

Katarzyna Ostaszewska

## A NEW WAY OF USING THE STATISTICAL INDICATORS IN LANDSCAPE RESEARCH

### INTRODUCTION

Statistical indicators have been used in the study of landscape for more than 20 years now (see, e.g., Richling, 1982; Bezkowska, 1986; Pietrzak, 1989; Grabowski, 1990). They are mainly used in the analysis of landscape structures, landscape being treated as a whole composed of spatial units (geocomplexes).

In practice to date the basis for the statistical analysis of landscape was constituted by the maps of geocomplexes, treated as the **source of data**. The respective assessments concerned, in particular, the magnitude of units, the way they were distributed, the degree of contrast of their boundaries, as well as the degree of integration of their components. On the other hand, the manner of distinguishing geocomplexes has not been subject to verification with the help of statistical indicators as the procedure antecedent with respect to statistical analysis.

The present paper is meant to show a new possibility of making use of the statistical measures. They will namely be used for the **optimisation of criteria of distinguishing the landscape units**.

The method proposed is founded on the definition of geocomplex as an organisational unity of the natural components, that is: the geological structure, the relief, waters, soil, vegetation cover, animal world, and the ground-adjacent layer of the atmosphere. The unity of a geocomplex is expressed through the fact that a change of features of one component entails the changes in **all** other components (see also Kondracki, 1976; Richling, 1982; Przewoźniak, 1987; Richling and Solon, 1996).

In connection with the definition provided, the proposed procedure of delimitation of geocomplexes will tend to achieve the possibly highest **agreement** among the divisions corresponding to individual components. This situation will be expressed through high values of the indicators of association of components. The regularity mentioned will make it possible to use the statistical measures for verification of the effects of individual stages in the elaboration of the map of geocomplexes. In this case, therefore, the image of the geocomplexes will be the **result** of application of the statistical measures.

---

MEASURES OF ASSOCIATION OF COMPONENTS  
USED IN LANDSCAPE GEOGRAPHY

The most popular methods of expressing the association between the components of landscape are:

- the association strength indicator,
- the multi-spatial indicator of association,
- the entropy indicator of dependence.

The association strength indicator is expressed with the formula (Richling, 1982):

$$W_{x,y} = P_{x,y}/P_x \text{ when } P_x < P_y, \text{ or}$$

$$W_{x,y} = P_{x,y}/P_y \text{ when } P_x > P_y,$$

where:  $P_{x,y}$  is the area within which the features  $x$  and  $y$  appear together (e.g. brown soils and clays),  $P_x$  is the area of appearance of the feature  $x$  (e.g. brown soils), and  $P_y$  is the area of appearance of feature  $y$  (e.g. clays).

The value of this indicator ranges from 0 (lack of association) to 1 (full association).

The multi-spatial indicator of association, presented in different variants by various authors, is expressed in several forms. Here, we will quote its definition after Pietrzak (1989):

$$\rho = (a - 1)/((k_1 - 1)(k_2 - 1))^{1/2},$$

where  $a = \Sigma (\Sigma (f^2/n_2)/n_1)$ ,  $f$  is the frequency of co-appearance of the components' features (expressed in % of total surface),  $n_1$ ,  $n_2$  are the frequencies of appearance of the individual components' features (expressed as before), and  $k_1$ ,  $k_2$  are the numbers of types of the components' features.

This indicator makes it possible to determine in general terms the association of two components (like, for instance, relief and soils). Its value ranges from 0 (lack of association) to 1 (full association).

The entropy-based indicator of dependence is calculated from the formula (here after Pietrzak, 1989):

$$K_{AB} = (\Delta H_{AB}/H(AB)) \cdot 100\%,$$

where  $K_{AB}$  is the indicator of dependence of the phenomena  $A$  and  $B$  (say, relief and lithology), expressed in percent, and  $H$  is the entropy of the phenomena.

The components of the formula are obtained in reference to the following assumptions:

1) If two phenomena ( $A$  and  $B$ ) are independent, then the fuzziness (entropy) of their distribution is equal the sum of entropy of distributions of each of them:

$$H(A + B) = H(A) + H(B) = - \sum_{i=1}^m \omega_i \log_2 \omega_i - \sum_{j=1}^m \omega_j \log_2 \omega_j,$$

where  $m$  is the number of features (classes) of the analysed phenomena  $A$  and  $B$ ;  $\omega_i, \omega_j$  are the frequencies of appearance of the features (classes) of the phenomena  $A$  and  $B$ , respectively; with  $\sum \omega_i = \sum \omega_j = 1$ .

