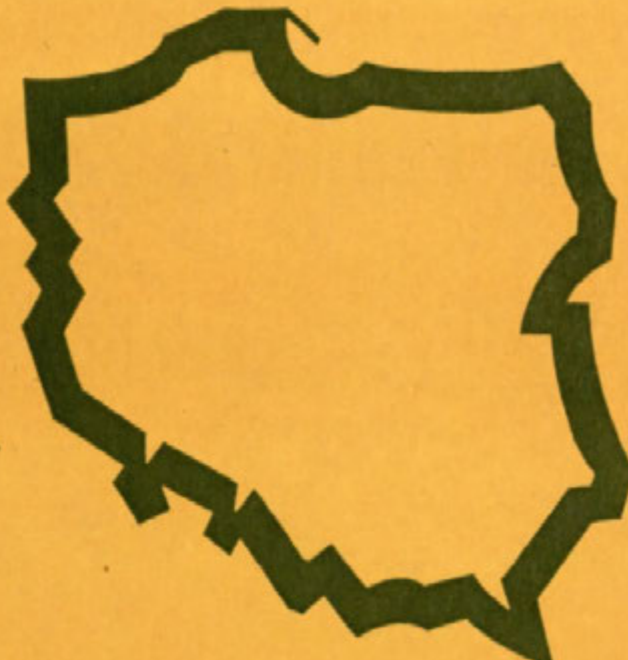
POLISH ACADEMY OF SCIENCES

GEOGRAPHIA POLONICA



31

PWN—POLISH SCIENTIFIC PUBLISHERS

Address of Editorial Board

KRAKOWSKIE PRZEDMIEŚCIE 30
00-927 WARSZAWA
POLAND

Printed in Poland

POLISH ACADEMY OF SCIENCES
INSTITUTE OF GEOGRAPHY

GEOGRAPHIA POLONICA

31

PWN — Polish Scientific Publishers • Warszawa 1975

CONTENTS

THE POPULATION POTENTIAL OF POLAND BETWEEN 1950 AND 1970

KAZIMIERZ DZIEWOŃSKI, PIOTR EBERHARDT, JERZY GAŹDZICKI, ELŻBIETA IWANICKA-LYRA, JACEK KROLSKI AND MAŁGORZATA ŻENIEWSKA

Research on population distribution in Poland has been carried on in the Institute of Geography of the Polish Academy of Sciences since its establishment in 1953. The research started with the analysis of the process of resettlement of the western and northern regions. It resulted in a monograph by L. Kosiński. Next the Institute undertook work connected with the preparation of population maps of Poland based on the model and methods developed by W. William-Olsson and approved by the Commission on the World Population Map of the International Geographical Union. These were dot maps which were prepared for the years 1950 and 1960. In their preparation there arose problems of comparability of the data compiled for different administrative divisions. Because of this the maps of distribution and density of population were prepared on the basis of the standard hexagonal grid.

In the nineteen-sixties other studies were undertaken on the evolution of the settlement network, dealing, for example, with the changes in the population size of cities, as well as with the process of urbanization. The analysis encompassed the entire 20th century in time intervals corresponding to the census years, i.e. 1897/1900, 1921/1925, 1931/1933, 1946, 1950, 1960. At the same time studies were undertaken of population potential, concentration and migrations. The results were published in synthetic form in 1967 as the first volume in the series *The Population of People's Poland*, under the title *The Development and Distribution of Population in Poland during the 20th Century*.

With the results of a new National Census for 1970 available it was decided to prepare a new version of the aforementioned publication. The work is carried on in the framework of research pertaining to the problem "*Basis of the National Physical Plan*". This was not merely a matter of expanding the entire analysis by adding another time cross-section, but also one of applying new theoretical constructions and methodological approaches. The present study is the basic publication of the new materials which were collected during this research and which are to be used for a new synthetic view of the distribution and migrations of population in Poland.

Such early, partial publication is justified by the innovative methodology, specifically the full mechanization of statistical computations and cartographic analysis. The results shed light in an extremely interesting way not only on the question of the distribution of population and changes occurring in that population, but also on the structure of socio-economic space in Poland. The goal of research dealing with population potential was to obtain a general picture of the distribution of population over the national territory at a given time cross-section. The mathematical side of the construction of such a picture is usually based on a synthetic measure which is derived from the family of gravitational models, a metric which takes into account both the size of the population at given points, as well as the distance between these points. Comparison of the maps of potential at different time cross-sections provides the basis for conclusions concerning the dynamics of change in the distribution of population, the shifts of gravitational poles as well as changes in the extent of areas of equal potential.

The analysis of potential was undertaken for the first time a few years ago in the Section on Population and Settlement Geography of the Institute of Geography of the Polish Academy of Sciences. The results were published in the form of articles and in the above-mentioned volume on the distribution of population in Poland.[1] The studies conducted at that time encompassed the period from 1910 to 1960. Population potential maps were compiled for the years 1910/1913, 1931/1933, 1950 and 1960, as were maps showing the changes for the periods under consideration. The length of the period of time as well as the fact that analysis encompassed years preceding the First and Second World Wars provided the opportunity to make interesting comparisons concerning the dynamics of the changes. The work carried out at that time, however, was based on a relatively small number of points. Thus, for example, in 1950 statistical data were compiled with reference to 271 points, and in 1960 the number of points was increased to 326. In this approach the point of reference was identified with the location of the largest city in a *poviat,* data being grouped by *poviats.*

In the preparation of the basic statistical data the value of the primary item — the number of people — was compiled by *poviats* and related to a point corresponding to the geographic location of the largest city or town in a *poviat.* Where the *poviat* surrounded a city excluded from its administrative boundaries, the population figures for the city and the *poviat* were added together. In this way the number of areal reference units was significantly reduced. As a result the cartographic representation of the population potential was strongly generalized.

The research described in the present article[2] encompassed a significantly greater period of time (1950–1970). The analysis was based on decennial censuses and a finer grid of reference points was used than in the studies discussed above, the total number of points exceeding 3200. Such richness of information was obtained by applying a national grid with standard squares with an area of 100 sq.km. A point in the centre of the basic field was taken as the zero point of reference. The use of a smaller operational unit yielded as a result a more precise picture of the population potential. The normalized shape of the basic grid and units was reflected in the geometric pattern of the boundaries of areas with equal potential.

The decision to base the analytic work on a geometric division of the national territory was taken because of the continuous changes in the boundaries of administrative units. Under such conditions comparative studies were practic-

---

[1] K. Dziewoński, L. Kosiński, *rozwój i rozmieszczenie ludności Polski w XX wieku* (The development and distribution of population in Poland during the 20th century), Warszawa 1963; L. Kosiński, Potencjał ludności jako miara jej rozmieszczenia (Sum.: Potential of population as a measure of its distribution), *Przegl. Geogr.*, 37 (1965), 2, pp. 355–368.

[2] This research was financed by the Committee for Space Economy and Regional Planning of the Polish Academy of Sciences.

ally impossible. Another reason for this decision was the great disparity among administrative units with respect to their size (the surface areas of *gromadas* in 1970 ranged from 6 to 360 sq.km) and shape. The advantages of a geometric grid of reference in the investigation of various phenomena have been discussed widely in the literature. In Poland a good deal of attention has recently been paid to these problems. For instance in 1972 a conference on "Problems of Basic Fields as a Framework for the Compilation and Processing of Data in Cartography" was organized in Lublin by the Institute of Geography of the Polish Academy of Sciences and the Department of Cartography of the Geographical Institute of the M. Curie-Sklodowska University.

If we omit the experimental works of F. Uhorczak and his students, a grid with units normalized both in area and shape was first applied in compiling the density maps that were based on the dot population maps according to the concepts of W. William-Olsson. The results were published in the aforementioned volume *The Development and Distribution of Population in Poland in the 20th Century* by K. Dziewoński and L. Kosiński. The basic unit of the system applied there was hexagonal in shape with an area of 100 sq.km. The imposition of the grid on the dot map was arbitrary, but the same for the analysis of the phenomena in 1950 and 1960. As a result fully comparable maps were obtained, and they were used as a basis for defining the changes which had taken place in the distribution of population between 1950 and 1960.

In the research presented here a geometrical grid of standard units has been applied for the first time to the analysis of population potential. This step was taken because of the positive features of normalized fields discussed above. In the light of the experience gained during the preparation of population density maps based on a hexagonal grid, a square corresponding to an area of 100 sq.km was assumed as the basic field in the new grid. The following substantive considerations justified the decision to construct a new grid even though a hexagonal grid was already available. With the use of squares it is possible to aggregate basic units by combining adjacent units into larger ones or to disaggregate them by dividing a basic unit into smaller ones of the same shape (squares). This property makes it possible to adjust the size of the basic unit to the specific features of various phenomena to be studied and to the degree of generalization needed for the results. For this reason a precisely constructed grid of basic squares which corresponds roughly to the number of the smallest administrative units — *gromadas* was used in the subsequent analysis. The hexagonal grid does not provide similar possibilities.

Another positive feature of the new system is the fact that it is possible to identify precisely the location of each basic unit by the use of geographical coordinates. On the one hand this permits the superimposition of the grid of squares on a map of Poland drawn in any projection and on the other hand it enables us to correlate the surface of basic units with the administrative units. As a result it is possible to compute the value of a given feature for each unit not only by counting dots in the dot map, but also on the basis of the numerical data derived directly from statistics.

It is logical to assume that basic units, identified by the rows and columns of the grid, should be considered as reference units in collecting or computing material for a data bank.

The grid of squares used here can also serve as the basis for feeding the data into computers for further transformations. In the analysis presented here this possibility was exploited with positive results. The ODRA 1204 computer not only performed all the mathematical operations, but also converted the potential values into a two-dimensional chart.

CONSTRUCTION OF THE GRID OF SQUARES

Dr. K. Michalik of the Institute of Geodesy and Cartography of Warsaw Technical University was responsible for the technical side of the construction of the geometric grid of normalized units. After consulting with the research group and the cartographers, he came to the conclusion that the best projection for the map of Poland — both in terms of the features being analysed and in terms of the grid — would be that of an equivalent cylindrical projection of a rotative ellipsoid. That projection is given by the following equations:

$$\bar{r}(\varphi,\lambda) = \begin{cases} x = \dfrac{a\cos\varphi\cos\lambda}{\sqrt{1-e^2\sin^2\varphi}} \\ y = \dfrac{a\cos\varphi\sin\lambda}{\sqrt{1-e^2\sin^2\varphi}} \\ z = \dfrac{a(1-e^2)\sin\varphi}{\sqrt{1-e^2\sin^2\varphi}} \end{cases} \tag{1}$$

$$\bar{r}' = \begin{cases} x = \dfrac{a(1-e^2)}{2}\left\{\dfrac{\sin\varphi}{1-e^2\sin^2\varphi} + \dfrac{1}{2e}\ln\dfrac{1+e\sin\varphi}{1-e\sin\varphi}\right\} \\ y = a\cdot\lambda \end{cases}$$

where:

$(\varphi\cdot\lambda) \in \Omega$
φ, λ — are the parameters of the grid of geographic coordinates on the rotative ellipsoid;
Ω — is the domain of the given class of projections;
$r(\varphi, \lambda)$, $r'(\varphi, \lambda)$ — are vectors of the functions;
$a, b,$ — are the lengths of the semiaxes of the ellipsoid:

$$e^2 = \frac{a^2-b^2}{a^2}$$

A characteristic feature of this projection is the isometric (true) projection of the equator $\varphi = 0$. In order to achieve isometric projection of latitude $\varphi = \varphi_0$, equation (1) was changed to:

$$\bar{r}(\varphi,\lambda) = \begin{cases} x = \dfrac{a\cos\varphi\cos\lambda}{\sqrt{1-e^2\sin^2\varphi}} \\ y = \dfrac{a\cos\varphi\sin\lambda}{\sqrt{1-e^2\sin^2\varphi}} \\ z = \dfrac{a(1-e^2)\sin\varphi}{\sqrt{1-e^2\sin^2\varphi}} \end{cases} \tag{2}$$

$$\bar{r}' = \begin{cases} x = \dfrac{a^2(1-e^2)}{2N_0\cos\varphi_0}\left\{\dfrac{\sin\varphi}{1-e^2\sin^2\varphi} + \dfrac{1}{2e}\ln\dfrac{1+e\sin\varphi}{1-e\sin\varphi}\right\} \\ y = (N_0\cos\hat{\varphi})\cdot\lambda \end{cases}$$

where: $N_0\cos\hat{\varphi} = \dfrac{a\cos\hat{\varphi}}{1-e^2\sin^2\hat{\varphi}_0}$ denotes the radius of the parallel of the rotative ellipsoid corresponding to the geographic latitude $\varphi = \hat{\varphi}_0$. This is a case of a so-called equivalent cylindrical projection of a rotative ellipsoid on a secant cylinder along a given parallel $\varphi = \hat{\varphi}_0$.

The projection (equation 2) and the cartographic grid of squares with basic fields of 100 sq.km. was adjusted to the territory of Poland by establishing the values of the parameters $\varphi = \hat{\varphi}_0$ and $\lambda = \hat{\lambda}_0$.

Since the territory of the country on the rotative ellipsoid encompasses the area between 49° and 55° north latitude and 14° and 25° east longitude an effort was made to minimize metric distortion in this area. Thus $\hat{\varphi}_0 = 52°$ was taken as the parallel to be projected isometrically.

The square cartographic grid was projected in such a way that the image of latitude $\varphi = \hat{\varphi}_0 = 52°$ and longitude $\lambda = \hat{\lambda}_0 = 19°$ constituted the axes of the grid.

Because of the shape of the basic fields the grid was constructed in terms of Cartesian coordinates in such a way that the dimensions of each field were 10 × 10 km. As a result of the area equivalence of the projection, both the fields of the grid (the squares) and the curvilinear trapezoid corresponding to them on the rotative ellipsoid have identical areas equal 100 sq.km. The form of the cartographic grid constructed according to projection (2) is illustrated in Figure 1. The darker parallel indicates the location of the area of minimal distortion.

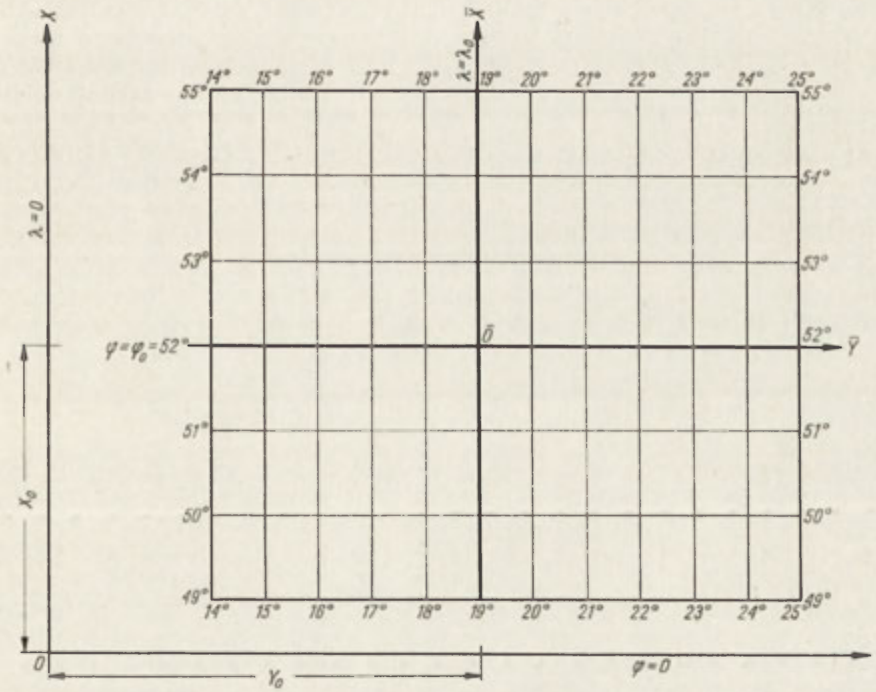


Fig. 1. Cartographic grid and the grid of squares used in research

The cartographic grid illustrated in Fig. 1 consists of selected meridians and parallels of latitude of Krasowski's rotative ellipsoid with dimensions $a =$ $= 6{,}378{,}245$ m, $b = 6{,}356{,}863$ m and $e = 0.0818133$ m. For the given values of

a and e and with the given value of $\varphi_0 = 52°$, the distances x_0 and y_0 were computed:

$$x_0 = x(\hat{\lambda}_0) = \frac{1}{N_0 \cos \varphi_0} S(\hat{\varphi}) = 8{,}114.775 \text{ km}$$
$$y_0 = (N_0 \cos \hat{\varphi}_0)\lambda_0 = 1{,}304.896 \text{ km} \tag{3}$$

where:

$$S(\varphi_0) = \frac{a^2(1-e^2)}{2} \left\{ \frac{\sin \varphi_0}{1 - e^2 \sin^2 \hat{\varphi}_0} + \frac{1}{2e} \ln \frac{1 + e \sin \varphi_0}{1 - e \sin \hat{\varphi}_0} \right\}$$

The values x_0 and y_0 are the coordinates of the centre of Poland in the equivalent cylindrical projection of the rotative ellipsoid onto a plane. These coordinates were taken as the origin of a set of Cartesian coordinates $(\bar{x}, \bar{y})$, in which the image of Krasowski's ellipsoid according to projection (2) is expressed by the equations

$$\bar{x} = x - x_0, \bar{y} - y - y_0. \tag{4}$$

The next step in the research was the computation of Cartesian coordinates $(\bar{x}, \bar{y})$ for all the key points of the basic fields of the grid. Next their geographic coordinates were marked on the rotative ellipsoid (φ, λ). This was done by an inverse transformation of the function specifying the ellipsoid according to equation (2). In this case the mathematical operations, carried out with the aid of a table of condensed interpolations of the first degree amounted to inverse linear interpolation. The values computed and included in the working materials make it possible to draw grids of basic fields on a map of Poland in any projection.

A grid of squares constructed in this way creates difficulties of interpretation when the values of a feature in the basic fields are to be illustrated graphically by means of isolines. The solution of this methodological problem is possible if the basic fields are shifted vertically by half the length of the side of the square. The resulting equilateral system of axes of interpretation (located between the centres of the basic fields) can serve as the basis for correct drawing of isolines.

STATISTICAL MATERIALS

The basic statistical information for research into potential is the population size at the points of reference. The relevant statistical data for the smallest administrative units (*gromadas*) considered in published sources are contained in the censuses for 1950, 1960 and 1970. In order to use these sources it is necessary to carry out the labourious task of identifying administrative units with basic fields of the grid, i.e. to determine which *gromadas*, towns and settlements belong to particular squares of the grid. Because of the frequent changes in the administrative division of the country, this identification procedure would have to be carried out for each of the three time periods under investigation. Since the authors had at their disposal a dot map of the distribution of Polish population in 1950[3], it was much easier to draw the grid of

[3] This map was prepared by the Institute of Geography of the Polish Academy of Sciences in cooperation with the Commission of the World Population Map of the International Geographical Union. It was published in a slightly different form in: K. Dziewoński and L. Kosiński, Rozmieszczenie ludności w Polsce w XX wieku

squares on that map and read off the values of the feature directly from the map. The dot map was executed in the scale 1 : 500,000 and the information recorded on it included: the rural population, represented by dots with the value of 1 point = 200 persons and distributed topographically on the basis of a map of settlements and built-up areas; and the urban population, represented by circles each of which was the projection of a sphere of volume proportional to the population of the given form of settlement. The administrative boundaries of the unit of settlement were taken into consideration in determining urban population figures and, if the unit lay on the boundary of two or more basic fields, the value of the feature in the squares was set in proportion to the area assigned to a given field. The rural population figures were determined by the product of the number of points and the value of each point (200). The total population of a basic field was the sum of its urban and rural population.[4]

Since the major numerical data of the 1960 Census, including the population figures for the *gromadas,* were computed by the Central Statistical Office (GUS) on the basis of 1970 administrative boundaries, the identification of basic fields had to be carried out only for one administrative division of the country, i.e. for that of 1970. This identification was carried out for the present research to an accuracy of 5% of the area of the *gromada.*

Because of the extent of the area under investigation, the analytic work was carried out in a relatively large scale (1 : 2,000,000) and the *gromadas* thus constituted small areas (the average in 1970 was 56 sq.km). This permitted the assumption of an even distribution of population with the boundaries of each administrative unit. Thus the population in each basic field was defined as being proportional to the percentage of the area of the given administrative unit contained in the square, i.e. if 60% of the area of a *gromada* constituted a given square of the grid, then that same percentage of the population of the *gromada* was used to compute the population figure for the basic field. Determining the total population for a square was a matter of a certain number of simple mathematical operations. This stage of the investigation is documented in a card-file of squares, in which each basic field is coded by means of a number indicating the row of the grid in which the square is located and two letters indicating the appropriate column of the grid. The following table illustrates the principle of the code:

AA AB AC ... AZ BA BB BC ... BZ CA CB CC ... CS
1
2
3
⋮
66

A field designated in this way, which could easily be identified, was characterized in terms of the administrative units constituting it, e.g. square 27 BN contains 60% of the area of the *gromada* of Dobrzyków, 60% of the area of the *gromada* of Juliszew, 10% of the area of the *gromada* of Słubice, 5% of the area of the *gromada* of Sanniki, 5% of the area of the *gromada* of Lipińskie (*poviat* of Gostynin) and 70% of the city of Gąbin.

(Sum: Distribution of population in Poland in the 20th century), *Przegl. Geogr.*, 36 (1964), 1, pp. 3–36. The investigation of potential was based on the first, unpublished projection of the map.

[4] W. Warntz, *Toward a geography of price. A study in geoeconometrics,* Philadelphia 1959; *Macrogeography and income fronts,* Philadelphia 1965.

Each basic field has its card, which lists, in addition to the administrative units, the population figures for the years 1960 and 1970, determined according to the principles described above, and the population figure computed from the dot map of population for 1950. The card also indicates the differences in population between the successive decades. The values of the basic feature thus

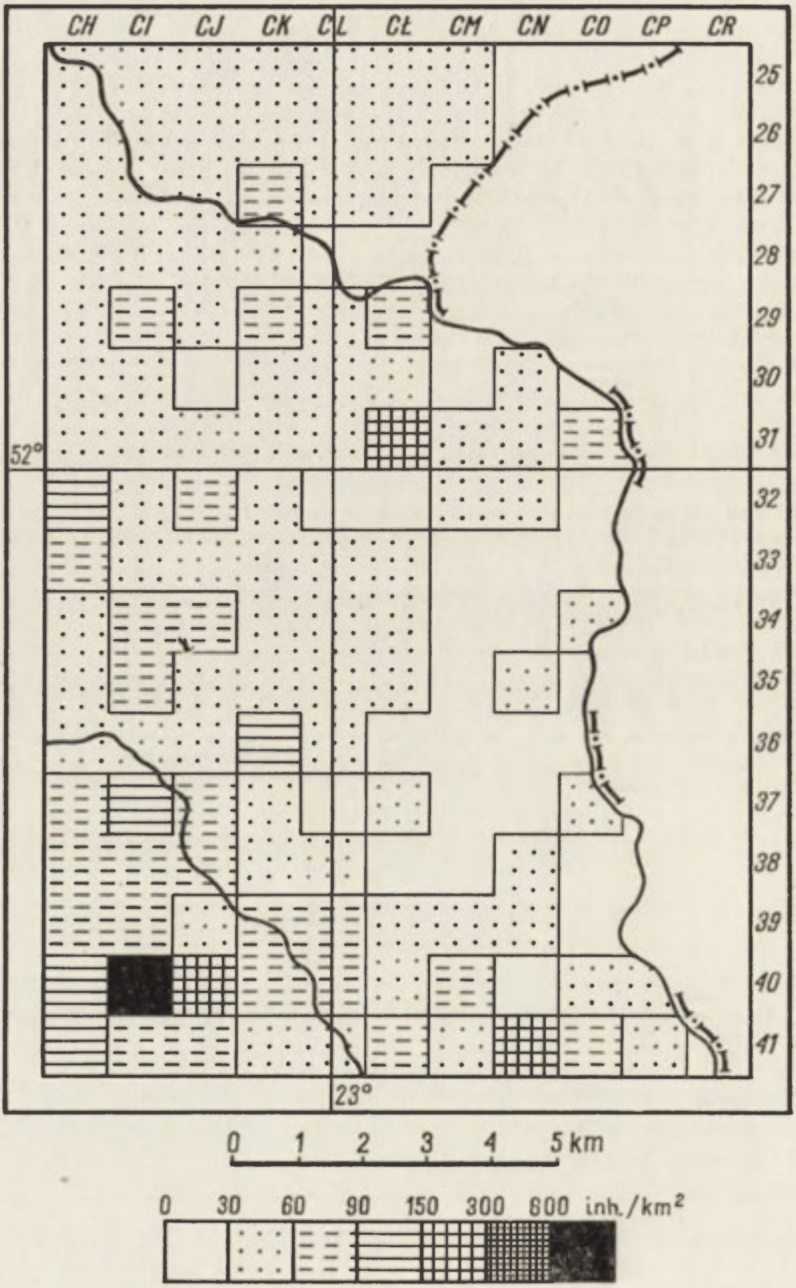


Fig. 2. Density of Poland's population in 1950 expressed by the method of squares (a section)

obtained were then entered on the map with the grid of squares and, in accordance with the method for computing the potential value, were connected with the reference point (the centre of the square). The map prepared in this way served as an input to the computer.

At the same time the resulting numerical data were used to prepare a series of population density maps. A fragment of such a map is presented in Fig. 2. This illustration is given as an example of the next stage in the effort to achieve a synthetic picture of population distribution — a picture that would be more general than the dot map and more detailed in comparison to the population potential map.

COMPUTATION OF THE POPULATION POTENTIAL

Potential in the point model of population distribution. The population potential P_{ij} at a point i resulting from the existence of a population L_j at a point j can be expressed by the equation:

$$P_{ij} = \frac{L_j}{D_{ij}} \tag{5}$$

where D_{ij} is the distance between point j and point i.

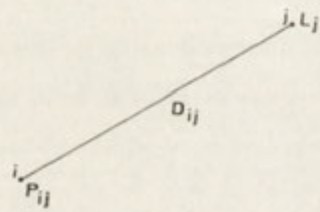


Fig. 3.

If the population is distributed among n points ($j = 1, 2, \ldots n$), then the potential P_i at the point i ($i \neq j$) is the sum of the potentials given by equation (5):

$$P_i = \sum_{j=1}^{n} P_{ij} = \sum_{j=1}^{n} \frac{L_j}{D_{ij}} \tag{6}$$

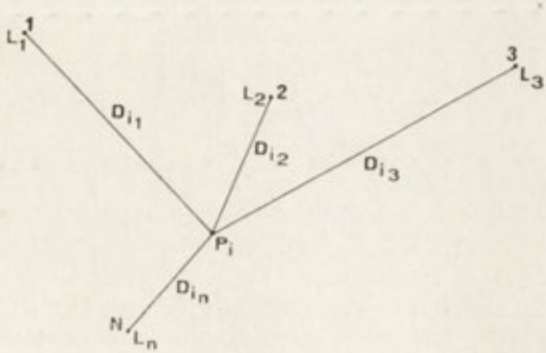


Fig. 4.

Potential in the planar model of population distribution. Applying the point model of population distribution presented above, we do not take into consideration the population at point i in computing the potential P_i.

Let us then consider the population potential P_{ij} at the point i resulting from the presence of a population L_j in the region s_j.

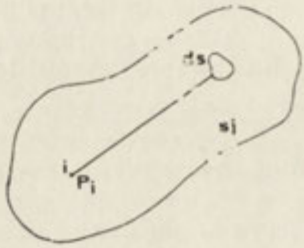


Fig. 5.

Assuming an even distribution of population in the region s_j, this potential can be given by the equation

$$P_{ij} = \delta_j \iint_{s_j} \frac{ds}{\sqrt{(x-x_i)^2+(y-y_i)^2}} \tag{7}$$

in which $\delta_j = \dfrac{L_j}{a_j}$ is the population density in the region s_j with area a_j and x_i, y_i are the coordinates of point i. This equation can be given in a form analogical to (5)

$$P_{ij} = \frac{L_j}{A_{ij}} \tag{8}$$

where:

$$A_{ij} = \frac{a_j}{\iint_{s_j} \frac{ds}{\sqrt{(x-x_i)^2+(y-y_i)^2}}} \tag{9}$$

Computing the potential P_i at the point i on the basis of a set of n areas s_j $(j = 1, 2, ..., n)$ we get

$$P_{ij} = \sum_{j=1}^{n} \frac{L_j}{A_{ij}} \tag{10}$$

Fig. 6.

Equation (10) makes possible the conceptually correct inclusion in the computation of potential of the influence of the population contained in the area in which point i is located.

Computation of potential in the basic fields. The grid of basic fields used in the present study consists of squares with sides $b = 10$ km.

The potential was computed for the centre of each square, using equation (10). In computing the potential for midpoint i of square s_i the values of the factors obtained from equation (9) for s_i and the eight neighbouring squares ($s_{i+1}, s_{i+2} \dots s_{i+8}$) were used. In particular for the square s_i the value $A_{ii} = 2.8365$ was obtained. For all other squares further removed from s_i the differences between the values of A_{ij} and the distances D_{ij} are negligible, so that it was assumed that $A_{ij} = D_{ij}$ in order to shorten the time of computation.

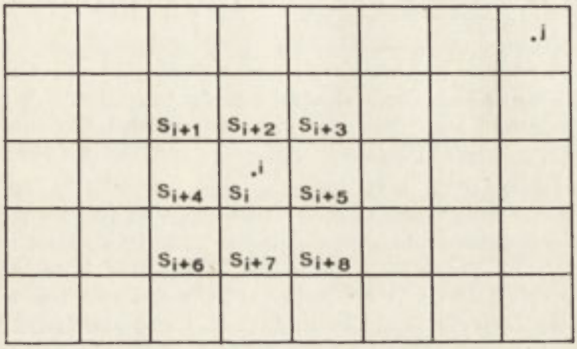


Fig. 7.

The computations were carried out on an ODRA 1204 computer, using a program (POTENCJAŁ) written in the algorithmic language ALGOL 1204 and partially in JAS. By using JAS for that part of the program that involved cyclically repeating computations, it was possible to reduce the time during which the computer was doing its calculation.

The output of the POTENCJAŁ program were numerical data characterizing the values of population potential within a basic field 100 sq.km in area. These data were then transformed, with the aid of a set of appropriate programs, into graphic form in such a way that each numerical value was represented graphically by means of a shading composed of typographic symbols.

A printer of the OPTIMA type was used to print the population potential maps. Because of the technical properties of this machine each map is printed in three parts, while each basic field is printed by means of 15 symbols (3 rows of 5 symbols).

Using this method population potential maps of Poland for the years 1950, 1960 and 1970 were produced in two graphic versions: 1) maps consisting of square basic fields printed according to the given shading and 2) maps consisting of square basic fields in which numerical values of the population potential are printed.

An additional program POTENCJAŁ RC was also used to compute changes in population potential within the basic fields. This made it possible to produce analogous maps of the differences in population potential for the years 1950–1960, 1960–1970, 1950–1970.

The maps of population potential and of differences in population potential obtained in this way were photographically reduced to a scale of 1 : 1,500,000 on a LITHOTEX copier and reproduced xerographically on a SAWA copier.

SPATIAL ANALYSIS OF THE POLISH POPULATION POTENTIAL IN THE YEARS 1950–1970

A. POPULATION CHANGES IN POLAND IN THE YEARS 1950–1970

In the period 1950–1970 very significant demographic changes took place in Poland. These changes were both quantitative and structural. A general description of the population changes during the twenty-year period under investigation will make it easier to interpret the population potential maps included in the article. (The maps have been graphically transformed for the purposes of reproduction in a smaller scale.)

In the years 1950–1970 the population of Poland increased by 7,581,000 i.e. by more than 30%. This increase is equal to the population of Sweden, and when this growth is put in terms of density of population per sq.km it gives an increase of 24 inhabitants per sq.km. In analysing this increase, which is very high for European standards, one should remember that the population in 1970 was close to that of 1939 and 1914. On the other hand the increase in 1950–1970 was similar to the population growth between 1921 and 1939. In both cases the increases involved offsetting war losses. In the eighteen-year interwar period the Polish population increased by 8.0 million (an increase of 21 inhabitants per sq.km). Despite the quantitative similarity, the two above-mentioned periods differ significantly. In the years 1921–1939 both the population in cities and towns and the population in the countryside increased by 4 million, while in the period 1950–1970 only the cities and towns showed an increase in population.

