ACTA SOCIETATIS
BOTANICORUM POLONIAE
Vol. 54, nr 4: 511-521
1985

Statistical analysis of the phytocoenose homogeneity. V. Frequency distributions of similarity and distance coefficients as a function of the area size

ANNA JUSTYNA KWIATKOWSKA, EWA SYMONIDES

Department of Phytosociology and Plant Ecology, Institute of Botany, Warsaw University, Al. Ujazdowskie 4, 00-478 Warszawa, Poland

(Received: April 1, 1985. Accepted: May 16, 1985)

Abstract

Homogeneity of the *Leucobryo-Pinetum* phytocoenose was assessed on the grounds of the type of frequency distributions of similarity and Euclidean distance coefficients. It was confirmed that: 1) the type of frequency distributions of similarity and distance coefficients, as well as values of their statistical characteristics depended on the area size; 2) for analysed phytocoenose the area size at which frequency distributions of both coefficients were symmetrical, i.e. fitted the normal distribution, could be determined; 3) at the same time such area determined the scale of spatial differentiation of the phytocoenose under which it was homogeneous.

Key words: phytocoenose homogeneity, similarity coefficient, distance coefficient

INTRODUCTION

Species composition and dominance structure have been for long subjects of comparative studies of plant (and animal) communities occuring in the same area, or following each other in the course of succession. However, in the analysis of similarity degree between different parts of the same patch these measures have been more rarely used. Although, it can led to a conclusion on the phytocoenose homogeneity (Gounot and Calléja 1962).

Random changes in the species composition in different parts of the patch, as well as agreement between the distribution of similarity coefficients and normal distribution testify to the relatively uniform floristic composition within the biochore. Similarly, small random changes in the biomass of individual species indicate constant dominance structure in various patch parts. Phytocoenose is considered homogeneous from the statistical point of view with respect to species composition and dominance structure if frequency distributions of similarity and distance coefficients fit the normal distribution (Kwiatkowska 1972).

The paper aims at analysis of the homogeneity of the model phytocoenose on the grounds of the type of frequency distributions of similarity and Euclidean distance coefficients as functions of the area size. Value of *P* coefficient, defined by the Jaccard-Steinhaus's (Jaccard 1932, Sławiński 1950) formula, also termed Jaccard-Sørensen's (Sørensen 1948) formula, i.e.: $\frac{2c \cdot 100\%}{a+b}$, where c — number of species common for both quadrats, a — number of species in the first quadrat, b — number of species in the second quadrat, has been assumed as a measure of quadrat floristic similarity.

Also coefficients of Euclidean distance (Legendre and Legendre 1979) have been calculated for the successive quadrat size. Here biomass of individual species serve as a basis for calculations. Similarity of two quadrats (i, j) is expressed by a vector of their distance in n-dimentional space which is related to the value of D coefficient. D is formulated as follows:

$D_{ij} = \sqrt{\sum_{p=1}^{n} (u_{pi} - u_{pj})^2}$, where n — species number in both quadrats, u_{pi} — biomass value of p species in i quadrat transformed into u variable, u_{pj} — analogously, biomass value in j quadrat.

Paper is the fifth in the series of works which compare results of studies with respect to qualitative and quantitative measures, as well as different homogeneity indices. In the earlier papers following mesures of the phytocoenose homogeneity have been discussed: 1) the type of distributions of total species diversity and evenness indices (Kwiatkowska and Symonides 1985a), 2) the type of species frequency distribution and frequency distribution of biomass (Kwiatkowska and Symonides 1985b), 3) the type of spatial distributions of species and their standing biomass (Kwiatkowska and Symonides 1985c), 4) species number and mean biomass value (Kwiatkowska and Symonides 1985d) — always related to the area size.

The studies were carried-out in the *Leucobryo-Pinetum* Mat. (1962) 1973 phytocoenose with the uniform physiognomy. It was composed of pure, even-aged, one-layered forest stand and floristicly poor ground layer (21 species of vascular and 7 species of sporogenous plants) dominated by dwarf-shrubs. The full floristic composition of phytocoenose under study and interpopulational quantitative relationships have been presented in the first paper of the series (Kwiatkowska and Symonides 1985a).