2) If the phenomena  $A$  and  $B$  are interdependent, then the entropy of their co-appearance is expressed by the formula

$$H(AB) = - \sum_{i=1}^m \sum_{j=1}^m \omega_{ij} \log \omega_{ij},$$

where  $\omega_{ij}$  is the frequency of co-appearance of the features (classes)  $i$  and  $j$ .

3) Since the entropy of the system of independent phenomena is not less than the entropy of the interdependent phenomena, hence

$$\Delta H_{AB} = H(A+B) - H(AB).$$

The value of the interdependence indicator ranges between 0 and 100%.

As mentioned at the beginning (Section 1), the indicators here outlined are used in the analysis of the maps of geocomplexes, treated as the source of data. In the further part of the present paper they will be used in a new role, namely for the verification of boundaries of the particular types of landscape settings (geocomplexes). Before that, though, a new, very simple indicator of association of components will be introduced. Its construction does not make reference to statistics, but only to the intuition of space as the set of points, quite appropriate for the geographers.

#### THE INDICATOR OF THE SPATIAL CONCENTRATION OF ASSOCIATION

The procedure here proposed is based upon the measurement of the area occupied by the particular forms of phenomena. It is composed of two stages (Ostaszewska, 2002).

The first stage consists in comparing the area of co-appearance of the two component **features** considered (like brown soils and clays) with the total area occupied by these two features. The geometrical sense of the operation can easily be represented in terms of the set theoretical calculus. The co-appearance of the features corresponds to the product of two sets, while the total area — to their sum.

In the calculation of the value of the indicator the same database can be used as in the case of other measures applied in landscape geography, presented before. The indicator proposed is expressed through the formula

$$W_{ij} = P_{ij} / (P_i + P_j - P_{ij}),$$

where  $P_{ij}$  is the area of co-appearance of the features indexed  $i$  and  $j$  (e.g. brown soils and clays),  $P_i$  is the area of appearance of feature indexed  $i$  (e.g. clays), and  $P_j$  is the area of appearance of the feature indexed  $j$  (e.g. brown soils).

If the areas  $P_i$  and  $P_j$  coincide fully, the indicator takes the value of 1. If the two areas have no common point, the indicator value is 0. In the situation, when at least one of the features co-appears not only with the feature considered, the indicator takes a value between 0 and 1. Because of the nature of information provided the indicator can be called the **measure of spatial concentration** of the association considered.

The second stage consists in the generalisation of the indicator describing the degree of concentration of the association of two features (say, brown soils and clays, or podzolic soils and sands), up to the determination of the indicator describing the association of two components (say, soils types and kinds of bedding). For this purpose the previously calculated indicators  $W_{ij}$  are treated as weight coefficients, informing of the degree of concentration of the particular partial associations. The spatial sense of the operation can be imagined as follows: by multiplying the area of co-appearance of the features,  $P_{ij}$ , by the coefficient  $W_{ij}$  we obtain a certain hypothetical area, over which the association in question fully concentrates. When  $W_{ij} = 1$ , the area of full concentration is simply equal  $P_{ij}$ . When  $W_{ij} < 1$ , then the hypothetical area of full concentration is smaller than  $P_{ij}$ . After summing the partial hypothetical areas we obtain the summary area, equal at most the entire surface area analysed. The overall indicator of concentration is calculated as the ratio of the summary hypothetical area and the area analysed:

$$W_{so} = (\sum_{i=1}^m \sum_{j=1}^n W_{ij} P_{ij}) / P,$$

where  $W_{so}$  is the general indicator of concentration of association,  $W_{ij}$  is the indicator of concentration of a pair of component features (e.g. brown soils and clays),  $P_{ij}$  is the area of co-appearance of this pair of features (e.g. brown soils and clays),  $m$  is the number of feature types of the first component (e.g.  $m$  types of soils),  $n$  is the number of feature types of the second component (e.g.  $n$  types of the bedding), and  $P$  is the surface area considered.

