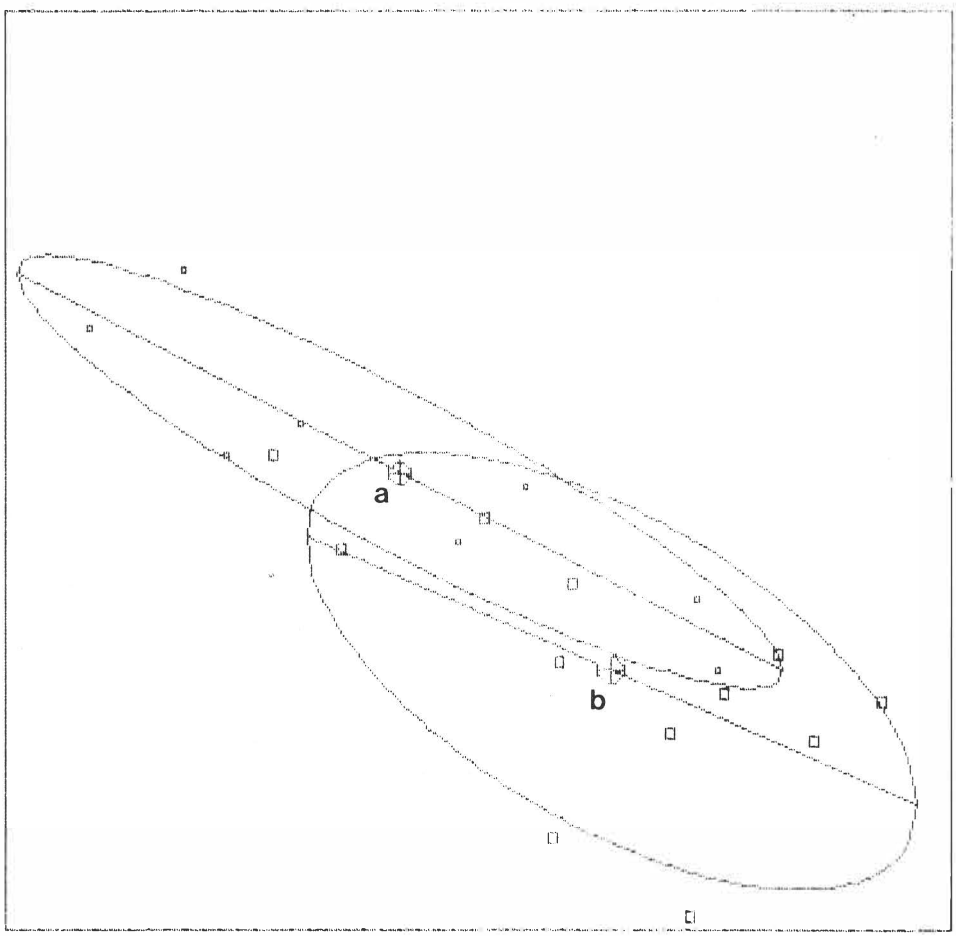

Fig. 4 — Ellipses of equal concentration, and regression lines, for two clusters of vegetation types a) and b) at the 5% livel. The centroids of the clusters are drawn; the distance between them is significant *($p < 0.02$)*.

two clusters are significant. The relationship between temperature and humidity is linear for cluster a, while it is non-linear for cluster b.

Instead of the individual environmental variables, it seemed useful to use the axes of a principal components analysis as axes. This is illustrated by Figure 7 and Figure 8. In Figure 7 the centroids of cluster a and cluster b are significantly separated. Also in this figure the deviation of the barycentre of suffrutescent chamaephytes from the centroid is significant for cluster b, but not for cluster a. The space defined by the first two components of Figure 7 can be considered homogeneous for rhizome geophytes, but not for scapose hemicryptophytes or suffrutescent camaephytes (see Table 1).

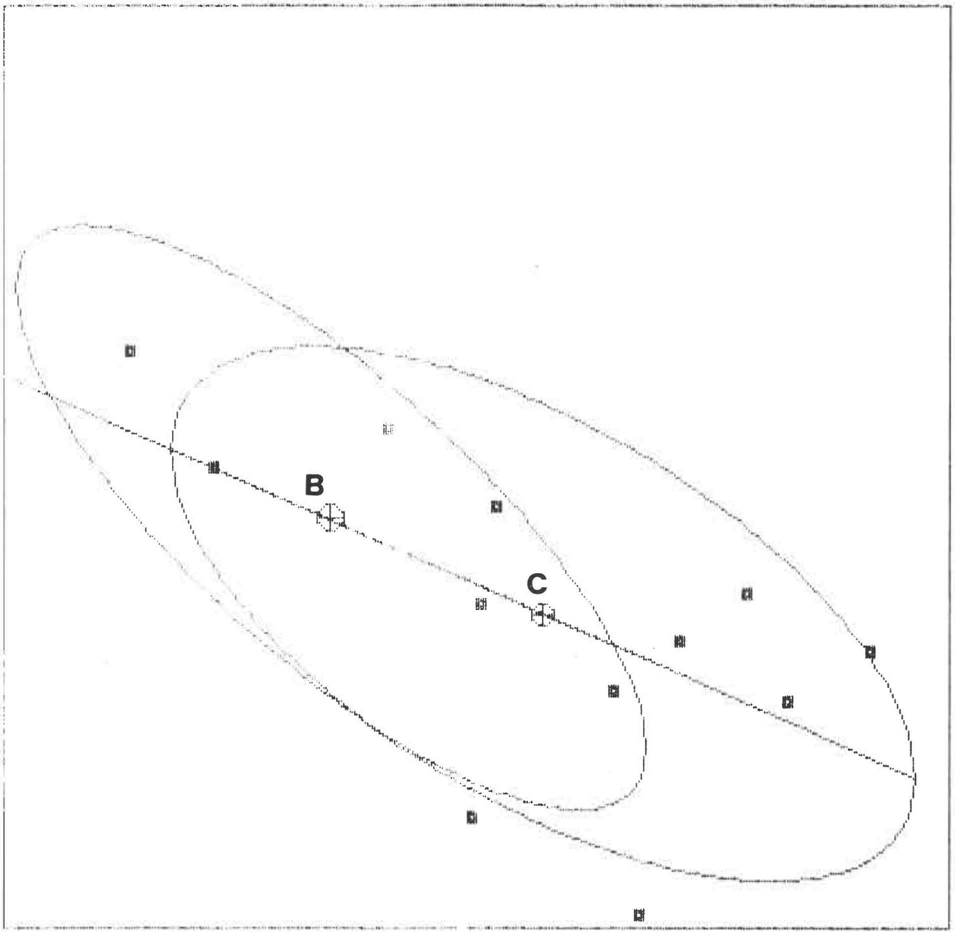

Fig. 5 — Ellipses of equal concentration around the barycentre B of suffrutescent chamaephytes of cluster b), and its centroid C, with reference to temperature and pH. The distance between the barycentre and centroid is significant *($p < 0.04$)*.