MATERIAL AND METHODS

In sampling Greig-Smith's (1952) grid was used. It was composed of 512 square quadrats (sample areas) of side 1 m each. In the period of maximum standing biomass (July) in each quadrat all species were recorded. Then, above-ground parts of all vascular plants in the ground layer were clipped and their dry-air biomass were assessed. The data were arranged in plans of species occurrence and cartogramms of their biomass in the successive quadrats. They served as a basis in the statistical analysis of the phytocoenose homogeneity.

For further analysis quadrats with sizes ascending with geometric progression, from 1 to 64 m^2 (10 quadrats for each smaller size and all — 8 — for 64 m^2), were sampled without replacement. Gounot and Calléja (1962) have proved that samples with 8 replications are representative from the statistical point of view.

For each sampled quadrat all species were recorded. Next, similarity coefficients were calculated for each quadrat pair in the successive samples. In further elaboration of the results distributive series were constructed, frequency of the class of similarity coefficient values was determined, and basic distribution characteristics, such as: mean ($\bar{x}$), standard deviation (S.D.), variance (S^2), and coefficient of variation (V) were calculated. At the significance level of 0.01 empirical and theoretical distributions were compared with Chi^2 test of goodness of fit.

Also analysis of the type of frequency distribution of distance coefficients was conducted analogously. Each sampled quadrat was characterized by standing biomass value. Mean ($\bar{x}$) and standard deviation (S.D.) were calculated. Due to the big differences in biomass of individual species, biomass value was transformed into standarized random u variable, according to the formula: $u_i = \frac{x_i - \bar{x}}{\text{S.D.}}$. On this basis distance coefficients (D) were calculated for each pair of quadrats in the successive samples.

Next, distributive series were constructed for distance coefficients and basic distribution characteristics were calculated, i.e. mean ($\bar{x}$), standard deviation (S.D.), variance (S^2), and coefficient of variation (V). Empirical and theoretical distributions were compared with Chi^2 test of goodness of fit at the significance level of 0.05.

RESULTS

The results indicate that both the type of similarity coefficient distribution and basic statistical characteristics depend on the area size (Figs. 1 and 2). Similarity of the floristic composition ($\bar{x}$) increases with the quadrat size growth. The highest increment of the mean value of similarity coefficient occurs between 1 and 4 m^2, then it becomes relatively constant (Fig. 1).

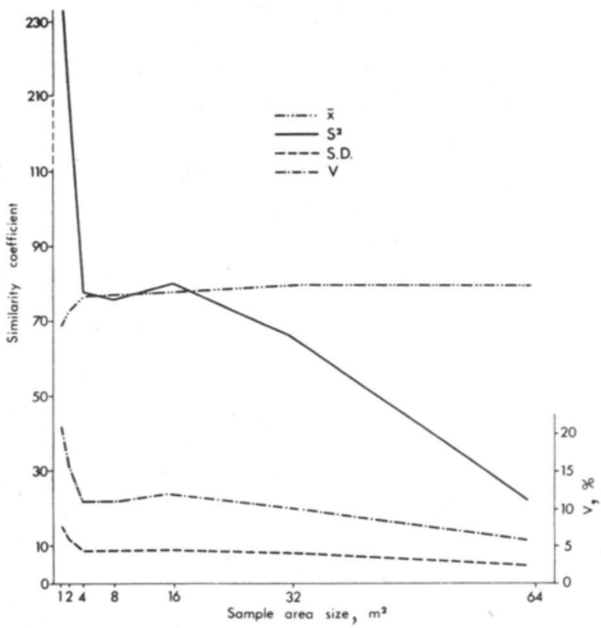

Fig. 1. Statistical characteristics of frequency distribution of similarity coefficients: mean ($\bar{x}$), variance (S^2), standard deviation (S.D.), and coefficient of variation (V) — as functions of the area size

Mean value of similarity coefficient is logarithmicly related to the quadrat size. Relations of other distribution parameters: standard deviation (S.D.), variance (S^2) and coefficient of variation (V) represent different functions (Fig. 1).