This obviously had completely new socio-economic and spatial implications. The years 1950–1970 were more than anything else a period of great structural changes when the percentage of non-agricultural and urban population grew rapidly. This fact is a significant feature of contemporary Polish demographic changes. The dynamics of change in the urban and non-agrarian population is important in that the growth of this population is reflected in the increasing concentration of population on a national scale and in the changes in the population potential of territorial units.

TABLE 1. Population of Poland in 1950–1960 and 1960–1970

Type of population	Population (in thousands)		Population increase (1950–1960)		Population (in thousands)		Population increase (1960–1970)	
	1950	1960	in 000's	%	1960	1970	in 000's	%
Urban	9,605	14,219	4,614	48.0	14,219	17,007	2,788	19.5
Rural	15,403	15,557	154	1.0	15,557	15,582	25	0.2
Non-agrarian	13,016	18,125	5,109	39.3	18,125	22,966	4,841	26.7
Agrarian	11,598	11,281	−317	−2.7	11,281	9,623	−1,658	−14.7
Total	25,008	29,776	4,768	19.0	29,776	52.589	2,813	9.4

Table 1 illustrates population changes in Poland for the two decades discussed in this article. From Table 1 one can conclude that the two decades 1950–1960 and 1960–1970 differ significantly in terms of both demographic development and structural changes. In 1960–1970, as a result of a decrease in the birth-rate, the rate of population increase slowed down. The increase in Polish population in this decade was half of the increase for the decade 1950–

1960. In the second ten-year period there occurred a sharp decrease in the size of the agrarian population, which had diminished insignificantly in the previous decade. The second decade also witnessed a slowing down of the rate of growth of the urban population. A common feature of both decades was the stabilization of the size of the rural population at the level of 15.5 million inhabitants and similar absolute increases in the non-agrarian population (approximately 5 million people in each decade).

The demographic trends outlined here (growth of the urban and non-agrarian population, decrease in the agrarian population etc.) were caused by the industrialization of the country and they brought about great changes in the settlement pattern of the country. The decisive factor in these changes was the growth of the urban population by 7.4 million people during the twenty-year period 1950–1970 (1950–1960: 4.6 million, 1960–1970: 2.8 million).

It should not be forgotten that contemporary urbanization processes in Poland consist not only in the growth of the population of cities and the increasing percentage of urban population, but also in the creation and development of functionally and spatially integrated groups of settlement units. The relatively evenly distributed, hierarchical point array typical for the early phases of urbanization is being replaced by large scale agglomerations with various territorial configurations. The processes of agglomeration are transforming previous settlement patterns. Their spatial results and consequences in Poland have been discussed in detail in the literature so there is no need to analyze the problem here. Cartograms prepared for the Committee on National Physical Planning of the Polish Academy of Sciences illustrate with considerable accuracy the distribution of population, the process of urbanization of the country and the creation of urban agglomerations.

The cartograms, however, showed these problems in a traditional way, i.e. they presented the degree of development of the phenomena, without dealing with the connections and functional interdependencies between individual units of settlement. As is well known, contemporary processes of functional-spatial integration lead to the growth of interdependencies in regional and national settlement patterns, with some evident feed-back effects as well as in units of settlement or spatially integrated groups of units of settlements, i.e., agglomerations. These processes require a systematic approach, i.e., settlement units or groups of units must be studied as sets of elements in constant interdependence. Mathematically speaking, these interdependencies do not have the character of linear functions.

For this reason the analysis of the population distribution of Poland by means of the potential metric can be considered the beginning of systemic investigation of the distribution of population and of the settlement grid.

B. POPULATION POTENTIAL IN 1950

The cartogram presents a relatively regular spatial pattern. Significantly higher potential values (more than 150,000 persons per km) are found in an approximately circular area with the centre in the region of Piotrków Trybunalski and radius of about 200 km which stretches from Włocławek and Pułtusk in the north, to Wrocław in the west, the middle Vistula in the east and the Polish-Czech border in the region of Cieszyn and Racibórz. Within this pattern are three distinct islands with high potential values: the Katowice, Warsaw and Łódź agglomerations. Even in these agglomerations the potential value does not exceed 350,00 persons per km. All of Western Pomerania, the greater part of the Lubuska region and the north eastern areas of Poland have a low potential (less than 100,000 persons per km).

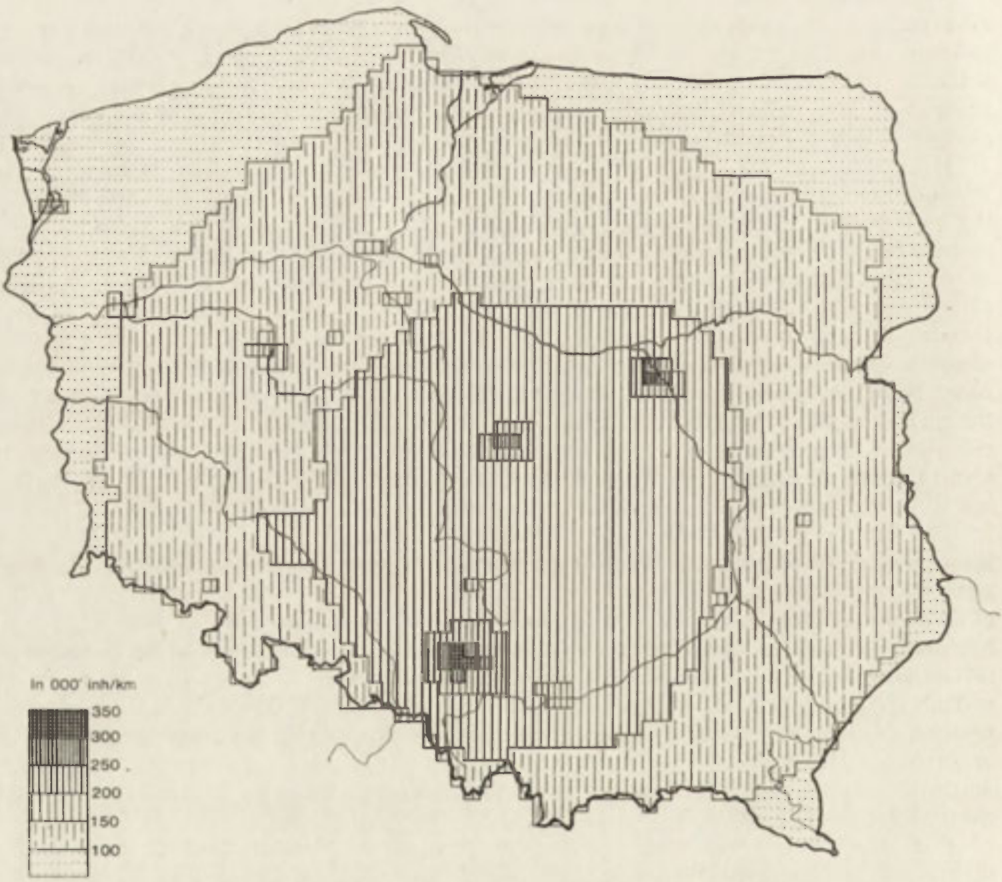


Fig. 8a. Population potential of Poland in 1950 *

* Note: The particular graphic possibilities of the computer used here have led to a certain insignificant distortion of the shape of Poland, while preserving accurate representation of areas. This deformation is not relevant to the question at hand (Ed.).

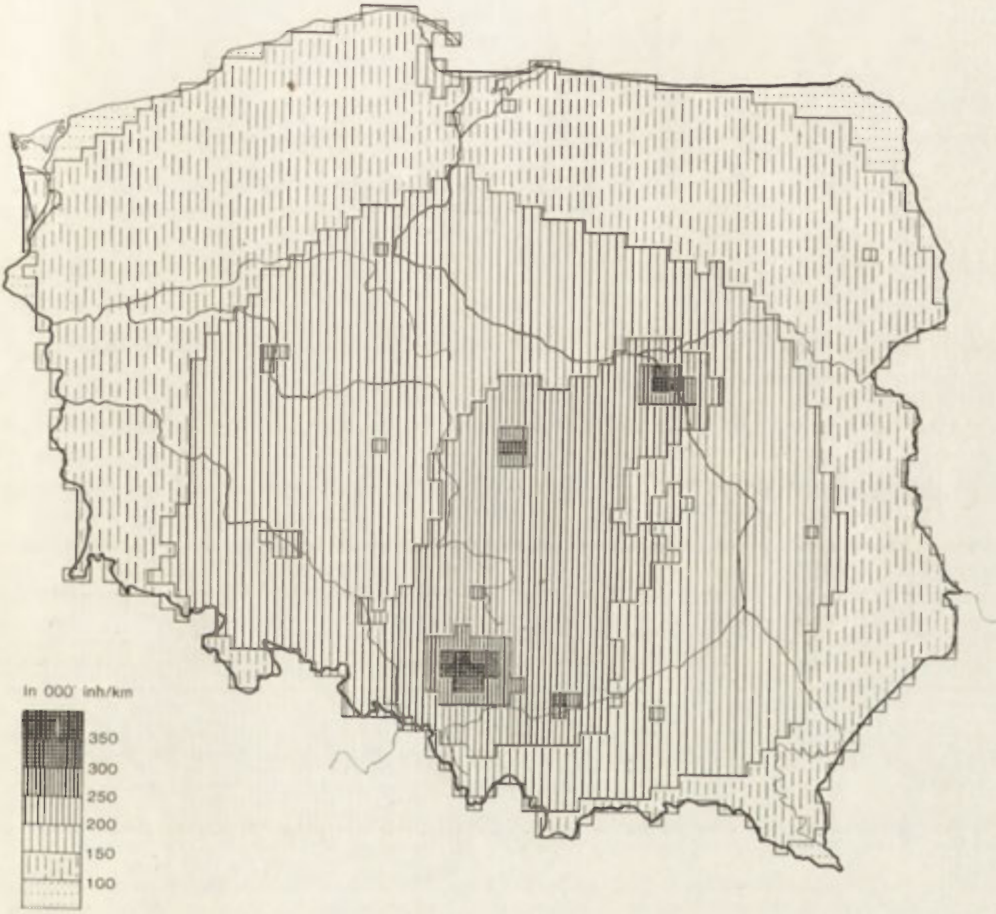


Fig. 8b. Population potential of Poland in 1960

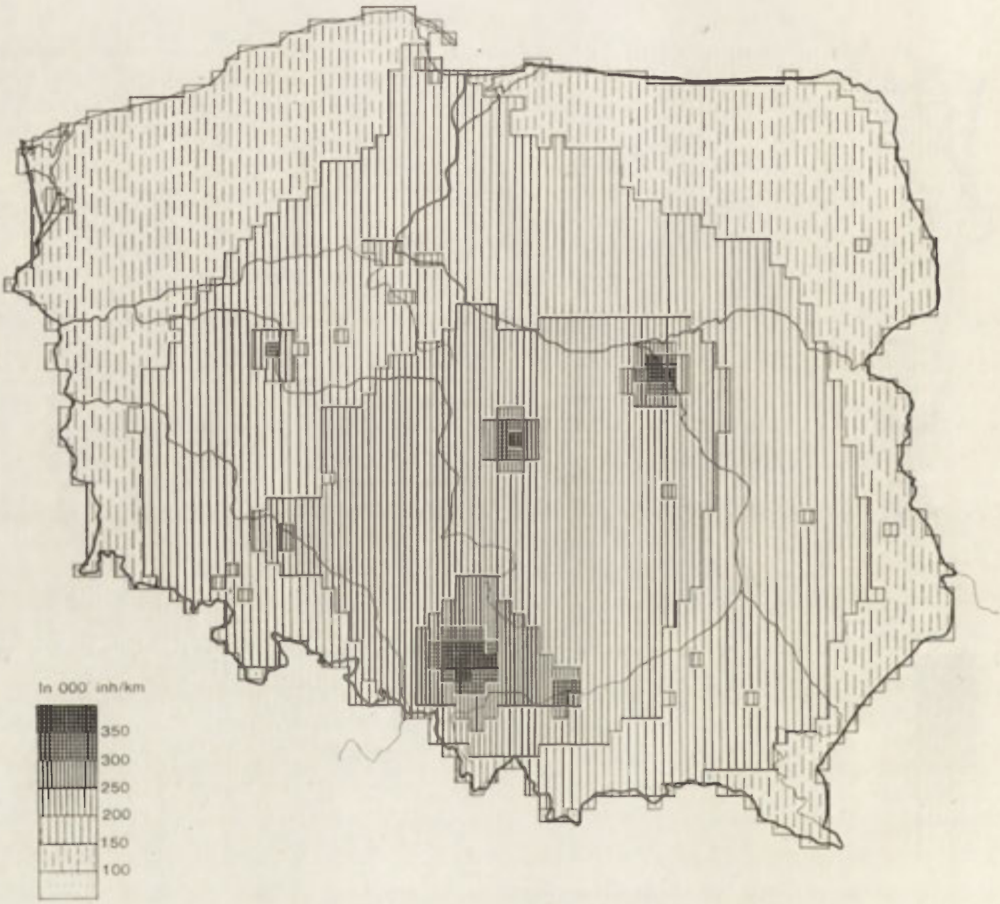


Fig. 8c. Population potential of Poland in 1970

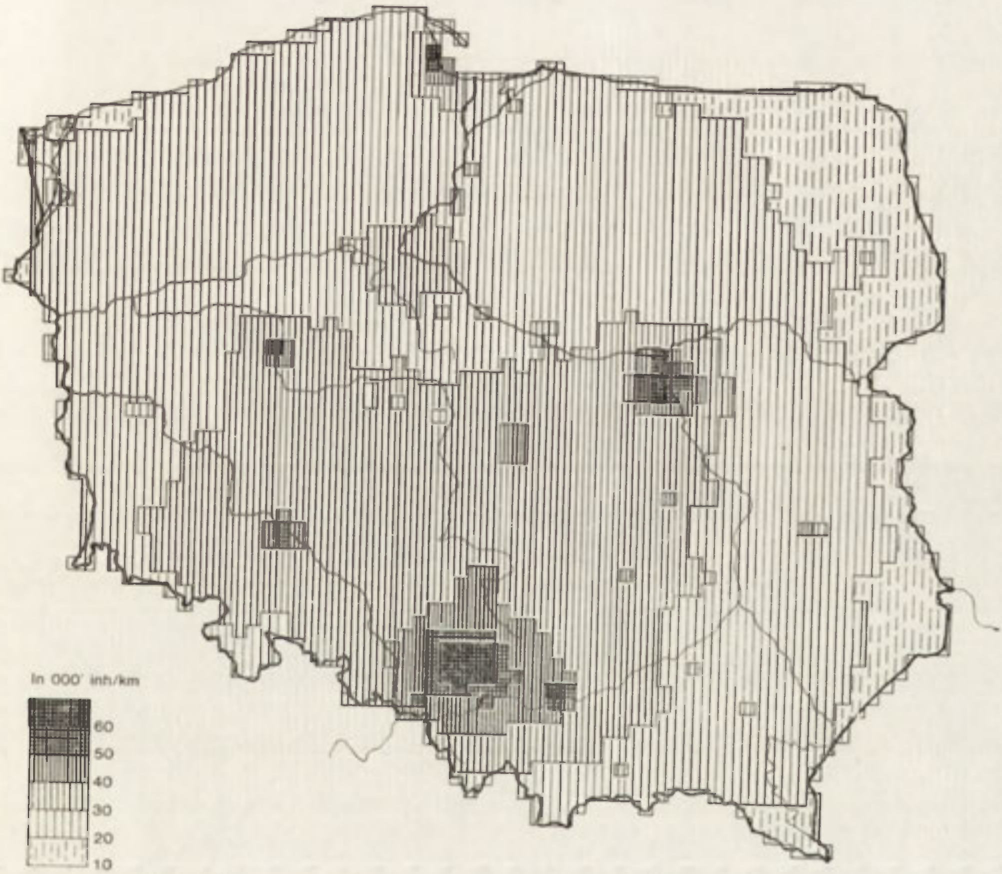


Fig. 8d. Changes of the population potential of Poland in 1950–1960

Fig. 8e. Changes of the population potential of Poland in 1960–1970

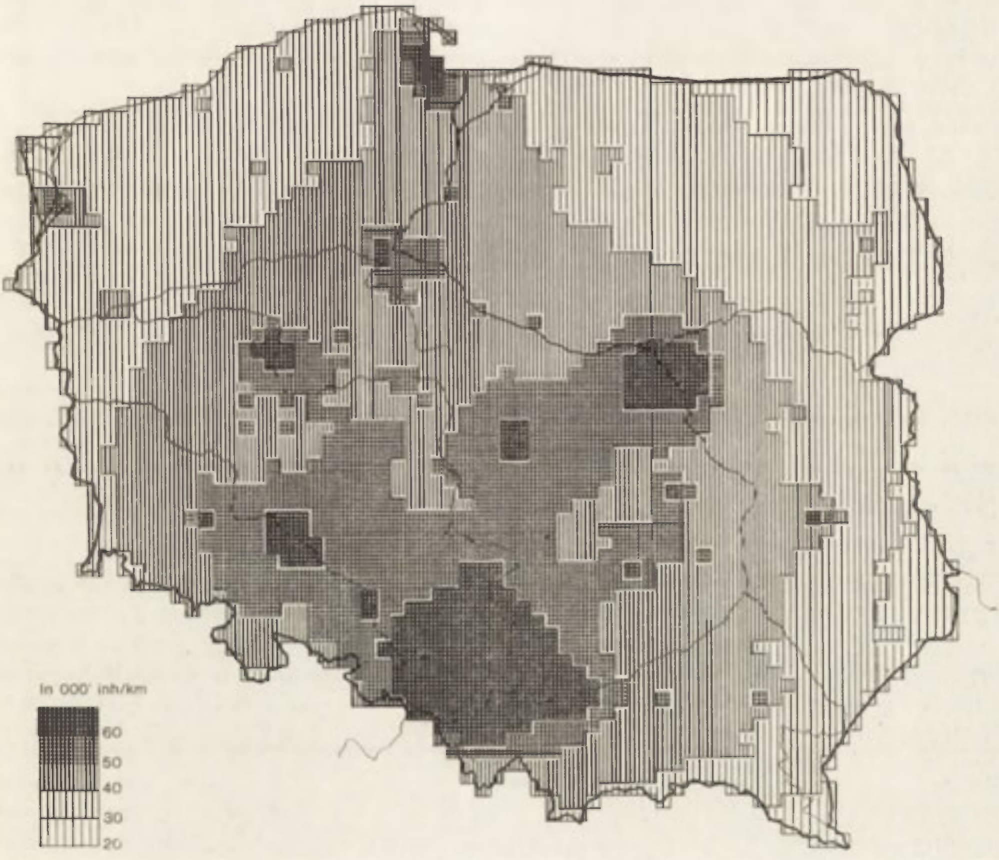


Fig. 8f. Changes of the population potential of Poland in 1950–1970

C. POPULATION POTENTIAL IN 1960

Except for the small border areas in the region of Suwałki and the island of Wolin the entire territory of Poland had in 1960 a potential of more than 100,000 persons per km. The circle enclosed by an isoline encompassing potential of more than 150,000 persons per km, which was mentioned in the description of the 1950 potential map, significantly expanded its area to contain some 2/3 of the territory of Poland. The circle contains a compact area with potential of more than 200,000 persons per km, bounded by Warsaw and Łódź on the north and Upper Silesia and Cracow on the south. The nearly square Upper Silesian area with its noticeably high potential deserves particular attention. In addition to the high-potential areas around Warsaw, Łódź and Cracow, the only other large urban areas to show up on the cartogram were Poznań and Częstochowa.

D. CHANGES IN POPULATION POTENTIAL BETWEEN 1950 AND 1960

In the period 1950–1960 the greatest changes in potential affected the voivodship of Katowice and the northwestern part of the voivodship of Cracow plus the city of Cracow, with the most intensive change affecting the Katowice agglomeration. A significant increase in potential was also observed for the Warsaw agglomeration. The map shows areas in which the increase in potential amounted to more than 40,000 persons per km. These are the clearly indicated agglomerations of Łódź, Wrocław, Poznań and Gdańsk; the cities of Radom, Kielce and Lublin appeared more as points than areas.

In the decade under consideration the whole territory of Poland underwent a growth in potential of more than 10,000 persons per km. Most of northern Poland was encompassed by the isolines of 20,000 and 30,000 persons per km, while central and southern Poland lay between the isolines of 30,000 and 40,000 persons per km.

E. POPULATION POTENTIAL IN 1970

The state of potential in 1970 is the product of the state in 1950 and of the changes of potential in the two successive decades, which were significantly different. The cartogram shows the coalescence of the Upper Silesia, Cracow and Częstochowa areas and the creation of a large area bounded by the isoline representing potential of 250,000 persons per km. The centre of this area is located in the Upper Silesian Coal Basin and has a potential of more than 350,000 persons per km. This area contains more than 3/4 of the territory of Poland. Remaining outside of it are the voivodships of Koszalin, Szczecin and Białystok; most of the voivodship of Zielona Góra and Olsztyn and several *poviats* from each of the following voivodships: Gdańsk, Wrocław, Lublin and Rzeszów.

The agglomerations of Warsaw and Łódź have increased in area and their central portions have a potential of more than 350,000 persons per km. The map also shows small areas with a potential of more than 300,000 persons per km in Wrocław and Radom. Those areas that were bounded in 1950 by the isoline of potential equal to 130,000 persons per km are in 1970 encompassed by the isoline representing potential of 200,000 persons per km. This is also the potential of the Bydgoszcz–Toruń and Gdańsk agglomerations and of the cities of Lublin and Rzeszów.

Finally, it should be noted that the potential map drawn on the basis of 1970 population data did not show any agglomerations in northern and western Poland. The determining factor here was the considerable distance from the areas of higher potential (Upper Silesia, Cracow, Warsaw and Łódź).

F. CHANGES IN THE POPULATION POTENTIAL BETWEEN 1960 AND 1970

The map of potential changes for the period 1960–1970 differs significantly from the previously described map of potential changes for 1950–1960. During the period 1960–1970 the rate of increase of potential slowed down as a result of the previously mentioned demographic changes, i.e., the slower increase in population in Poland. The fundamental difference in spatial configurations is, however, more interesting. During the decade 1960–1970 potential grew evenly over the whole country. It is difficult to find on the map any areas with a significantly more rapid rate of change. This is a reversal of the tendency of the years 1950–1960, when territorial disproportions were significant. Whereas in the first decade the significant increase in potential affected the areas of the greatest agglomerations — primarily the Upper Silesia-Cracow agglomeration — the second decade showed a surprisingly even distribution of increases in potential. This confirms the effectiveness of the policy of deglomeration of the largest urban centres that was carried out in Poland during this time.

G. CHANGES IN THE POPULATION POTENTIAL OF POLAND BETWEEN 1950 AND 1970 AND THE PATTERN OF URBAN AGGLOMERATIONS

The possibility of using population potential measures to evaluate the size and distribution of settlement units or groups of such units has been considered only to a minimal degree in analyses of Polish settlement patterns. The only such attempt was the aforementioned monograph by Leszek Kosiński.

The present study provides much broader possibilities for analysis in this area. It is interesting to compare the map for 1960 published by Kosiński with the present map for the same year. Kosiński's map was based on potential computations for 326 points, while the new map encompasses the relations between about 3200 points. The general picture is similar, which seems to speak well for the simplified assumptions adopted by Kosiński. Nonetheless, the present picture is more accurate and at the same time more transparent and easier to interpret.

The three potential maps (for 1950, 1960 and 1970) and the three maps of population potential change (for 1950–1960, 1960–1970 and 1950–1970), all of which are included in the present article, constitute rich material for comparative study. These cartograms should first of all be compared with maps of population density and dynamics and with maps of the main areas of migration or of socio-economic activity. This kind of correlative analysis will reveal the influence of the distance parameter on the degree of concentration of a number of spatial phenomena.

As an example we shall compare the cartogram illustrating population potential changes in Poland during the years 1950–1970 with a cartogram of the pattern of urban agglomerations in Poland published by S. Leszczycki, P. Eberhardt and S. Herman.

These two maps, despite certain essential territorial analogies, give rather distinct spatial pictures, characterized by a number of differences in territorial patterns. The two cartograms are most similar in their representation of the spatially integrated set of agglomerations of southern Poland, namely the agglomerations of Katowice, Cracow, Częstochowa, Bielsko-Biała and Opole.

Corresponding to the three agglomerations of central Pland (Warsaw, Łódź and Kamienna Valley) the map of changes in population potential shows a huge area with large changes in potential. This area includes Warsaw and the surrounding *poviats* of the voivodship of Warsaw, the major part of the voivodship of Łódź and extends directly in a broad belt to the Katowice region, from which there is a further, narrower extension in the direction of Kielce and Radom.

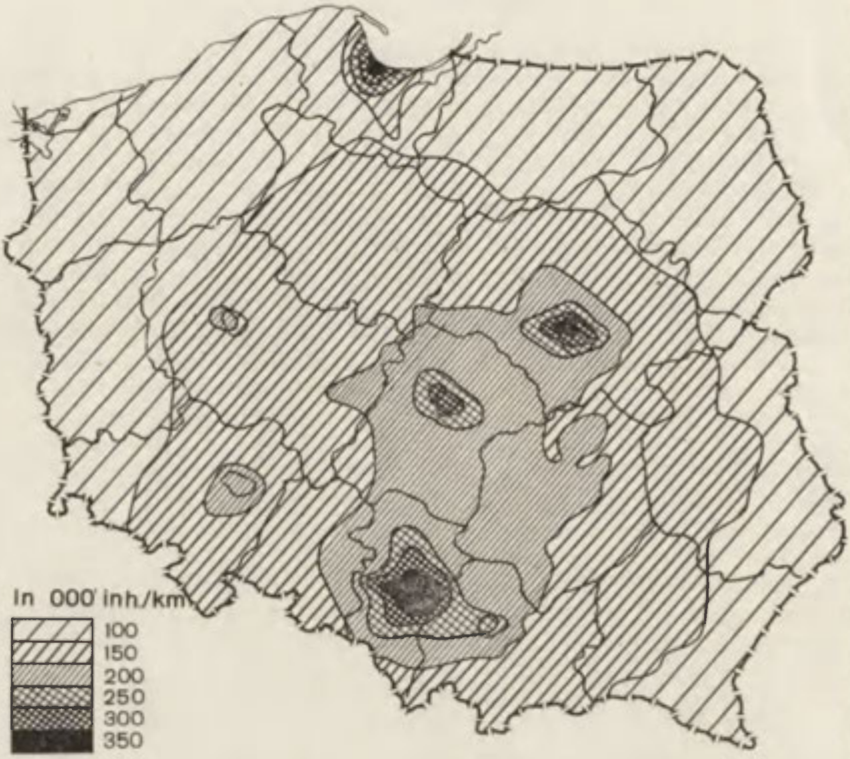


Fig. 9. Population potential of Poland in 1960 after L. Kosiński

The remaining agglomerations show up less clearly on the cartogram of potential changes. The Poznań, Bydgoszcz–Toruń, Gdańsk and Szczecin agglomerations are visible, but not those of the Sudetes Mts., Lublin and Białystok. The two cartographic approaches differ primarily in the fact that the map of agglomerations shows a polycentric spatial pattern made up of the number of unevenly distributed centres, while the map of potential changes for the years 1950–1970 shows two large areas with large (albeit different) changes in potential. One of these areas occupies a territory in southern Poland that stretches from Legnica to Tarnów, including the voivodships of Katowice and Opole, the western and central portions of the voivodship of Cracow and the eastern portion of the voivodship of Wrocław. An integral part of this area is constituted by the voivodship of Łódź together with the Warsaw agglomeration, defined broadly, plus two smaller areas: first, the Kalisz–Ostrów area and the Poznań agglomeration and secondly, the central part of the voivodship of Kielce. The southern and central agglomerations contain areas that show the greatest change in potential in Poland (more than 60,000 persons per km).

The second area is bounded by the isoline of 40,000 persons per km. It contains 2/3 of the territory of Poland. It is wide in southern and central Poland, but narrows significantly to the north in the region of the Bay of Gdańsk.

Spatial relations on the map of potential changes were unavoidably distorted by the distance parameter, but this distortion made more explicit the processes of integration that are usually overlooked or barely visible when traditional cartographic methods are applied.

The results of the comparative analysis suggest that in future investigations of potential the distance parameter might be treated in various ways with a view to concrete theoretical or practical goals. Geometrically expressed distance might be replaced by a time measure or an economic parameter.

It should be emphasized that the analysis of potential changes introduces for the first time the dynamic element of change and growth into the range of problems dealing with the identification and delimitation of urban agglomerations. This property of the analysis was responsible for the non-appearance (within the scale of analytic divisions applied) of the Lublin, Białystok and especially the Sudetes Mts. agglomerations.

The map of population potential in 1970 together with the map of potential changes for the years 1950–1970 do, however, make possible a rather precise delimitation of fully developed and developing agglomerations.

It should be emphasized that the cartographic analysis only partially revealed the differences in potential changes on a regional and local scale that were connected with the concentration of population in the cities and the beginning of the process of depopulation of the countryside in large areas of the country.

Fig. 10. The net balance of population migrations in 1966–1970 after A. Gawryszewski

The map of changes in population distribution and density in the years 1960–1970 show only the beginnings of this process. A. Gawryszewski's map of the net balance of population migrations in 1966–1970 (Fig. 10), shows that as far as net balance is concerned, this process encompasses almost the whole country. Unfortunately, the lack of precise data for 1965 in contrast to the census data for 1950, 1960 and 1970, makes it impossible to prepare maps of potential and its changes for intermediate periods of less than ten years. It also makes it impossible to prepare the time series that would be necessary for mathematical analysis and evaluation of surface trends. Such analysis will have to wait at least until the next census.

Finally it should be emphasized that the potential changes for the period 1960–1970 show for the first time since the war a faster growth in potential (and in size of population) on the outskirts of the largest urban agglomerations than in their central regions. This is a well-known phenomenon in all large urban agglomerations in the world, but in Poland it was not visible after the war because of the rebuilding of urban centers destroyed during the war and the resulting marked growth in the population of those areas.

The investigation of integrative processes and the desire to generalize the most important spatial patterns are playing an ever more important role in studies carried out for physical planning at the national level. For this reason potential should be broadly applied as a spatial measure especially in studies that deal structurally with spatial and economic phenomena and processes.

Maps of population potential and its changes can in fact be interpreted as a first approximation to a synthetic map of the socio-economic space of the country.

The significance from this point of view of the maps already prepared will be more apparent when additional potential maps are compiled, such as maps of national income as created and divided, a map of fixed assets and new investments, as well as a map of the manpower with their professional qualifications taken into account, all computed *per capita*. The works of W. Warntz should be mentioned in this connection: he compiled maps of population potential combined with the size of national income as created by all the countries of the whole world.

A simple comparison of these maps would show their agreement or lack of agreement with respect to the spatial structure of the respective population potentials. This, in turn, would permit the possible choice of one of these maps as a second, better approximation to a synthetic map of the socio-economic space of the country or to the compilation of further, derived maps, based on the potential of values combined in one way or another, which would give an even more precise picture of that space.

Thus the approach that began simply as research in the geography of population would become an essential research tool in the geography of settlement (the delimitation of large urban agglomerations) and even in all of economic geography and would be useful in physical planning. It is to be hoped that in the course of research carried out in the next few years in the framework of the project "*Basis of the National Physical Plan*" these possibilities will be exploited fully.

Institute of Geography
Polish Academy of Sciences, Warsaw

GEOGRAPHIA POLONICA 31, 1975

AN ECONOMETRIC MODEL OF INDUSTRIAL DEVELOPMENT IN POLAND, 1950–1970 *

Lawrence W. Murray, Jr. and Gerald J. Karaska

Geographers and economists have generally concluded that the process of socio-economic growth over space can be expressed as a dynamic set composed of: (1) a subset of growth points,[1] or centres, at which the growth process is initiated; and (2) a subset of dependent areas, or regions, where the growth process has not yet been initiated. Such spatial patterns are well documented in capitalistic economies, and reflect economic and social necessities (i.e., production economies, consumption patterns, the quality of infrastructure, raw material locations, etc.) of concentrating factors of production at specified points in space. Over time, it is theorized, the impetus for growth supplied by the growth centres may spread (trickle down) to the periphery, thereby "flattening" the cost curves over space by improving the infrastructure of the periphery to support the expanded production capacities now located there (Hirschman, 1958; Myrdal, 1957). But as Friedmann (1956) and Hilhorst (1971) note: (1) those dependent areas in closest proximity to the growth centres have the highest probability of benefiting from the spread effects; and (2) the process of extending industrialization and urbanization to the periphery is inextricably interwoven within the decision-making model of the national economy.