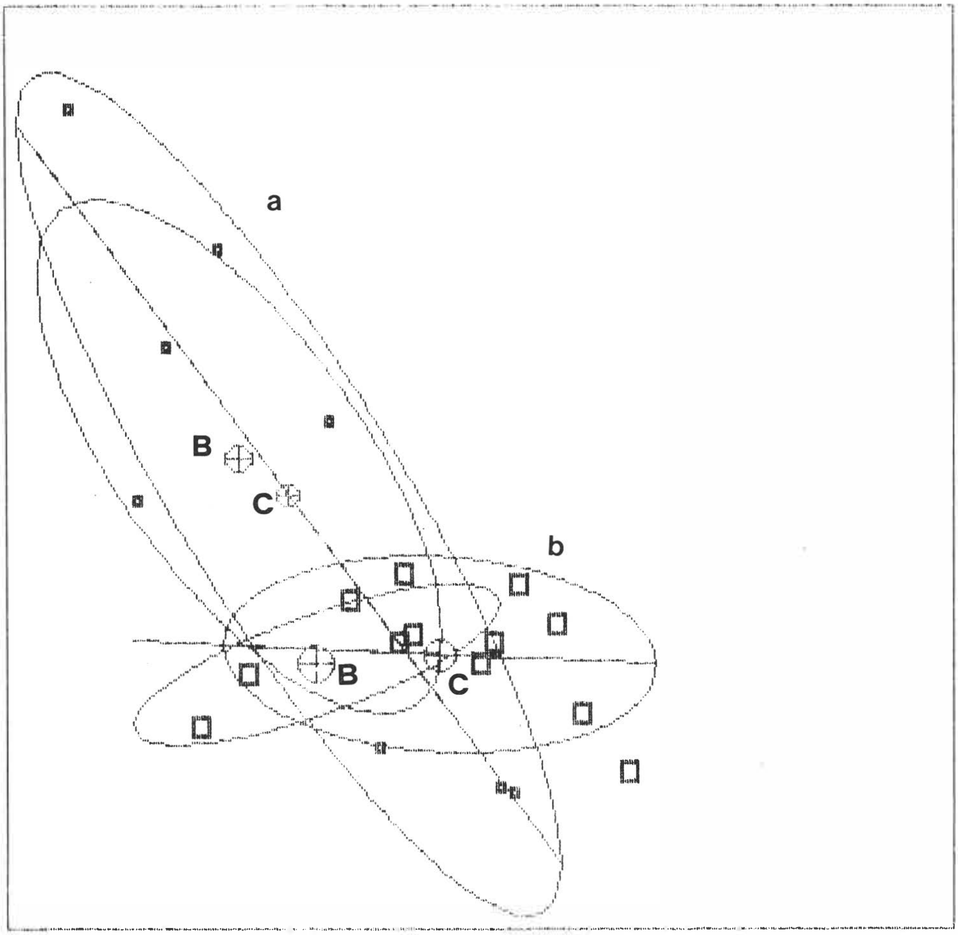

Fig. 6 — Ellipses of equal concentration at the 5% level around the centroids (C) and barycentres (B) of suffrutescent chamaephytes of clusters a) and b). The variables are humidity and temperature. The distances are significant between the barycentres of clusters a) and b), the centroids of cluster a) and b), the centroid of cluster a) and the barycentre of cluster b) and the centroid and barycentre of cluster b). The slopes of the regression lines are significantly different, while the intercepts are not.

In respect of the three life-growth forms, one can conclude that, in the space defined by the environmental variables and the 20 vegetation types, their niches (sensu Hutchinson, see Hurlbert, 1981) are shifting in an orderly fashion as one passes from the relatively, driest and most basic soils to the most humid and acid soils according to the sequence: suffrutescent chamaephytes, scapose hemicryptophytes and rhizome geophytes (Figure 8). In this sequence the barycentre of suffrutescent chamaephytes is significantly distant from the barycentres of the other two life-growth forms. The sequence corresponds also with decreasing

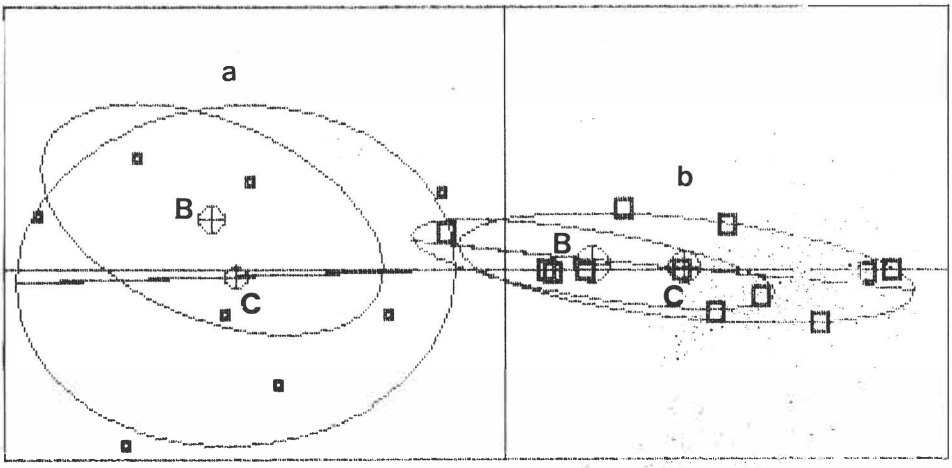

Fig. 7 — Ellipses of equal concentration at the 5% level around the centroids (C) and barycentres (B) of the clusters a) and b). The barycentres have been computed for suffrutescent chamaephytes. The variables are the first two components of the matrix describing the vegetation types by the eight average indicator values (see text).