In all distributions there is a tendency to decrement of variation with the increase in the area size. Lower and lower values of all parameters indicate that. Also a decrease in range values goes in the same direction. For successive quadrats of: 1, 2, 4, 8, 16, 32 and 64 m^2 they equal: 80, 60, 40, 40, 50, 40 and 35, respectively. It indicates a reduction in the variation of floristic composition and increase in the quadrat similarity. It can be concluded from Figure 1 that for quadrats bigger than 4 m^2

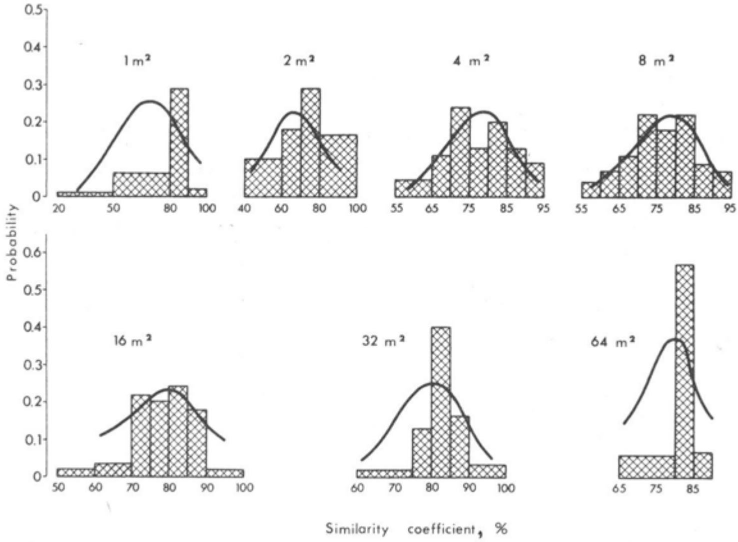

Fig. 2. Effect of the area size on frequency distribution of similarity coefficients

distribution variation is fairly constant and similarity coefficients equal at least 50 per cent.

The type of frequency distribution of similarity coefficients changes with an increase in the area size. First it is close to symmetrical, then transforms into, negatively skewed, *J*-shaped distribution (Fig. 2).

Small quadrats (1 and 2 m^2) markedly differ with respect to floristic composition. It results in the high variation of similarity coefficient values. Their distributions do not fit the normal distribution due to the „surplus" in high and low value classes, in comparison with expected values.

Bigger quadrats are characterized by different relationships. Nearly half of all quadrats of 32 and 64 m^2 are very similar in their relatively constant floristic composition. Moreover, probability of falling of the coefficient in the 80-85% class equals 0.40 and 0.59, respectively. It results in the considerable „surplus" of observed over expected values in this frequency class, and at the same time "deficiency" in the lower classes. The empirical distribution does not fit the normal distribution, it assumes *J* shape and is negatively skewed, changing towards the uniform distribution.

Empirical frequency distributions of similarity coefficients for 4, 8 and 16 m^2 fit the normal distribution at the significance level of 0.01 (Fig. 2). Therefore, according to the assumed criterion, quadrats 4-16 m^2 can be considered homogeneous with respect to the floristic composition.

Investigations have showed that mean value of Euclidean distance coefficient grows with an increase in the quadrat size (Fig. 3). Therefore, the bigger quadrat size, the smaller quadrat similarity with respect to the species dominance structure. Indeed, variation of the distance coefficient values, high for 1 m^2, rapidly decreases at 2 m^2, but further courses of curves of standard deviation, variance and coefficient of variation are ascending. Variation becomes relatively constant for the mean quadrats: 8 and 16 m^2. Quadrats bigger than 16 m^2 differ apparently with respect to inter-species quantitative relationships. Increase in values of S.D., S^2 and V shows that to be true (Fig. 3).