Recent investigations (Mihailovic, 1972; Kukliński, 1972; Leszczycki, 1971; Dawson, 1970; Hamilton, 1970; Kruczała, 1970; Róg, 1972) have formalized the theoretical constructs of the socialist development model, and have attempted to empirically support those projections by denoting the socio-economic importance of: (1) industry as a growth pole (Dawson, 1970); (2) the interdependency of the urbanization and industrialization processes (Dziewoński, 1964; Regulski, 1972); (3) preserving natural and human resources (Wróbel and Zawadzki, 1966); and (4) growth centres and industrial regions as the means by which the equalization of socio-economic activities over space, the protection of the natural environment, and a realization of agglomeration economies might be achieved simultaneously (Gokham and Karpov, 1972; Zaremba, 1966; Wojtasiewicz, 1972).

The purpose of this paper is to, first, briefly describe the recent changes in industrial employment in post-war Poland, particularly as these changes refer to the theoretical generalizations which constitute the socialist spatial development model, and, secondly, to present an econometric analysis which models the developmental dynamics of Polish industry over space during 1950–1970.

* We wish to acknowledge the assistance of Professor F. W. Puffer, Clark University.

[1] There has been considerable controversy about the use of the terms "growth centres", "growth poles", and "growth points". For the purposes of this paper, growth "points" refer specifically to points in space where growth has occurred, growth "centres" describe areas which are created for the purpose of growth, while growth "poles" describes economisectors from which growth emanates.

REVIEW

Many attempts have been made to investigate industrialization in Poland so as to evaluate the results of economic development over space during the period 1946–60. Widely varying in time dimension, areal differentiation, and methods of analysis, five of these studies demand review here for they offer a series of empirical results and methodological demarcations in the measurement of Polish industrial development.

Of a series of studies done on industrialization in Poland using voivodship (regional) data, two recently published studies by S. Leszczycki (1970, 1971) represent the most complete analysis of the national economy for the period 1961–1965. The relative magnitude of Leszczycki's general synthetic index, representing an unweighted combination of synthetic indices for fixed assets, national income distributed, suggests that there was wide variation in the degree of industrialization. Although Leszczycki found that index values for the most concentrated areas of industrialization (i.e., the voivodships of Warsaw, Cracow, Katowice, and Wrocław) were substantially higher than for the four eastern voivodships — Rzeszów, Lublin, Olsztyn, and Białystok — the changes in the indices were greatest for the least developed, eastern voivodships. These studies provided the evidence to support Leszczycki's conclusion that economic development was spreading from the highly developed voivodships to the areas of least development.

This conclusion was further supported in a recent study by S. Róg (1972). By ranking voivodships by levels of industrialization, and then by extending his analysis to industrialization by *poviats*, Róg found that by 1968 there were different "sets" of highly industrialized areas as identified by the *poviat* and voivodship analyses. The most striking difference was found in the identification of a few very highly industrialized *poviats* located in Rzeszów, Kielce, Poznań, Bydgoszcz, and Szczecin voivodships. Interestingly, all of these exceptions were located within the boundaries of newly created industrial districts. Therefore, Róg suggested that the spatial distribution of industrialization in Poland had shifted significantly during 1946–1966 due to the more rapid development of industrial production and employment within these industrial growth centres, as compared with the older, traditional industrial centres of Silesia.

The importance of these formalized industrial districts — growth centres — to the spread of industrialization in Poland was the subject of a recent article by A. Kukliński (1972). By examining the degree of industrialization in 33 areal units — 7 industrial districts, 7 industrial areas, 6 industrialized regions, and 13 unindustrialized areas — Kukliński found that the spread of industrialization to the least developed areas was very evident. The unindustrialized areas had the highest percentage growth of industrial employment during 1950–1965 while the growth of industrial employment in the more highly developed areas was much lower.

The dynamics of the changes in the levels of industrialization over space in Poland was the subject of a recent study by A. Dawson (1970). He used multifactor analysis to combine both the levels of industrial employment *per capita* and the changes in *per capita* industrial employment during 1949–1965, and deduced that by 1965 two distinct levels of industrialization were evident: a high degree of industrialization in the urban-industrial centres and their adjacent areas, and a much lover level of industrialization in the rural, peripheral areas. Further, Dawson asserted that this evidence suggested that the level of industrialization of the two groups was merging; there was a very evident spreading of industrialization from the urban-industrial growth centres, through their adjacent areas, to the rural areas as a direct result of the conscious efforts of the Polish planners to equalize the local industrial bases over space.

The importance of the methods used to delimit these changes in industrialization over space as a basis for policy decisions was noted in a study by J. Kruczała (1970). Kruczała suggested that synthetic measures employed to measure spatial disproportions might mislead national and regional planners unless the fundamental purpose of such indices was to show differences in personal income *per capita* over space. His detailed analysis of the Cracow voivodship showed that considerable disproportions in *per capita* personal income did exist in 1960, disproportions which were not generally delimited by past investigations employing more sophisticated analytical tools, and disproportions which should be seriously considered in the formulation of future regional investment plans. At the same time, the Kruczała study also contained a sharp retort to those who used synthetic measures which did not consider *per capita* measures. The full equalization of personal income *per capita* over space, as measured by wages and salaries paid to the employees at his place of employment, disregarded workers who commuted to work. Thus, the data favoured those locations where employment took place rather than considering where the worker lived. Kruczała further suggested that while improved transportation facilities and increased employment opportunities would help to correct real disproportions in personal incomes over space, the use of broader areal units as data bases, including both the urban-industrial centre and its commuting field, would enable the planner to discover the actual spatial distribution of *per capita* incomes and make recommendations for future action considerably more appropriate with respect to national goals.

A DESCRIPTIVE MODEL

A synthesis of these investigations of industrialization in Poland suggests that three considerations are critical to the analysis of the spatial structure of economic development. First, the geographic unit selected and the subsequent scale of analysis may reveal very different sets of conclusions. *Poviat* data must be aggregated into meaningful sets which are functionally integrated. Second, the use of *per capita* data is especially meaningful to the analysis because of socialist development goals. Third, Polish planning appears to be guided by development themes which emphasize growth centres as the motor mechanisms of industrial growth; hence, the analysis of industrial change should strive to identify these growth centres.

In the analysis presented here industrial employment *per capita* was tabulated for *poviats* in 1949, 1960, and 1970. To the 314 administrative units (*poviats*) an additional 82 were added; the latter resulted from the disaggregation of certain *poviats* into their urban and rural components. These 396 areal units were next classified on the basis of geographic location with respect to characteristics of industrial location in the context of development policies in post-war Poland. These principles include: a contrasting history of economic development between the older, major industrial centres *vis-à-vis* the underdeveloped, eastern parts of the nation; the differential rates of investment in the five largest cities as compared to their surrounding areas, especially in comparison to the large cities in these adjacent areas; and, the growing industrial importance of the medium-sized cities in Poland, reflecting developmental principles of agglomeration or economies of scale (Kukliński, 1965).

In this descriptive analysis *poviats* have been grouped as follows:

(1) Eastern *poviats*; defined as all of the *poviats* in the voivodships Białystok, Lublin, Olsztyn, and Rzeszów.

(2) Eastern cities; medium-sized cities (i.e., urban populations between 25,000 and 75,000) in the four eastern voivodships.

(3) All medium-sized cities in Poland.

(4) The five largest cities (Warsaw, Cracow, Łódź, Poznań, and Wrocław).

(5) The "first ring" of *poviats* surrounding the five cities.

(6) The "second ring" of *poviats* surrounding the five cities.

(7) The "adjacent cities"; medium-sized satellite cities surrounding the 5 cities.

(8) All *poviats* in the voivodship of Katowice.

The levels of industrial employment per 1000 inhabitants for these classified *poviats* are presented in Table 1, together with the changes in this statistic by decades during 1950–1970. It is clear from this table that the largest concentrations of industrial employment existed in Katowice, the five largest cities, and in all medium-sized cities. On the other hand, the most significant changes in industrial employment occurred in the eastern voivodships, and particularly in the eastern and adjacent (satellite) cities.

TABLE 1. Industrial employment per 1000 inhabitants

	1949	1960	1970	1960/50	1970/60	1970/49
Poland	81.91	101.28	140.3	23.65	38.53	71.13
Katowice	208.79	223.34	233.67	6.97	4.63	11.92
All Medium-Sized Cities	167.62	209.08	223.22	24.73	6.76	33.17
Five Cities	164.68	182.45	207.76	10.79	13.87	26.16
Warsaw	87.06	146.86	178.32	68.69	21.42	104.82
Cracow	123.5	173.39	195.38	40.28	12.68	58.20
Łódź	296.03	267.88	289.37	−9.51	8.02	−2.25
Poznań	167.18	176.43	195.70	5.53	10.92	17.06
Wrocław	141.59	150.69	187.03	6.43	24.12	32.09
Poviats 1st Ring	49.33	81.52	90.75	65.25	11.32	83.97
Poviats 2nd Ring	102.59	115.23	141.70	12.32	22.97	38.17
Adjacent Cities	118.87	226.21	245.62	90.30	8.58	106.63
Eastern Cities	46.93	104.56	137.43	122.80	31.44	192.84
Eastern Voivodships	26.53	41.27	76.40	59.83	58.83	168.91
Białystok	22.68	38.73	59.24	70.77	52.96	161.20
Lublin	19.41	39.17	63.02	101.91	60.72	224.68
Olsztyn	31.20	40.26	63.27	29.04	57.15	102.79
Rzeszów	34.94	61.81	102.31	76.90	65.52	192.82

Further, Table 1 reveals numerous complexities in the trends or dynamics of industrial employment over time. For example, the first-ring *poviats* show significant rates of growth in comparison to the five largest cities during the period of 1950—1960, while in the next decade the more significant growth occurred in the second-ring *poviats*.

This descriptive analysis suggests that the spread of industrialization from the traditional industrial centres of Katowice and the five major cities to the least developed regions of Poland, particularly to the eastern voivodships, has not been a uniform shift. Rather, the importance of the growth centres — the medium-sized cities — and the interaction of the urban complexes — the major urban centres and their satellite centres — in the process of spreading industrialization to the lesser developed areas of Poland are paramount. On this bases, we present the following econometric model as a means of examining the interaction between these areal groupings.

A DYNAMIC MODEL OF INDUSTRIAL GROWTH

The complexities of the spatial interaction between separate areal units over time requires that a formalized model consider pooled cross-sectional and time-series observations on industrial employment *per capita* in Poland during 1950–1970. We wish to specifically expand upon the themes of the previous descriptive analysis by modelling changes in industrial employment *per capita* with regard for apparent planning policies and objectives. The purpose of this model is to measure the rates of change of industrial development by "regions"; regions which are delineated by their importance to the spread of socio-economic benefits over space in Poland.

It is our contention that the following regression equation adequately explains industrial development in post-war Poland and measures the evolution of the spread of industrialization through planned growth centres. The methodology employed is especially significant since it incorporates, *a priori*, regional units and time into one equation permitting the evaluation of simultaneous, statistical interactive effects between time and the regional variables, and at the same time, measures the separate characteristics of each regional class. This final composite regression equation, a special case of the covariance model, considers the effect of time and the class to which the *poviat* belongs, as represented by a series of dummy variables, on industrial employment per 1000 inhabitants.

We begin by first considering the rate of growth of industrial employment in Poland during 1950–1970:

$$\text{Log } IE_{tn} = f(T_{tn}) \qquad \text{Where} = T = 1, 3, 4, 5 \quad n = 1, 2, \ldots, 396 \qquad (1)$$

The use of all observations on each poviat at four time intervals[2] (i.e., N = 1548) produces:

$$\text{Log } IE = 1.015 + 743 \text{ Time}^{**} \underset{(.0188)}{} \qquad R^2 = .497 \quad F = 1564.99^{**}$$

Therefore, the growth of industrial employment per 1000 inhabitants for all *poviats* in Poland during 1950–1970 is estimated to be approximately 74 per cent per time period.[3]

We may then examine the rate of growth of industrial employment *per capita* for each spatial combination separately. We first set the null hypothesis

[2] We have used observations on industrial employment per 1000 inhabitants for each of the 396 *poviats* in 1949, 1960, 1965, and 1970. To preserve the linear trend over time, the value of time as an independent variable omits the value T2 because data for 1955 were not available.

[3] We denote significance at the 0.05 level by ** and significance at the 0.01 level by *a* *. The numbers in paranthesis below the coefficients are the standard error of the coefficient.

that the rate of growth of industrial employment per 1000 inhabitants in each geographic grouping considered separately is not significantly different from the rest of Poland. For example, we form a regression equation:

$$\text{Log } IE_{tn} = f(T_{tn}, D_m, TD_m) \tag{2}$$

Where: $T = 1, 3, 4, 5$, $n = 1, 2, ..., 396$, $D = 1$ for the m *poviats* ($m < n$) in the group of *poviats* examined and $D = 0$ for all other $(n-m)$ *poviats*.

To determine the rate of growth for the five cities (i.e., $D = 1$) as compared to the rate of growth of industrial employment per 1000 inhabitants is not significantly different for each of these separate geographic groupings from the rest of Poland, we employ an appropriate test of significance to evaluate the entire equation by an F-test.[4] That is, we wish to determine whether the addition of the explanatory variables D and TD has reduced the unexplained variation of the regression. The *F-test* employed is:

$$F_0 = \left(\frac{R_Q^2 - R_K^2}{1 - R_Q^2}\right) \left(\frac{N-Q}{Q-K}\right)$$

Where: $R^2 = R^2$ from expanded equation $(2-a)$
$R_K = R^2$ from equation (1)
N = number of observations (i.e., 1584)
Q = number of variables employed in equation $(2-a)$
K = number of variables employed in equation (1)

The F_0 test of significances (see Table 2) permits the rejection of the null hypothesis in the case of all but three of the groups of *poviats* considered by

TABLE 2. Results of Regressions on Log of Industrial Employment per 1000 Inhabitants ($n = 1584$)

Regression	a	b_1	b_2	b_3	R^2	F	F_0
1	1.015	.7426** (.0188)	—	—	.497	1564.99**	—
2-*a* (5 cities)	.998	.7431** (.0188)	1.324**	−.032 (.5962)	.505 (.1670)	537.31**	12.77**
2-*b* (lst Ring)	.972	.7410** (.0193)	.0274 (.2989)	.0304 (.0838)	.498	522.48**	.60
2-*c* (2nd Ring)	.953	.7444** (.0193)	.3694 (.2862)	−.301 .0801	.499	524.57**	2.65
2-*d* (Eastern Voivodships)	.975	.7375** (.0207)	−.0047 (.1758)	.0291 (.0492)	.498	522.47**	.94
2-*e* (Katowice)	.875	.7502** (.0183)	.2131** (.0295)	−.159 (.0834)	.548	638.53**	87.67**
2-*f* (Adjacent Cities)	.910	.7406** (.019)	.8612** (.2508)	.0276 (.0703)	.523	576.99**	41.98**
2-*g* (Eastern Cities)	.978	.7360** (.0193)	−.005** (.2754)	.1033 (.0769)	.501	526.13** •	5.126*
2-*h* (Other Cities)	.851	.7455** (.0206)	.6637** (.1604)	.0141 (.0449)	.524	579.78**	43.31**

[4] See J. Kmenta (1971), p. 371.

this analysis. We may, therefore, conclude that the addition of the dummy variable (D) and the "interaction" variable (TD) to regression model has contributed to the explanation of the variation in industrial employment per 1000 inhabitants for this period. As examined separately against the rates of growth of industrial employment in *all other poviats* in Poland, the rate of growth in the five major cities (70 per cent), in Katowice (59 per cent), in the adjacent cities (77 per cent), in the cities of the eastern voivodships (84 per cent), and in all other medium-sized cities (73 per cent [5]) would appear to be significantly different from the rate of growth in all other *poviats* in Poland (74 per cent). But in the case of the other three geographic groupings examined separately — the first-ring *poviats,* the second-ring *poviats,* and all *poviats* in the eastern voivodships — the results of the F_0 test do not permit the rejection of the null hypothesis. For these three groupings, the variance in the rates of growth of industrial employment of *poviats* within these groups is too large to permit us to generalize about these combinations. Since the level or aggregation is apparently too gross for *poviats* combined for these three groups, we must reject their use for the purpose of describing the significant changes in industrial employment *per capita* in Poland.

But are the rates of growth for each of these geographic groupings significantly different from each other e.g., the rate of growth of industrial employment *per capita* in the five cities as compared to the rate in Katowice, etc.? To attempt an answer requires that the model be further expanded to include all of the groups together within the regression model. We form a composite equation:

$$\text{Log } IE_{tn} = f(T_{tn}, D1_r, D2_s, D3_u, D4_v, D5_w, TD1_{tr}, TD2_{ts}, TD3_{tu}, TD4_{tv}, TD5_{tw}) \quad (3)$$

Where: $T = 1, 3, 4, 5$

$n = 1, 2, ..., 396$

$D1 = 1$ for the five (r) cities, $D1 = 0$ otherwise

$D2 = 1$ for the s *poviats* in Katowice, $D2 = 0$ otherwise

$D3 = 1$ for the U adjacent cities, $D3 = 0$ otherwise

$D4 = 1$ for the v city *poviats* in the eastern voivodships, $D4 = 0$ otherwise

$D5 = 1$ for the w other city *poviats*, $D5 = 0$ otherwise

and

$$(r + s + u + v + w)\, n$$

This composite equation captures the structural interaction between each of the groupings while differentiating the rate of growth for each of the groups separately. The results of this composite model approximates those of the separately examined geographic regions, [6] thereby maintaining the spatial structuring of the rates of growth of industrial employment *per capita* that have been identified in equation (2) above. The results of this regression show:

---

[5] Note that we have considered three unique groups of medium-sized cities in this analysis. Of all medium-sized cities we are examining separately: (1) adjacent cities; (2) eastern cities; and (3) all other cities not included in groups 1 or 2.

[6] The inclusion of the "interaction terms", TD, suggests that we expect to find a difference in the rate of growth of industrial employment *per capita* when comparing each region against all others, and that magnitude of these differences in the growth rates are not the same for all regional classifications. Note that this formulation is the equivalent of performing separate regressions on each of the regions separately. See Kmenta, (1971, p. 421) and Goldberger (1964, pp. 224–26).

$$\text{Log } IE = \underset{(.0272)}{.505 + .746T^{**}} + \underset{(.503)}{1.782D1^{**}} + \underset{(.2655)}{2.504D2^{**}} + \underset{(.2190)}{1.273D3^{**}} + \underset{(.2341)}{.47D4^{**}} + \underset{(.1421)}{1.0D5^{**}}$$
$$- \underset{(.1495)}{.046TD1} - \underset{(.0743)}{.155TD2^{**}} + \underset{(.0614)}{.022TD3} + \underset{(.0656)}{.093TD4} - \underset{(.0398)}{.015TD5}$$

where: $R^2 = .652$
$F = 267.732^{**}$

The significance of the regression coefficients suggests that these geographic groupings have been correctly specified, thereby permitting the composite model to identify differentiated rates of growth of industrial employment for each of these regions. From the results of equation (3) we identify the following rates of growth by combining like terms:

$$\begin{aligned}
\text{Log } IE(\text{5 cities}) &= .505 + .7467 + 1.782D1 - .046TD1 \\
&= 2.287 + .70T \qquad (\text{where: } D1 = 1) \\
\text{Log } IE(\text{Katowice}) &= .505 + .7467 + 2.504D2 - .155TD2 \\
&= 3.009 + .591T \qquad (\text{where: } D2 = 1) \\
\text{Log } IE(\text{adj. cities}) &= .505 + .7467 + 1.273D3 + .022TD3 \\
&= 1.778 + .7687 \qquad (\text{where: } D3 = 1) \\
\text{Log } IE(\text{east cities}) &= .505 + .7467 + .47D4 + .093TD4 \\
&= .975 = .839T \qquad (\text{where: } D4 = 1) \\
\text{Log } IE(\text{other cities}) &= .505 + .7467 + 1.0D5 - .015TD5 \\
&= 1.505 + .731T \qquad (\text{where: } D5 = 1)
\end{aligned}$$

and

$$\text{Log } IE(\text{all other } poviats) = .505 + .746T \qquad (\text{where: } D1 = D2 = D3 = D4 = D5 = 0)$$

Employing the results of regression (3) above, we may now test the hypothesis that the rate of growth of any one of these regions is significantly different from the rate of growth of industrial employment *per capita* in any other of the regions. We first set the null hypothesis:

$$H_0 : b_j = b_k \qquad j \neq k$$

that the regression coefficient for one region equals the regression coefficient for another. In equation (3) above, where:

$$\text{Log } IE = a + b_1T + b_2D1 + b_3D2 + b_4D3 + b_5D4 + b_6D5 + b_7TD1 + \\ + b_8TD2 + b_9TD3 + b_{10}TD4 + b_{11}TD5$$

we test the null hypothesis for the equality between the regression coefficients of the interaction terms, TD (i.e., $b_7 = b_8$; $b_7 = b_9$; etc.), until all possible pairs of combinations of regression coefficients in equation (1) have been tested by: [7]

$$t_{n-k} = \frac{b_j - b_k}{s_{b_j} - b_k} \qquad \text{where } s_{b_j} - b_k - \text{Var}(b_j - b_k) = \text{Var}(b_j) + \text{Var}(b_k) - 2\text{ Cov}(b_j, b_k)$$

The results of the t-test for each pair of interaction regression coefficients in equation (3) are shown in Table 3. The results of this analysis indicate that the rates of growth for the five geographic groupings under examination are significantly different from each other, except for the case of Katowice and the group identified as "all other cities". In comparing the rates of growth of industrial employment per 1000 inhabitants for these two regions, we may therefore conclude that they are statistically not different from one another,

[7] See J. Kmenta (1971), pp. 371–72.

while for all other regions there exists a significant difference between their growth rates.

The generalized model has permitted a simultaneous examination of the growth rates of industrial employment *per capita* for functionally related geographic regions. As revealed by this model (see Figure 1) the largest concentration of employment initially was found in the oldest industrial centres in

TABLE 3. Tests of Significance of Regression Coefficients

Coefficients	b_7	b_8	b_9	b_{10}
b_7				
b_8	−4.218			
b_9	−2.991**	−20.776**	– –	
b_{10}	−5.983**	−27.533**	−9.753**	– –
b_{11}	−1.528	−22.206**	7.978**	−21.165**

Katowice and in the five largest cities. However, the greatest rates of growth occurred in the medium-sized cities in the eastern voivodships, and in the large satellite cities adjacent to the five cities. The most significant centres of industrial growth are the adjacent cities, since these reveal strong rates of growth together with significant initial concentrations of industrial employment.

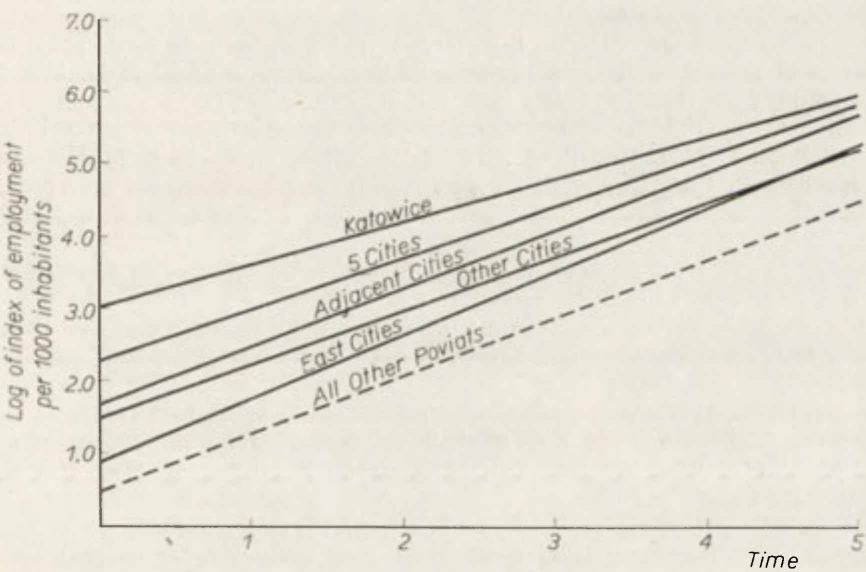


Fig. 1.

CONCLUSION

In the context of the model presented here, we conclude that industrial development in Poland during 1950–1970 has followed a distinctive path of growth. While the traditional industrial concentrations — Katowice and the five largest cities — have maintained their primacy and have experienced only limited rates of growth of industrial employment during this period, the most intensive rates of growth have occurred in the growth centres (medium-sized cities), in the eastern voivodships, and in those cities adjacent to the five largest cities.

Clark University, Worcester, Massachusetts

BIBLIOGRAPHY

Dawson, A. H., 1970, The changing distribution of Polish industry, 1949–1965, *Trans. Inst. Brit. Geogr.*, 50, pp. 177–196.

Domański, R., 1971, Co-variables and synthesis in management of the Konin–Łęczyca–Inowrocław industrial region, in: *Regions under industrialization,* Warsaw, Polish Academy of Sciences, pp. 282–287.

Dziewoński, K., 1962, Theoretical problems in the development of economic regions, *Papers. Reg. Sci. Ass.*, 13, pp. 43–54.

Dziewoński, K., 1964, Urbanization in contemporary Poland, *Geogr. Pol.*, 3, pp. 37–56.

Feiwel, G. R., 1971, Poland's industrial policy: a current analysis, vol. 1, in: *Industrial and planning under Polish socialism,* New York, Praeger.

Feiwel, G. R., 1971a, Problems in Polish economic planning: continuity, change, and prospects, vol. 2, in: *Industrialization and planning under Polish socialism,* New York, Praeger.

Friedmann, J. P. R., 1956, Locational aspects of economic development, *Land Economics,* 32, 3, pp. 213–227.

Gokham, V. M. and Karpov, L. N., 1972, Growth poles and growth centers, in: *Growth poles and growth centers in regional planning,* A. Kukliński (Ed.), The Hague, Mouton, pp. 125–134.

Goldberger, A. S., 1964, *Econometric theory,* New York, Wiley.

Hamilton, F. E. I., 1967, Models of industrial location, in: *Socio-economic models in geography,* R. Chorley and P. Haggett (Eds.), London, Methuen, pp. 361–424.

Hamilton, F. E. I., 1970, Aspects of spatial behavior in planned economies, *Papers. Reg. Sci. Ass.*, 25, pp. 83–105.

Hamilton, F. E. I., 1971, Decision-making and industrial location in Eastern Europe, Trans. Inst. Brit. Geogr., 52, pp. 77–94.

Hamilton, F. E. I., 1971a, The location of industry in East-Central and Southeast Europe, in: *Eastern Europe: Essays in Geographical Problems,* New York, Praeger.

Hilhorst, J. G. M., 1971, Spatial structure and decision-making, in: *Issues in regional planning,* D. Dunham and J. G. M. Hilhorst (Eds.), New York, Humanities Press.

Hirschman, A. O., 1958, *The strategy of economic development,* New Haven, Yale University Press.

Kmenta, J., 1971, *Elements of econometrics,* New York, Macmillan.

Kruczała, J., 1970, Differentiation of subregions and the possibility of their economic equalization, in: *Proceedings. Second Poland-Norden Regional Science Seminar,* Warsaw, pp. 13–36.

Kukliński, A., 1965, The interregional differentiation of Poland's national economy, *Geogr. Pol.,* 7, pp. 49–57.

Kukliński, A., 1966, Progress and change in the industrialization of Poland, *Geogr. Pol.,* 3, pp. 57–71.

Kukliński, A., 1966a, Research activity in the Committee for Space Economy and Regional Planning, in: *City and Regional Planning in Poland,* J. C. Fisher (Ed.), Ithaca. N. Y., Cornell University Press, pp. 57–71.

Kukliński, A., 1967, Regional differentiation of the Polish national economy, in: *Proceedings. First Scandinavian–Polish Regional Science Seminar*, Warszawa, pp. 51–70.

Kukliński, A., 1970, Regional development, regional policies, and regional planning, *Reg. Stud.*, 4, pp. 269–278.

Kukliński, A., 1972, *Regional development and planning in Poland*, Warsaw (mimeographed).

Lange, O. and Taylor, F. M., 1938, *The economic theory of socialism*, New York, McGraw-Hill.

Leszczycki, S., 1965, Problems of post-war industrial concentration and decentralization in Poland, *Geogr. Pol.*, 7, pp. 29–48.

Leszczycki, S., 1970, Spatial macrostructure of Poland's national economy between 1961–1965, in: *Proceedings. Second Poland–Norden Regional Science Seminar*, Warsaw, pp. 13–36.

Leszczycki, S., 1971, *Long-term planning and spatial structure of Poland's national economy*, Accademia Polacca delle Science, Fasc. 51, Warszawa.

Mihailovic, K., 1972, *Regional development. Experiences and prospects in Eastern Europe*, The Hague, Mouton.

Montias, J. M., 1962, *Central planning in Poland*, New Haven, Yale University Press.

Myrdal, G., 1957, *Economic theory and under-developed regions*, London, Duckworth.

Regulski, J., 1972, Development of pole theory and its application in Poland, in: *Growth poles and growth centers in regional planning*, A. Kukliński (Ed.), The Hague, Mouton, pp. 207–226.

Róg, S., 1972, An attempt to determine the degree of industrialization of voivodships and *poviats*, *Geogr. Pol.*, 20, pp. 85–99.

Wojtasiewicz, L., 1972, Konin industrial region in the economic and social development of Poland, in: *Growth poles and growth centers in regional planning*, A. Kukliński (Ed.), The Hague, Mouton, pp. 221–230.

Wróbel, A. and Zawadzki S. M., 1966, Location policy and the regional efficiency of investments, in: *City and Regional Planning in Poland*, J. C. Fisher (Ed.), Ithaca, N.Y., Cornell University Press, pp. 433–440.

Zaremba, J., 1966, Regional planning in Poland: Theory, methods and results, in: *City and Regional Planning in Poland*, J. C. Fisher (Ed.), Ithaca, N. Y., Cornell University Press, pp. 271–297.

GEOGRAPHIA POLONICA 31, 1975

CHOSEN ELEMENTS OF THE THEORY ON THE HINTERLAND AND THE FORELAND OF SEA PORTS RELATING TO RESEARCH ON POLISH PORTS

Witold Barczuk

Modern sea transport is recently, more than previously, concerned with the geography of the land adjacent to the port which forms the hinterland and the foreland. This is indispensable from the point of view of the proper organization of transport and also the acquisition of cargoes for sea transport. This is, above all, connected with the modern technology of reloading and transporting for the reasons given below:

— Acquisition of cargoes for container-ships or LASH-system often takes place outside the ports, at their hinterland or foreland and not in the port as was previously practiced in the traditional system. Consequently, cargoes are often transferred for sea transportation far away from the port. [1]

— Proper organization of modern sea transport makes the sea carrier concerned in land transport and its conditions in the hinterland and the foreland.

— At the hiterland and the foreland special conditions are created to collect cargoes designed for sea transport and there is also the demand for cargoes. All the problems involved, are of interest to those who organize sea transport mainly because of the necessity of planning and forecasting cargo turnover for the future.

Theoretical works on the hinterland and the foreland are so far insufficient for this purpose and they do not clearly explain their role in relation to the needs of transport and the terminology used in those works is not precise enough or clearly understood.

The author of the present work suggests the acceptance of certain formulations relating to the meaning and function of the hinterland and foreland of ports in sea transport.

THE HINTERLAND AND FORELAND OF PORTS

The hinterland and foreland are spatial conceptions [2]. Both the hinterland and foreland relate to definite geographical regions organized so as to serve the cargo designed for ports.

[1] Container cargo acquisition is often done by the charterer himself, who also owns the containers and means of land transport. This has now been introduced in Poland. Polish Ocean Lines administrate their own containers on land without the shipper's intermediary. An example of contacts of Polish Ocean Lines with the hinterland and the foreland.

[2] Steen de Geer Eneberg states that hinterland means only cargo-neglecting the geographical region. In his work *Ostersjoe* (Stockholm, 1926) he states that the hinterland is not the region but rather the cargo existing on a certain territory. This is the so-called material or immaterial conception of the port's hinterland.