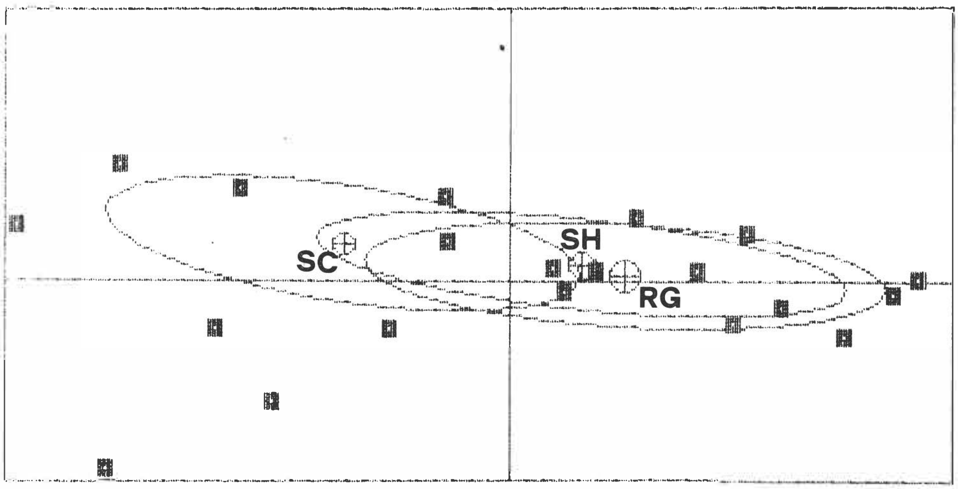

Fig. 8 — Ellipses of equal concentration at 31.7% around the barycentres of suffrutescent chamaephytes (SC), scapose hemicryptophytes (SH) and rhizome geophytes (RG). The variables are the same as in Figure 7.

temperature, light, and continentality and with increasing nutrients, the proportion of humus and silt and clay. This is proved by the correlation coefficients of the environmental variables with the first principal component of Figures 7 and 8. They are as follows: humidity 0.978, nutrients 0.93, humus 0.98, soil texture 0.93, pH −0.85, light −0.96, temperature −0.58, continentality −0.60.

Aknowledgements. The work has been supported by a CNR grant for the project "Structure and Dynamics of forests". We are grateful to prof. D.W. Goodall for reading and correcting the text.

References

Cliff A.D. and Ord J.K. (1981) - *Spatial process*. Models and applications. Pion, London.

Daboni L. (1967) - *Calcolo delle probabilità*. Boringhieri, Torino.

Feoli E. and Ganis P. (1986) - *Autocorrelation for measuring predictivity in community ecology: An example with structural and chorological data from mixed forest types of NE Italy*. Coenoses 1: 53-56.

Hurlbert S.H. (1981)-*A gentle depilation of the niche:Dicean resource sets in resource hyperspace*. In May, R.M. (ed.) Evolutionary theory, 5: 177-184. The University of Chicago, Chicago.

Lagonegro M. e Feoli E. (1985) - *Analisi multivariata di dati*. Manuale d'uso di programmi Basic per personal computers. Libreria Goliardica, Trieste.

Landolt E. (1977) - *Oekologische Zeigerwerte zur Schweizer Flora*. Ber. Geobot. Inst. ETH, 64: 64-207.

Mardia K.V., Kent J.T. and Bibby J.M. (1979) - *Multivariate analysis*. Academic Press, London, New York.

Orloci L. (1973) - *Ranking characters by a dispersion criterion*. Nature 244: 371-373.

Poldini L. (1982) - *Ostrya carpinifolia-rich woods and bushes of Friuli-Venezia Giulia*. Studia Geobotanica 2: 69-122.

Zar J.H. (1974) - *Biostatistical Analysis*. Prentice-Hall. Englewood Cliff, N.J.

Authors' address:
Prof. Dr. E. Feoli
Prof. Dr. M. Lagonegro
Dipartimento di Biologia, Università degli Studi di Trieste, 34100 Italia

Appendix.

The data for the program must be prepared in a sequential file in which the information is contained in blocks of records, each block containing in order:

1) the alphanumerical code for the variable for which the analysis is required (type wanted);

2) the score for the variable (prod);

3) the scores of the variables which, two by two, define the space (variables); these may be several and their number is placed after the "prod" variable.

The following is an example of the data structure as printed:

(type wanted, prod, variables)					
casu	20	2.5	2.7	9.8	10.75
casu	10	0.9	1.3	6.9	9.88
casu	9	0.7	1.3	5.3	6.97

In this case there are 3 objects described by variable casu (type wanted), the scores of casu for the three objects, and the four variables which, two by two, define the spaces. The file must be typed as follows for both the BASIC and the FORTRAN versions:

casu
20
2.5
2.7
9.8
10.75
casu
10
0.9
1.3
6.9
9.88
casu
9
0.7
1.3
5.3
6.97

The listings of the program are given below, both in BASIC and FORTRAN.

```
10 REM-Program ELLIPLOT II (M.Lagonegro-1985)
20 OPTION BASE 1
30 ON ERROR GOTO 2010
40 INPUT"give n.lines,n.ellipses,n.vars in data table ";NL,NTY,NV
50 DIM TIPO$(NL),PROD(NL),VARS(NL,NV),PESI(NL),ELI(100,3,NTY),SLO(NTY),ES1(NTY),ES3(NTY),ES5(NTY),ELV(NTY)
60 DIM X(NL,NTY),Y(NL,NTY),R(NL,NTY),XG(NTY),YG(NTY),IKS1(NTY),IKS2(NTY),UN(NTY),S2(NTY),ES2(NTY)
70 INPUT"give filename for data ";DAF$
80 OPEN"i",#1,DAF$
90 FOR I=1 TO NL
100  INPUT#1,TIPO$(I)
110 INPUT#1,PROD(I)
120 FOR J=1 TO NV
130  INPUT#1,VARS(I,J)
140  NEXT J
150 NEXT I
160  CLOSE#1
170 INPUT"print data-y/n ";SN$
180  IF SN$="n" THEN 270
190 LPRINT"    --- data table ---"
200  FOR I=1 TO NL
210 LPRINT TIPO$(I);PROD(I);
220  FOR J=1 TO NV
230 LPRINT VARS(I,J);
```

```
240  NEXT J
250  LPRINT
260  NEXT I
270 INPUT"give labels of couple of wanted vars-to stop give 0,0 ";N,M
280 K=1
290 IF M*N=0 THEN END
300 INPUT"type wanted/no ";TY$
310  IF TY$="no" THEN 270
320 INPUT"weighting-unif/prod ";TI$
330 PETOT=0
340 NU=0
350 FOR I=1 TO NL
360  A$=TIPO$(I)
370  IF A$<>TY$ THEN 440
380 NU=NU+1
390  PESI(NU)=PROD(I)
400 X(NU,K)=VARS(I,N)
410 Y(NU,K)=VARS(I,M)
420 R(NU,K)=I
430 PETOT=PETOT+PROD(I)
440 NEXT I
450 UN(K)=NU
460 XG(K)=0
470 YG(K)=0
480 FOR I=1 TO NU
490 PESI(I)=PESI(I)/PETOT
500 IF TI$="unif" THEN PESI(I)=1/NU
510 XG(K)=XG(K)+X(I,K)*PESI(I)
520 YG(K)=YG(K)+Y(I,K)*PESI(I)
530 NEXT I
540 SX2=0
550 SY2=0
560 SXY=0
570 FOR I=1 TO NU
580 PES2=PESI(I)
590 SX2=SX2+PES2*(X(I,K)-XG(K))^2
600 SY2=SY2+PES2*(Y(I,K)-YG(K))^2
610 SXY=SXY+PES2*(X(I,K)-XG(K))*(Y(I,K)-YG(K))
620 NEXT I
630 SX=SQR(SX2)
640 SY=SQR(SY2)
650 S2(K)=(SX2+SY2)/2
660 RO=SXY/(SX*SY)
670 IF SN$="y" THEN LPRINT"  Ellipse n.=";K
680 IF SN$="y" THEN LPRINT"sx=";SX;" sy=";SY;" ro=";RO
690 INPUT"give residual % ";REP
700 IF SN$="y" THEN LPRINT"residual %=";REP
710 REP=-LOG(REP/100)
720 IF RO>.99999 THEN RO=.99999
730 IF RO<-.99999 THEN RO=-.99999
740 X1=XG(K)-SQR(REP)*SX
750 X2=XG(K)+SQR(REP)*SX
760 DX=(X2-X1)/99
770 IKS1(K)=X1
780 IKS2(K)=X2
790 FOR I=1 TO 100
800 ELI(I,1,K)=DX*(I-1)+X1
810 XS=(ELI(I,1,K)-XG(K))/SX
820 B=-2*RO*XS
830 C=XS^2+REP*(RO^2-1)
840 BC=ABS(B^2-4*C)
850 YS1=.5*(-B-SQR(BC))
860 YS2=.5*(-B+SQR(BC))
870 ELI(I,2,K)=YS1*SY+YG(K)
880 ELI(I,3,K)=YS2*SY+YG(K)
890 NEXT I
900 IF SN$="n" THEN 1090
910 LPRINT"mode is ";TI$
920 LPRINT"x is ";N"-th var. and y is ";M;"-th"
930 LPRINT"x from ";X1;" to ";X2;"  y from ";ELI(1,2,K);" to ";ELI(100,2,K)
940 LPRINT" centroid has X=";XG(K);" Y=";YG(K)
950 LPRINT"  points have
960 FOR I=1 TO NU
970 LPRINT R(I,K);"  x=";X(I,K);" y=";Y(I,K);" weight=";PESI(I)
```

```
980 NEXT I
990 NSCB=0:FOR I=1 TO NU-1:IF X(I,K)<=X(I+1,K) THEN 1040
1000 ACI=X(I,K):X(I,K)=X(I+1,K):X(I+1,K)=ACI
1010 ACI=Y(I,K):Y(I,K)=Y(I+1,K):Y(I+1,K)=ACI
1020 ACI=R(I,K):R(I,K)=R(I+1,K):R(I+1,K)=ACI
1030 NSCB=1
1040 NEXT I
1050 IF NSCB<>0 THEN 990
1060 LPRINT"horizontal sequence"
1070 FOR I=1 TO NU:LPRINT R(I,K);
1080 NEXT I:LPRINT
1090 K=K+1:IF K<=NTY THEN 300
1100 INPUT"plot-scale factors diff/equal ";DIEQ$
1110 CLS:SCREEN 3
1120 MX=-1E+30
1130 MY=MX
1140 MNX=1E+30
1150 MNY=MNX
1160 FOR K=1 TO NTY
1170 NU=UN(K)
1180 FOR I=1 TO NU
1190 IF MX<   X(I,K) THEN MX=   X(I,K)
1200 IF MNX>   X(I,K) THEN MNX=   X(I,K)