Comparison of empirical frequency distribution of distance coefficients and the normal distribution has showed that null hypothesis of goodness

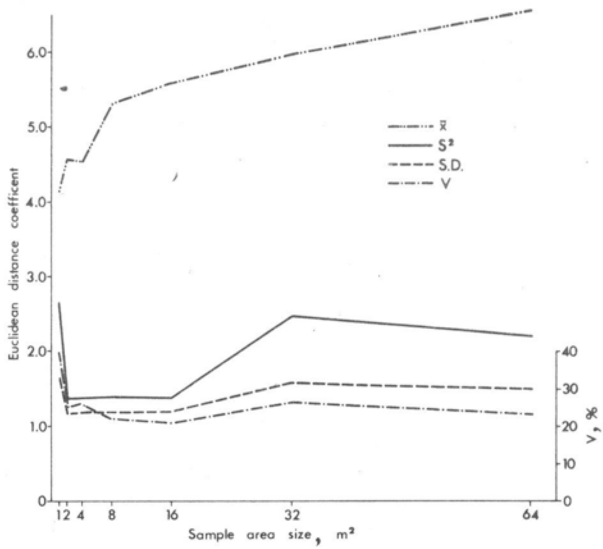

Fig. 3. Statistical characteristics of the frequency distribution of Euclidean distance coefficients: mean ($\bar{x}$), variance (S^2), standard deviation (S.D.), and coefficient of variation (V) — as functions of the area size

of fit is true at 0.05 significance level only for 2 m^2. For the remaining quadrats (1, 4, 8, 16, 32, 64 m^2) likelihood of hypothesis is smaller than 0.05.

The type of frequency distribution of Euclidean distance coefficients is a function of the area size (Fig. 4). It gets normal and then passes into *J*-shapped distribution with an increase in the area size. For the latter, value of coefficient is the highest. Only small quadrats (2 m^2) are statistically similar to each other with respect to the standing biomass of individual species growing on them. Large quadrats resemble each other to the similar, small degree.

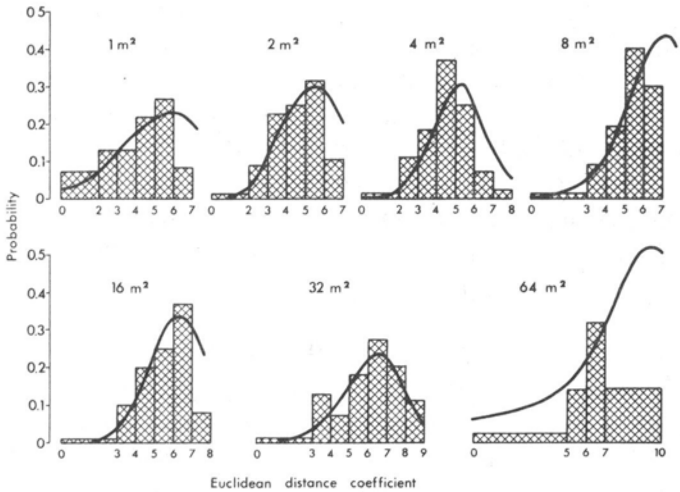

Fig. 4. Effect of the area size on frequency distribution of Euclidean distance coefficients

From the type of frequency distribution of Euclidean coefficients it may be inferred that with respect to its dominance structure analysed phytocoenose is homogeneous only under microscale (2 m^2).

DISCUSSION

About 20 different coefficients have been devised to assess similarity of the analysed units (Legendre and Legendre 1979). Coefficient value depends upon assumed mathematical definition of similarity. Therefore, its formula plays an essential role, especially when threshold association similarity is determined conventionally (cf. Romaniszyn 1953, 1970). On the other hand, it does not affect the type of distribution of similarity coefficients.