The word "organized" means, that certain geographical regions have been prepared by human intentional and conscious activity to serve a definite economic purpose, and with regard to the hinterland and foreland to serve the purpose of sea transport.[3]

Since this is a function consisting mainly in transport services to be rendered by the hinterland and foreland. S. Berezowski[4] is right in stating that both the hinterland and foreland of ports are the territories of activity of the freight-transport markets. This has been confirmed as the result of a more thorough examination of their function. Just like other kinds of markets the hinterland and the foreland are organized in a certain way. What the author of the present work calls "organized regions" is somehow more generally treated because it was his intention to touch on the problem of the forming of cargo mass which problem does not fall under the term "freight market".

To organize a region to serve sea transport means to accumulate a certain cargo mass or else to create the demand for the cargo while the transporting of the cargo to or from the port is the next stage.

The hinterland, the foreland and also the port itself are motions determined by sea transport. They all form a set of spatial elements which strictly depend on each other and are complementary to one another in forming one system in which any of these elements, individually treated is of no significance.

The hinterland cannot exist without either the port or the foreland and likewise the foreland without either the hinterland or the port is meaningless and also the port could not exist without the remaining two elements — of course with the exception of some unusual, very rare or irregular cases. However, the modern technology of sea transportation, may in future eliminate the port from this complex system. For example containers can be reloaded from a harboured ship by means of helicopters not only to the port, but to any place situated in the hinterland or foreland. LASH-system are already becoming completely independent of ports.

It is quite the same if we call the hinterland and the foreland the freight-cargo market or more generally the "organized regions" for the needs of sea transportation, we will still have to deal with definite spatial structures of special importance. These, however, are the regions where cargo materializes, is then transported and finally ceases to exist in its original form.[5]

The flow of cargo through the three aforesaid elements of the system i.e.: hinterland-port-foreland encloses the whole of the problem of research on the geography of sea transport.[6] Practically, the term "sea transport" with regard to the transportation of cargoes through all the elements of the system becomes more and more common, even if sea transport was only one of the elements in the flow of cargoes through the ports. This is all the more reasonable because it relates to the same fact of transporting the cargoes which are through-

---

[3] Cf. W. Barczuk, Funkcje zaplecza i przedpola w transporcie morskim (The function of hinterland and foreland in sea transport), *Technika i Gospodarka Morska* 1973, 7, p. 395.

[4] S. Berezowski, Zaplecze i region (The hinterland and the region), *Gospodarka morska* 1949, 2, and personal communication.

[5] Decline of cargo in its original meaning when handed over to the new owner, i.e., when it passes from the shipper's control to somebody elses controls, e.g., the merchant's.

[6] Sea transport treated in the traditional way relates to transportation by sea only which arises from the word "sea". In modern navigation there are cases in which the use of this expression may be difficult or quite impossible. In the case of LASH-system it is impossible to tell the difference between sea and land transport.

out controlled by the same shipper. This definitely emphasizes the uniformity of the complex: hinterland–port–foreland.

When we analyse the function of the hinterland and the foreland we come across the problem of special structures — and this has been already emphasized. Hence the necessity of research on the hinterland and foreland and also on the flow of cargoes through these elements of a complex system, to be undertaken by a geographer by means of geographic methods such as an examination of the relationship between different regions, cartography and regionalization.

CARGOES AND THEIR ROLE IN THE HINTERLAND AND FORELAND

It is generally acknowledged that port cargo means any goods which have passed through the port to its hinterland or to its foreland, or which are addressed to the seaport (or rather designed for sea transport). However, it does not necessarily have to be actually in the port.

Cargo, is indubitably the most important factor which provides the existence of the hinterland or foreland of the port. All problems connected with sea transport are related to the cargo because this kind of transport has been developed to shift cargo.

The cargoes determine the size of the hinterland and of the foreland and the differentiated rate of its concentration in different regions ranges from those which are of either bigger or smaller importance for the turnover of goods which pass through the port. Depending on the kind and quantity of the cargo appropriate means of transport as well as techniques of transport and reloading are applied.

TABLE 1. The general cargo on the foreland of Polish ports divided into groups according to the techniques of reloading and kind of transportation (in thousands of tons)

Specification	Kind of general cargo					
	Heavy	Packed	In containers, acc. to the rate of applicability for containers[a]		Cooled[b]	Refrigerated[c]
			medium[a]	high		
Import						
Local cargo from foreign trade enterprises	213.6	273.2	643.5	322.2	79.3	54.5
Transit cargo	40.4	133.7	445.3	109.6	9.2	26.8
Export						
Local cargo from foreign trade enterprises	920.7	1,056.0	200.5	277.1	135.9	134.7
Transit cargo	415.3	540.9	184.3	82.2	32.4	—

Source: According to information received from *Centralny Zarząd Spedycji Międzynarodowej* in Gdynia (Central Board of International Transport).

[a] General cargo suitable for container transport, not all of which was transported that way in 1970.

[b] General cargo which requires cooled transport.

[c] General cargo which requires refrigerated transport.

All goods become cargo, once they have been designed for sea transport and they remain cargo until the moment of takeover by the new owner and when the shipper has no more control over them.

The big variety of cargoes as a result of the complexity of economic life calls for the necessity to form a certain system in order to classify it into groups and categories etc. With regard to sea transport facilities cargo has been classified so as to meet certain specific criteria. In *The Foreland of Polish Ports*[7] the cargo was classified according to the criteria of techniques relating to the reloading and means of transport and especially with regard to general cargoes which are always a problem from the point of view of reloading and transportation. This is shown in Tables 1 and 2.

TABLE 2. Bulk cargo in the foreland of Polish ports in 1970 divided into groups including cargo from Polish foreign trade enterprises and transit cargo

Specification	Coal and coke	Ores and concentrated metals	Chemical raw materials	Mineral raw materials	Wood	Corn and fodder	Liquid fuels	Other liquid cargo
				in thousands of tons				
				Import				
Polish foreign trade enterprises	—	2,187.0	2,291.0	298.0	99.5	1,014.0	475.0	159.2
Transit cargo	—	830.4	527.6	472.0	64.0	71.4	86.0	—
				Export				
Polish foreign trade enterprises	16 154.0	—	1,238.0	608.7	463.0	117.0	555.0	167.3
Transit cargo	82.0	—	6.4	3.6	105.0	120.0	—	1.6

Source: According to information given by *Centralny Zarząd Spedycji Międzynarodowej* in Gdynia, 1971.

The cargoes determine the size of the hinterland and of the foreland where the area is not the important factor but rather the rate of saturation of cargoes. This is so important, that Steen de Geer Eneberg[8] states that apart from the surface on which the cargoes appear it is the only factor which creates the hinterland. The author of the present paper assumes in another of his works *The domestic hinterland of Polish ports*[9], that the cargo is of significant importance when linked with the space on which it appears in the hinterland or in the foreland thus forming the third dimension of its substance. When talking of variations of size of the hinterland or of the foreland we generally have in mind rather the variations of the total of the cargoes more than the territory, which is much less important. Likewise with regard to the port its transit capacity is important and not the area it covers.

[7] W. Barczuk, *Przedpole portów polskich* (The foreland of Polish ports), Gdańsk 1973, Inst. Morski.

[8] Steen de Geer, op. cit.

[9] W. Barczuk, *Krajowe zaplecze portów polskich* (The domestic hinterland of Polish ports), Gdynia 1966, Wyd. Morskie.

THE ORIGIN OF CARGO

Sea transport takes an interest in the cargo long before it materializes, namely in how and under what conditions it originates and above all, what are the prospects for the future. True enough, foreign trade is in charge of that problem but sea transport will not be able to ignore the conditions of transport in the hinterland or the foreland in the future, to take only the problem of the appropriate means of transport to be adapted to certain kinds of cargo to be handled in the future and this is a problem which foreign trade will not be able to solve.

BOUNDARIES OF THE HINTERLAND AND THE FORELAND

The very conception of the space of the hinterland and foreland calls for the necessity of defining the boundaries of these territories which is not simple or easy. The transportation of cargoes takes places along definite streams while it is loaded or discharged in places marked as spots. The principle of defining boundaries according to the functional reach of these spot-marked places (localities) is not applicable here, at all. Some of these localities are located too far from the ports and are somehow isolated from them; just to mention that in 1970 certain consignments of goods were dispatched from Polish ports to Mongolia and Italy by land transport which

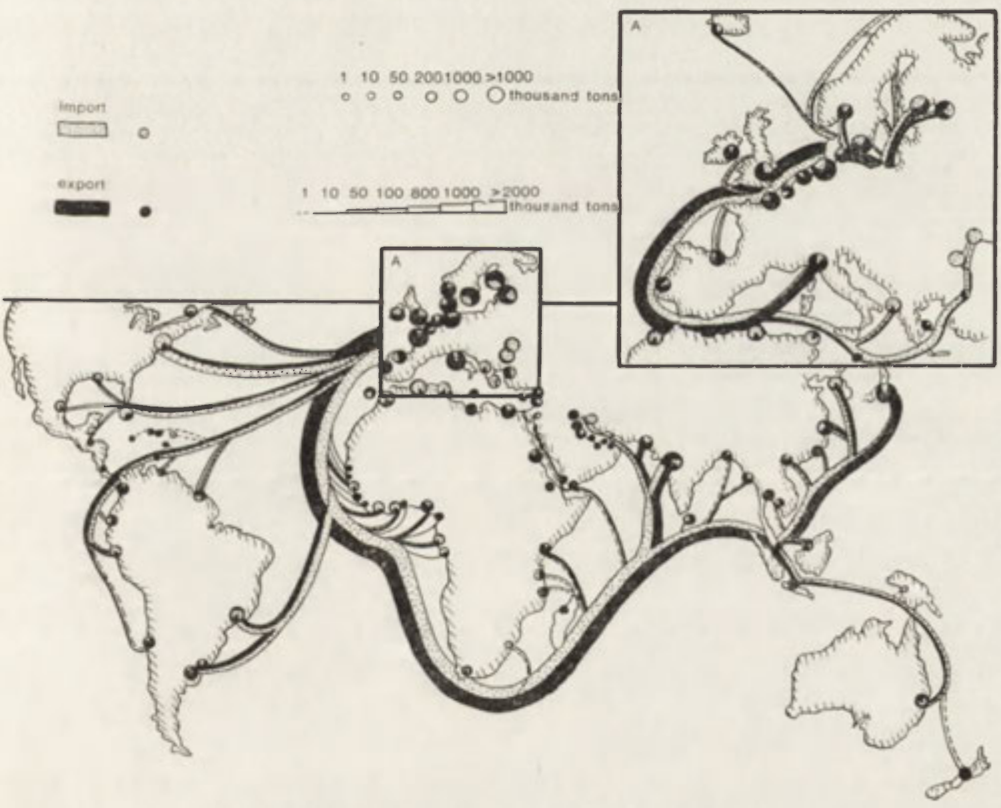


Fig. 1. The foreland of the Polish ports, 1970

was an example of links of Polish ports with these countries. It is however impossible to consider these countries as the hinterland of the Polish ports, mainly because these were exceptional cases and also because the consignments dispatched were too small. In the container transportation system such rare cases may become more frequent. Sometimes, even domestic territories cannot be included in the hinterland although they are situated not far away, owing to lack of contact between these territories and the ports.

Certain conventions should be assumed and the author suggests the acceptance of administrative limits e.g. of the *poviat* or voivodship as limits of the hinterland, provided however, that a certain minimum quantity of cargoes e.g. 15 tons has been reloaded on the territory of such an administrative unit. With regard to small neighbouring countries which are included in the hinterland, their frontiers should also be the limits of the hinterland. With regard to big neighbouring countries the same should relate to the frontiers of republics or states.

The same problem with regard to the foreland is much more difficult because there are two zones to cope with, namely the sea zone and the land zone. Some authors suggest that the frontier should run along the sea shore and that

Fig. 2. The hinterland of Gdynia port with regard to the export of food products (1970)
Poviats which supply cargoes to Gdynia port: 1 — 0–50 t, 2 — 50–100 t, 3 — 100–1000 t, 4 — 1000–10,000 t, 5 — above 10,000 t, 6 — no cargo supply

the foreland should be limited to the sea zone. This however is unacceptable if we assume that the cargo, being the factor which determines the hinterland or the foreland, no matter whether in the sea zone or the land zone still belongs to the same sea transport shipper and still remains "cargo".

The borders of the foreland are generally the frontiers of overseas countries and in the case of bigger countries, the borders of republics, federal republics or states.

In 1970 the foreland of Polish ports consisted of 74 overseas countries to which cargo was exported or imported by sea, even if they were direct neighbours which is seen in Fig. 1. The reach of Polish activities as drawn on this map is purely theoretic and based on approximate data of Polish foreign trade.

All European countries, except Austria, Czechoslovakia and Hungary belong to the foreland of Polish ports. The borders (limits) of the hinterland are shown in Fig. 2.

Both, the hinterland and the foreland of the port may be considered in fractions thereof which means that one can make an analysis of a certain region of either of them or of a part of the total of cargo or of certain consignments or cargo. This complies with F. Morgan's[10] principle, which says that every port may have a number of hinterlands depending on its actual position. And so every kind of cargo determines its own hinterland of the chosen port.

Both the boundaries and the surface of the hinterland and foreland are not the most important problems in the research on sea ports. Much more important is the extent of links they maintain with the ports and this can vary largely with regard to different hinterlands or forelands and their size as determined by the rate of cargo saturation of a certain territory or else by the rate of concentration of cargo in different geographical regions.

Using a certain scale of cargo concentration with regard to a definite area one can easily show the extent of links between every part of the hinterland or the foreland with the port, hence the function of a certain territory with regard to sea transport. Based on the above we can make a regional division of the hinterland[11] depending on the links between individual regions and the port. This gives us information as to which part of the hinterland or of the foreland should be better looked after by the forwarding agents. Regions of high concentration of cargo require better transport organization to serve the hinterland or the foreland.

TRANSPORT AND ITS FUNCTION IN SHAPING THE HINTERLAND AND THE FORELAND

The flow of cargoes in the hinterland or the foreland is actuated by transport organizations which supply the port with the bulk of cargo which in turn is the ports basic function without which neither the hinterland nor the foreland nor the port could exist. The efficiency of the transport organizations in servicing the hinterland and the foreland has an influence upon the efficient functioning of the port.

Transport organizations perform their transporting job which is mainly of interest to economists and geographers. The economists appraise its functions from the point of view of expenditures and economic effects. For the geographers transport and its performance, and especially its main routes of activity

[10] F. Morgan, Observations on the study of hinterland, *Tijdschrift voor Economische en Sociale Geographie* 1951, 12.

[11] Cf. footnote 9, pp. 78–85.

TABLE 3. Distances to cargo concentration points on the hinterland of Polish ports in 1970 according to 100 km sectors by rail

With regard to import and export	0–100 km	100–200 km	200–300 km	300–400 km	400–500 km	500–600 km	600–700 km	700–800 km	above 800 km
	in thousands of tons								
	From Gdańsk								
Import	151.5	695.0	645.8	1,370.6	2,716.9	2,200.5	163	45.5	—
Export	405.7	510.1	1,270.6	1,039.0	2,651.5	15,957.2	319.6	78.4	—
Total	557.2	1,205.1	1,916.4	2,409.6	5,368.4	18,157,7	482.6	123.9	
	From Szczecin								
Import	412.4	233.2	1,233.5	1,021.8	3,206.5	1,357.5	528.4	94.1	44.0
Export	187.3	350.9	1,300.5	1,143.3	5,822.9	11,618.2	600.5	1,299.0	21.0
Total	599.7	584.1	2,534.0	2,165.1	9,029.4	12,975.7	1,128.9	1,393.1	65.0

Source: Calculations of the author.

are the best evidence of links existing between a certain region of the hinterland with the respective port.[12]

The results of transport activity in the hinterland depends on two factors: the quantity of cargo and the distance i.e. the mileage from the region to the port. The managers of transport organizations are interested in, what is called the dispersion of cargo in the hinterland and the foreland, namely with the influence of distance upon the transport costs and the efficiency of transport

TABLE 4. Distances to cargo concentration points on the foreland of Polish ports in 1000 km zone

Zone in thousand km	Through Suez Canal			Round the Cape		
	Total	Import	Export	Total	Import	Export
	in thousands of tons					
Total	34,836	10,733	24,104	34,836	10,733	24,103
0–1	9,929	1,437	8,492	9,929	1,437	8,492
1–2	9,378	1,902	7,476	9,378	1,902	7,476
2–3	1,738	577	1,161	1,738	577	1,161
3–4	1,495	1,104	391	1,495	1,104	391
4–5	1,575	1,109	466	1,575	1,109	466
5–6	2,883	797	2,086	2,883	797	2,086
6–7	501	174	327	466	152	314
7–8	944	702	242	979	724	255
8–9	224	64	160	125	55	70
9–10	512	307	206	508	307	201
10–11	142	56	86	128	56	72
11–12	526	453	73	525	453	72
12–13	1,902	1,280	622	98	90	8
13–14	236	68	168	15	2	13
14–15	497	121	376	507	130	377
15–16	16	14	2	—	—	—
16–17	62	10	52	14	12	2
17–18	8	1	7	98	17	81
18–19	308	43	265	53	4	49
19–20	31	1	30	—	—	—
20–21	292	81	211	77	17	60
21–22	3	3	—	1,684	1,078	606
22–23	386	284	102	183	44	139
23–24	112	92	20	34	2	32
24–25	—	—	—	329	137	192
25–26	1,136	52	1,084	34	3	31
26–27				138	11	127
27–28				292	81	211
28–29				386	284	102
29–30				—	—	—
30–31				—	—	—
31–32				1,126	48	1,078

Source: Calculations of the author.

[12] Cf. footnote 9, p. 86.

activities. The author of the present work applied in his research on Polish ports a graphic method based on calculating distances with regard to both — the hinterland and the foreland.[13] This method gives information on the whole territory of the hinterland and the foreland with regard to distance to cargo concentration locations. With regard to the hinterland 100 km long sectors were assumed and with regard to the foreland — 1000 km sectors.

Another method applied in the research on transportation was the flow and stream method which informs us about the quantity and the directions of the flow of mass cargoes between the ports and their hinterland or foreland. This method is shown in Fig. 3.[14] By using this method we can also obtain other infomation such as the rate of charge of the respective transportation routes, the demand for means of transport etc.

Fig. 3. The cargo streams of the Gdynia port hinterland, 1970
1 — import, 2 — export, 3 — coal export

For geographers this method is valuable because it provides quick appraisal of different regions or other parts of a geographic area from the point of view of links between them and sea transport. In other words, it shows the rate of spatial influence of the ports upon different parts of the territory of the hinterland and of the foreland.

[13] Cf. W. Barczuk, op. cit. and Fig. 2.
[14] Cf. W. Barczuk, op. cit. and Fig. 2.

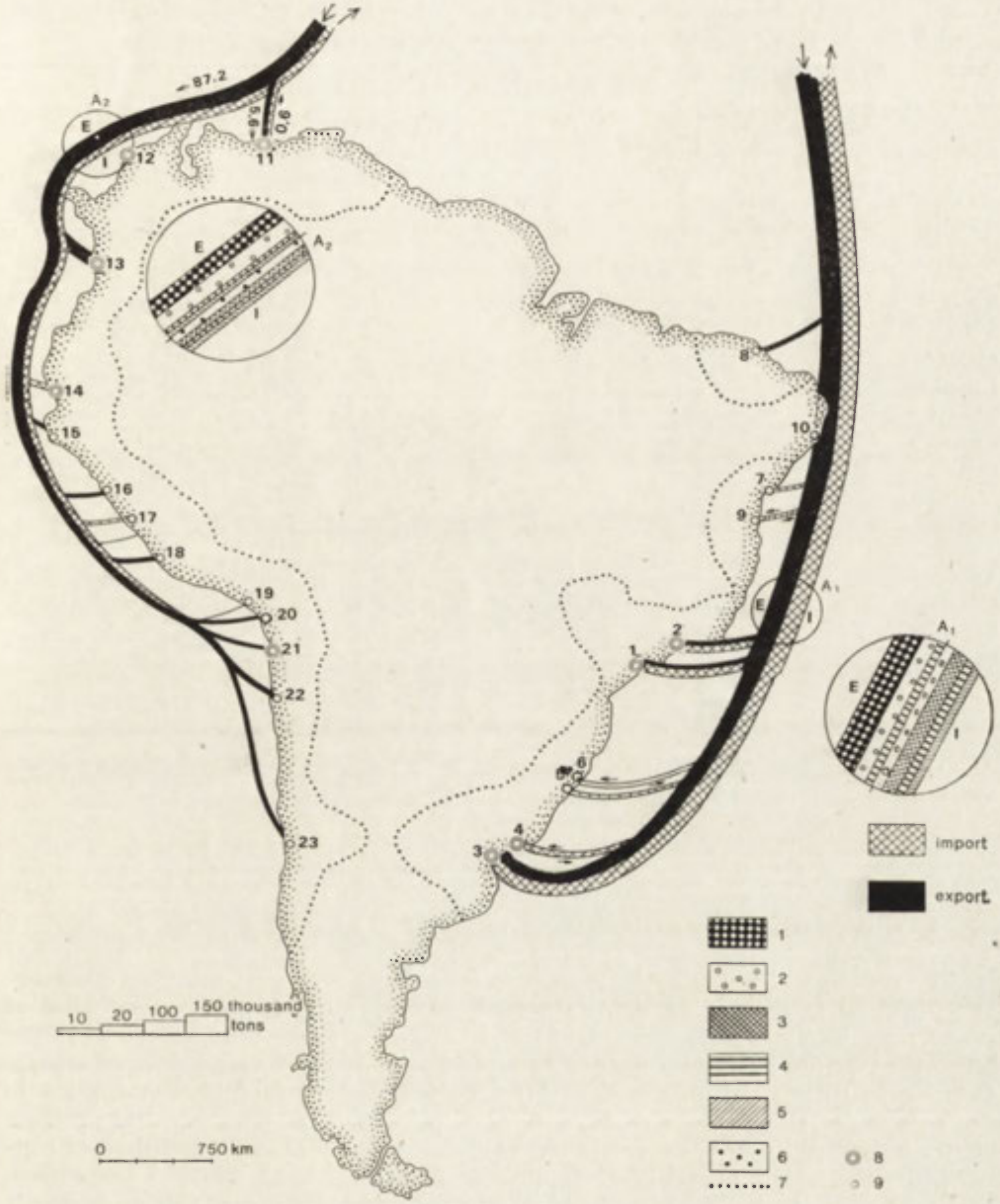


Fig. 4. South America as a part of Polish ports foreland with regard to general cargo

General cargo: 1 — heavy, 2 — packed, 3 — medium applicability for containers, 4 — high applicability for containers, 5 — cooled, 6 — refrigerated, 7 — inland border of the foreland (theoretical one

The ports of the foreland which have a turnover of goods with Polish ports equal to: 8 — less than 10,000 t, 9 — above 10,000 t

Ports which belonged in 1970 to the foreland of Polish ports:

1 — Santos Brazil, 2 — Rio de Janeiro Brazil, 3 — Buenos Aires Argentina, 4 — Montevideo Uruguay, 5 — Rio Grande Brazil, 6 — Pelotas Brazil, 7 — Salvador Brazil, 8 — Fortaleza Brazil, 9 — Ilheus Brazil, 10 — Maceio Brazil, 11 — Caracas Venezuela, 12 — Cartagena Columbia, 13 — Buenaventura Columbia, 14 — Guayaquil Ecuador, 15 — Talara Peru, 16 — Trujillo Peru, 17 — Huacho Peru, 18 — Callao Peru, 19 — Arica Chile, 20 — Iquique Chile, 21 — Tocopilla Chile, 22 — Antofagasta Chile, 23 — Valparaiso Chile

TABLE 5. South American zone — West Coast (Pacific) as part of the foreland of Polish ports in 1970 with regard to general cargo

Countries	Kind of general cargo						
	Total	Heavy	Packed	In containers acc. to applicability for containers[a]		Cooled[b]	Refrigerated[c]
				medium	high		
			Import				
Total	88.4	—	77.3	6.2	2.4	4.0	2.0
Venezuela	1.8	—	—	—	0.1	0.5	1.2
Columbia	1.6	—	1.5	—	0.1	—	—
Ecuador	3.4	—	1.5	1.7	0.1	0.2	—
Peru	74.9	—	73.4	—	1.4	—	—
Chile	5.4	—	0.7	4.5	0.3	—	—
Other	1.3	—	—	—	—	0.6	0.7
			Export				
Total	84.5	45.2	29.8	6.0	1.43	1.1	—
Venezuela	4.1	1.1	2.2	0.7	0.1	—	—
Columbia	48.9	37.0	11.7	1.1	0.4	—	—
Ecuador	9.6	4.5	4.1	0.4	0.6	—	—
Peru	4.5	1.7	—	2.0	0.1	0.6	—
Chile	12.6	0.1	10.7	1.8	0.1	—	—
Other	3.5	1.0	1.8	—	0.2	0.5	—

[a] General cargo suitable for container transport, not all of which has been transported in that way this year.
[b] General cargo which requires cooled transport.
[c] General cargo which requires refrigerated transport.

DETAILS ON THE HINTERLAND AND THE FORELAND OF PORTS

The principle defined by F. Morgan [15] says that every port may have numerous hinterlands depending on the actual position and likewise every kind of cargo may also have its hinterland. This allows a more detailed look at the hinterland of the port with regard to the definite kind of cargo or a small-size area i.e. to treat the hinterland as divided. A similar principle was applied in the analysis of the foreland of Polish ports. [16] A detailed look upon the problem of the hinterland or of the foreland enables a more thorough and deeper consideration of the topic which is appreciated by transport routine employees. The map (Fig. 4) is an example of a tabular presentation of the problem. It shows the foreland of the Latin American East Atlantic Coast region and the West (Pacific) Coast [17] — for the latter Table 5, which is an example of a tentative analysis.

Gdańsk University

[15] F. Morgan, op. cit.
[16] W. Barczuk, *Przedpole portów*...
[17] W. Barczuk, op. cit.

POSSIBILITIES OF DETERMINING THE FACTORS THAT AFFECT URBAN LAND USE. CASE-STUDY OF WARSAW

JULITTA GROCHOLSKA

For quite a long time the problems of land use have been discussed in various geographical papers. It would hardly be possible to name all of them, therefore we will mention only the most important ones. The leading work in this respect seems to be that of Sir L. Dudley Stamp who initiated and carried out the investigations in England, Scotland and Wales in 1939–1949. The effect of those investigations was that he presented a definite method of research and made a land use survey which showed the distribution of the main forms of land use (L. D. Stamp, 1948). In Poland, research of this type was initiated in the mid-fifties by J. Kostrowicki. It was an attempt to work out a detailed land use survey that would fit the classification of the International Geographical Union. In comparison with the previous ones, the Polish survey was richer by additional information on the ways and trends of the utilization of the main categories of land use, as can be seen in the map (J. Kostrowicki, 1961, 1964).

As for settlement land use, which is the subject of the present paper, current investigations are a continuation of what are called landscape studies, initiated at the beginning of the 20th century. The adherents of this trend were of the opinion that a town differs from a village first of all by its landscape. Thus their investigations were concentrated on the analysis of the urban landscape.

In post-war studies one can notice a more thorough approach to the problems and a search for more sophisticated methods of analysis. In the case of urban studies, this is revealed by combining the research on land use with functional analysis or with research on the spatial structure.

The studies on land use in towns and cities which have appeared during the past 30 years, can be roughly divided into four groups. Group one — studies of a monographic character which contain some information on land use, yet treat it rather marginally. Group two — works of an historical character in which information can be found on the formation of the structure of land use in the past, this, however, proves insufficient to the analysis of transformations in land use. Group three — works of a plannistic character. These afford much concrete information on the current and forecasted use of urban areas, the scope of area covered by the different categories of land use, etc. However, they are not pure geographical studies and thus the information they contair are as a rule insufficient for more detailed investigations. Group four includes on the one hand studies whose main aim is land use analysis, and on the other — works of a general character, for instance those which present the current state of research in this field, certain instructions, etc.[1] All these works are character-

[1] Some more information on this subject can be found in two other publications by the same author (J. Grocholska, 1970; 1974).

ized by different approaches to the subject, from the point of view of both methods and research work, which is well seen, among other things, in the fact that they generally point to quite different factors which in their opinion affect the manner of land use.

When taking J. Kostrowicki's definition of land use, according to which it is "the use of the geographical environment, i.e., of the power and resources of nature used by all sections of human activity", [2] one must notice the different character of the various forms of land use. Thus, for instance, if we compare settlement land use with other forms (sections) of human economy, we can see that all these different forms are differently related to the environment as far as its utilization is concerned, and also differently dependent on it. It is generally known that the effectiveness of the use of agricultural or forests areas depends to a large extent on the geographical environment with which it is directly connected. On the other hand, settlement land use is much less dependent on existing conditions. It should be remembered, though, that the term "settlement land use" is a very general one. It can be used for "macro" investigations, i.e. those carried out for the whole country or a particular region. In research into a selected town, city, or hamlet, it is indispensable to introduce a further division of settlement land use. In such cases, the manner of land use should be adjusted to the existing conditions, so as to ensure the rational management of both natural resources and man-introduced investments.

In order to know the influence of various factors on the categories of land use in the town, it was decided to carry out a detailed analysis of a selected subject. For these investigations we have chosen the capital city, Warsaw; the analysis was made for the year 1965. This choice was made for several reasons. Firstly, so far there is little geographical literature dealing with Warsaw. Secondly, the studies on land use in cities are also rather scarce in Poland, although, as it should be recalled, such studies are *par excellence* geographical: they add to a better knowledge of the subject of research, which for settlement and population geography is the distribution and structure of historically shaped territorial human groups (in principle, housing estates), and of permanent facilities connected with them. Investigations in this field should include not only man and permanent facilities, but also the forms of utilization of both: permanent facilities as well as the area they cover (K. Dziewoński, 1956).

Thus, in research on cities, the subject of investigation should be the analysis and estimation of land from the point of view of a rational utilization of the existing resources of the geographical environment with the social interest broadly understood. From this formulated subject of investigations ensues the research postulate, which is the search for the factors that affect the manner of land use.

These studies on Warsaw have been aimed to find these factors. They will be described later on, together with the research procedure which was accepted for their identification. Here it should only be emphasized that both the factors themselves and the force of their influence, can be different in different cities. The aim of this publication, which is an abridged version of a larger work by the same author, is to show a certain procedure as well as the possibilities of investigation resulting from it. The essential aim of this kind of study in general is to know more about the very nature of land use in large urban agglomerations.

[2] The definition is included in the study on the current state of research on land use in Poland (J. Kostrowicki, 1959).

In this investigation, which was aimed to determine the importance of the influence of the factors under consideration, the material for a detailed statistical analysis was collected by representative sample technique. This method, apart from its other merits, allows to make the analysis more detailed and less labour-consuming. Besides representative sample technique is a generalizing method: conclusions are drawn out of a certain number of observations which have been selected for the investigated population. This was an additional argument for the application of this particular method.