Similarly as the partial indicators, the general indicator takes values not bigger than 1 (full concentration over the entire area), and not smaller than 0. Its values are higher when there are numerous partial indicator values equal 1 and when the partial areas, over which the association reached high concentration (high weight), are large. The advantage of the indicator seems to consist in the facility of interpretation through the „stepwise” method (from the specific to the general), and also in the possibility of application in cases, when the use of other general measures is technically infeasible. The simplicity and the geometric principle of construction allows, as well, for the conduct of respective operations with the GIS (e.g. the product of sets corresponds to „and”, while the sum — to „or”).

---

AN EXAMPLE OF APPLICATION OF STATISTICAL INDICATORS  
IN THE PROCEDURE OF OPTIMISATION OF LANDSCAPE CLASSIFICATION

The use of the indicators in the verification of the classification of landscape will be presented on a concrete example.

Fig. 1 shows the setting of components typical for the upland landscapes of Poland. The solid bedding composed of the carbonate (e.g. limestone) rocks is shaped into hills with various gradients of slopes. The convex forms are separated by dry valleys. The flatter slopes are covered with sand, whose thickness increases towards the valleys. The surface is overgrown with the deciduous forests (on the limestones), mixed forests (on sandy slopes), and coniferous forests (on deep sands in the valleys). Associations of the steppe vegetation appear in some places. The soils display a significant dependence upon the rocks of the bedding. On the carbonate rocks they are mainly the rendzinas, on the sandy covers — the brown soils, and on deep sands — podzolic soils. The crevice groundwater appears at important depths and does not influence the quality of the habitats considered. The plants are being supplied with water in this case from the so-called suspended capillary water, originating from precipitation and thawing.

The data of this type can be taken from the commonly accessible source materials, including maps, for instance topographic, geological, and soil-and-agricultural maps.

The purpose of the work is to determine the general naturalist typology of the area considered. The analysis of the spatial associations of the components will be performed with the in-house method. It will be understood that „association” denotes the co-appearance of the **classes** of components' features, indicated in the legend, this co-appearance assumed to be an expression of existence of a relation between them. The relation will be apparent through the situation, in which a change in one of the distinguished classes will entail an **observable** or **demonstrable** change in the remaining ones.

The first step of the method consists in the initial division of the area. It appears that the simplest solution is to make use of one of the so-called formalised methods (see Richling, 1982).

The division of morphology is done according to the criterion of morphometry. Distinction is made of ridges, steep slopes, flat slopes, valley bottoms. The classification of the bedding is done according to the kinds of rocks: limestone (fertile, solid bedding), sands on limestone (permeable, medium fertile bedding), deep sands (permeable, infertile bedding). Vegetation is classified into deciduous, mixed and coniferous forests, and the grassy (steppe) associations, while soils — into rendzinas, brown and podzolic soils.

The second step consists in the elaboration of the databases containing information on the co-appearance of the classes distinguished. The areas occupied by the particular features are shown in Fig. 1.

The third step of the method consists in calculation of the concentration indicator for the particular pairs of features, and calculation of the general