In this paper Jaccard-Steinhaus's coefficient has been chosen due to the many-years' tradition of Polish phytosociological studies. It has been used by: Kulczyński (1928), Matuszkiewicz W. (1952), Matuszkiewicz A. (1955), Balogh (1958), Faliński (1958, 1960, 1964), Traczyk (1960), Jabłoński (1964), Matuszkiewicz W. and Matuszkiewicz J. (1973), Matuszkiewicz J. (1976, 1977), Symonides (1979) and others.

Choice of Jaccard-Steinhaus's coefficient has also enabled comparison of own results with those Gounot and Calléja (1962) have obtained for grass community with *Brachypodium ramosum*. Although these authors have used systematic sampling and different number of quadrats have been analysed, results are comparable. In both cases it has been determined that: 1) mean value of coefficient of similarity between quadrats of the specified runs is a function of the area size and grows with its increase; 2) variation in values of similarity coefficient (S.D., V) is a function of the area size and decreases with its growth; 3) the type of frequency distribution of similarity coefficients is a function of the quadrat size; 4) phytocoenose homogeneity with respect to the floristic composition depends on the scale of analysis, therefore, it is a function of the area size. It is really interesting that in both investigations the quadrat size at which frequency distribution of similarity coefficients fits the normal distribution equals the size of representative area.

The representative area size of the analysed *Leucobryo-Pinetum* phytocoenose, calculated according to the Calléja's method, amounts to 3 m^2 (Kwiatkowska and Symonides 1985d). The phytocoenose can be considered homogeneous with respect to species spatial distribution under the analogous scale of its differentiation. The quadrat of 4 m^2 is characterized by significantly smaller fraction of species with contagious distributions and dignificantly smaller fraction of "contagious tests", in comparison with quadrats of 1 and 2 m^2 (Kwiatkowska and Symonides 1985c).

The type of frequency distribution of similarity coefficients has got its adventages as a homogeneity measure, namely, it is directly related to the statistical criterion of floristic homogeneity of the phytocoenose and it is of synthetic nature. This index enables to verify hypothesis of similar floristic composition of different fragments of biochore, i.e. homogeneity degree of the analysed patch.

From the results it may be concluded that phytocoenose is homogeneous with respect to the species composition in a scale larger than 2 m^2. The bigger quadrat size, the more similar floristic composition of quadrats. It results in the frequency distribution of similarity coefficients getting first normal and with further increase in the quadrat size — uniform. According to the assumed criterion also areas bigger than 16 m^2 can be considered homogeneous with respect to the analysed measure. As, regarding sample homogeneity uniform distribution meets condition of the statistical homogeneity better than the normal distribution.

Although the type of frequency distribution of Euclidean distance coefficients has not been so far used in the analysis of phytocoenose homogeneity, some trends can be deduced on the grounds of obtained results. They are analogous to those characteristic of frequency distribution of similarity coefficients, namely: 1) the type of frequency distribution of distance coefficients, and values of its statistical characteristics are functions of the area size; 2) mean value of distance coefficient grows with an increase in the area size; 3) for analysed phytocoenose the area size can be determined at which biomass values of individual species in any point within biochore are similar.

The fact that with an increase in the area size type of frequency distribution of Euclidean coefficients first gets normal and then changes into *J*-shaped distribution (with the highest mean value of the coefficient), indicates that large quadrats are similar in little. It probably results from transformation of biomass values into normalized variable and causes that statistical values of populations with high and low standing biomass are the same. For large quadrats the result is governed by species with low frequency in the patch, the chance of finding them increases with the growth of area size. Such species are characterized by high value of variance in the sample, which gives an increase in value of Euclidean distance coefficient.

Summing-up, phytocoenose under study is homogeneous with respect to its dominance structure only under the small scale (quadrat size — 2 m^2) of its differentiaion. At this size different fragments of the patch are similar to each other from the statistical point of view. They represent one structural type with respect to the species composition and biomass of individual populations.

REFERENCES

Balogh J., 1958. Lebensgemeinschaften der Landtiere. Ihre Erforshung unter besonderer Berücksichtigung der zoozönologischen Arbeitsmethoden. Budapest.