The selection of samples was carried out by laying a square geometrical grid of 1 km side in the system of Cartesian coordinates (Fig. 1) upon a plan of the city scaled 1 : 10,000. Such a formalized frame of reference enables the researcher to disregard the administrative divisions, as well as to have the same size

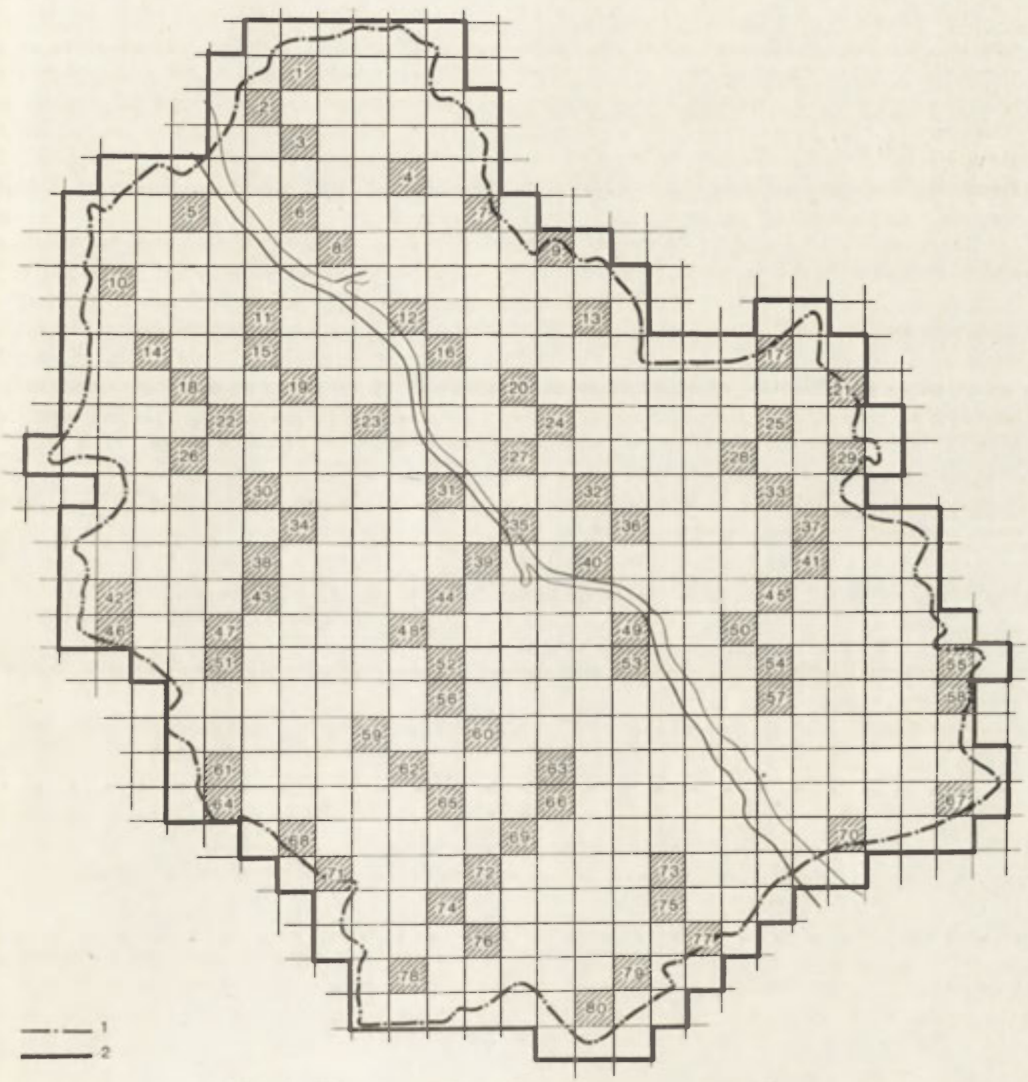


Fig. 1. Warsaw. The scheme of sample selection
1 — the administrative boundary, 2 — the boundary of the analysed population

of particular elements (squares) under consideration. The latter fact creates the possibility of comparing these elements, and makes calculation much easier.

A special table of random numbers (called iron numbers), worked out by H. Steinhaus[3] was taken as the basis of selection. In view of the fact that the aim of quantitative analysis was to estimate the parameters of linear multiple regression with many independent variables, it was decided that 80 elements would be enough for the sample. The sample comprised about 16% of the investigated population, which consisted of 508 squares.

Having chosen the sample of 80 squares, the values of particular characteristic features of every square, as a specimen of the investigated population, could be established. Those features were selected based on the supposition that there exists a certain dependence between the size of the area, covered by a given type of land use, and the factors which are the characteristic features of this area. Moreover, the analysis was based on the assumption that in the spatial structure of land use, the probable dependence between land use, as expressed in its various forms, and certain factors which were to be determined is reflected. This dependence was presented by means of many statistical models. They were the linear models of the function of regression of many variables devised separately for each kind of land use. The analysis was made for all basic kinds of land use, whose character indicated that their localization was due to some definite reasons.

Each of the investigated types of land use was treated as dependent variables and denoted by symbols $y_1, y_2, ..., y_l$. The value of the analysed variable y was calculated as a fraction of the area occupied by the given type of land use within each of the 80 squares; the area of the square was assumed to be equal to 1 (100%). Individual y variables were recognized on the basis of the function performed by the given area within the general socio-economic spatial pattern of the city. On the basis of the similarities of these functions 10 fundamental classes of land use were distinguished which within the assumed model were treated as dependent variables $y_1, y_2, ..., y_{10}$.

Last of all were distinguished three groups of factors responsible for the formation of a given kind of land use.

Group I — natural conditions, including the effect of such factors as ground water, ground carrying capacity, soil grades, climate and sites of natural quality.

Group II — technical infrastructure; the access of the examined area to the existing water, sewage, heating and gas networks.

Group III — the situation of the area, considering its distance to the city centre and the type of land ownership.

These factors were denoted by appropriate independent variables, $x_1, x_2, ..., x_k$.

Apart from the above-mentioned variables, there are many others, which have not been included in the analysis, although there is no doubt as to the influence they exert on the type of land use. They are, for instance:

— the personal ideas of the town planners responsible for designing the city's spatial pattern,

— the general socio-economic directives,

— the various socio-cultural factors, including the significant factor of the population's preferential attitudes,

— historical factors, etc.

---

[3] Cf. W. Sadowski, *Tablice statystyczne. Liczby żelazne* (Statistical Tables. Iron Numbers), Warszawa 1951, Table 31, p. 123.

The influence of all these variables, which were not analysed in the models because they proved to be difficult to measure or of less importance, has been taken into account by introducing an appropriate random factor into the regression model.

When compiling a list of independent variables, which were taken into account in the statistical analysis, efforts were made to get the specificity of every factor, with special regard to the influence it can (or should) exert on the manner of land use. Finally a list of 9 independent variables was accepted.

For each dependent variable, $y_1, y_2, ..., y_{10}$, was formulated one multiple regression equation which defined its dependence on the independent variables. The latter (x_1, $x_2, ..., x_9$) were treated as the variables providing information as to the values of dependent variables which measure various categories of land use. Thus it can be said that the accepted models have been formulated in view of the assumptions that there exists a dependence between certain categories of land use on the one hand, and the definite factors characteristic of the given dependent variable, on the other. The random factor of the model can be interpreted as a joint influence of all other variables, not included in the model. These variables are difficult to define and to measure, yet indubitably they do affect land use. It has been assumed that the random factor, denoted by ξ, will be treated as a random variable, which is assumed to have normal distribution with a mean equal to zero and variation σ^2. The general notation of the linear model constructed for each analysed variable has the form:

$$y = \beta_1 x_1 + \beta_2 x_x + \dots + \beta_k x_k + \beta_{k+1} + \xi$$

where $\beta_1, \beta_2, ..., \beta_{k+1}$ are unknown parameters of regression equations, and ξ is the random variable.

Based on the general form of the model, particular models for each analysed variable y were built. All independent variables were selected in such a way as to take into account all probably existing dependences.

Thus, for instance, building model No. 1 for multi-family dwellings, the following factors were taken into account: ground water (variable x_1), ground carrying capacity (variable x_2), sites of natural quality (x_5), the existence of at least one of three essential types of technical infrastructure, i.e. water, sewage, or gas networks (x_6), central heating networks (x_7), distance to the city centre (x_8), as well as the occurrence of municipal land ownership (x_9). The above variables were included in the model in view of the possible dependence between the spatial arrangement or multi-family dwellings, and the definite factors expressed by these variables.

Model No. 1 — for multi-family dwellings — could thus be written as follows:

$$y_1 = f(x_1, x_2, x_5, x_6, x_7, x_8, x_9, \xi).$$

Models for the remaining nine dependent variables y were built in the same way.

A comparison of these models with the information about dependent and independent variables obtained from the sample has shown that some models may be badly estimated due to a small number of observations with variable values different from zero. Thus new models had to be built. This was done by combining some categories of land use with similar functions and similar localization requirements. In the end of the following dependent variables have been assumed as integrated variables (aggregates):

z_1 — fraction of the areas occupied by housing development,
z_2 — fraction of the areas occupied by administration and services,

z_3 — fraction of the areas occupied by industry and stores,
z_4 — fraction of the areas occupied by public open spaces, forests and sports,
z_5 — fraction of the areas occupied by the internal and external transport network,
z_6 — fraction of the areas occupied by agricultural areas.

For these variables new models have been built which took into account all those variables which were included in the models built for appropriate variables y. Thus the list of models for integrated variables z is as follows:

Model No. 11 for housing development:

$$z_1 = f(x_1, x_2, x_4, x_5, x_6, x_7, x_8, x_9, \xi).$$

Model No. 12 for administration and services:

$$z_2 = f(x_1, x_2, x_4, x_5, x_6, x_7, x_8, x_9, \xi).$$

Model No. 13 for industry and stores:

$$z_3 = f(x_1, x_2, x_6, x_8, x_9, \xi).$$

Model No. 14 for public open spaces, forests and sports:

$$z_4 = f(x_1, x_3, x_4, x_5, x_8, x_9, \xi).$$

Model No. 15 for the internal and external transport network:

$$z_5 = f(x_1, x_5, x_8, x_9, \xi).$$

Model No. 16 for agricultural areas:

$$z_6 = f(x_1, x_3, x_4, x_8, x_9, \xi).$$

In comparison with the previously described model for multi-family dwellings, the new model for the housing development (combining multi-family and one-family dwellings), additionally takes into account the influence of climate (variable x_4). Variable x_4, which occurs in models No. 14 and 16, denotes soil fertility.

Finally, the models for 16 dependent variables, have been obtained, from among which models for variables z (Nos. 11–16) were assumed as the most important ones. These models were estimated on the basis of statistical data obtained from the sample. In all these models f denotes the linear function together with the free expression.

The model of the regression function of many variables, has been taken as the basis for calculations in view of the assumption that there exists a correlation between the dependent variables y and independent variables x, the latter being treated as informational variables. In such cases, the exactness of the correlation is determined by the values obtained of the multiple correlation coefficient of the particular regression equations, or by its square, known as the determination coefficient. For the problem under consideration the model of linear regression has been assumed, for which numerical calculations are much simpler than in non-linear models.

The spatial pattern of various categories of land use is affected by many different factors. Therefore the aim of the investigations was to determine the influence (or lack of it) of each of the examined factors on the given dependent variable. Such an analysis would permit the verification of the existing views on statistical relationships which accompany the causalities. It should be emphasized that the determination of correlation itself does not bring any information about the causality. The fact that we know the result does not allow us to draw unequivocal conclusions as to the reason. In our work statistical analysis has been used as a tool for detecting probable existing relationships between certain phenomena. The investigation of correlation is merely enough for the first step of the analysis. The lack of correlation means that there is no causality and further inquiries would be futile. If, however, correlation is found, this may mean that either there exists a causality, or else that there exists what is called "apparent correlation", caused by a third reason which acts simultaneously on variables x and y.

Each of the 16 models, devised for the particular variables, was estimated by the same procedure. Unknown parameters, $\beta_1, \beta_2, ..., \beta_{k+1}$, were estimated by the classical least-squares method. This method allows the estimation of the parameters of the model that would be best adjusted to the empirical data, with the assumed criterion. In the models in which some random factors occur in addition to independent variables, any method of estimation will always give the values of parameters with some random error. Therefore in such cases what are called the "mean errors" of estimation for particular parameters are sought.

In view of the great labour consumption of calculations for multi-dimensional models, calculations have been performed on an Odra-1204 computer, using the standard programme of regression analysis.

The basic data necessary for computer calculations were the observation matrices on independent variables, and observation vectors on the dependent variable in the given model. The values of dependent variables for each square were obtained by measuring the area occupied by the given type of land use (e.g. housing land use) and calculating its per cent share in the whole square area. In the case of immeasurable features, the values of independent variables have been introduced into the matrix as zero-one variables, assuming the value of 1 in the case when the given factor does occur, and the value of 0 — when it does not occur. Because there exist two kinds of models (for variables y and for integrated variables z), two observation matrices have also been constructed: Final calculations gave the following results:

— estimation of parameters $\beta_1, \beta_2, ..., \beta_{k+1}$,

— the value of multiple correlation coefficient, indicating the degree of adjustment of the model to the empirical data,

— the value of t-Student's statistics, used for testing hypothesis concerning the importance of the obtained multiple regression coefficient (parameters $\beta_1, \beta_2, ..., \beta_{k+1}$).

On the basis of the values obtained of t-Student's statistics all insignificant variables could be removed from the model. After all insignificant independent variables had been removed, the calculations were repeated in order to estimate the final model with all statistically significant independent variables. The accepted procedure has made it possible to compile an accurate list of those variables which in the course of analysis were found to have influence on the manner of land use in 1965. The spatial arrangement of particular forms and kinds of land use is illustrated in the map (Fig. 2).

Our investigation has not solved all problems, yet it was significant as an example of the application of statistical methods in spatial research. Owing to the method of regression analysis of many variables it can be said now in which cases a correlation was observed between the variables that express the manner of land use (variables y and z) and those defining the characteristic features of the area (variables x).

The investigations have revealed a discrepancy between the fully rational form of land use in a big city, as was looked for, and the real picture of this city at a given moment of its historical development.

The results obtained have shown that the land use existing in Warsaw in 1965 can be only partly explained in terms of the dependence on the above-mentioned factors. As it has been characterized based on different, appropriately arranged classes, it is to these that the results obtained were ascribed in the first place. Thus, according to the different degree of dependence which were found, certain classes of areas with similar localization features have been distinguished.

Class A — the categories of land use characterized by the least interdependence on the distinguished factors. They were: housing development, industry and stores. Only one factor — technical infrastructure — has proved to be significant for them. Apart from that, no regularities in the spatial arrangement of these areas have been observed.

Class B — the categories of land use characterized by a significant dependence on several factors. It included: administration and services, and recreational and agricultural areas. The spatial arrangement of these areas exhibits some regularities and a smaller influence of random factors is observed.

Class C included only one category of land use — public transport, which exhibits the greatest dependence on the distinguished factors. However, in view of the specific character of this manner of land use, the results obtained should be treated with particular thoroughness.

The areas occupied by the internal and external transport network have not been included in the final classification, as it was considered impossible to disclose any regularities of their spatial arrangement, unless extending the analysis beyond the administrative boundaries.

When discussing the results of our studies, it should be recalled that all fundamental conclusions have been drawn on the basis of a 16% sample. In general, the representativeness of such a sample is checked by adequate tests. However, in our case this could not be done due to the lack of materials. Therefore all the values for the observation matrix had to be calculated, which means that the results obtained should be treated with a certain amount of criticism.

Generally it can be said that not enough attention has been paid to the areas of natural conditions, which could be seen in frequent inadequate adjustment of land use to the existing conditions. On the other hand, a modifying influence of the city has been observed.

Moreover, the influence of the technical infrastructure has proved to be relatively large. In other words, the outlays that had once been put into it, raise the value of the area so much as to be decisive in its further management. This can explain why for instance when restoring land use on the area of pre-war Warsaw, the old pattern was generally preserved.

The influence of the distance from the city centre has not appeared everywhere, which indicates that Warsaw has preserved only to some limited extent the features of a concentrated lay-out. Contemporary Warsaw is a city of relatively little compactness, stretched out along main lines of communication, with enclaves of agricultural areas coming near to the city centre. The lay-out of 1965 indicates the superior role of administration and services centres and the formation of districts centres.

Also the factor of land ownership proved to be of little importance. This means that the nationalization of land within the boundaries of pre-war War-

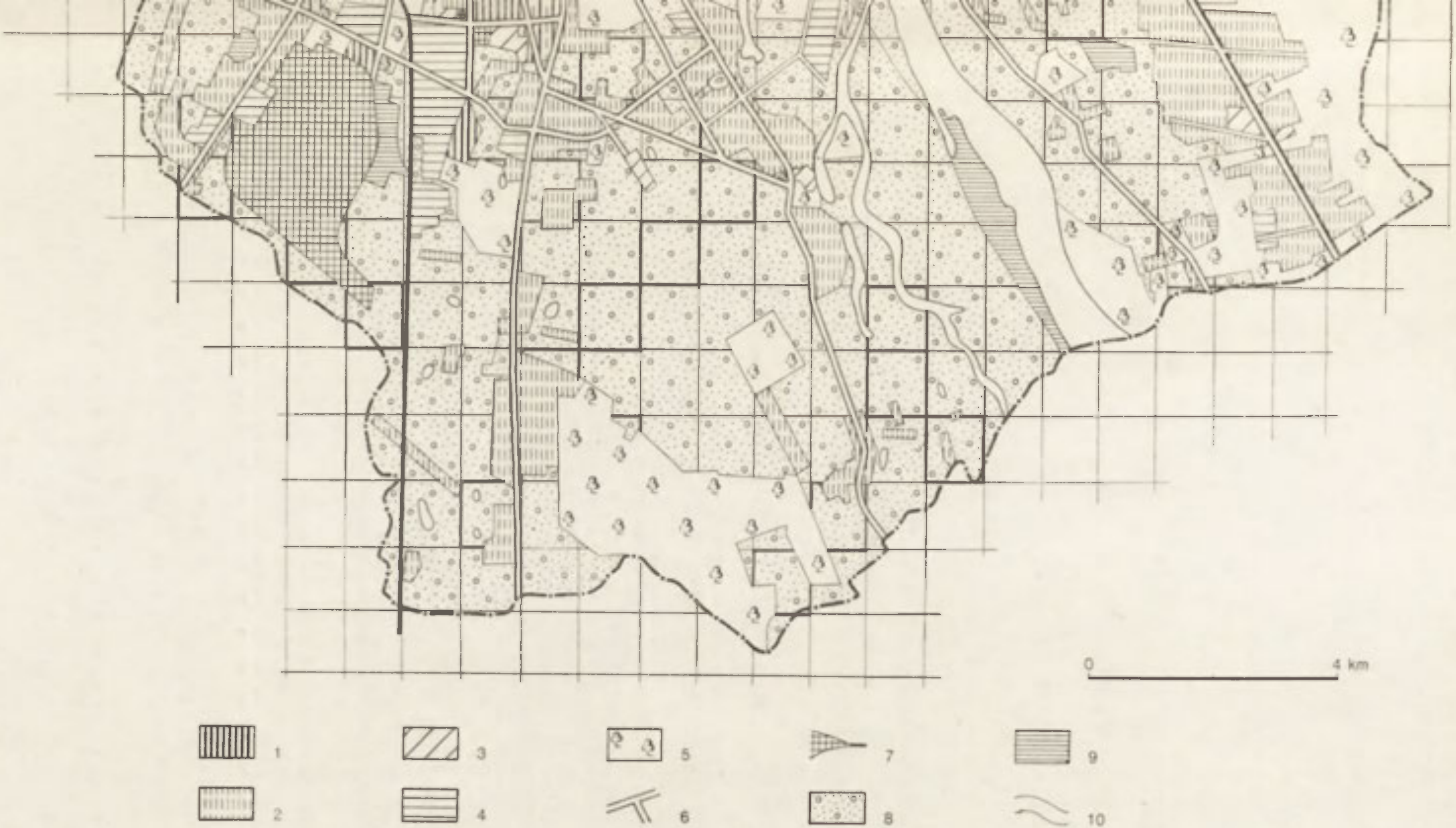


Fig. 2. Warsaw 1965. Land use

1 — multi-family dwellings, 2 — one-family dwellings, 3 — administration and services, 4 — industry and stores, 5 — public open space, 6 — main roads, streets and squares, 7 — national and international transport, 8 — agricultural areas, 9 — other uses, 10 — surface water
Note: Mixed uses are marked by the combination of the above signs.

saw, an unprecedented fact in the country, has not exerted any significant influence on the subsequent manner of land use.[4]

Of course, on the basis of the above studies alone it is extremely difficult to estimate the land use in the Warsaw of 1965, as it would mean great simplification. However, it seems that some of the inconsistent decisions and changes in the principles of realization of the spatial policy in the past, have found their expression in the present utilization of the areas.

The chief aim of this publication has been to show the possibilities of determining the factors that influence land use in cities. Therefore, equally important as the empirical results obtained is the fact that the usefulness of the applied procedure for the studies on spatial pattern of land use in cities has been demonstrated.

First of all, it should be emphasized that the spatial pattern of land is a very complex problem. It is the result of the impact of a great number of factors which frequently overlap each other, thus giving rise to a chain reaction or feed-back. In this situation it is extremely difficult to establish the regularities or the impartially existing interdependences.

Because of this entanglement of factors one would not be able to obtain any significant results, either methodical or cognitive, unless some reduction was made in the method. Therefore statistical formulation was accepted which meant that for the time being we have had to be satisfied with the diagnosis of the currently existing state. Perhaps in the future it will be possible to improve the procedure so as to procure such tools of analysis that would permit the study of the transformations as well. However, this would require further investigation and more precise determining of the variables taken into account. It also seems necessary to try again to express the analysed factors in measurable characteristics and to introduce a new variable — the factor of time.

In the course of research on the pattern of land use in Warsaw at a given moment of its historical development it has been found that some of the factors have a systematic, and others — a random character. Systematic factors should be characterized by some regularities. Based on the theorems of the calculus of probability it can be assumed that between the factors of systematic character and some categories of land use there can exist a certain dependence, or the phenomenon of co-existence. This assumption has made it possible to present all possible dependences in the form of regression equations of y in relation to x — with a large number of variables and a random factor that disturbs the regularities. These equations have been presented in the form of mathematical models and have been treated as a simplified reflection of reality. All this has made possible the analysis of the current state of land use from the point of view of definite assumptions.

As has already been said, the basis for statistical inference has been a sample of 80 elements. This was done in an attempt to check the usefulness of representative sample technique in this kind of research. It has been shown that in the cases where it is impossible for one person to carry out a detailed research, or when the researcher has not got sufficient empirical material, this method brings satisfactory results. However, when analysing a larger number of categories of land use (the author has assumed six of them, as she had to do the grouping), the numerical force of the sample should be increased.

---

[4] The decree on the nationalization of land (so-called communalization) was passed on October 26, 1945. It liquidated private land ownership on the area of pre-war Warsaw.

The problem of selection of the elements for a detailed analysis remains an open question. Perhaps one could obtain more interesting results and a better adjustment of the model to the empirical data using another, more elaborate scheme of selection for the sample.

This research has been the first attempt to check certain hypotheses about the spatial arrangement of various forms of land use by means of the analysis of regression of many variables. Hence also the procedure of quantitative formulation of the factors chosen for the analysis should be treated as some kind of methodical suggestion. For instance, for some time spatial planners have agreed that in the optimalization calculus elements of natural conditions should be taken into account. However, so far no unequivocal method has been proposed to do it. On the contrary, their immeasurability is frequently mentioned.

This study comprised both the factors belonging to a group of natural (i.e. primary) conditions, and those created by man (i.e. secondary, such as the technical infrastructure, the situation of the area) which could not be expressed in measurable characteristics.[5] All of them have been introduced into statistical analysis in the form of appropriately constructed variables zero-one. This was a certain step forward, as it permitted the estimation of them objectively, and the adjustment of the accepted convention to the requirements of different modes of land use and to the natural conditions of the given area. It seems that in such a way their introduction into information matrix can be applied also in other studies of this kind.

As a role of the geographer consists not only in describing the phenomena and looking for their reasons, but also in forecasting, it seems that in this field too the accepted procedure can bring some creative inspiration. Hitherto performed studies were directed towards formulating models of a standard character, which would permit the finding of the best solution, i.e. the maximal effectiveness of a given manner of land use. This effectiveness should be adequate to the existing conditions, both natural and those created by man. It should also help to realize the aims of the socio-economic policy. We do hope that the experience and results obtained in the course of this study will make the formulation of models that will permit the determination of the best structures under given conditions easier.

Institute of Geography
Polish Academy of Sciences, Warsaw

BIBLIOGRAPHY

Dziewoński, K., 1956, Geografia osadnictwa i zaludnienia. Dorobek, podstawy teoretyczne i problemy badawcze (Sum.: Geography of population and settlement — Achievements, theoretical bases and plan for research), *Przegl. Geogr.*, 28, 4, pp. 723–762.

Grocholska, J., 1970, Investigations on urban land use in Polish geographical studies, *Geogr. Pol.*, 18, pp. 85–92.

Grocholska, J., 1974, *Czynniki wpływające na użytkowanie ziemi w Warszawie* (Sum.: The factors influencing land use in Warsaw), Studia KPZK PAN, 46 128 pp.

[5] The readers' attention should be called to the fact that taking the average values of results (with lack of appropriate data) does not afford any new cognitive values and therefore has been rejected as a method in this study.

Kostrowicki, J., 1959, Badania nad użytkowaniem ziemi w Polsce (Sum.: Research studies on land utilization in Poland), *Przegl. Geogr.*, 31, 3–4, pp. 517–530.

Kostrowicki, J., 1961, Polish land utilization survey, in.: *Problems of applied geography, Prace Geogr. IG PAN*, 25, pp. 45–68.

Kostrowicki, J., 1964, Geographical typology of agriculture in Poland. Methods and problems, *Geogr. Pol.*, 1, pp. 111–146.

Stamp, L. D., 1948, The land of Britain — its use and misuse, London 507 pp.

A MODEL OF RESIDENTIAL STRUCTURE IN A SOCIALIST CITY. A CASE STUDY OF WARSAW *

PEGGY A. LENTZ

THEORETICAL FRAMEWORK

A number of theories and models have been formulated to deal with urban land use development and the organization of the urban landscape. They include three major, interrelated contributions in the field which have been made by Alonso (1964), Muth (1969), and Wingo (1961). Among these authors only Muth explicitly considers the housing sector of the land market. Because the theoretical and empirical analyses for socialist economies, which follow in succeeding sections, are oriented to the residential sector of the urban area, it seems that Muth's model may provide a good theoretical framework. Thus, selected portions of Muth's formulations will be re-examined for implications relating to a socialist economy. There are a number of assumptions used to derive urban location theories that are applicable to socialist urban areas: (1) All workplaces are located at a central area on a homogeneous plain — while the center of a socialist city may not be a central *business* district, it is as much the central point for retailing, administrative offices, light manufacturing, cultural functions, etc., as it is in any capitalist city. For the city of Warsaw, a national capital, this centrality assumption is perhaps even more reasonable since a high proportion of employment is in administrative services. (2) Households are far likelier to be identical in incomes in a socialist country, and the prevailing philosophy of socialism encourages collective attitudes — which may be regarded as similar preferences, and family compositions. (3) The costs of commuting are probably more uniform per individual in a socialist context, since public transportation predominates. As for marginal costs of transport, these are as likely to be a function of straight-line distance as anywhere else. It is still a reasonable assumption that locations distant from the central area have less value to the individual, because the costs of transport, even if only measured by the time spent in transit, is a valid consideration. And finally, (4) consumers in socialist countries are not different from non-socialist consumers; they too wish to maximize their satisfaction in regards to all commodities given a budget constraint.

However, some of the assumptions are by definition not applicable. Since socialist countries operate under planned economies, the assumptions related to a free market may not be relevant. It is true that those individuals who

* This article is an abridged version of Chapters: III, and VII from: P. A. Lentz, *The influence of a socialist economic environment on intra-urban residential spatial structure*. Ph. D. dissertation, Department of Georgaphy, Clark University, Worcester, Mass., 1973.

make the necessary locational decisions have as perfect knowledge as possible about the prices of commodities, the quantities available, and the amount of funds they have to invest. Due to the numerous cost efficiency studies made by planning research institutes, they are also aware of the currently cheapest choice of factors to achieve their objectives. It is certainly true that an individual cannot influence the prices of available commodities. Socialist economies have traditionally not considered economic rents for land. All locations, including those in the central districts, are allocated to uses on the basis of considerations other than rent pay ability, but the utility and scarcity value of land are taken into account. Finally, it is fairly obvious that a socialist economy, in its institutions and characteristics, does affect the location of urban land uses, and it may be possible to understand in some manner how these locational decisions occur.

A reformulation of the residential model. Muth's model deals with the location of residences in a city. He analyzes how both the consumers and the producers of housing achieve their equilibrium positions. The utility function Muth ascribes to consumers:

$$U = U\ (x, q) \tag{1}$$

where: x is expenditures on all commodities including leisure but excluding housing and transportation, and q is quantity of housing, is not quite true for consumers in a socialist society. The quantity of housing consumed anywhere and by anybody can be assumed to be a constant; housing is mainly allocated on the basis of family size, and in the present analysis family size is assumed to be constant. Thus, individuals attempt to obtain the maximum possible amount of all commodities available except housing, given a fixed income

$$U = U(x) \tag{2}$$

Additionally, the budget constraint:

$$y = r(s)q + T(s, y) + x \tag{3}$$

where $r(s)$ is defined as the price per unit of housing which is a function of distance from CBD, and T, the cost per trip to the CBD, is a function of y, income, and distance from the CBD, will require some rewritting since the price of housing is not a function of distance in a socialist city but a constant, r_0.

$$y = r_0 q_0 + T(s; y) + x \text{ and also}$$

$$g = r_0 q_0 + T(s, y) + x - y = 0 \tag{4}$$

In order to find the optimal amount of x from equations (2) and (4), first-order conditions for household equilibrium can be derived using the method of Lagrangean multipliers.

$$Z = U(x) + \lambda[r_0 q_0 + T(s, y) + x - y] \tag{5}$$

Taking first partial derivatives and setting them to zero

$$\partial Z/\partial x = U_x + \lambda = 0 \tag{6a}$$

$$\partial Z/\partial \lambda = Y - [x + r_0 q_0 + T(s, y)] = 0 \tag{6b}$$

$$\partial Z/\partial s = \lambda(T_s) = 0 \tag{6c}$$

It should now be possible to obtain an optimal solution for x and s, but upon inspection it becomes clear that equation (6c) prevents this. Consider that in order for this relationship to be true, either λ or T_s must be equal to zero.

But we have already indicated that transportation costs must be a function of distance, if only because the individual's time spent in transit has a value; and the greater the distance, the more time must be spent. By equations (2) and (6a), λ cannot equal zero. Therefore, there must be a missing variable in equation (6c) which will be labelled c. But this implies that the same variable is missing in either equations (2) or (3) and in equation (5).

How is this new variable c to be defined? Recall that $r(s)$ was defined as the price per unit of housing, inversely related to distance from the CBD. At the city center $T(s, y)$ equals zero and $r(s)$ equals some value r_0', which is a bid rent. With a unit increase in distance, the value of $r(s)$ declines, but at the same time the value of $T(s, y)$ increases. In case of the capitalist city (Lentz 1973, p. 32) r_s was directly related to $-T_s$ and inversely to q. It can be seen immediately for the socialist city that r_s only depends on T_s since q is constant for every individual at any location in the city. Thus, c may be defined as:

$$c = r(s) - r_0' = c(s) = -T(s, y) \tag{7}$$

where $r(s)$ is taken to be the price of housing that would have occurred in a free market at distance s from the CBD, with quantity of housing held constant.

In effect, any location except the central one is a diseconomy to the individual, and the maximum utility function for the consumer occurs in the central area. Thus, equations (2) and (5) can be rewritten as

$$U = U[x, c(s)] \tag{2'}$$

and

$$Z = U[x, c(s) + \lambda[r_0 q_0 + T(s, y) + x - y] \tag{5'}$$

Thus, for individuals living at the center, $c(s) = r(0) - r_0' = 0$ and total income is $y + c(s) = y$. For individuals not at the center, $c(s) = r(s) - r_0' = -T(s, y) < 0$ and total income is $y + c(s) < y$. The variable $c(s)$ represents a loss of real income for individuals living at distance s from the center; this loss equals zero at the city center and increases linearly with distance.

Given the conditions defined in the above analysis, it is clearly to the advantage of every individual to try to live as close to the city center as possible. Of course, planning and land use controls prevent people from locating in any manner they might find advantageous. Housing is allocated very carefully, and the intensity of development is subject to planning commission decisions. However, any housing in Warsaw, if put on the market, will command a price larger than the nominal price; and housing near the center of Warsaw commands the highest price of all.

Continuing with the main analysis, equation (6c) can be rewritten as

$$\partial Z / \partial s = \lambda T_s + U_c = 0 \tag{6c'}$$

It is now possible to obtain an optimal solution for x and s, but it should be obvious that the individual wishes to locate as near as possible to the center. The optimum amount of x the individual can obtain depends on the amount of transport costs incurred by traveling distance s. This is clearly an unstable equilibrium except that socialist governments try to counteract this instability through residency regulations and planning controls.

Casetti (1967) has suggested that there may exist another consideration affecting the distribution of residential population in a city. Concurrent with the centrality preference discussed here, there may be a strong preference for noncongested sites. Each site may be evaluated by the individual in terms of not only centrality, but also congestion. If this hypothesis is accepted, then the unstable locational situation indicated in the above analysis may not be as serious as it would otherwise appear. It is certainly true that the planners-technicans-leaders of socialist countries are very much aware of possible dis-economies of congestion and try to prevent it as much as possible. [1] Even if individual households do not recognize the problem of congestion, the fact that the decision makers do would provide enough constraints on the location of residences.