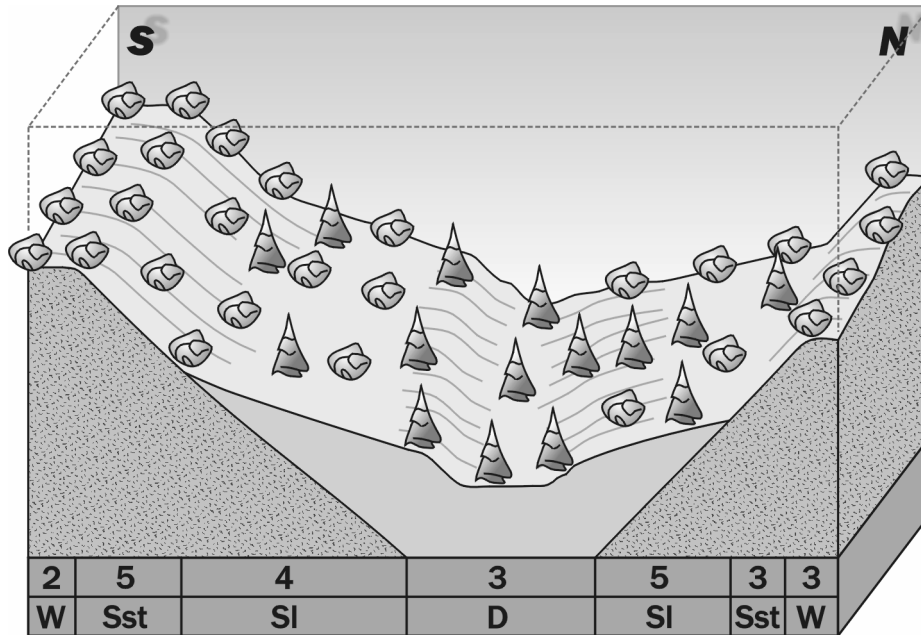


Fig. 1. Schematic cross-section of the upland landscape. W — ridge area; Sst — steep slope; Sl — flat slope; D — valley, 2, 5, 4,... — areas occupied by the relief types. Remaining explanations in the text.

indicator for each of the component pairs. In the case here considered the following values of the association concentration indicator have been obtained: relief-bedding — 0.75, relief-soil — 0.75, relief-vegetation — 0.70, bedding-soil — 1, bedding-vegetation — 0.81, soil-vegetation — 0.81. The values calculated, though high, are not fully satisfactory. Attention should especially be paid to the relatively low concentration of some associations between relief and vegetation, as well as between vegetation and bedding. And so, in particular, the steep slope form is associated with vegetation in a clearly dispersed manner (at 0.38 for both the steppe and the forest), while the concentration of the association of the ridges with the forest attains merely 0.5. The association indicator attains an even lower value for the limestone bedding and steppe (0.23). Worse yet, the same, extremely low value of the concentration indicator characterises the association between steppe and soil. These observations incline to conclude that the criteria of division of the area can (and should) be corrected, so as to yield more consistent landscape settings.

The search for the new landscape division consists in the analysis of the initial material from the point of view of the information given „indirectly”. All kinds of reasoning are admitted (like analogy and deduction), as well as intuition and perceptiveness. The result of the search is subject to verification with the help of the indicator. If the new division of landscape yields a higher

concentration of the associations analysed, then the direction of search can be considered promising. If, however, the association of components turns out to be weaker in the framework of the new division, the attempt should be considered failed.

In the example here presented the method of „critical observation” was applied. Using this method changes were made in the initial division of relief. The changes consisted in:

— linking into a single class the steep slopes with cool exposure and the ridges; their distinction does not find any reflection neither in lithology, nor in the soil type, nor in vegetation cover;

— establishment of the separate class of steep slopes with warm exposure; distinction of this class is justified by the different character of the natural vegetation cover.

The operation described increases the value of concentration of the associations considered. Thus, for relief and vegetation the maximum value (1) is attained, while for relief and bedding, and for relief and soil — the value higher than before (0.81 for both).

In the subsequent stage of work the method of deduction was applied. Reasoning concerned the features of soils and waste, appearing below the steppe plant association. The starting point for considerations was constituted by relief, and more precisely — by the slope and exposure, decisive for the topo-climatic specificity of the steppe habitat.

The reasoning starts from the — anyway obvious — observation that the solar radiation reaches a steep slope of warm exposure at a bigger angle than that resulting from the latitude. Owing to the significant supply of heat the surface layer of the soil warms up strongly. This entails intensified evaporation, stimulating the uptake of the capillary soil water. Since the soil is formed on the limestone rock, we can expect the soil solution to be rich in calcium kations and the acid carbonates. Intensive evaporation is conducive to the migration of these elements towards the humus horizon. A measurable effect of this process ought to consist in the enrichment of the upper part of the waste and soil with calcium (and carbonates), compensating for the effect of washing away of the bases by the percolating precipitation waters. This phenomenon ought to be limited to the topo-climatic conditions of a steep slope with warm exposure, since only there the conditions exist for a seasonal intensive migration of solutions upward. Thus, on other calciferous areas one should expect the rendzinas, preserving more humidity, and somewhat more strongly acid (especially at the humus horizon). As we accept this conclusion, we can perform the classification of soils and limestone waste into the less acid (secondarily enriched with carbonates) and the more acid (carbonate starting with a certain depth). The conclusion presented can also be reached through analogy, although this way of reasoning is formally weaker: the steppe plant associations grow in Polish conditions on the soils rich in carbonates, and so it can be expected that there is more of them in the bedding of the steppe than under the forests.