Faliński J. B., 1958. Nomogramy i tablice współczynników podobieństwa między zdjęciami według wzoru Jaccarda i Steinhausa. Acta Soc. Bot. Pol. 27: 115-130.

Faliński J. B., 1960. Zastosowanie taksonomii wrocławskiej do fitosocjologii. Acta Soc. Bot. Pol. 29: 333-361.

Faliński J. B., 1964. O różnych sposobach rozumienia pojęcia typu w fitosocjologii. Ekol. Pol. Ser. B., 10: 297-306.

Gounot M., Calléja M., 1962. Etude statistique d'une pelouse à *Brachypodium ramosum*. VIII. Coefficient de communaute, homogeneite et aire minimale. Bull. Serv. Carte Phytogéogr. Sér. B. 2: 181-200.

Greig-Smith. P., 1952. The use of random and contigious quadrats in the study of the structure of plant communities. Ann. Bot. 16: 293-316.

Jabłoński B., 1964. Uwagi na temat stosowania wzoru Jaccarda w badaniach ornitologicznych. Ekol. Pol. Ser. B, 10: 307-317.

Jaccard P., 1932. Die statistich-floristiche Methode als Grundlage der Pflanzen Soziologie. Handb. Biol. Arbeitsmeth. 11: 165-202.

Kulczyński S., 1928. Die Pflanzenassoziationen der Pieninen. Bull. Int. Acad. Polon. Sci. Lettres Classe Sci. Math. Nat., Ser. B. Suppl. II: 57-203.

Kwiatkowska A. J., 1972. Analiza homogeniczności runa fitocenozy przy zastosowaniu nieparametrycznego testu serii. Phytocoenosis 1: 37-77.

Kwiatkowska A. J., Symonides E., 1985a. Statistical analysis of the phytocoenose homogeneity. I. Distribution of the total species diversity and evenness indices as a homogeneity measure. Acta Soc. Bot. Pol. 54: 449-463.

Kwiatkowska A. J., Symonides E., 1985b. Statistical analysis of the phytocoenose homogeneity. II. Species frequency distribution and frequency distribution of the standing biomass as a function of the area size. Acta Soc. Bot. Pol. 54: 465–475.

Kwiatkowska A. J., Symonides E., 1985c. Statistical analysis of the phytocoenose homogeneity. III. Spatial distributions of species and their standing biomass as a function of the area size. Acta Soc. Bot. Pol. 54: 477-491.
Kwiatkowska A. J., Symonides E., 1985d. Statistical analysis of the phytocoenose homogeneity. IV. Species number and mean biomass value as functions of the area size. Acta Soc. Bot. Pol. 54: 493-509.
Legendre L., Legendre P., 1979. Ecologie numérique. 2. La structudes donnes ecologique. Collection d'écologie 13, Masson, Paris et les Press de l'Universitató du Québec.
Matuszkiewicz A., 1955. Stanowisko systematyczne i tendencje rozwojowe dąbrów białowieskich. Acta Soc. Bot. Pol. 24: 459-494.
Matuszkiewicz J., 1976. Przegląd fitosocjologiczny zbiorowisk leśnych Polski. Cz. 3. Lasy i zarośla łęgowe. Phytocoenosis 5: 3-66.
Matuszkiewicz J., 1977. Przegląd fitosocjologicznych zbiorowisk leśnych Polski. Cz. 4. Bory świerkowe i jodłowe. Phytocoenosis 6: 149-277.
Matuszkiewicz W., 1952. Zespoły leśne Białowieskiego Parku Narodowego. Ann. UMCS, Sec. C, Supl. 6: 1-216.
Matuszkiewicz W., Matuszkiewicz J., 1973. Przegląd fitosocjologiczny zbiorowisk leśnych Polski. Cz. 2. Bory sosnowe. Phytocoenosis 2: 273-356.
Romaniszyn W., 1953. Analiza ilościowa fauny ochotkowatych (*Tendipedidae, Diptera*) w strefie brzeżnej jeziora Charzykowo. Pol. Pismo Entomol., Wrocław 23: 1-51.
Romaniszyn W., 1970. Próba interpretacji tendencji skupiskowych zwierząt w oparciu o definicję podobieństwa i odległości. Wiad. Ekol. 14: 306-327.
Sørensen T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetatio on Danish commons. Biol. Skr. Kopenhavn 5: 1-34.
Sławiński W., 1950. Podstawy fitosocjologii. Cz. III. Lublin.
Symonides E., 1979. The structure and population dynamics of psammophytes on inland dunes. II. Loose-sod populations. Ekol. Pol. 27: 191-234.
Traczyk T., 1960. Badania nad strefą przejścia zbiorowisk leśnych. Ekol. Pol. Ser. A. 8: 85-125.