1210 IF MY<   Y(I,K) THEN MY=   Y(I,K)
1220 IF MNY>   Y(I,K) THEN MNY=   Y(I,K)
1230 NEXT I:NEXT K
1240 FOR K=1 TO NTY
1250 FOR I=1 TO 100
1260 IF ELI(I,1,K)<MNX THEN MNX=ELI(I,1,K)
1270 IF ELI(I,1,K)>MX THEN MX=ELI(I,1,K)
1280 IF ELI(I,2,K)<MNY THEN MNY=ELI(I,2,K)
1290 IF ELI(I,2,K)>MY THEN MY=ELI(I,2,K)
1300 IF ELI(I,3,K)<MNY THEN MNY=ELI(I,3,K)
1310 IF ELI(I,3,K)>MY THEN MY=ELI(I,3,K)
1320 NEXT I: NEXT K
1330 LPRINT"x from ";MNX;" to ";MX;"  y from ";MNY;" to ";MY
1340 FX=400/(MX-MNX)
1350 FY=320/(MY-MNY):IF DIEQ$="diff" THEN 1380
1360 IF FX<=FY THEN FY=FX
1370 IF FY<=FX THEN FX=FY
1380 LINE (0,49)-(420,399),,B
1390 LPRINT"scale factors: fx=";FX;"  fy=";FY
1400 X0=-MNX*FX+5
1410 Y0=399+MNY*FY-5
1420 LINE (X0,49)-(X0,399)
1430 LINE (0,Y0)-(420,Y0)
1440 FOR K=1 TO NTY
1450 NU=UN(K)
1460 XX=(XG(K)-MNX)*FX+5
1470  YY=399-(YG(K)-MNY)*FY-5
1480 AR=K+4:UR=.83*AR:CIRCLE(XX,YY),AR
1490 LINE (XX-AR,YY)-(XX+AR,YY):LINE (XX,YY-UR)-(XX,YY+UR)
1500 FOR I=1 TO 100
1510 ELI(I,1,K)=(ELI(I,1,K)-MNX)*FX+5
1520 ELI(I,2,K)=399-(ELI(I,2,K)-MNY)*FY-5
1530 ELI(I,3,K)=399-(ELI(I,3,K)-MNY)*FY-5
1540  NEXT I
1550 LINE (ELI(1,1,K),ELI(1,2,K))-(ELI(1,1,K),ELI(1,3,K))
1560 FOR I=2 TO 100
1570 X1=ELI(I-1,1,K)
1580 X2=ELI(I,1,K)
1590 Y1=ELI(I-1,2,K)
1600 Y2=ELI(I,2,K)
1610 Y3=ELI(I-1,3,K)
1620 Y4=ELI(I,3,K)
1630 LINE (X1,Y1)-(X2,Y2)
1640 LINE (X1,Y3)-(X2,Y4)
1650 NEXT I
1660 LINE (ELI(100,1,K),ELI(100,2,K))-(ELI(100,1,K),ELI(100,3,K))
1670 FOR I=1 TO NU
1680  XX=(X(I,K)-MNX)*FX+5
1690 YY=399-(Y(I,K)-MNY)*FY-5
1700 LINE (XX-K,YY-K)-(XX+K,YY+K),,B
```

```
1710 NEXT I
1720 S1=0:S2=0:S3=0:S4=0:S5=0
1730 FOR I=1 TO NU:S1=S1+X(I,K)^2:S2=S2+X(I,K):S3=S3+X(I,K)*Y(I,K):S4=S4+Y(I,K):S5=S5+Y(I,K)^2:NEXT I
1740 PEND=(NU*S3-S2*S4)/(NU*S1-S2^2):TNOT=(S4-S2*PEND)/NU
1750 X=IKS1(K):Y=PEND*X+TNOT:XX=(X-MNX)*FX+5:YY=399-(Y-MNY)*FY-5:X=IKS2(K):Y=PEND*X+TNOT
1760 XXX=(X-MNX)*FX+5:YYY=399-(Y-MNY)*FY-5:LINE (XX,YY)-(XXX,YYY)
1770 IF SN$="y" THEN LPRINT"regression line: ";PEND;"*x ";CHR$(44-SGN(TNOT));ABS(TNOT)
1780 SLO(K)=PEND:ES1(K)=S1:ES2(K)=S2:ES3(K)=S3:ES5(K)=S5:ELV(K)=TNOT:NEXT K
1790 LOCATE ,55:INPUT"plot on printer-y/n ";MAH$
1800 IF MAH$="y" THEN LCOPY 0
1810 IF NTY=1 THEN 1980
1820 LPRINT:LPRINT" *** Comparison among centroids,slopes & intercepts ***":LPRINT
1830 FOR K=1 TO NTY-1:FOR J=K+1 TO NTY:N=UN(K):M=UN(J)
1840 ST2=((N-1)*S2(K)+(M-1)*S2(J))/(M+N-2)
1850 D=SQR((XG(K)-XG(J))^2+(YG(K)-YG(J))^2):TI=D*SQR(M*N/(ST2*(M+N)))
1860 T=TI:DF=M+N-4:GOSUB 2030
1870 LPRINT"pair ";K;" and ";J;" --- t=";TI;" deg.of freedom=";DF;" prob(%)=";A
1880 RESS1=ES5(K)-(ES3(K)^2)/ES1(K):RESS2=ES5(J)-(ES3(J)^2)/ES1(J)
1890 ES2XY=(RESS1+RESS2)/(N+M-4):SB1B2=SQR(ES2XY/ES1(K)+ES2XY/ES1(J))
1900 TI=ABS(SLO(K)-SLO(J))/SB1B2:T=TI:DF=M+N-4:GOSUB 2030
1910 LPRINT" slopes   t'=";TI;"  d.freedom=";DF;" with prob(%)=";A
1920 BICI=(ES3(K)+ES3(J))/(ES1(K)+ES1(J)):ME1=ES2(K)/N:ME2=ES2(J)/M:DIME=ME1-ME2
1930 BOH=DIME^2/(ES1(K)+ES1(J)):TI=(ELV(1)-ELV(2)-BICI*(DIME))/SQR(ES2XY*(1/N+1/M+BOH)):TI=ABS(TI)
1940 DF=M+N-3:T=TI:GOSUB 2030
1950 LPRINT" intcpts  t''=";TI;"  d.freedom=";DF;" with prob(%)=";A
1970 NEXT J:LPRINT: NEXT K
1980 CLS
1990 GOTO 270
2000 END
2010 PRINT"kapelmeister n.";ERR;"  line n.";ERL
2020  END
2030 REM-Subroutine Probt
2040 T=ATN(T/SQR(DF)) : IF DF=1 THEN A=.5-T/3.14159
2050 IF DF=2 THEN A=.5*(1-SIN(T))
2060 IF DF=3 THEN A=.5-(1/3.14159)*(T+COS(T)*SIN(T))