If we now accept the result of this reasoning as the premise for the classification of soils and wastes into four types, we obtain a new list of the pairs of components. The classification of features shows four types of units, characterised in the following manner:

- type 1: ridges and steep limestone slopes, with soils of the type of rendzinas, overgrown with deciduous forest;
- type 2: flat slopes, made of limestone covered by sand, with brown soils, overgrown with mixed forest;
- type 3: valleys made of deep sands, with podzolic soils, overgrown with pine forests;
- type 4: steep slopes with warm exposure, made of limestone, with rendzinas secondarily enriched with carbonates, occupied by the steppe (grassy) plant associations.

The measure of spatial concentration for each of the newly established pairs of features is equal 1. The general indicator of concentration of the association among all the component pairs of the area in question is also equal 1. Hence, the division determined satisfies the definition of the geo-complex as a consistent system, that is — the one, in which a change in one of the classes analysed causes a change in all the other (see Section 1 of the paper).

In order to control the correctness of the here obtained criteria of delimitation of the holistic landscape units we can calculate other indicators of association between the components (see Section 2). In the case here considered the following values were obtained:

- the indicator of the strength of association of individual pairs of features is equal in each case 1,
- the multi-spatial association indicator is equal for each pair of components 1,
- the entropy-based indicator of interdependence is equal for each pair of components 100%.

Hence, optimisation of the criteria for the division of the landscape in question into the geocomplexes can be considered confirmed. Since optimisation was carried out with the in-house method, and the premise of the division was the result of the deductive reasoning, the effect of the work must be subject to **in-field** verification. Special attention ought to be paid to the forested steep slopes, on which erosion brought — perhaps — formation of the soil profiles and the variants of plant associations, which cannot be identified just on the basis of the maps available.

## CONCLUSIONS

The study of the vertical structure of landscape with the use of indicators of component association is a simple procedure, relatively fast, and easily verifiable. From the cognitive point of view the biggest advantage of this



group of methods consists in the possibility of in-house establishment and optimisation of the criteria of mapping the natural wholes. The areas satisfying the condition of the maximum concentration of association might serve as the initial **patterns (types)** of the vertical landscape structures, analogous to the patterns — types of soils or the geo-chemical units. An essential limitation in this kind of work is constituted by the scale of the input material. The weakest link is always represented by the least detailed image. As a rule, alas, this applies to the identification of the bedding and the climatic conditions, in which a given system develops. That is why it seems better to carry out the in-house naturalist studies for the large areas, considered on the scale of 1:50,000 or less.

From the point of view of the facility of calculations the simplest methods appear to be worth recommending. Their use guarantees, namely, obtaining **the same results of optimisation** of the geocomplex image as in case of application of the more complicated measures.

#### REFERENCES

- Bezowska G., 1986, Struktura i typy geokompleksów w środkowej części Niziny Południowowielkopolskiej [The structure and the types of geocomplexes in the central part of the Southern Greater Polish Lowland; in Polish], *Acta Geographica Lodzense*, 54.
- Grabowski T., 1990, The indices of force of interdependence in physical geography, [in:] *Ecological Management of Landscape*, Akapit DTP, Warszawa.
- Kondracki J., 1976, *Podstawy regionalizacji fizycznogeograficznej* [Foundations for the physico-geographical regionalisation; in Polish], PWN, Warszawa.
- Ostaszewska K., 2002, *Geografia krajobrazu. Wybrane zagadnienia metodologiczne* [Landscape geography. Selected methodological problems; in Polish], Wyd. Naukowe PWN, Warszawa.
- Pietrzak M., 1989, *Problemy i metody badania struktury geokompleksu* [Problems and methods in the study of the structure of geocomplexes; in Polish], Wyd. Naukowe UAM, Poznań.
- Przewoźniak M., 1987, *Podstawy kompleksowej geografii fizycznej* [Foundations for the comprehensive physical geography; in Polish], Uniw. Gdański, Gdańsk.
- Richling A., 1982, *Metody badań kompleksowej geografii fizycznej* [Methods of study of the comprehensive physical geography; in Polish], PWN, Warszawa.
- Richling A., Solon J., 1996, *Ekologia krajobrazu* [Landscape ecology; in Polish], Wyd. Naukowe PWN, Warszawa.