In analyzing Muth's model with regard to a socialist economy, it is worthwhile to note that the elasticity of price with respect to quantity for housing can be assumed to equal zero, i.e., it is completely inelastic. There are a number of reasons to explain this phenomenon. Foremost is the critical housing shortage which was caused by widespread war destruction in the 1940's. Following the war, investment was channeled mainly to industrialization and not enough housing was produced compared to the existing and growing need. As long as control of the housing sector rests with the state, prices will remain relatively constant even if the quantity of housing eventually meets demand.

Socialist housing producers. When a socialist economy is assumed, an analysis of Muth's model for housing producers is not particularly interesting. Socialist housing producers operate neither as free market competitors nor as monopolists. Whether public or cooperative, housing producers are not interested in maximizing their profits, since according to the law, housing is obtained by the user for cost or less. In essence, the nature of their concern is different from that of a capitalist housing producer. Housing producers are not trying to achieve an equilibrium condition with regard to price, quantity, and location; and traditional economic analysis of producers' behavior is inappropriate.

Because they act on behalf of the entire public, socialist housing producers wish to maximize the amount of housing that can be constructed. This goal is restricted by the amount of investment allocated to housing by the government. In addition, prices are set by the government for all materials necessary for building. Therefore, to maximize housing production, costs per unit of housing produced must be minimized. Cost minimization is constrained by standards for minimum (and maximum) size and for type of facilities provided per unit, which are determined from national spatial planning acts.

The only variables housing producers can control are the intensity of residential land use and its location in the city. Variable costs in constructing housing are associated with the provision of public utilities. These costs vary directly with distance from the center and directly with the amount of housing built in any location. Thus, a trade-off exists in the construction of urban housing between the quantity built in any one district and the distance of the district from the center. Because there are scale economies connected with building large amounts of housing at one location in one time period, the tendency should be in favor of building high density housing as near to the center as possible (Fisher 1962, p. 253). Once a district has as much housing as planning standards allow, construction should move to the next district nearest the center. Development proceeds most inexpensively along existing utility lines; these are normally dense networks near the center but probably assume radial patterns as distance from the center increases.

---

[1] A bulletin issued by the Polish Town Planning Institute lists thirteen recently published studies dealing with various aspects of this topic: *Notes on Selected Institute Papers Dealing with Regional Planning Problems,* Institute for Town Planning and Architecture, Warsaw 1971.

A reasonable assumption then is that the maximum intensity of residential land use allowed by planning standards occurs in the central area of the city; and intensity decreases with utility capacity. If a dense utility network serving equally all city areas (the same as assumed for the central area) is built with each unit increase of distance from the center, then the costs of providing utilities would increase geometrically with distance, and intensity of land use would remain constant. This would result in a very dense, compact city. If a utility network is built to cover all space, but the capacity of the system is allowed to decrease with distance from the center according to a -1 power function, costs of provision would increase linearly with distance, and intensity of residential land use would decrease in a manner similar to the capacity of the utility system. The above possibilities are simpler than the third general case. If the utility network is extended along radials with the interstitial areas not serviced, intensity of residential land use would continue to vary directly with the capacity provided in the system but costs would depend on the amount of capacity provided at increasing distances from the center, on the width of the utility network along the radials, and on the number of radials.

The above discussion of possible relationships between utility capacity, residential land use intensity, and costs can be summarized as

$$D = P/L = f(B) \tag{8}$$

Residential population density is D; population is P; land area is L; and the capacity of the utility system is B. Capacity within the city might be a constant, B_0, or else a decreasing function of distance, $B(s)$. The costs associated with the provision of utilities can be represented by $E(B)$. Provision cost for utilities in the central area is assumed constant, E_0. Distance, s, is measured from the boundary of the central area.

$$\text{Case 1: If } B_0, E(B) = f(s^2)+E_0 \tag{9a}$$

$$\text{Case 2: If } B(s) = f(1/s) \text{ and } s > 0,$$

$$E(B) = f(s)+E_0 \tag{9b}$$

$$\text{Case 3: if on a radial, } B > 0\text{; if not on a radial, } B = 0,$$

$$(1)\ B_0,\ E(B) = f(s^2m)+E_0 \tag{9c}$$

$$(2)\ B(s) = f(m/s),$$

$$E(B) = f(sm)+E_0 \tag{9d}$$

where m is the percentage of the city area serviced by the radials (the city is assumed to be circular).

These considerations regarding the economics of building urban housing are well-known to socialist city planners. There have been numerous studies dealing with the costs of developing new urban areas, with threshold capacities of utility systems, and with how these thresholds may be expected to constrain urban development (Malisz 1966).[2] Urban master plans consider not only how much new housing to build and where to build it, but also the variable costs involved in developing any given site (*Techniczno-ekonomiczne studium*... 1966).

Urban Population densities in a socialist city. The last section of Muth's model deals with the joint behavior of housing consumers and producers. Given

[2] Valuable information on this topic was obtained in numerous conversations with Polish urban planners.

his previous analyses, Muth derives some theoretically expected patterns for urban residential location and intensity of land use. His equation:

$$D = P/L = (P/M)\ (rQ/L)\ (rQ/H) \tag{10}$$

where D is population density, P — population, and H — households, relating population density to people per household, expenditures on housing per household, and the value of housing per unit of land, can probably be regarded as true for a socialist urban economy. However, Muth's assumptions for this part of his formulation cannot all be considered applicable; without a market, a demand curve for consumers is irrelevant; and producers' production functions are likely to be of the fixed input-output coefficient type for materials and land, i.e., a zero elasticity of substitution between materials and land (Mills 1969, p. 252).

However, the means Muth uses to derive a negative exponential function for the relationship between population density and distance are inappropriate for a socialist city. From the discussions in this section, there are some indications of the types of relationships that may exist between population density and urban structure. Real income for households was found to be highest if residential location was at the city center; and it was shown to decrease linearly with distance. Also, a strong tendency favoring constant, high density residential construction anywhere in the city was shown to exist for housing producers due to scale economies for construction.

When the carrying capacity of public utility networks was considered, the intensity of residential land use became an indeterminant question. Provision and expansion of utility networks depends on previous network patterns and planning decisions regarding: (1) how present resources can be best utilized, and (2) how socialist goals for urban areas can be best realized. Since socialist principles require that each urban resident have equal access to urban amenities, and since the minimum cost solution for the socialist city favors compact development, it may be possible to extend Casetti's suggestion for discounting over space (Casetti 1967, p. 97) to a socialist urban environment. In particular,

$$u(s) = he^{-vs} \tag{11}$$

The centrality value of the central area is represented by h, which may be valued in terms of ease in providing all urban amenities; s is distance from the center; v is the marginal cost associated with extending development one additional unit of distance from the center; and $u(s)$ is the centrality value associated with sites at distance s from the center.

Thus, a theoretically expectable residential pattern of population density for socialist cities is not immediately obvious. Possible patterns include: (1) residential densities declining linearly with distance from the center, (2) constant residential densities at any distance, (3) residential densities declining exponentially with increasing distance from the city center, and (4) other density patterns depending on the capacity and location of public-utility networks or on other similar socialist planning considerations.

FACTORS INFLUENCING URBAN STRUCTURE

The essence of the earlier analysis in this study indicated that current and past socialist planning and decision making are based on two premises: (1) the need to create an urban-industrial society, and (2) the intent to achieve equal possibilities for all workers. Both of these premises have implications for the location of housing within an urban area and for locational processes in a socialist society.

Limited resources. Following the Second World War, most East European countries were faced with very limited resources for development. Thus, cost efficiency became an important consideration. Warsaw at the end of the war not only had the problem of limited resources, but also the need to replace large amounts of capital stocks such as industrial plants, housing, infrastructure, etc. which were destroyed during the war. Cost efficiency in rebuilding and development was an absolute necessity; even the careful rebuilding of Warsaw's Old Town served the functional purpose of quickly providing a large amount of housing and creating the basis for a tourist industry.

A number of premises concerning the effective use of limited investment in the construction of housing have become prominent in Poland. First consideration was given to achieving the greatest possible economies of scale. Thus, massive apartment blocks in large housing estates were built first by public housing authorities and then by cooperative associations in Warsaw, resulting in high intensity in residential land use. Concomitant with this was the standardization of housing units. Standards were set to not only assure minimal levels of facilities and space but also to obtain the greatest amount of housing given these minimums. Standardization also encouraged the development of unit construction and industrialized housing.

While extreme spatial concentration of housing did not last very long, the occurrence of large scale urban sprawl has never been considered feasible. Limited resources are perceived not only in terms of investment capital but also in terms of productive agricultural land. As an example, Warsaw, similar to many other large cities, has a viable market gardening ring, which is important to the economy of the city. Limited capital resources also constrain the character of urban transportation; public transportation has been favored, because of economies of scale in construction of facilities, in the number and maintenance of vehicles, in fuel usage, etc.

Housing fulfills a definite societal need in socialist countries. Regardless of other factors, because the allocation of housing is controlled by the state, housing is maintained and utilized until it no longer has functional utility. In addition, the policy of eliminating differentials implies that no residential area should become an undesirable location in a socialist city. This means that central residential densities should decline only if the amount of housing allocated to every individual increases, and not because central areas have become undesirable residential locations.

Thus, the major implication of limited resources for spatial patterns in a socialist city is compactness of development. Table 1 showing densities for six American cities and for Warsaw is indicative of this. The American cities have declining average population densities through time, and, in contrast, Warsaw exhibits rising density levels. Tremendous pressures for maximum utilization of residential land in Warsaw have been documented in this study. City areas, especially those close to the center, can be expected to be intensely developed. Intensive development is also encouraged along major public transportation lines and in the case of Warsaw along the tram lines. Compactness in Warsaw has the additional impetus of the central districts (within the pre-war boundaries) being municipally owned while outer districts are not. Public ownership implies cheaper initial development costs for public or cooperative builders.

Equal possibilities. The principle of equal possibilities has similarly important ramifications for socialist urban development. The concept of centrality becomes significant here, because access to urban goods is not spatially equal; the best access can be assumed to be in the center of the city.

Conversely, this same principle motivates the goal to eliminate large differentials in living standards, and in particular, the differentials between urban and

TABLE 1. A Comparison of population, land area, and density for six American cities and Warsaw, Poland[a]

City and Year	Central City			Urbanized Area			SMSA[b]		
	Pop. (1,000's)	Area (sq.mi.)	Density (1,000's sq.mi.)	Pop. (1,000's)	Area (sq.mi.)	Density (1,000's sq.mi.)	Pop. (1,000's)	Area (sq.mi.)	Density (1,000's sq.mi.)
Baltimore									
1950	949.7	78.7	12.7	1161.9	151.8	7.7	1337.4	1106.0	1.2
1960	939.0	78.3	12.0	1418.9	219.6	6.5	1803.7	1807.0	1.0
1970	905.8	78.3	11.6	1579.8	309.6	5.1	2070.7	2259.0	0.9
Boston									
1950	801.4	47.8	16.8	2233.5	344.8	6.5	2370.0	770.0	3.1
1960	697.2	46.0	15.2	2413.2	514.0	4.7	2595.5	969.0	2.7
1970	641.1	46.0	13.9	2652.6	664.4	4.0	2753.7	1043.0	2.8
Cleveland									
1950	914.8	75.0	12.2	1383.6	300.1	4.6	1465.5	688.0	2.1
1960	876.1	75.9	11.5	1783.4	581.4	3.1	1909.5	688.0	2.6
1970	750.9	75.9	9.9	1959.9	646.1	3.0	2064.2	1519.0	1.4
Pittsburgh									
1950	678.8	54.2	12.5	1533.0	253.6	6.0	2213.2	3053.0	0.7
1960	604.3	55.2	10.9	1805.3	526.1	3.4	2405.4	3051.0	0.8
1970	520.1	55.2	9.4	1846.0	596.4	3.1	2401.2	3049.0	0.8
St. Louis									
1950	856.8	61.0	14.0	1400.0	227.8	6.1	1681.3	2520.0	0.7
1960	750.0	61.2	12.3	1667.7	323.4	5.2	2104.7	3187.0	0.6
1970	622.2	61.2	10.2	1882.9	460.6	4.1	2363.0	4119.0	0.6
Washington D.C.									
1950	802.2	61.4	13.1	1287.3	178.4	7.2	1464.1	1488.0	1.0
1960	764.0	61.4	12.4	1808.4	340.7	5.3	2076.6	1485.0	1.3
1970	756.5	61.4	12.3	2481.5	494.5	5.0	2861.1	2353.0	1.2
Warsaw[c]									
1950	—	—	—	822.0	174.0	4.2	1298.0	1659.0	0.8
1960	904.8	49.9	18.1	1136.0	174.0	6.5	1763.3	1659.0	1.0
1970	1059.9	49.9	21.2	1308.9	174.0	7.7	1997.5	1659.0	1.2

[a] *Sources*: U.S., Bureau of the Census, *Eighteenth Decennial Census of the United States*: 1960. *Population*, *1970. Number of Inhabitants*, Final Report PC(1) — Al, United States Summary, 74–86, 171–178, and 187–188; and Warsaw, Warsaw Town Planning Office, data sheets.

[b] Data for the American cities in 1950 were given by the census for Standard Metropolitan Areas, which were defined as slightly different from Standard Metropolitan Statistical Areas.

[c] The pre-1951 boundary of Warsaw has been used to define "central city". The post-1957 boundary has been used to define "urbanized area". The Warsaw conurbation or Warsaw and the five poviats immediately surrounding the city have been used to define an equivalent to SMSA. Data were not obtained for Warsaw central city in 1950.

rural areas. One means of achieving this goal is through urban deglomeration of new industry and of controlled population growth. This sort of development is directed in Poland to centers beyond the major conurbation areas and is not of direct interest in this study. The socio-economic costs of over-development within individual city areas are another type of differential. Trying to eliminate these costs would lead to the spread of development into underutilized areas.

In general, the principle of equal possibilities implies concern for minimization of total social diseconomies within a city. The implications for urban development are in favor of relative compactness rather than sprawl. Consequently, major concerns are for: cheap mass transportation, housing located near major transport lines, and residential construction concentrated into large estates to provide as much centrally located housing as possible.

In calculating the social costs to be minimized within the context of compact city development, there also is recognition of the diseconomies associated with locating residential uses together with industrial ones. Thus, socialist planning principles favor developing separate industrial districts within cities. This is actually a question of scale. The data in Table 1 suggest that Warsaw is more spatially compact than comparable American cities. Thus, while American cities also have industrial parks, these are normally located in suburban areas, while industry within the American central city has often not been regulated in order to separate it spatially from housing.

Other factors. There are, obviously, other factors considered in decisions about the location of housing within a socialist city. As an example, the physical characteristics of a site are a major component in the cost calculations for defining feasible new development areas.[3] While attached to each site, these costs are random in the sense that their importance would be unpredictable from city to city. If a site is extremely desirable for development, the additional costs of an unfavorable physical environment may be regarded as necessary. Thus, five items have been discussed as having an influence on the amount and location of housing in Warsaw: strong pressures for compact city development, public transportation, distinct intra-city industrial park areas, housing estates, and public versus private ownership of land.

HYPOTHESES CONCERNING RESIDENTIAL LOCATION IN WARSAW

Using the data available, a number of hypotheses can be formulated and tested in a multiple regression analysis. The intent is to test whether residential population densities in Warsaw are systematically influenced by (or related to) these specific variables. The first hypothesis is that there is no relationship between the variables and residential population densities, i.e., the null hypothesis. The alternative hypothesis for each is that there is a relationship. The regression equation is of the form

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_n X_{in}$$

Distance. Compactness in city development is not a readily measurable quantity. However, given equation (11), which is a modified version of Casetti's suggestion for discounting the value of land over space (Casetti 1967, p. 97), the level of population density of a city area can be related to the distance that the area is from the city center.

[3] Examples of both theoretical and applied concern for these costs are: *Techniczno-ekonomiczne studium...* 1966, and W. Różycka 1972.

Equation (11) is similar to the standard negative exponential density function used in empirical analyses of American cities, except that it is motivated on the basis of centrality considerations for residential development and not on the basis of direct location costs. The theoretical analysis also suggested that the same concerns for centrality in a socialist city may produce a simple negative linear relationship between population densities and distance. In either model, the assumption for the relationship between population density and distance is that population densities should be highest in the center of Warsaw and decline away from the center. A negative exponential relationship implies a more centrally compact distribution of population than the alternative model of population densities declining linearly with increasing distance. As it is well known, the negative exponential model yields only positive expected values for population density. A linear model might produce negative expected values, but these could be justified on the basis that areas with negative values currently lack the potential for urban development. The spatial extent of the city's direct territorial dominance determines the relevant range of either model's estimated values.

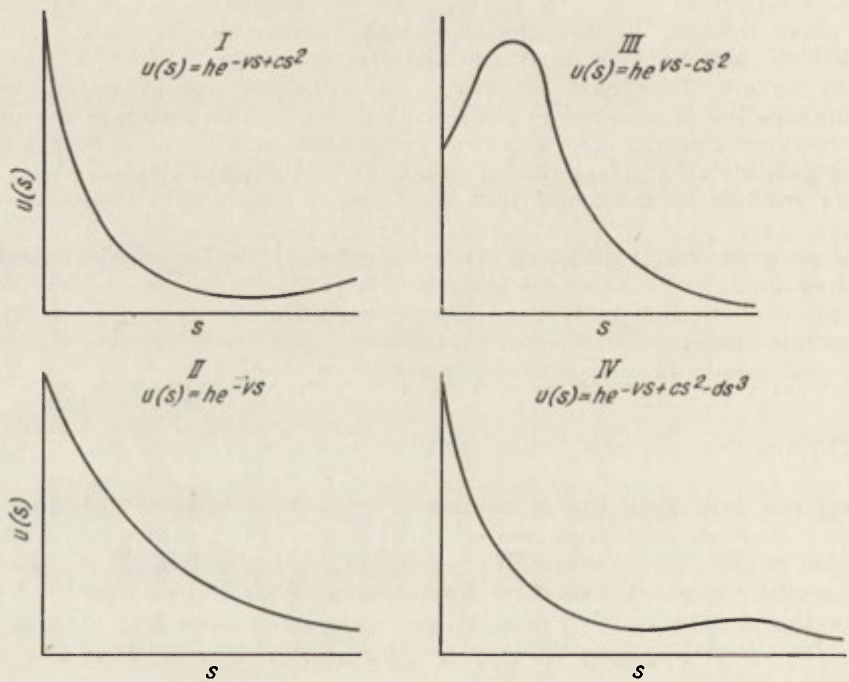


Fig. 1. Hypothetical density curves for second- and third-degree equations

A quadratic equation in distance was suggested by a number of authors. The usual formulation gives the first-degree term as a positive number ($+v$) and the second-degree term as a negative ($-$c); see Figure 1, function III. If, instead, the first-degree term remained negative as in function II and the second-degree term was found to be positive (See Figure 1, function I), this would imply more centrally located residential development than function III. Additionally, as compared to function II, function I shows a decline in the density grandient for areas not close to the city center. Thus, function I implies that since the absolute value of c is less than the absolute value of v, a sharp decrease and then a leveling occur in the amount of development within a short distance of the central area.

Analysis of these curves beyond the actual data points would not be fruitful. Extrapolation of functions II and III would imply that no other center of development can occur, while extrapolation of function I would quickly produce large unlikely estimates of population density. However, if the data base is extended to include Warsaw's metropolitan area, function I with the addition of a third-degree term would account for the development in central Warsaw and also the ring of urban development around Warsaw in the metropolitan area. (Figure 1, function IV). Thus, the upturning of function I in the boundary region of Warsaw is also anticipatory of the suburban ring development.

Land ownership. Public ownership of land within Warsaw's pre-war boundaries and private ownership for the land included by the post-war boundaries may also be assumed to influence the distribution of population. This hypothesis will be expressed in both the negative exponential and linear models as a dummy (0-1) variable labelled *Pi*. Districts within the old boundary are assigned a zero value and districts outside this boundary a value of one. A discontinuity effect resulting from the boundary would imply an overall decrease in population density beyond it. As illustrated in Figure 2, this would mean that the

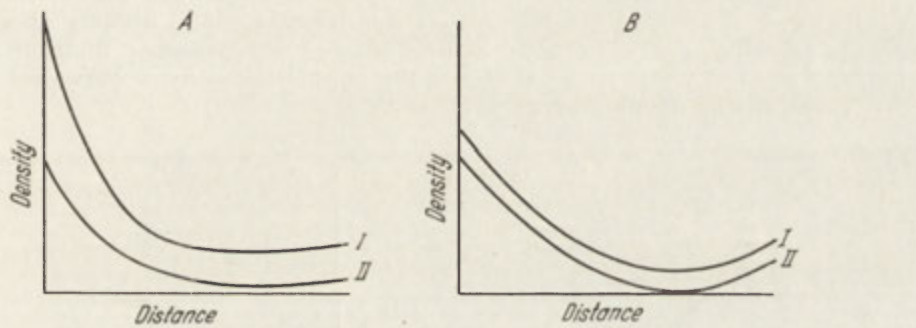


Fig. 2. The hypothetical effect of the presence (curve II) or absence (curve I) of the condition specified by either P_i, T_i, or I_i on the intercept of the negative exponential (A) and simple (B) density decline curves

correct estimator of population density for districts with private land ownership should be curve II in either the negative exponential or simple density decline model. Curve I represents the districts with public land ownership. In Warsaw's case, all districts would be on curve I until approximately six kilometers from the center; districts farther than six kilometers would be on curve II.

Other dummy variables. The last three location items to be considered in this analysis are the influence of public transportation, of separate industrial areas, and of housing estates. These will also be investigated through the use of dummy variables. For public transportation (the dummy is labelled T_i), the expectation is that districts not serviced by tram lines will have significantly lower residential population densities than the districts with lines crossing through them or running along a major portion of a boundary.[4] In a similar

[4] "The principal means of transportation are the tramways, transporting ca. 60 per cent of the total number of passengers." (S. Jankowski, J. Wilski, and B. Wyporek 1971, p. 16).

manner, districts with major industrial development can also be expected to have significantly less residential population than those without such development. (The industrial area dummy is labelled I_i.) Thus, negative coefficients are expected for both of these variables. Figure 2 can again be used to illustrate the hypothetical curves that represent the expected effect of the tram line dummy and industrial area dummy on the density-curves of the two models. is used here only as an indication of the effect that these intercept dummies should have in the models. It is not meant to be implied here that these binary variables, P_i, T_i, and I_i, will have an equal effect on the intercepts of the empirically derived curves.

The effect of housing estates on the location of residential population within the city may appear at first to be a result of, rather than a factor in, the location processes occurring in a socialist city. Clearly, strong pressures have been identified for locating residences, and therefore population, near the city center; the amount of housing should decline in a manner similar to the decline in population densities. It is hypothesized, however, that the intensity of development associated with the construction of a housing estate should lessen the impact of distance from the city center on the location of residential population, and a positive regression coefficient can be expected. That is, those districts with housing estates, which are located at the same distance from the center as those without estates, should have higher densities per unit area.

Figure 3 illustrates the effect that this housing slope dummy should have on the density curve in both models. Curve I represents districts without housing estates, while curve II shows the population density curve for districts with such estate development.

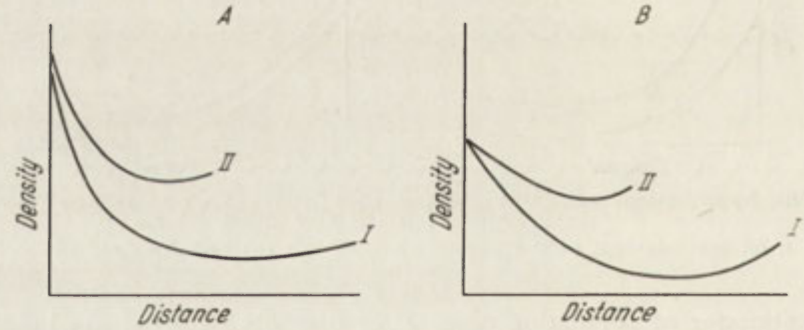


Fig. 3. The hypothetical effect of the presence (curve II) or absence (curve I) of the condition specified by H_i on the slope of the negative exponential (A) and simple (B) density decline curves

EMPIRICAL ANALYSES

The Warsaw city analysis. Five major hypotheses have been stated in the above discussion concerning the influence of particular independent factors on the location of population within Warsaw. In essence, it is hypothesized that: (1) population density should decline with distance from the city center, and the decline may exhibit an extremely convex curvature, which is specified by a quadratic distance term; (2) private ownership of land in the outlaying districts of Warsaw will result in an overall downward shift of the density curve for these areas; (3) lack of the most important mode of public transportation in a district, the tram, will result in an overall downward shift in the density curve;

Fig. 4. The Warsaw study area showing planning districts, 1960–1975
1 — city district boundary, 2 — city planning area boundary, 3 — city boundary

(4) the presence of a large amount of industrial development in a district will also result in an overall downward shift of the density curve; and (5) the presence of large housing estate developments in a district will shift the density gradient, i.e., the slope of the curve, upward. Due to the lack of clear *a priori* evidence in the theoretical analysis, two models have been suggested, in which to test these five hypotheses: (1) population density as the dependent variable,

$$Y_i = Y_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 P_i + \beta_4 T_i + \beta_5 I_i + \beta_6 H_i X_i + \varepsilon_i$$

and (2) the natural logarithm of population density as the dependent variable.

$$Y_i = Y_0 \varepsilon - \beta_1 X_i + \beta_2 X_i^2 + \beta_3 P_i + \beta_4 T_i + \beta_5 I_i + \beta_6 H_i X_i + \varepsilon_i \quad \text{or}$$

$$\ln Y_i = \ln Y_0 - \beta_1 X_i + \beta_2 X_i^2 + \beta_3 P_i + \beta_4 T_i + \beta_5 I_i + \beta_6 H_j X_i + \varepsilon_i \qquad (12)$$

The above hypothesis concerning the location of population within Warsaw for the years 1960, 1965, 1970, and 1975 have been formalized in the following regression equation. [5]

---

[5] All measurement units for data are metric in this study.

$$Y_i = \hat{\alpha} + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \hat{\beta}_3 P_i + \hat{\beta}_4 T_i + \hat{\beta}_5 I_i + \hat{\beta}_6 H_i X_i + \varepsilon_i \quad (13)$$

Where: α is the estimated central density.

$\hat{\beta}_1$ through $\hat{\beta}_6$ are the estimated regression coefficients.

X_i is distance measured in kilometers between the city center and centroid of each planning district in Warsaw.

P_i is a binary variable; districts outside the pre-war boundary have a value of 1, all other districts assume a value of zero.

T_i is a binary variable; districts without tram lines have a value of 1; the rest zero.

I_i is a binary variable; districts with primarily industrial development have a value of 1; the rest zero.

H_iX_i is a slope dummy variable; districts having housing estates assume a value of 1 for H_i and the dummy takes the value of X_i; all other districts assume a value of zero for H_i and the dummy must also take a zero value.

ε_i represents the least squares residuals.

Y_i is defined in two ways: (1) for a negative exponential density function, Y_i, is the natural logarithm of the population density of district i, $i = 1, 2, ..., 76$;[6] and (2) for a linear function, Y_i is simply the population density of each district.

Camparing the results of the two models. Table 2 gives the regression coefficients for the logarithmic model, and Table 3 shows the regression coefficients for the simple density model. In all cases, the F tests for the analyses were significantly different from zero at less than the 0.01 level, and the percentages of explanation indicated by the R^2's are satisfactory.

The results of the two models are to a large extent quite similar. As can be seen in Tables 2 and 3, changes in the values of the various regression coefficients occur in the four time periods in the same direction for both models. In the logarithmic model, the estimated regression coefficients generally indicate that the hypothesized relationships between the independent variables and population density were strongly supported. The exceptions are the boundary dummy in all time periods and the quadratic term in the 1975 equation, for which the null hypotheses were not disproven; and also the transportation dummy, which was significant between the 0.01 and 0.5 levels in 1960; see Table 2. In the simple density model, hypotheses for the distance and distance squared variables were strongly supported (significant at less than the 0.01 level) in all four years and for the housing slope dummy in three of the four years. Unlike the logarithmic model, the industrial area and transport dummies in the simple model were somewhat less significant; see Table 3.

The coefficients for the binary variable, which differentiates between public and private land ownership, do not show a significant relationship to the level of population density in any time period in either model.

As anticipated, the coefficient of the distance variable in both the logarithmic and simple density models indicates a significant inverse relationship with population density. That is, the density surface declines away from the city center. In addition, both models produced a significant positive coefficient for

[6] For those districts with less than 50 residents, a value of zero was recorded in the original data gathering. This was done because population in each district was rounded to the nearest hundred and all values were expressed in thousands. For the exponential analysis, districts with zero values were reassigned values of 8. This value was chosen because it was approximately one-tenth of the smallest recorded value.

TABLE 2. Regression coefficients of the logarithmic density decline model for Warsaw city[a]

	$\hat{\alpha}$ Intercept (1000's/ /sq.km)	$\hat{\beta}_1 X_i$ (Distance)	$\hat{\beta}_2 X_i^2$ (Dist.Sq.)	$\hat{\beta}_3 P_i$ (Boundary Dummy)	$\hat{\beta}_4 T_i$ (Transport Dummy)	$\hat{\beta}_5 I_i$ (Industrial Dummy)	$\hat{\beta}_6 H_i X_i$ (Housing Dummy)	
1960 Regression Coefficients	3.950 (52)	−0.742	0.035	−0.098[c]	−0.932[b]	−1.190	0.286	$R^2 = 0.60$
Standard Errors	0.631	0.206	0.013	0.482	0.429	0.438	0.115	
1965 Regression Coefficients	3.738 (42)	−0.589	0.029	−0.605[c]	−1.197	−1.413	0.181	$R^2 = 0.70$
Standard Errors	0.544	0.186	0.012	0.418	0.404	0.385	0.080	
1970 Regression Coefficients	3.525 (34)	−0.555	0.027	−0.550[c]	−1.163	−1.196	0.251	$R^2 = 0.72$
Standard Errors	0.511	0.175	0.011	0.393	0.379	0.361	0.075	
1975[d] Regression Coefficients	3.418 (30.5)	−0.448	0.017[c]	−0.466[c]	−1.059	−1.617	0.230	$R^2 = 0.70$
Standard Errors	0.539	0.185	0.012	0.416	0.393	0.380	0.080	

[a] Unless otherwise indicated coefficients are significant at the 1-tail, 0.01 level or less. All R^2's are significant at less than 0.01 levels; $n = 76$ for all regressions.

[b] Significance level between 0.01 and 0.05 levels, 1-tail test.

[c] Significance level greater than 0.05, 1-tail test.

[d] 1975 analysis based on estimated data supplied by the Warsaw Town Planning Office.

the second degree term, which reinforces the concept of compaction being an important factor in Warsaw's development.

Binary variables in the logarithmic and simple density models. For the logarithmic model, the coefficients of the industrial area dummy are significant in all four time periods; See Table 2. As hypothesized, the presence of an industrial area reduces the average level of population density likely to occur in a district. The coefficients of the transportation dummy also show declines in average population densities in districts without tram lines. The significance level of the 1960 transport dummy coefficient in the logarithmic model falls between the 0.01 and 0.05 levels. This is probably due to the dummy being defined as districts in 1970 with tram lines, because the 1960 pattern of tram lines in Warsaw was slightly different.

As shown in Table 3, the influence of the industrial area dummy and of the tram line dummy is not as clear in the simple model; these variables are significant at the 0.01 level or less only for the 1975 estimated data, and the transport dummy is not significant at the 0.05 level in 1965. Thus, the logarithmic model is better able to separate the effects of these variables from the overall trend of densities declining with distance.

The positive impact of the presence of housing estates within districts is also as expected. These housing dummy coefficients indicate an important shift upwards in the slope of the distance variable, i.e., the presence of housing estates lessens the importance of distance in the location of residential population. All of the housing slope coefficients are significant at the 0.01 level or less, with the exception of the 1965 coefficient in the simple density model; this has a significance level between 0.05 and 0.01.