2070 IF DF<4 THEN GOTO 2180
2080 IF DF/2>INT(DF/2) THEN 2140
2090 C=1 : Q=2
2100 U=1 : FOR I=2 TO Q STEP 2 : U=U*(I-1)/I
2110 NEXT I : C=C+U*(ABS(COS(T)))^Q
2120 Q=Q+2 : IF Q<DF THEN 2100
2130 A=.5*(1-(SIN(T))*C) : GOTO 2180
2140 C=1 : Q=2
2150 U=1 : FOR I=2 TO Q STEP 2 : U=U*I/(I+1) : NEXT I
2160 C=C+U*(ABS(COS(T)))^Q : Q=Q+2 : IF Q<DF-1 THEN 2150
2170 A=.5-(1/3.14159)*(T+COS(T)*SIN(T)*C)
2180 A=A*200
2190 RETURN
```

```
$LARGE
      PROGRAM ELLIDIS
C---DRAWS EQUI-CONCENTRATION ELLIPSES FOR OBJECTS DESCRIBED
C---BY PAIRS OF VARIABLES-COMPARES THEN REGRESSION LINES
C---WRITTEN BY  M.LAGONEGRO-1986 (FROM BASIC VERSION)
      COMMON SLO(25),ELV(25),ES1(25),ES2(25),ES3(25),ES5(25),NL,NTY
     1,NV,PROD(200),VARS(200,25),PESI(200),ELI(100,3,25),
     2X(200,25),Y(200,25),IR(200,25),XG(25),YG(25),EKS1(25)
     3,EKS2(25),KUN(25),SS2(25)
      CHARACTER *4TIPO(200),ADO,TIDO,SNDO,TISO,DIEQS
      CHARACTER *10DAFDO
      WRITE(*,*)' GIVE N.LINES,N.ELLIPSES,N.VARIABLES'
      READ(*,*)NL,NTY,NV
      WRITE(*,*)' GIVE FILENAME OF DATA'
      READ(*,'(A10)')DAFDO
      OPEN(1,FILE=DAFDO,STATUS='OLD')
      DO 10120  I=1,NL
       READ(1,'(A4)')TIPO(I)
      READ(1,*)PROD(I)
      DO 10120 J=1,NV
10120  READ(1,*)VARS(I,J)
       CLOSE(1)
      WRITE(*,*)' PRINT DATA-Y/N'
      READ(*,'(A4)')SNDO
       IF(SNDO.EQ.'N')GOTO 10260
      WRITE(*,*)'     --- DATA TABLE ---'
      DO 10200 I=1,NL
10200 WRITE(*,1)TIPO(I),PROD(I),(VARS(I,J),J=1,NV)
1     FORMAT(1X,A4,G11.4,(1X,10G10.3))
10260 WRITE(*,*)' GIVE INDEXES OF WANTED VARS-TO STOP GIVE 0,0'
      READ(*,*)N,M
      K=1
      IF(M.EQ.0)STOP
10290 WRITE(*,*)' TYPE WANTED/N'
      READ(*,'(A4)')TIDO
       IF(TIDO.EQ.'N')GOTO 10260
      WRITE(*,*)' WEIGHTING-UNIF/PROD'
      READ(*,'(A4)')TISO
      PETOT=0
      NU=0
      DO 10430 I=1,NL
       ADO=TIPO(I)
       IF(ADO.NE.TIDO)GOTO 10430
      NU=NU+1
       PESI(NU)=PROD(I)
      X(NU,K)=VARS(I,N)
      Y(NU,K)=VARS(I,M)
      IR(NU,K)=I
      PETOT=PETOT+PROD(I)
10430 CONTINUE
      KUN(K)=NU
      XG(K)=0
      YG(K)=0
      DO 10510 I=1,NU
      PESI(I)=PESI(I)/PETOT
      IF(TISO.EQ.'UNIF')PESI(I)=1./FLOAT(NU)
      XG(K)=XG(K)+X(I,K)*PESI(I)
10510 YG(K)=YG(K)+Y(I,K)*PESI(I)
      SX2=0
      SY2=0
      SXY=0
      DO 10600 I=1,NU
      PES2=PESI(I)
      SX2=SX2+PES2*(X(I,K)-XG(K))**2
      SY2=SY2+PES2*(Y(I,K)-YG(K))**2
10600 SXY=SXY+PES2*(X(I,K)-XG(K))*(Y(I,K)-YG(K))
      SX=SQRT(SX2)
      SY=SQRT(SY2)
      SS2(K)=(SX2+SY2)/2
      RO=SXY/(SX*SY)
      IF(SNDO.EQ.'Y')WRITE(*,*)'  ELLIPSE N.=',K
      IF(SNDO.EQ.'Y')WRITE(*,*)' SX=',SX,' SY=',SY,' RO=',RO
      WRITE(*,*)' GIVE RESIDUAL(%)'
```

```
      READ(*,*)REP
      IF(SNDO.EQ.'Y')WRITE(*,*)' RESIDUAL %=',REP
      REP=-LOG(REP/100)
      IF(RO.GT.(.99999))RO=.99999
      IF(RO.LT.(-.99999)) RO=-.99999
      X1=XG(K)-SQRT(REP)*SX
      X2=XG(K)+SQRT(REP)*SX
      DX=(X2-X1)/99.
      EKS1(K)=X1
      EKS2(K)=X2
      DO 10870 I=1,100
      ELI(I,1,K)=DX*(I-1)+X1
      XS=(ELI(I,1,K)-XG(K))/SX
      B=-2.*RO*XS
      C=XS**2+REP*(RO**2-1.)
      BC=ABS(B**2-4.*C)
      YS1=.5*(-B-SQRT(BC))
      YS2=.5*(-B+SQRT(BC))
      ELI(I,2,K)=YS1*SY+YG(K)
10870 ELI(I,3,K)=YS2*SY+YG(K)
      IF(SNDO.EQ.'N')GOTO 11180
      WRITE(*,2)TISO
 2    FORMAT(' MODE IS:',A4)
      WRITE(*,3)N,M
 3    FORMAT(' X IS ',I4,'-TH VAR.  Y IS ',I4,'-TH')
      WRITE(*,4)X1,X2,ELI(1,2,K),ELI(100,2,K)