Statystyczna analiza jednorodności fitocenozy. V. Typ rozkładu wartości współczynników: podobieństwa i odległości jako funkcja wielkości powierzchni

Streszczenie

Praca jest częścią szerszych studiów nad problemem homogeniczności fitocenozy, mających na celu paralelizację wyników uzyskanych dla cech jakościowych i ilościowych, z zastosowaniem różnych wskaźników homogeniczności. Do badań wytypowano fizjonomicznie jednorodną fitocenozę *Leucobryo-Pinetum* Mat. (1962) 1973, szczegółowo opisaną w pierwszej pracy cyklu (Kwiatkowska i Symonides 1985a).

We wcześniejszych opracowaniach miarą jednorodności był: typ rozkładu wskaźników ogólnej różnorodności gatunkowej i równomierności (Kwiatkowska i Symonides 1985a), rozkład frekwencji i rozkład wartości stanu biomasy jako funkcja wielkości powierzchni (Kwiatkowska i Symonides 1985b), typ rozkładu przestrzennego gatunków i stanu ich biomasy jako funkcja wielkości powierzchni (Kwiatkowska i Symonides 1985c) oraz liczba gatunków i średnia wartość stanu biomasy jako funkcja wielkości powierzchni (Kwiatkowska i Symonides 1985d).

W niniejszej pracy analizę homogeniczności badanej fitocenozy przeprowadzono na pod-

stawie typu rozkładu wartości współczynników podobieństwa i współczynników odległości Euklidesa, jako funkcji wielkości powierzchni. Miarą podobieństwa florystycznego powierzchni była wartość współczynnika *P*, określonego wzorem Jaccarda-Steinhausa (Jaccard 1932, Sławiński 1950), zaś miarą podobieństwa powierzchni pod względem struktury dominacji — wektor ich odległości w przestrzeni *n*-wymiarowej, odpowiadający wartości współczynnika odległości Euklidesa *D* (Legendre i Legendre 1979).

Przy wzrastającym w postępie geometrycznym rozmiarze powierzchni podstawowej stwierdzono: 1) wzrost podobieństwa składu florystycznego, przy czym wartość średnia jest logarytmiczną funkcją wielkości powierzchni (rys. 1); 2) spadek zmienności w obrębie rozkładu wartości współczynników podobieństwa; S.D., S^2 i V maleją wraz ze wzrostem wielkości powierzchni (rys. 1); 3) zmianę typu rozkładu wartości współczynników podobieństwa: początkowo na zbliżony do symetrycznego, następnie na lewoskośny, typu *J* (rys. 2); 4) wzrost średniej wartości współczynników odległości Euklidesa i po początkowym spadku wzrost wartości S.D., S^2 i V, a zatem spadek podobieństwa poletek pod względem stuktury dominacji gatunków (rys. 3); 5) zmianę typu rozkładu wartości współczynników odległości Euklidesa: najpierw jego normalizację, a następnie przejście w rozkład typu J (rys. 4).

Badania wykazały, że jeśli pod względem składu florystycznego fitocenoza *Leucobryo-Pinetum* jest jednorodna w skali większej niż 4 m^2, to pod względem struktury dominacji — przeciwnie, wyłącznie w mikroskali, odpowiadającej 2 m^2.