An analysis of time trends in the data. In Tables 2 and 3, not only do the coefficients indicate the expected relationships between the independent variables and density, but these relationships are consistently similar for the four years in both models. As an example, the coefficients for the industrial area dummy over the four years fluctuate between −1.190 and −1.617 in the

TABLE 3. Regression coefficients of the simple density decline model for Warsaw City[a]

	α	$\hat\beta_1 X_i$	$\hat\beta_2 X_i^2$	$\hat\beta_3 P_i$	$\hat\beta_4 T_i$	$\hat\beta_5 I_i$	$\hat\beta_6 H_i X_i$	
1960								
Regression coefficients	21.942	−4.004	0.201	−0.011[c]	−2.529[b]	−2.854[b]	1.747	$R^2 = 0.66$
Standard errors	2.313	0.755	0.048	1.768	1.572	1.607	0.421	
1965								
Regression coefficients	23.383	−4.037	0.202	−0.892[c]	−2.878[c]	−3.724[b]	0.744[b]	$R^2 = 0.64$
Standard errors	2.463	0.842	0.053	1.894	1.827	1.741	0.363	
1970								
Regression coefficients	20.520	−3.291	0.179	−2.736[c]	−3.319[b]	−3.346[b]	1.505	$R^2 = 0.67$
Standard errors	2.453	0.838	0.053	1.886	1.819	1.734	0.361	
1975[d]								
Regression coefficients	19.275	−2.737	0.133	−1.777[c]	−3.268	−4.079	1.363	$R^2 = 0.73$
Standard errors	2.024	0.694	0.044	1.562	1.477	1.429	0.299	

[a] Unless otherwise indicated, coefficients are significant at 1-tail, 0.01 level or less. All R^2's are significant at less than 0.01 levels; $n = 76$ for all regressions.

[b] Significance level between 0.01 and 0.05 levels, 1-tail test.

[c] Significance level greater than 0.05, 1-tail test.

[d] 1975 analysis based on estimated data supplied by the Warsaw Town Planning Office.

logarithmic model with no obvious trend to the changes. However, it can be seen in both models that the coefficients for the intercept, distance, and distance squared do seem to show a slight trend occurring in the changes. The question is whether these apparent declines in value over time are significant.

Thus, another regression analysis for both the simple and logarithmic density models was computed using all the data for the four years, i.e., $n = 4 \times 76 = 304$. In addition to the significant variables listed in equation (13), this analysis specified dummies, which tested for changes in the intercept and in the coefficients for distance and distance squared over time. The new specification is

$$Y_i = \hat{a} + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \hat{\beta}_3 T_i + \hat{\beta}_4 I_i + \hat{\beta}_5 H_i X_i + \hat{\beta}_6 D_1 + \hat{\beta}_7 D_2 + \hat{\beta}_8 D_3 + \hat{\beta}_9 D_1 X_i + \hat{\beta}_{10} D_2 X_i + \hat{\beta}_{11} D_3 X_i + \hat{\beta}_{12} D_1 X_i^2 + \hat{\beta}_{13} D_2 X_i^2 + \hat{\beta}_{14} D_3 X_i^2 + \varepsilon_i \quad (14)$$

Where: $\hat{a}$ is the estimated central density for 1960.

$\hat{\beta}_1$ through $\hat{\beta}_{14}$ are the estimated regression coefficients.

Y_i, X_i, T_i, I_i, and H_i are defined exactly the same as in equation (13).

D_1 is an intercept time dummy equal to 1 if the year is 1965 and zero otherwise.

D_2 is an intercept time dummy equal to 1 if the year is 1970 and zero otherwise.

D_3 is an intercept time dummy equal to 1 if the year is 1975 and zero otherwise.

$D_1X_i, D_2X_i, D_3X_i, D_1X_i^2, D_2X_i$, and $D_3X_i^2$ are slope time dummies, which assume the values of X_i or X_i when their time dummies have a value of 1, or otherwise are equal to zero.

ε_i represents the least squares residuals.

As shown in Table 4, none of the time dummies, neither intercept nor slope, are significantly different from zero at an acceptable level for either model. In fact, the trends of these dummies appear to be random and occasionally contradictory to the trends observed in Tables 2 and 3. As an example, in the logarithmic model of Table 4 the estimated intercept for 1965 would be $\hat{a} + \hat{\beta}_6$ or 3.980+1.018, which indicates an increase in value. In Table 2, however, there is a decrease in the size of the intercept estimate from 3.95 in 1960 to 3.738 in 1965. Since the coefficients of D_2 and D_3 are also positive, the analysis implies that the intercept estimates for 1970 ($\hat{a}+\beta_7$) and 1975 ($\hat{a}+\hat{\beta}_8$) are also larger than the 1960 estimate; this is contradictory to the intercept estimates in Table 2. In a similar manner, the dummy slope estimates for both models have a random pattern in relationship to the apparent trends in Tables 2 and 3.

In contrast, the estimated coefficients for the intercept, X_i, X_i^2, T_i, I_i and H_i in Table 4 are strikingly similar to the estimates in Tables 2 and 3; the intercept, X_i, and X_i^2 have values quite close to their 1960 estimates, while T_i, I_i, and H_i have values about equal to the average of their values over the four separte year analyses. The major implication of this time trend analysis is that the apparent trends in the coefficients of the intercept, X_i, and X_i^2 in the logarithmic and simple density models are not truly important to the overall analysis of structure.

Estimated curves of the logarithmic and simple density models. Table 5 gives the various combinations of intercept and slope estimates determined by the binary variables in the two models. That is, for each year and in both models, there are six relevant combinations of coefficients from Tables 2 and 3

TABLE 4. Regression coefficients of the time trend models for Warsaw City, 1960–1975

	$\hat{\alpha}$	$\hat{\beta}_1 X_i$	$\hat{\beta}_2 X_i^2$	$\hat{\beta}_3 T_i$	$\hat{\beta}_4 I_i$	$\hat{\beta}_5 H_i X_i$	$\hat{\beta}_6 D_1$	$\hat{\beta}_7 D_2$	$\hat{\beta}_8 D_3$	$\hat{\beta}_9 D_1 X_i$	$\hat{\beta}_{10} D_2 X_i$	$\hat{\beta}_{11} D_3 X_i$	$\hat{\beta}_{12} D_1 X_i^2$	$\hat{\beta}_{13} D_2 X_i^2$	$\hat{\beta}_{14} D_3 X_i^2$
Logarithmic Density Model															
Regression Coefficients	3.980 (53.5)	−0.749	0.034	−0.854	−1.467	0.323	1.018	0.844	0.699	−0.144	−0.966	0.025	0.004	−0.001	−0.007
															$R^2 = 0.64$
Standard Errors	0.552	0.178	0.012	0.378	0.201	0.041	0.901	0.901	0.901	0.245	0.245	0.245	0.017	0.017	0.017
Significance Levels[a]	0.000	0.000	0.004	0.023	0.000	0.000	0.258	0.352	0.465	0.564	0.785	0.916	0.801	0.955	0.675
Simple Density Model															
Regression Coefficients	22.163	−3.959	0.195	−2.439	−3.801	1.556	4.003	2.664	0.940	−0.829	−0.243	0.350	0.041	0.002	−0.035
															$R^2 = 0.65$
Standard Errors	2.249	0.725	0.048	1.542	0.821	0.166	3.673	3.673	3.673	0.999	0.999	0.999	0.068	0.068	0.068
Significance Levels[a]	0.000	0.000	0.000	0.111	0.000	0.000	0.276	0.476	0.794	0.413	0.803	0.727	0.557	0.974	0.612

[a] Two-tail test.

TABLE 5. Regression, coefficients of the estimated curves for the logarithmic and simple models of density decline for Warsaw City

I. $E(Y_i|T_i = 0,\ I_i = 0,\ H_i = 0) = \alpha+\beta_1X_i+\beta_2X_i^2$

$$\ln Y_i \left\{\begin{array}{l} 1960 = 3.950-0.742X_i+0.035X_i^2 \\ 1965 = 3.738-0.589X_i+0.029X_i^2 \\ 1970 = 3.525-0.555X_i+0.027X_i^2 \\ 1975 = 3.418-0.448X_i+0.017X_i^2 \end{array}\right.$$

$$Y_i \left\{\begin{array}{l} 1960 = 21.942-4.004X_i+0.201X_i^2 \\ 1965 = 23.383-4.037X_i+0.202X_i^2 \\ 1970 = 20.520-3.291X_i+0.179X_i^2 \\ 1975 = 19.275-2.737X_i+0.133X_i^2 \end{array}\right.$$

II. $E(Y_i|T_i = 0,\ I_i = 0,\ H_i = 1) = \alpha+(\beta_1+\beta_6)X_i+\beta_2X_i^2$

$$\ln Y_i \left\{\begin{array}{l} 1960 = 3.950-0.456X_i+0.035X_i^2 \\ 1965 = 3.738-0.408X_i+0.029X_i^2 \\ 1970 = 3.525-0.304X_i+0.027X_i^2 \\ 1975 = 3.418-0.218X_i+0.017X_i^2 \end{array}\right.$$

$$Y_i \left\{\begin{array}{l} 1960 = 21.942-2.257X_i+0.201X_i^2 \\ 1965 = 23.383-3.283X_i+0.202X_i^2 \\ 1970 = 20.520-1.786X_i+0.179X_i^2 \\ 1975 = 19.275-1.374X_i+0.133X_i^2 \end{array}\right.$$

III. $E(Y_i|T_i = 1,\ I_i = 0,\ H_i = 0) = (\alpha+\beta_4)+\beta_1X_i+\beta_2X_i^2$

$$\ln Y_i \left\{\begin{array}{l} 1960 = 3.028-0.742X_i+0.035X_i^2 \\ 1965 = 2.541-0.589X_i+0.029X_i^2 \\ 1970 = 2.362-0.555X_i+0.027X_i^2 \\ 1975 = 2.359-0.448X_i+0.017X_i^2 \end{array}\right.$$

$$Y_i \left\{\begin{array}{l} 1960 = 19.413-4.004X_i+0.201X_i^2 \\ 1965 = 20.505-4.037X_i+0.202X_i^2 \\ 1970 = 17.201-3.291X_i+0.179X_i^2 \\ 1975 = 16.007-2.737X_i+0.133X_i^2 \end{array}\right.$$

IV. $E(Y_i|\ T_i = 1,\ I_i = 0,\ H_i = 1) = (\alpha+\beta_4)+(\beta_1+\beta_6)X_i+\beta_2X_i^2$

$$\ln Y_i \left\{\begin{array}{l} 1960 = 3.028-0.456X_i+0.035X_i^2 \\ 1965 = 2.541-0.408X_i+0.029X_i^2 \\ 1970 = 2.362-0.304X_i+0.027X_i^2 \\ 1975 = 2.359-0.218X_i+0.017X_i^2 \end{array}\right.$$

$$Y_i \left\{\begin{array}{l} 1960 = 19.413-2.257X_i+0.201X_i^2 \\ 1965 - 20.505-3.283X_i+0.202X_i^2 \\ 1970 = 17.201-1.786X_i+0.179X_i^2 \\ 1975 = 16.007-1.374X_i+0.133X_i^2 \end{array}\right.$$

V. $E(Y_i|T_i = 0,\ I_i = 1,\ H_i = 0) = (\alpha+\beta_5)+\beta_1X_i+\beta_2X_i^2$

$$\ln Y_i \left\{\begin{array}{l} 1960 = 2.760-0.742X_i+0.035Y_i^2 \\ 1965 = 2.325-0.589X_i+0.029X_i^2 \\ 1970 = 2.329-0.555X_i+0.027X_i^2 \\ 1975 = 1.801-0.448X_i+0.017X_i^2 \end{array}\right.$$

$$Y_i \begin{cases} 1960 = 19.088 - 4.004X_i + 0.201X_i^2 \\ 1965 = 19.659 - 4.037X_i + 0.202X_i^2 \\ 1970 = 17.174 - 3.291X_i + 0.179X_i^2 \\ 1975 = 15.196 - 2.737X_i + 0.133X_i^2 \end{cases}$$

VI. $E(Y_i|T_i = 1, I_i = 1, H_i = 0) = (\alpha+\beta_4+\beta_5)+\beta_1 X_i+\beta_2 X_i^2$

$$\ln Y_i \begin{cases} 1960 = 1.828 - 0.742X_i + 0.035X_i^2 \\ 1965 = 1.128 - 0.589X_i + 0.029X_i^2 \\ 1970 = 1.166 - 0.555X_i + 0.027X_i^2 \\ 1975 = 0.742 - 0.448X_i + 0.017X_i^2 \end{cases}$$

$$Y_i \begin{cases} 1960 = 16.559 - 4.004X_i + 0.201X_i^2 \\ 1965 = 16.781 - 4.037X_i + 0.202X_i^2 \\ 1970 = 13.855 - 3.291X_i + 0.179X_i^2 \\ 1975 = 11.928 - 2.737X_i + 0.133X_i^2 \end{cases}$$

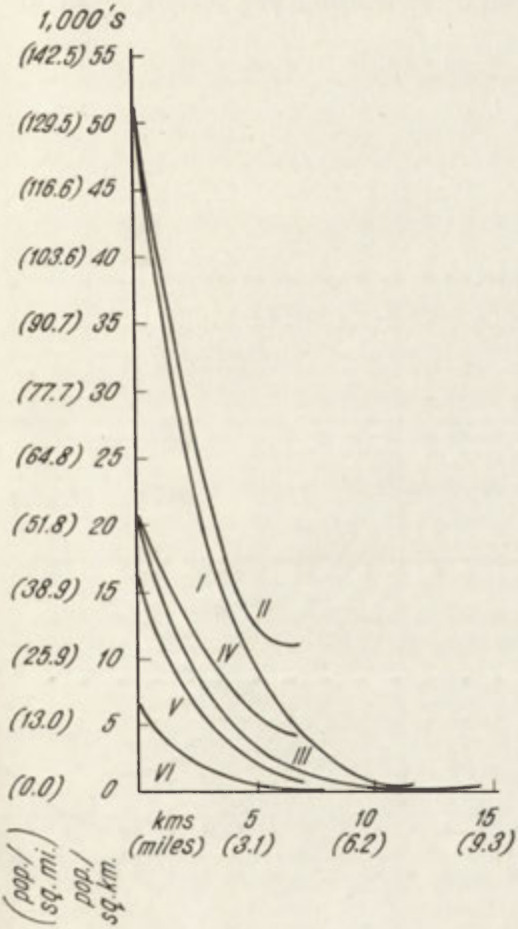
for $\hat{\alpha}$ and its dummies, $\hat{\beta}_4 T_i$ and $\hat{\beta}_5 I_i$, and for $\hat{\beta}_1 X_i$ and its dummy, $\hat{\beta}_6 H_i X_i$. These combinations produce six functions for determining expected values of Y_i, which are labelled I, II, III, IV, V, and VI in Table 5. Thus, function I represents the base hypothesis of density declining with distance for those districts having tram lines ($T_i = 0$), not having major industrial development ($I_i = 0$), and not having major housing estates ($H_i = 0$); other possible combinations for the hypotheses are given by functions II through VI.

Figures 5 through 12 graphically illustrate these estimated functions for the logarithmic and simple density models. In the models, the transport and industrial area dummies have been defined as intercept dummies, i.e., the influence of these conditions shifts the entire density function downwards. In contrast, the housing dummy has been defined as a slope dummy, i.e., this condition lessens the rate of population density decline. Thus, in every graph, six curves are presented, and each model has a graph for each of the years giving a total of eight graphs. Each curve represents different possible combinations of dummy variables as indicated in Table 5. In particular, curve I in each of the figures and in Table 5 represents the base condition of density declining with distance, which includes the presence of a tram line and the absence of both major industrial development and major housing estate development. Curve II shows the influence of housing estate development on the density gradient of curve I. Curve III represents the impact of the absence of a tram line; and curve IV shows the absence of a tram line but the presence of housing estate development. Curve V represents the impact of a major industrial area on density within city districts; and curve VI shows the combined impact of no tram line and the presence of a major industrial area on residential population density. By definition of the variables, the occurrence of both major industrial and housing estate developments within a district is not of interest to this analysis.

Curves I through VI represent hypothetical distributions of population density. If an imaginary line were drawn from the center of Warsaw to any city boundary, the actual population densities along that line would not be described by any particular curve. As an example, consider such a line with respect to Figure 5, which gives the curves for the logarithmic model in 1960. In the central district, curve I is the appropriate estimator, but within one kilometer, curve II may be the population density estimator. At two kilometers, it may become curve IV, and then at four kilometers, curve III. At five kilometers, the conditions specified by curve II may again occur, and at six, the conditions for curve V. After seven kilometers, only the conditions specified for curves I, III, and VI can occur; after nine kilometers, only curves I and III are applicable population density estimators; and beyond eleven kilometers, only the conditions for curve III exist.

The sequence of curves from the top of the graph downward is the same in all years in both models. The influence of industrial area development and of the lack of a tram line can be seen by these curves to be quite similar; in 1970, the difference between their coefficients is so small as to justify their representation by one common curve.

The simple versus the logarithmic density model. As shown above, the results of the simple density model are quite similar to the logarithmic one. In Figures 5 through 12, not only do the curves look very similar for the same model in each of the four years, but the resemblance is obvious between the two models in any one year. The similarity through the years within a model



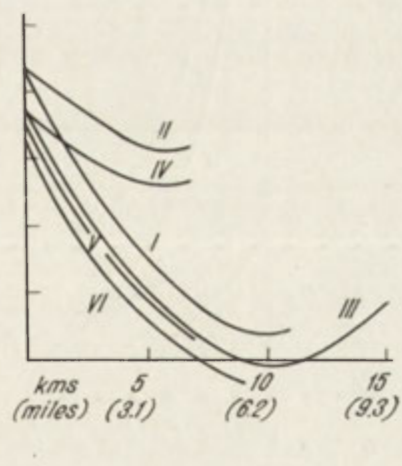


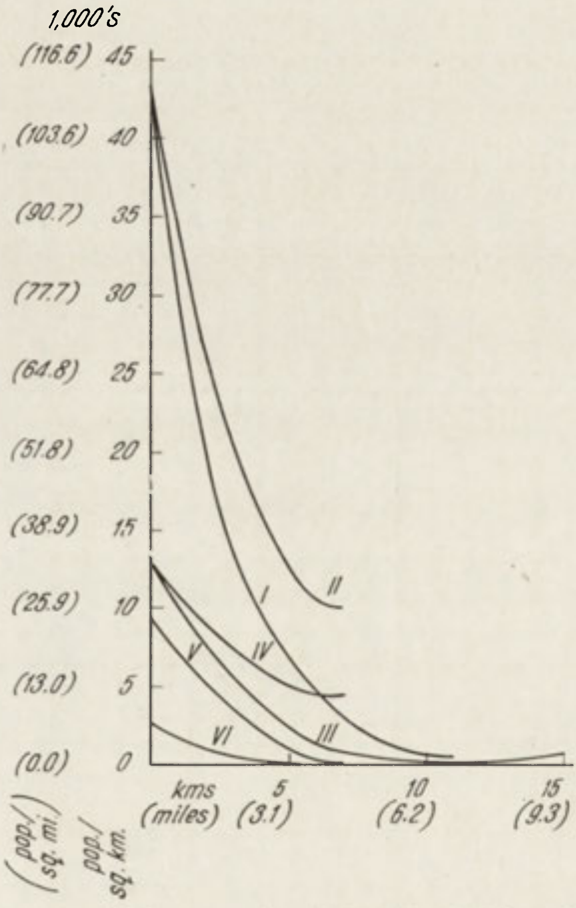
Fig. 5. Curves estimated in the 1960 logarithmic regression analysis (natural logarithms of population density expressed in linear scale)

Fig. 6. Curves estimated in the 1960 simple density regression analysis

was shown in the time trend analysis. The resemblance between the models is shown by the similarly high percentages of explanation in every year and by almost identical relationships (in terms of signs and significance levels) between the independent variables and population density.

This similarity may seem somewhat perplexing, because the implied relationships between the dependent variable, population density, and the independent variables in the two models are quite different. The results, however, do provide support to the idea that both models are reasonably good approximations to the actual distribution of population in Warsaw but may reflect different processes in socialist city development.

In general, the simple model better reflects overall development trends in Warsaw. The value of its intercept increases between 1960 and 1965, indicating continued development in the central area. The drop in 1970 to approximately the 1960 level probably indicates slackening in central area development, and the further drop for the 1975 estimate implies the movement of residential construction out of the central area. A second interpretation for the fluctuating behavior of this intercept is that the amount of housing per person began to



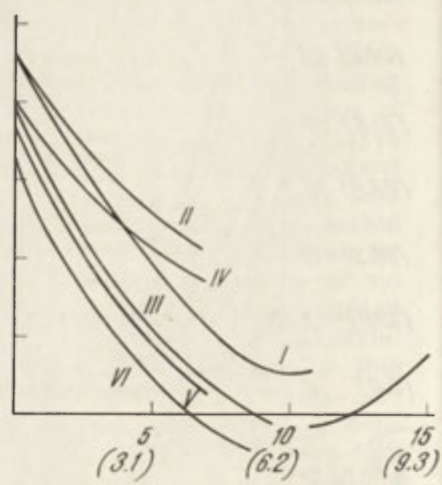


Fig. 7. Curves estimated in the 1965 logarithmic regression analysis (natural logarithms of population density expressed in linear scale)

Fig. 8. Curves estimated in the 1965 simple density regression analysis

increase sufficiently after 1965 to allow a decline in density, not only in the city center but throughout the city. This interpretation is supported by the marked flattening of the density gradient after 1965. While the simple density model does predict some negative population densities, these occur in the city boundary region for districts lacking tram lines or also having industrial areas. (See Figures 6, 7, 10, and 12).

The logarithmic density model is the better descriptive model for the general distribution of population within Warsaw. A major feature is the empirical nicety of never predicting negative population densities. In addition, this model is better able to distinguish the various policy variables from the overall trend in each year, while also increasing its percentage explanation. Finally, the logarithmic model is reflective of the strong forces for central city development and the limitation of urban sprawl. A major difference between the two models is the estimate of central density. The simple model produces a very good estimate of the true central density. The considerable overestimation of the central density by the logarithmic model is indicative of the potential level of centrally located residential development. This potential is unrealized due to enforcement of planning standards. See Figures 5 through 12.

The metropolitan results. As an extension of the above, the logarithmic and simple density analyses were also applied to 1960 data for Warsaw's metropolitan area. Unfortunately, 1965, 1970, and 1975 data were not obtainable for the metropolitan area. The models were slightly modified to include a cubic

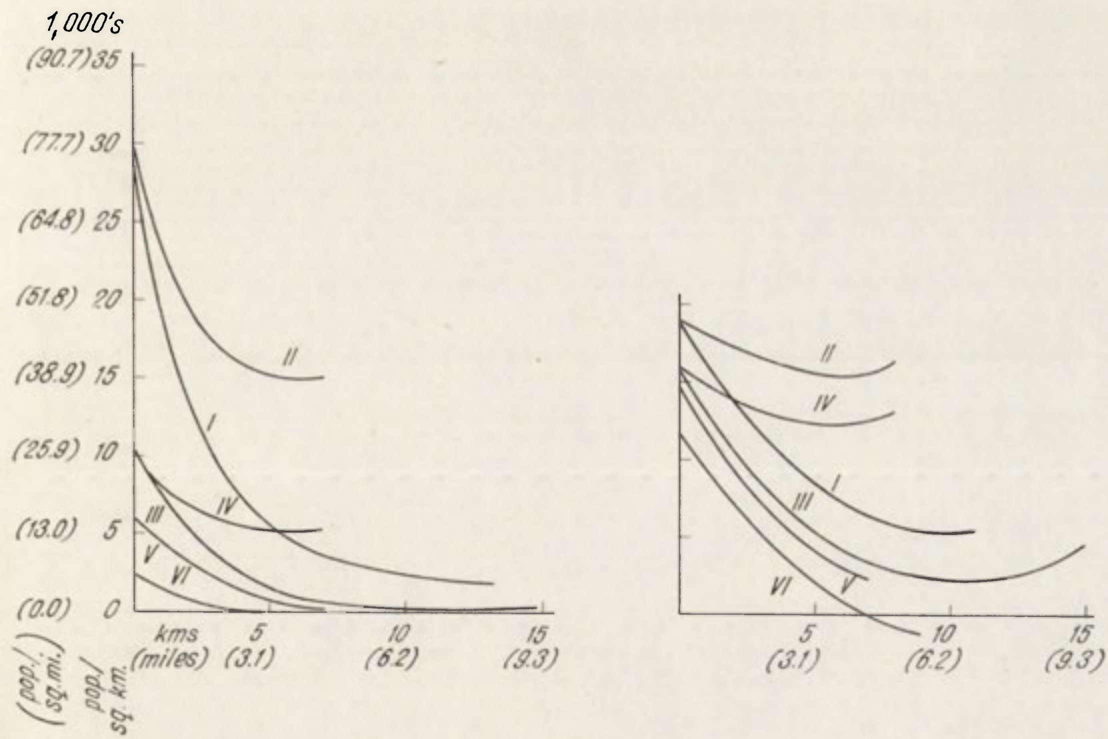
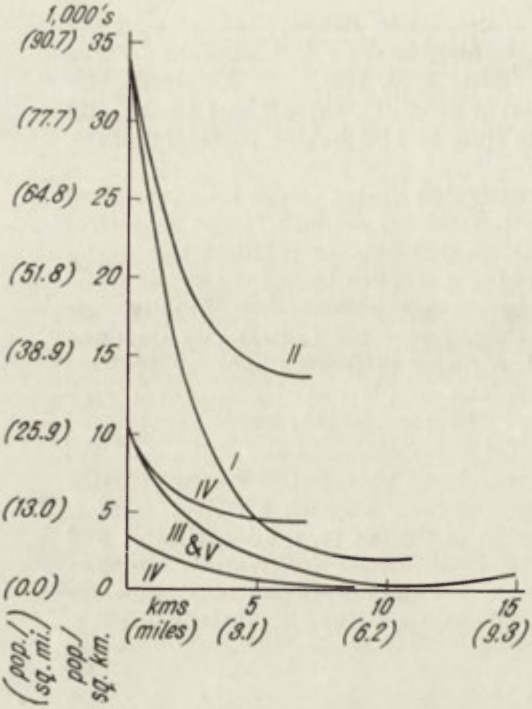


Fig. 9. Curves estimated in the 1970 logarithmic regression analysis (natural logarithms of population density expressed in linear scale)

Fig. 10. Curves estimated in the 1970 simple density regression analysis



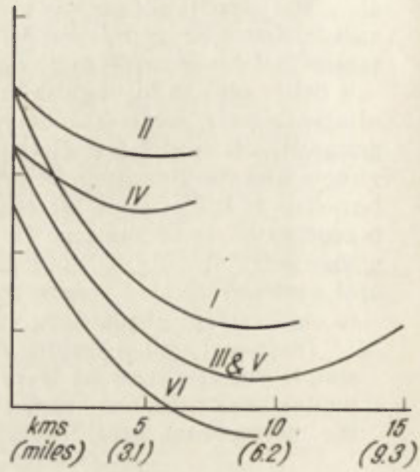


Fig. 11. Curves estimated in the 1975 logarithmic regression analysis (natural logarithms of population density expressed in linear scale)

Fig. 12. Curves estimated in the 1975 simple density regression analysis

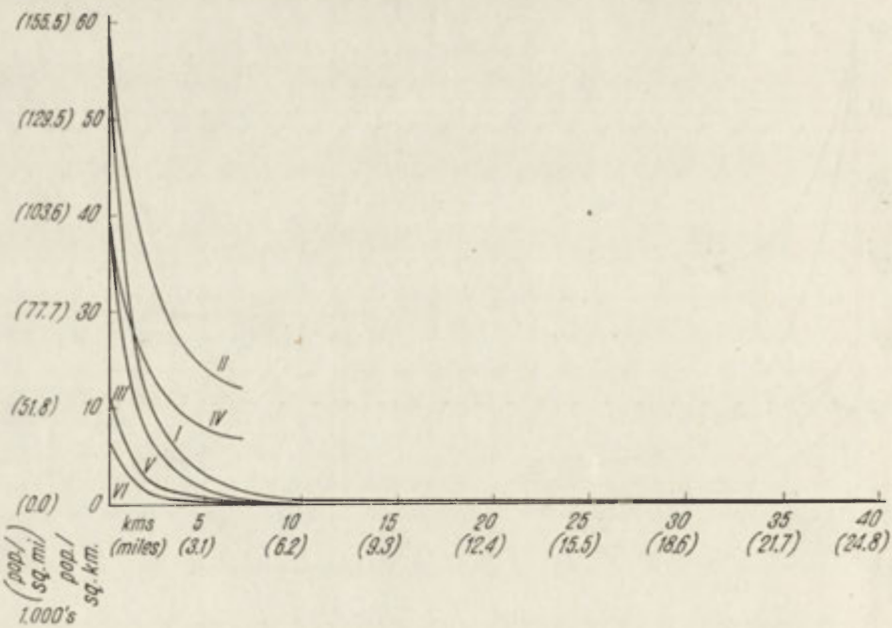


Fig. 13. Curves derived by the 1960 logarithmic density analysis of the Warsaw metropolitan region (natural logarithms of population density expressed in linear scale)

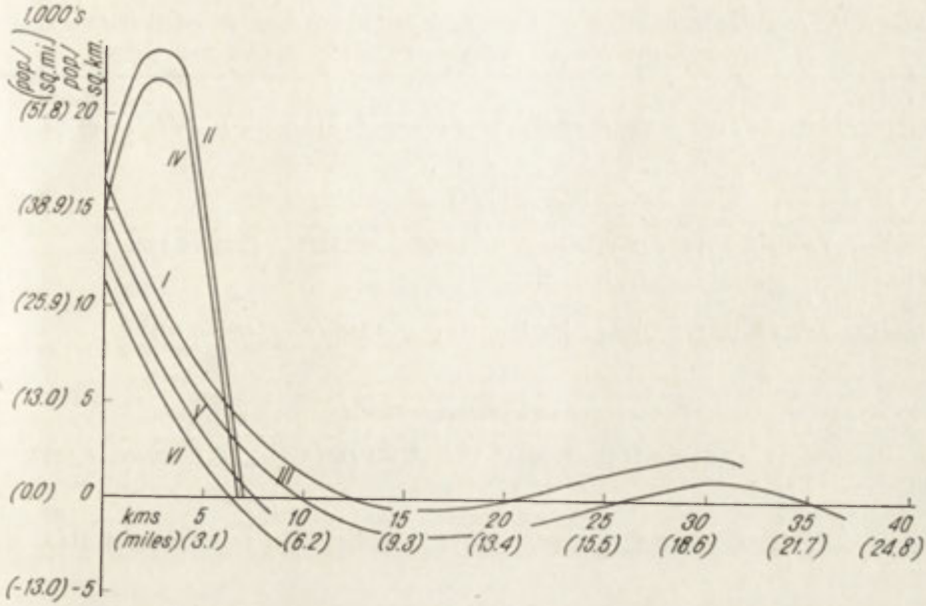


Fig. 14. Curves derived by the 1960 simple density analysis of the Warsaw metropolitan region

term for distance, X_i^3; a second housing slope dummy related to the quadratic term, $H_i X_i$; and a dummy variable to differentiate between Warsaw city areas and metropolitan areas, S_i. The results of the analyses are presented in Table 6 and in Figures 13 and 14.

In Table 6, the coefficient of the cubic term is significant at less than the 0.01 level. This term was included, because the quadratic analysis for Warsaw had only been valid to approximately the city boundary. In the discussion of the quadratic term for the Warsaw city model it was shown that in order to estimate accurately residential population densities for the metropolitan area, another term for the functions was necessary. The metropolitan dummy was included to determine if there were any sharp differences in land use development between Warsaw, in which development is mainly public, and its metropolitan area, which has mostly private residential development. Since this coefficient is not significantly different from zero, it appears that there is little difference in public and private development. This lack of difference can be attributed to the governmental controls on private construction, which is mainly limited to single-family, owner-occupied housing (Andrzejewski 1966, pp. 165–167).

An initially striking difference between the 1960 Warsaw city and metropolitan simple analyses is the complete reversal of the housing dummy curves (II and IV) in Figures 6 and 14. This can be attributed to two causes. First, in addition to the distance term, a quadratic term for the housing dummy was added to the metropolitan analysis. Second, there is the strong influence of all land in the central part of Warsaw being owned communally and not elsewhere in the metropolitan areas.