 4    FORMAT(' X FROM:',G11.4,' TO:',G11.4,' Y FROM:',G11.4,' TO:',
     1G11.4)
      WRITE(*,5)XG(K),YG(K)
 5    FORMAT(' CENTROID HAS X=',G11.4,' Y=',G11.4)
      WRITE(*,*)'   POINTS HAVE'
      DO 10960 I=1,NU
10960 WRITE(*,6)IR(I,K),X(I,K),Y(I,K),PESI(I)
 6    FORMAT(' POS.IN TAB.=',I4,' X=',G11.4,' Y=',G11.4,' WEIGHT='
     1,G11.4)
10980 NSCB=0
      DO 11110 I=1,NU-1
      IF(X(I,K).LE.X(I+1,K))GOTO 11110
      ACI=X(I,K)
      X(I,K)=X(I+1,K)
      X(I+1,K)=ACI
      ACI=Y(I,K)
      Y(I,K)=Y(I+1,K)
      Y(I+1,K)=ACI
      IACI=IR(I,K)
      IR(I,K)=IR(I+1,K)
      IR(I+1,K)=IACI
      NSCB=1
11110 CONTINUE
      IF(NSCB.NE.0)GOTO 10980
      WRITE(*,*)' HORIZONTAL SEQUENCE'
      DO 11150 I=1,NU
11150 WRITE(*,7)IR(I,K)
 7    FORMAT(1X,20I4)
11180 K=K+1
      IF(K.LE.NTY)GOTO 10290
      WRITE(*,*)' GIVE PLOT-SCALE FACTOR:DIFF/EQUA'
      READ(*,'(A4)')DIEQS
      RMX=-1.E+30
      RMY=RMX
      RMNX=1.E+30
      RMNY=RMNX
      DO 11350 K=1,NTY
      NU=KUN(K)
      DO 11340 I=1,NU
      IF(RMX.LT.X(I,K)) RMX= X(I,K)
      IF(RMNX.GT.X(I,K)) RMNX= X(I,K)
      IF(RMY.LT.Y(I,K)) RMY= Y(I,K)
      IF(RMNY.GT.Y(I,K)) RMNY= Y(I,K)
11340 CONTINUE
11350 CONTINUE
      DO 11450 K=1,NTY
      DO 11440 I=1.100
```

```
      IF(ELI(I,1,K).LT.RMNX)RMNX=ELI(I,1,K)
      IF(ELI(I,1,K).GT.RMX)RMX=ELI(I,1,K)
      IF(ELI(I,2,K).LT.RMNY)RMNY=ELI(I,2,K)
      IF(ELI(I,2,K).GT.RMY)RMY=ELI(I,2,K)
      IF(ELI(I,3,K).LT.RMNY)RMNY=ELI(I,3,K)
      IF(ELI(I,3,K).GT.RMY)RMY=ELI(I,3,K)
11440 CONTINUE
11450  CONTINUE
      WRITE(*,9)RMNX,RMX,RMNY,RMY
 9    FORMAT(' X FROM',G11.4,' TO',G11.4,' -- Y FROM',G11.4,' TO'
     1,G11.4)
      FX=100/(RMX-RMNX)
      FY=100/(RMY-RMNY)
      IF(DIEQS.EQ.'DIFF')GOTO 11520
      IF(FX.LE.FY) FY=FX
      IF(FY.LE.FX) FX=FY
11520 CALL PLONPA(RMNX,RMX,RMNY,RMY,FX,FY,SNDO)
      IF(NTY.EQ.1)GOTO 12640
      WRITE(*,*)'0 * COMPARISON FOR CENTROIDS,SLOPES,INTERCEPTS'
      DO 12630 K=1,NTY-1
      DO 12610 J=K+1,NTY
      N=KUN(K)
      M=KUN(J)
      ST2=(FLOAT(N-1)*SS2(K)+FLOAT(M-1)*SS2(J))/FLOAT(M+N-2)
      D=SQRT((XG(K)-XG(J))**2+(YG(K)-YG(J))**2)
      TI=D*SQRT(FLOAT(M*N)/(ST2*FLOAT(M+N)))
      T=TI
      DF=M+N-4
      CALL PROBT(T,DF,A)
      WRITE(*,10)K,J,TI,DF,A
 10   FORMAT(' PAIR ',I4,' & ',I4,' -- T=',G11.4,' D.FR.=',F6.0,
     1' HO PROB.(%)=',G11.4)
      RESS1=ES5(K)-(ES3(K)**2)/ES1(K)
      RESS2=ES5(J)-(ES3(J)**2)/ES1(J)
      ES2XY=(RESS1+RESS2)/FLOAT(N+M-4)
      SB1B2=SQRT(ES2XY/ES1(K)+ES2XY/ES1(J))
      TI=ABS(SLO(K)-SLO(J))/SB1B2
      T=TI
      DF=M+N-4
      CALL PROBT(T,DF,A)
      WRITE(*,11)TI,DF,A
 11   FORMAT('      SLOPES   T=',G11.4,' D.FR.=',F6.0,' HO PROB.(%)='
     1,G11.4)
      BICI=(ES3(K)+ES3(J))/(ES1(K)+ES1(J))
      AME1=ES2(K)/FLOAT(N)
      AME2=ES2(J)/FLOAT(M)
      DIME=AME1-AME2
      BOH=DIME**2/(ES1(K)+ES1(J))
      TI=(ELV(1)-ELV(2)-BICI*(DIME))/SQRT(ES2XY*(1./FLOAT(N)+
     11./FLOAT(M)+BOH))
      TI=ABS(TI)
      DF=M+N-3
      T=TI
      CALL PROBT(T,DF,A)
      WRITE(*,12)TI,DF,A
 12   FORMAT('      INTCPTS  T=',G11.4,' D.FR.=',F6.0,' HO PROB.(%)='
     1,G11.4)
12610 CONTINUE
      WRITE(*,*)'
12630 CONTINUE
12640 CONTINUE
      GOTO 10260
      END
      SUBROUTINE PROBT(T,DF,A)
      T=ATAN(T/SQRT(DF))
       IF(DF.EQ.1.) A=.5-T/3.14159
      IF(DF.EQ.2.) A=.5*(1.-SIN(T))
      IF(DF.EQ.3.) A=.5-(1./3.14159)*(T+COS(T)*SIN(T))
      IF(DF.LT.4.) GOTO 12970
      IF((DF/2.).GT.INT(DF/2.))GOTO 12870
      C=1.
       Q=2.
12780 U=1.
```

```
      IQU=Q+.5
       DO 12800 I=2,IQU,2
12800  U=U*FLOAT(I-1)/FLOAT(I)
       C=C+U*(ABS(COS(T)))**Q
      Q=Q+2.
       IF(Q.LT.DF)GOTO 12780
      A=.5*(1.-(SIN(T))*C)
       GOTO 12970
12870 C=1.
       Q=2.
12890 U=1.
      IQU=Q+.5
       DO 12910 I=2,IQU,2
12910  U=U*FLOAT(I)/FLOAT(I+1)
      C=C+U*(ABS(COS(T)))**Q
       Q=Q+2.
       IF(Q.LT.(DF-1.))GOTO 12890
      A=.5-(1./3.14159)*(T+COS(T)*SIN(T)*C)
12970 A=A*200.
      RETURN
      END

      SUBROUTINE PLONPA(RMNX,RMX,RMNY,RMY,FX,FY,SNDO)
      COMMON SLO(25),ELV(25),ES1(25),ES2(25),ES3(25),ES5(25),NL,NTY
     1,NV,PROD(200),VARS(200,25),PES1(200),ELI(100,3,25),
     2X(200,25),Y(200,25),IR(200,25),XG(25),YG(25),EKS1(25)
     3,EKS2(25),KUN(25),SS2(25)
      CHARACTER PLOT(102,102)
      CHARACTER *4SNDO
      DO 100 I=1,102
      DO 100 J=1,102
 100  PLOT(I,J)=CHAR(32)
      DO 101 I=1,102
      PLOT(1,I)=CHAR(45)
      PLOT(102,I)=CHAR(45)
      PLOT(I,1)=CHAR(33)
 101  PLOT(I,102)=CHAR(33)
      WRITE(*,1)FX,FY
 1    FORMAT(' SCALE FACTORS - FX=',G11.4,' FY=',G11.4)
      XO=-RMNX*FX+1.
      YO=102.+RMNY*FY-1.
      CALL LINE(XO,1.,XO,102.,PLOT,33)
      CALL LINE(1.,YO,102.,YO,PLOT,45)
      DO 12220 K=1,NTY
      NU=KUN(K)
      DO 11710 I=1,100
      ELI(I,1,K)=(ELI(I,1,K)-RMNX)*FX+1.
      ELI(I,2,K)=102.-(ELI(I,2,K)-RMNY)*FY-1.
11710 ELI(I,3,K)=102.-(ELI(I,3,K)-RMNY)*FY-1.
      DO 11820 I=1,100
      X2=ELI(I,1,K)
      Y2=ELI(I,2,K)
      Y4=ELI(I,3,K)
      IX2=X2+.5
      IY2=Y2+.5
      IY4=Y4+.5
      PLOT(IY2,IX2)=CHAR(42)
11820 PLOT(IY4,IX2)=CHAR(42)
      S1=0.
      S2=0.
      S3=0.
      S4=0.
      S5=0.
      DO 12000 I=1,NU
      S1=S1+X(I,K)**2
      S2=S2+X(I,K)
      S3=S3+X(I,K)*Y(I,K)
      S4=S4+Y(I,K)
12000 S5=S5+Y(I,K)**2
      PEND=(FLOAT(NU)*S3-S2*S4)/(FLOAT(NU)*S1-S2**2)
      TNOT=(S4-S2*PEND)/FLOAT(NU)
      AX=EKS1(K)
      AY=PEND*AX+TNOT
      XX=(AX-RMNX)*FX+1.
```

```
      YY=102.-(AY-RMNY)*FY-1.
      AX=EKS2(K)
      AY=PEND*AX+TNOT
      XXX=(AX-RMNX)*FX+1.
      YYY=102.-(AY-RMNY)*FY-1.
      CALL LINE(XX,YY,XXX,YYY,PLOT,43)
      IF(SNDO.EQ.'Y') WRITE(*,*)' REGRESSION LINE:'
      WRITE(*,2)PEND,TNOT
 2    FORMAT(' SLOPE=',G11.4,' INTERCEPT=',G11.4)
      SLO(K)=PEND
      ES1(K)=S1
      ES2(K)=S2
      ES3(K)=S3
      ES5(K)=S5
      ELV(K)=TNOT
      DO 11880 I=1,NU
       XX=(X(I,K)-RMNX)*FX+1.
      YY=102.-(Y(I,K)-RMNY)*FY-1.
      IXX=XX+.5
      IYY=YY+.5
11880 PLOT(IYY,IXX)=CHAR(64+K)
      XX=(XG(K)-RMNX)*FX+1.
       YY=102.-(YG(K)-RMNY)*FY-1.
      IXX=XX+.5
      IYY=YY+.5
      PLOT(IYY,IXX)=CHAR(35)
12220 CONTINUE
      WRITE(*,*)' PLOT ON PRINTER-Y=1/N=D'
      READ(*,*)MAH
      IF(MAH.EQ.D)RETURN
      DO 3 I=1,102
 3    WRITE(*,4)(PLOT(I,J),J=1,102)
 4    FORMAT(1X,102A1)
      RETURN
      END
      SUBROUTINE LINE(X1,Y1,X2,Y2,PLOT,NC)
      CHARACTER PLOT(102,102)
      DX=X2-X1
      DY=Y2-Y1
      IF(DX.EQ.D.)DX=1.E-9
      A=DY/DX
      B=Y1-A*X1
      PA=DX/100
      DO 5 I=1,101
      XX=X1+(I-1)*PA
      YY=A*XX+B
      IF(X1.EQ.X2)YY=I
      IF(Y1.EQ.Y2)XX=I
      IXX=XX+.5
      IYY=YY+.5
      IF(IXX.GT.102.OR.IXX.LT.1)IXX=1
      IF(IYY.GT.102.OR.IYY.LT.1)IYY=1
 5    PLOT(IYY,IXX)=CHAR(NC)
      RETURN
      END
```