The quadratic housing dummy is not shown for the logarithmic metropolitan model due to lack of statistical significance. Given this one difference in the two models of the metropolitan region, the results of the logarithmic model are quite similar to the simple density model. The minimum points on the

TABLE 6. Regression coefficients of the logarithmic and simple density models and the associated curves for the Warsaw metropolitan density decline analysis[c]

	α	$\hat{\beta}_1 X_i$	$\hat{\beta}_2 X_i^2$	$\hat{\beta}_3 X_i^3$	$\hat{\beta}_4 S_i$	$\hat{\beta}_5 T_i$	$\beta_6 I_i$	$\beta_7 H_i X_i$	$\beta_8 H_i X_i^2$	
Logaritmhic Density Model										
Regression Coefficients	3.9005 (49.5)	−0.7194	0.0326	-0.005^a	0.5300^b	-0.5387^a	−1.5289	0.3106	—	$R^2 = 0.59$
Standard Errors	0.5081	0.1195	0.0084	0.001	0.5101	0.250	0.4088	0.1057	—	
Simple Density Model										
Regression Coefficients	16.3686	−2.5100	0.1162	-0.0016^a	1.3058^b	-1.4676^a	−3.3641	7.9940	−1.1657	$R^2 = 0.74$
Standard Errors	1.5085	0.3631	0.0247	0.004	1.4491	0.7085	1.1658	1.1880	0.2147	

Estimated Curves

I. $E(Y_i|T_i = 0,\ I_i = 0,\ H_i = 0) = \alpha+\beta_1 X_i+ +\beta_2 X_i^2+\beta_3 X_i^3$

$\ln Y_i = 3.9005-0.7194X_i+0.0326X_i^2- -0.0005X_i^3$

$Y_i = 16.3686-2.5100X_i+0.1162X_i^2- -0.0016X_i^3$

II. $E(Y_i|T_i = 0,\ I_i = 0,\ H_i = 1) = \alpha+(\beta_1+ +\beta_7)X_i+(\beta_2+\beta_8)X_i^2+\beta_3 X_i^3$

$\ln Y_i = 3.9005-0.4088X_i+0.0326X_i^2- -0.0005X_i^3$

$Y_i = 16.3686+5.4840X_i-1.0495X_i^2- -0.0016X_i^3$

III. $E(Y_i|T_i = 1,\ I_i = 0,\ H_i = 0) = (\alpha+\beta_5)+ +\beta_1 X_i+\beta_2 X_i^2+\beta_3 X_i^3$

$\ln Y_i = 3.3618-0.7194X_i+0.0326X_i^2- -0.0005X_i^3$

$Y_i = 14.9010-2.5100X_i+0.1162X_i^2- -0.0016X_i^3$

IV. $E(Y_i|T_i = 1,\ I_i = 0,\ H_i = 1) = (\alpha+\beta_5)+ +(\beta_1+\beta_7)X_i+(\beta_2+\beta_8)X_i^2+\beta_3 X_i^3$

$\ln Y_i = 3.3618-0.4088X_i+0.0326X_i^2- -0.005X_i^3$

$Y_i = 14.9010+5.4840X_i-1.0495X_i^2- -0.0016X_i^3$

V. $E(Y_i|T_i = 0,\ I_i = 1,\ H_i = 0) = = (\alpha+\beta_6)+\beta_1 X_i+\beta_2 X_i^2+\beta_3 X_i^3$

$\ln Y_i = 2.3716-0.7194X_i+0.0326X_i^2- -0.005X_i^3$

$Y_i = 13.0045-2.5100X_i+ +0.1162X_i^2-0.0016X_i^3$

VI. $E(Y_i|T_i = 1,\ I_i = 1,\ H_i = 0) = = (\alpha+\beta_5+\beta_6)+\beta_1 X_i+\beta_2 X_i^2+\beta_3 X_i^3$

$\ln Y_i = 1.8329-0.7194X_i+ +0.0326X_i^2-0.005X_i^3$

$Y_i = 11.5369-2.5100X_i+ +0.1162X_i^2-0.0016X_i^3$

[a] Significance level between 0.01 and 0.05 levels, 1-tail test.
[b] Significance level greater than 0.05, 1-tail test.
[c] Unless otherwise indicated coefficients are significant at the 1-tail 0.01 level or less. R^2's are significant at less than 0.01 levels; $n = 123$ in all regressions.

logarithmic curves II and IV in Figure 13 occur at approximately the same points as where the simple density curves II and IV in Figure 14 intersect the distance axis.

Otherwise the metropolitan analyses strongly resemble the city analyses for 1960 and, therefore, will not be extensively analyzed. The extent of Warsaw's metropolitan dominance is indicated in the simple analysis by the downward turning of all the curves after the thirty kilometer point in Figure 14.

COMPARING THE WARSAW EMPIRICAL RESULTS TO MUTH'S EMPIRICAL RESULTS

The tantalizing question is how similar or different are the empirical results for Warsaw compared to results derived for American cities of similar population size. The theoretical analyses of this paper clearly indicated that different processes operate in the two economic systems. However, one particular expected pattern of population distribution for socialist cities was not derived; instead, a number of patterns seemed to be possible, depending on which factors are stressed within a socialist urban planning context. The empirical results presented in Tables 2 and 3 are not directly comparable to previous empirical studies, because of differences in the measurement units for distance and area applied.

As a result, a second set of regression analyses for Warsaw were computed using miles and square miles as the measurement units.[7] The regression coefficients from these new analyses are presented in Table 7 for the logarithmic and simple models. The intercept is influenced by areal measurements, and the coefficients associated with distance, i.e., distance, distance squared, and the complex housing-distance dummy, are dependent on linear measurements. Comparing the logarithmic model results in Tables 2 and 7, the intercept and distance-related coefficients have different values, which is the expected result of changing the measurement units. The values of the transport and industrial area dummies have not exhibited similar large changes, because these are average shifts in the intercept; such shifts should be reasonably constant in a logarithmic model regardless of any changes in measurement units. All of the coefficients of the simple density model (see Tables 3 and 7) exhibit considerable change as a result of the change in measurement units, since this model inherently emphasizes absolute changes in value.

Population density analyses for Warsaw and American cities. Among the existing empirical studies, Muth's analysis provides the best results for comparison to the empirical results presented in this paper. Of the forty-six American cities, which he used in his analysis of population density, seven are in the population size range of Warsaw. In his more detailed analysis of Chicago, the hypotheses Muth tested are quite similar to those proposed in this study. Besides, a comparison of American cities after the Second World War to Warsaw in approximately the same time period should provide an adequate basis for comprehending any differences that exist in Warsaw's urban structure.

Table 8 gives the estimated central densities and density gradients from Muth's study for thirteen of the forty-six cities that he analyzed. The first

---

[7] The original analyses were performed using metric units, because of the international and scientific preference for the metric system. It is hoped that these metric analyses will be of greater value to future research than many of the previous analyses have been in terms of comparability.

TABLE 7. Regression coefficients for Warsaw city following a change in measurement units from metric to U.S. (or English) system[d]

Year	$\hat{\alpha}$ Intercept (1000's/sq.mi.)	$\hat{\beta}_1 X_i$ (Distance)	$\hat{\beta}_2 X_i^2$ (Distance squared)	$\hat{\beta}_3 T_i$ (Transport Dummy)	$\hat{\beta}_4 I_i$ (Industrial Area Dummy)	$-\hat{\beta}_5 H_i X_i$ (Complex Housing Dummy)
		Logarithmic Density — Regression Coefficients				
1960	($\ln\hat{\alpha}$) 4.988	−1.270	0.099	−0.987[a]	−1.275	0.478
	($\hat{\alpha}$) 147					
1965	($\ln\hat{\alpha}$) 4.927	−1.106	0.083	−1.357	−1.519	0.301[a]
	($\hat{\alpha}$) 138					
1970	($\ln\hat{\alpha}$) 4.691	−1.041	0.078	−1.282	−1.236	0.418
	($\hat{\alpha}$) 109					
1975[c]	($\ln\hat{\alpha}$) 4.552	−0.844	0.052[b]	−1.178	−1.656	0.379
	($\hat{\alpha}$) 95					
		Simple Density — Regression Coefficients				
1960	56.840	−16.734	1.356	−6.552[a]	−7.392[a]	7.300
1965	61.398	−17.408	1.379	−7.654[a]	−9.681[a]	3.279[a]
1970	55.712	−15.415	1.256	−9.212[a]	−8.778[a]	6.425
1975[c]	51.558	−12.471	0.926	−8.991	−10.630	5.763

[a] Significance level between 0.01 and 0.05 levels, 1-tail test.
[b] Significance level greater than 0.05, 1-tail test.
[c] 1975 analysis based on estimated data supplied by the Warsaw Town Planning Office.
[d] Unless otherwise indicated coefficients are significant at 1-tail, 0.01 level or less. See Tables 2 and 3 for a more complete presentation of results.

seven cities were chosen on the basis of similarity in population size to Warsaw; five of the last six have density gradients somewhat similar to Warsaw's and Cincinnati's estimated central density is close to those estimated for Warsaw.

Some caution is required in comparing the results in Tables 7 and 8. The simple density results for Warsaw are not comparable to Muth's results and were only included to indicate the magnitude of influence that the data measurement units have on the size of the estimated values of Y_i. As noted above, the relevant comparison for the simple model results is between Tables 3 and 7. Also, while Muth computed regressions for both a linear and a quadratic model, only twelve of the forty-six cities produced significant second-degree terms; Muth indicated the signs of the significant quadratic terms in his results but did not include the relevant coefficients. For the thirteen cities shown here, only Cleveland has a significant second-degree coefficient. Thus, a major difference between the American cities Muth analyzed and Warsaw is the positive coefficients for the quadratic term in the Warsaw analysis. As a result of this difference, the coefficients of the density gradients are not directly comparable. That is, the β_1 parameter for Warsaw measures the instantaneous rate of change of the density gradient at the center of the city, and the β_2 parameter measures the change of density away from the city center. The β_1 parameter for the American cities is the rate of density decline throughout the city. On the basis of these density gradients, it can be seen in Tables 7 and 8 that the relative rate of decline near the central area is not as steep for the seven large American

TABLE 8. Regression results from Muth's analysis of population density decline for American cities in 1950[a]

City	α (1000's/sq. mi.)	$\hat{\beta}_1 X_i$	$\hat{\beta}_2 X_i^2$
	Cities with Population Size Similar to Warsaw's		
Baltimore	69	−0.52	...
Boston	78	−0.30	...
Chicago	60	−0.18	...
Cleveland	22	−0.13[c]	—
Pittsburg	17	−0.091[b]	...
St. Louis	47	−0.28	...
Washington, D.C.	20	−0.27	...
	Cities with Coefficients Similar to Warsaw's		
Akron	38	−0.84	...
Cincinnati	120	−0.69	...
New Haven	46	−0.99	...
Richmond	41	−0.82	...
Syracuse	48	−0.92	...
Utica	51	−1.2	...

[a] *Source*: Richard F. Muth, *Cities and Housing*: *The Spatial Pattern of Urban Residential Land Use* (Chicago: University of Chicago Press, 1969), p. 142.

[b] The sign of the second-degree term is shown where significant at the 0.05 level.

[c] Not significantly greater than zero at the 0.10 level.

cities, even when the second-degree term is taken into account for Warsaw's gradient. Also, it is rather striking that all seven large American cities in Table 8 have smaller density gradients than the six other cities.

On the basis of the estimated central densities in Tables 7 and 8, it is clear that the American cities have far less central city development of housing than Warsaw. Trends previously identified in the literature for American cities are decreasing central densities in the twentieth century and steadily dropping values for density gradients (Berry, Simmons, Tennant 1963). Since Muth's analysis used 1950 data, these general trends imply that even greater differences might have been found between urban structure in Warsaw and in American cities for the 1960's and 1970's. The much lower level of urban residential development indicated by Muth's results explains the much larger land areas of the American cities in Table 1.

Further empirical comparisons: Warsaw versus Chicago. Muth's analysis for the south side of Chicago included a number of binary variables, which were defined quite similarly to the binary variables used in the Warsaw analysis. Similarly to the aim of this study, Muth wanted to consider questions about how population density is related to local variations in accessibility to industrial centers and to differences in marginal transportation costs. In Table 9, regression coefficients are given for six of the fifteen variables that Muth included in his Chicago regression analysis; the other nine variables are excluded here, because they are inappropriate with regard to Warsaw and, thus, have no counterparts in the Warsaw analysis, e.g., median income of family units, areas with a black population majority, etc. The Warsaw analysis did include one variable, the dummy for major housing estate development, that Muth

TABLE 9. Selected regression coefficients from Muth's analysis of population density decline on the south side of Chicago[a]

	$\hat{\beta}_1 X$ (Distance)	$\hat{\beta}_2 X_i^2$ (Distance Squared)	$\hat{\beta}_3 T_i$ (Transport Dummy)	$\hat{\beta}_4 T_i X_i$ (Complex trans.-Dist. Dummy)	$\hat{\beta}_5 T_i X_i^2$ (Complex trans.-D^2 Dummy)	$\hat{\beta}_6 I_i$ (Industrial Area Dummy)
			Gross Residential Densities			
1950						
Regression	A-0.10[b]	—	0.59[b]	−0.02[b]	—	−0.19[b]
Coefficients	B-0.46[b]	0.020[b]	0.10[b]	0.06[b]	−0.002[b]	−0.19[b]
Standard	A 0.064	—	0.58	0.07	—	0.20
Errors	B 0.34	0.018	1.4	0.35	0.020	0.20
1960						
Regression	A-0.070[b]	—	0.66[b]	−0.041[b]	—	−0.38[b]
Coefficients	B 0.015[b]	−0.005[b]	1.2[b]	−0.20[b]	0.009[b]	−0.38[b]
Standard	A 0.049	—	0.46	0.052	—	0.24
Errors	B 0.26	0.014	1.1	0.27	0.015	0.24
			Net Residential Densities			
1950						
Regression	A-0.11	—	−0.32[b]	0.073	—	−0.23[b]
Coefficients	B-0.59	0.027	−2.2	0.56	0.027	−0.23[b]
Standard	A 0.049	—	0.44	0.050	—	0.15
Errors	B 0.25	0.014	1.0	0.26	0.015	0.15
1960						
Regression	A-0.14	—	−0.27[b]	0.067[b]	—	−0.28[b]
Coefficients	B-0.28[b]	0.008[b]	−1.3	0.37	−0.018[b]	−0.28[b]
Standard	A 0.039	—	0.37	0.941	—	0.19
Errors	B 0.20	0.011	0.84	0.21	0.012	0.19

[a] Source: Richard F. Muth, *Cities and Housing: The Spatial Pattern of Urban Residential Land Use* (Chicago: University of Chicago Press, 1969), pp. 217, 223, 224, and 228.
[b] Not significantly greater than zero at the 0.05 level, 1-tail test.

chose to exclude from analysis by deleting those census tracts with public housing from his sample. Note that Table 9 does not give estimated intercepts (central densities); this is because Muth did not supply any intercept values for his analysis of Chicago. It must be assumed here that the intercept estimated for Chicago in the more general analysis of American cities is the relevant value.

Upon inspection of Tables 7 and 9, two things become clear. First, Muth's results simply do not provide the same strong evidence for the hypothesized relationships that the Warsaw results provide. Muth analyzed both net and gross residential densities to determine the difference in density variation when variation in the proportion of total land area used for residential purposes is removed. None of the coefficients are significant at the 0.05 level in the 1950 and 1960 analyses of gross densities. Only half are significant at the 0.05 level for the net densities analyses, and most of these are in the 1950 analysis; see Table 9.

Quadratic equations were estimated by Muth for both the gross and net density analyses. In 1950, the quadratic equation (B) for net densities has significant coefficients for both the first and second-degree terms. Thus, this equation can be directly compared to the Warsaw results in Table 7. As was generally true of the results for the American cities presented in Table 8, the gradient coefficients for the south side of Chicago are considerably smaller than the ones for Warsaw.

Muth defined his transport dummy in a different manner from the definition used in the Warsaw analysis; census tracts within one mile of rapid transit facilities received a value of 1, and all other tracts received a value of zero. Because this is exactly the opposite definition from that used for the Warsaw transport dummy, the significant negative coefficient in Table 9 for this intercept dummy is surprising; the negative coefficient implies an average drop in net population densities near transport facilities in 1950, which is exactly the opposite result obtained for Warsaw. However, in the same equation, the transport slope dummy, which Muth uses but does not occur within the Warsaw analyses, indicates a less rapid decline in density along rapid transit routes. In contrast to the Warsaw logarithmic model results, the results Muth obtained for his two transport dummy variables are quite mixed. In the analysis of gross densities, none of the transport coefficients are significantly different from zero at the 0.05 level. The signs of the transport coefficients are not as expected for either the slope dummies in the gross density analyses or the intercept dummies in the net density analyses; see Table 9.

The coefficients for the other dummy variable in Table 9, industrial development, do not show significant relationships to density in any of the Chicago analyses. The negative signs for these coefficients, however, are according to expectation, indicating declines in average densities for districts with major industrial developments. This agrees with the significant negative relationships derived for Warsaw.

The second important conclusion to be drawn from Tables 7 and 9 is that the intensity of urban land development is far higher in Warsaw than in Chicago. This conclusion is based not only on the steeper density gradients for Warsaw but also on the significance levels of these coefficients estimated in Muth's Chicago analysis. Muth included the analyses of net residential densities to determine the influence of variations in total land used for residential purposes. Yet, it was only in the net analyses that any of the coefficients achieved significance. The Warsaw analysis, with almost all of the coefficients being significant, is based entirely on gross densities. Thus, the amount of land devoted to residential development in Warsaw is much higher than in Chicago.

As previously noted, the dummy variable for tracts with industrial development does not obtain a value significantly different from zero in any of the Chicago analyses, while it is strongly significant in the comparable Warsaw analysis. This implies that industrial development is not strongly segregated from residential areas in Chicago but is in Warsaw. Rather interestingly, this segregation of industrial development in Warsaw occurs concurrently with a high percentage of land devoted to residential use.

Thus, a number of obvious differences in urban structure have been found on the basis of empirical evidence for Warsaw compared to some American cities, and in particular, Chicago. In general, there is a much steeper decline in density for Warsaw and a higher amount of average and central city development. These results, in turn, provide an understanding for the much larger land areas American cities occupy compared to Warsaw. It could be argued that residential attraction to sites with good transportation facilities and repulsion from those with major industrial development are naturally occurring preferences, which would be found in any land market and do not need a strong planning influence to affect development levels. However, the empirical evidence in this chapter indicates the opposite, at least for Chicago and Warsaw. That is, a planned urban environment obtains these results more easily and noticeably.

SUMMARY

Clark's belief that a negative exponential function describes the distribution of urban population densities for all times and places studied has not been disproven here. As shown in this study, urban development in capitalist and socialist cities shows a similarity in overall pattern. A negative exponential function provides an equally good description of the pattern of population densities for Warsaw as well as for Chicago, Boston, and other American cities. Nevertheless, the expectation was that the differences in economic processes should produce differences in structures for capitalist and socialist cities. The answer to this dichotomy is in the type of results derived from the negative exponential model and in the motivating forces producing the results. Because of the differences in the way land is valued in socialist and capitalist societies the values for the parameters of the model have been shown to be radically different. As a result, there are striking differences in residential structure between capitalist and socialist cities but the observable differences existing in space are in the degree and kind of development and not in the general pattern.

University, New Orlean
Louisiane State, U.S.A.

BIBLIOGRAPHY

Alonso, W., 1964, *Location and land use: Toward a general theory of land rent,* Harvard University Press, Cambridge, Mass.

Andrzejewski, A., 1966, Post-war housing development in Poland, in: *City and Regional Planning in Poland,* Jack C. Fisher (ed.), Cornell University Press, Ithaca, N.Y., pp. 153–182.

Berry, B. J. L., Simmons, J. W., Tennant, R. J., 1963, Urban population densities: Structure and change, *Geogr. Rev.,* 53, 3, pp. 389–405.

Casetti, E., 1967, Urban population density patterns: An alternative explanation, *Canadian Geogr.,* 11, 2, pp. 96–100.

Fisher, J. C., 1962, Planning the city of socialist man, *J. Amer. Inst. Planners,* 28, pp. 251–265.

Jankowski, S., Wilski, J., Wyporek, B., 1971, *Warsaw, 1970–1985*; Transl. T. Świecka, Warsaw Town Planning Office, Warsaw.

Lentz, P. A., 1973, *The influence of a socialist economic environment on intra-urban residential spatial structure,* Ph. D. dissertation, Department of Geography, Clark University, Worcester, Mass.

Malisz, B., 1966, Urban planning theory; Methods and results, in: *City and Regional Planning in Poland,* Jack C. Fisher (ed.), Cornell University Press, Ithaca, N.Y., pp. 57–84.

Mills, E. S., 1969, The value of urban land, in: *The quality of the urban environment*, Harvey S. Perloff (ed.), Resources for the Future, Inc., Washington, D. C.
Muth, R. F., 1969, *Cities and housing: The spatial pattern of urban residential land use*, University of Chicago Press, Chicago.
Techniczno-ekonomiczne studium kierunków rozwoju przestrzennego m.st. Warszawy w okresie planu perspektywicznego, 1960–1980 (A technical-economic study of directions of spatial development of Warsaw. The period of longterm plan, 1960–1980), vol. 1: Metody i wyniki (Methods and results), Warsaw Town Planning Office, Warszawa 1966.
Wingo, L., Jr., 1961, *Transportation and urban land*, Resources for the Future, Inc., Washington, D.C.

GEOGRAPHIA POLONICA 31, 1975

THEORY OF INTRA-URBAN STRUCTURE: REVIEW AND SYNTHESIS. A CROSS-CULTURAL PERSPECTIVE

Piotr Korcelli

(I) INTRODUCTION

The literature on intra-urban structure and growth has been expanding rapidly during recent years, which is a typical feature in the case of complex, interdisciplinary problems. This expansion is well testified by the increasing number of reviews and theoretical syntheses such as those by B. J. Garner (1967), Ch. L. Leven (1969), R. Colenutt (1970), R. J. Johnston (1972), B. Greer-Woottle (1972) and M. L. Senior (1973).

The present article [1] has two specific objectives. The first aim is to reduce the multitude of generalizations found in the contemporary literature on intra-urban structure into a set of basic theoretical statements and postulates. The second aim is to demonstrate to what extent individual theories and statements reflect particular social, political and cultural contexts and to show which approaches, and why, may be regarded as of a more or less universal range of applicability.

It is possible to identify six major approaches which have contributed to the existing body of theory on urban spatial structure and growth. These include:

(1) ecological concepts,
(2) theories of urban land market and land use,
(3) urban population density models,
(4) models of intra-urban functional patterns (or spatial interaction models),
(5) settlement network (or system) theories,
(6) models of spatial diffusion on an intra-urban scale.

Such a division seems to be becoming conventional as, if account is taken of some differences in terminology, it is rather similar to the classifications recently employed by other authors, particularly Colenutt (1970) and Senior (1973).

In terms of origin, those theoretical approaches are related to various disciplines, such as sociology (1), economics (2), demography (3), urban planning (4) and geography (5, 6). Until very recently they developed largely independently of each other. The trends towards an integration are mainly a product of the present decade. Even as late as 1969/1970 some leading authors could deplore the state of urban and regional research which was "more of a market place where various ideas are traded than a building ground where brick is placed on brick in an orderly fashion" (Hägerstrand, 1969, p. 62), and

[1] It is based on a larger study: Korcelli (1974).

was characterized by a lack of an "interdisciplinary framework for thinking let alone theorizing" and split into a number of "confined conceptual worlds" (Harvey, 1970, p. 47).

With the above opinions still largely holding true, it may be noted, however, that the late 60s and early 70s have witnessed a growing interaction between various concepts and theories. The integration of classical ecological models, via factorial ecology, with the social area theory, and the application of land rent concepts to an interpretation of population density patterns within cities are but two leading examples. They are characteristic of an advanced stage in theory formulation when findings of certain approaches are interpreted in terms of, or justified by, generalizations developed within other approaches. The systems framework, introduced to urban spatial studies by B. J. L. Berry (1964), is believed to offer still greater potential for integrating seemingly divergent streams of theoretical thought. The already established links between the six approaches earlier identified are traced throughout the paper, and some other links are also suggested.

(II) ECOLOGICAL CONCEPTS

There are two reasons for starting the discussion with the urban-ecological school. First, this approach has so far commanded great interest among representatives of "urban science", a fact which is reflected by the immense and still rapidly expanding literature. Second, the question of cross-cultural differences comes out in ecological studies more explicitly than elsewhere. The crux of the matter lies in whether or not within the urban residential space there are pronounced variations in terms of median family income, level of education, demographic characteristics and other related variables. In case of an affirmative answer ecological forces are assumed to operate and made responsible for the variations identified. The next question is whether those variations tend to become more intensive with the passage of time or, otherwise, tend to level-out. The third basic question refers to the spatial pattern. This can be: (a) concentric with a positive or negative relationship between the distance from the city centre and neighbourhood quality, (b) sectoral, (c) mixed (some variables showing a concentric, others — a sectoral pattern), (d) mosaic-like, or (e) other. Comparative factorial ecology (B. J. L. Berry, 1971), seeks answers to the above questions and it has produced a number of important generalizations (Timms 1971, Rees 1970, Robson 1969). These, however, when seen in the world-wide scale, are still quite partial and sometimes inconclusive.

Thus, according to B. J. L. Berry (1971, p. 219) the findings of factorial urban ecology demonstrate that detailed comparisons based on uniform methods and concepts, appear appropriate only within: (a) Western, industrial societies, (b) societies of planned development, (c) traditional, or pre-industrial societies. Other evidence suggests the actual divisions to be still more complicated and related to:

(1) level of economic development and technological advancement,
(2) social and economic systems,
(3) cultural systems,
(4) degree of ethnic heterogeneity,
(5) size, age, rate of growth, and functions of the city.

In R. E. Pahl's (1967) opinion the dividing lines can be found even within individual urban communities as various social groups differ in respect to the level of spatial mobility and the degree of freedom in selecting residential location. Similar notions are also expressed by J. E. Vance, Jr. (1969).

One of the problems which bear heavily on factor ecological studies and conclusions reached in such studies, is that the principal theoretical basis of urban factorial ecology i.e. the social area theory, explicitly pertains to contemporary Western societies. It should be emphasized that this fact is well reflected in the results of empirical work. Thus, the social area model which postulates the three basic dimensions of differentiation of the urban community: socio-economic status, family status, and ethnic statuts, has been generally confirmed by factor ecological studies of those Western cities with a high degree of ethnic heterogeneity (see Murdie 1969, Rees 1970, Salins 1971, Timms 1971), as well as by studies of some other Western industrialized cities (Pedersen 1967, Janson 1971) in which factors corresponding to the first two axes have been identified. In the case of Mediterranean and Latin American cities (McElrath 1962, Morris and Pyle 1971) these factors could also be derived, though they seldom formed distinctly separate spatial patterns. On the other hand, analyses for Asian and African cities (Abu-Lughod 1969, Linday 1971) show a high correlation between socio-economic and demographic variables and a clear presence of the ethnic status factor.

Few factor ecological studies exist for the cities of socialist countries and therefore it is not easy to discuss their patterns in terms of the social area model. However, empirical work based on other concepts and methods largely suggests that the Shevky-Bell hypothesis may be invalid in this case. For instance, V. V. Pokshishevsky (1973) has demonstrated a gradual increase in the spatial integration of various ethnic groups in the large cities of the USSR. Planning in a factor which was in fact disregarded by the social area theory. B. T. Robson (1969) has already noted that in his study of Sunderland. When equipped with such tools as housing assignment policies and the control over a bulk of housing construction, planning has proved to be a very effective vehicle in eliminating sharp differences within urban residential space.

Studies of Polish cities (Pióro 1962, Jałowiecki 1968), although displaying a persistence of spatial differentiation in a number of ecological variables, show a substantial decrease in such differentiation since World War II. A recent factor ecological study of Warsaw (Węcławowicz 1975) has proved that while in the 1930s residential variations within the city closely corresponded to the one postulated by the social area model, the contemporary (1970) structure was much less pronounced and could yield to an interpretation mostly in terms of the evolving housing policies. Although the image of a high or low prestige, attached to individual districts, has not fully disappeared, the motivations of individuals to change residence concern, in most cases, the apartment itself (i.e. its size) rather than its locations within the urban social space. This is an effect of the basic policy idea, followed since World War II, which has been to achieve a more or less homogeneous socio-economic structure and a similar standard of services, especially of educational, transportation, and recreational facilities throughout the city.

In contrast to the problem of dimensions and intensity of residential differentiation, the question of its spatial patterns has been subject to hypotheses whose scope extends to a cross-cultural scale. There exist in fact two competing concepts, both formulated in quasi-dynamic terms. Reference is made here to L. F. Schnore's (1955) model and to the scheme developed by R. J. Johnston (1972). The first one combines the "reverse Burgess" and "Burgess-type" patterns into an evolutionary sequence where transition from one into the other is attributed to the change in economic and technological conditions which in turn determine locational preferences of those social groups which are characterized by the highest degree of spatial mobility. Johnston's concept which stems from a criticism of Schnore's model, sees the emergence of a mosaic-like pattern as an outcome — both in developed and developing countries — of the growing role of the intermediate social strata.

It has been proved that certain elements of the concentric pattern, together with some sectoral components, can be found in cities representing all the major contemporary socio-political and cultural systems (see Tanabe 1970), although in the socialist city they are more closely related to the arrangement of land uses than to the differentiation of residential space (Werwicki 1973). One may agree with R. J. Johnston that present patterns are reflecting a stage of transition, and that they are, at least temporarily, becoming more complex or "mosaic-like". (A good example of this trend is Warsaw; see Węcławowicz 1974). It seems to be a general rule that when technological conditions (mostly transportation) allow, groups with growing incomes tend to seek residence near the edge of a city. Nevertheless, proportions between those moving towards an outer zone and those selecting a central location greatly differ. This rule, therefore, is not strong enough to suggest whether the existing patterns are now converging or whether they are still evolving in divergent directions.

The question is closely related to the ecological succession which remains one of the least developed aspects in comparative urban ecology. Most studies pertain to the succession of ethnic groups and, more recently, to the life-cycle succession (Simmons 1968). They all are characteristic of American cities and, in the case of the latter, also of some other large Western cities (Johnston 1969). Other types of succession, such as those connected with urban redevelopment, are probably more universal in scope and they should be subjected to comparative study on an international scale. The existing evidence allows one to suspect that different types of succession may possess numerous common properties. One such property is the cyclical nature of succession, another — a path of movement from the city centre towards the peripheries. At this point social ecological concepts meet spatial diffusion models which are treated in one of the later sections.

Succession (along with competition in space and dominance) belongs to those basic notions of human ecology which were directly borrowed from plant ecology. In fact, the definition of human succession by R. Park [2] closely resembles a standard biological definition (see, for example, Odum 1959). The notion of invasion–succession lay also at the base of the dynamic interpretation of the Burgess model. There it referred both to population change and the evolution of urban landscape. Later on this notion became chiefly restricted — at least in sociological literature — to the movement of social groups, particularly ethnic minorities, over space and over time. On the other hand, geographers developed a parallel concept of sequent occupance (Whittlesey, 1929) which covered a rather wide array of phenomena but primarily referred to the change in area physiognomy and in land use. Within the former approach, i.e., the study of human succession proper, it is possible to identify a traditional interpretation

[2] "... an orderly sequence of changes through which a biotic community passes in the course of its development from a primary and relatively unstable, to a relatively permanent or climax stage..." In the course of this development the community moves through a series of more or less clearly defined stages (after Timms 1971, p. 87).

emphasizing the antagonistic character of succession, being an outcome of the expansion of the area dominated by a particular group, more dynamic than its neighbours; and a more recent, and quite detached from plant-ecological origins, behavioural interpretation.

In the traditional interpretation, phases of succession were determined in reference to stages of community organization and the intensity of change. P.E. Cressey (1938), for example, divided the succession cycle into the phases of invasion, conflict, recession and reorganization, while O.D. Duncan and B. Duncan (1957) identified the phases of penetration, invasion, consolidation, and saturation. It was established empirically that the succession cycle might be depicted by a normal, unimodal curve if its change of intensity over time was taken as a measure, and by a S-shaped curve, well known in innovation diffusion studies, if the degree of its accomplishment was considered. A high regularity of the succession cycle led sometimes to a narrow, deterministic interpretation (and related action) often referred to as a "self-fulfilling prophecy" (Wolf 1957, quoted by Bourne, 1971).

The more recent, behavioural interpretation of ecological succession is mainly based upon the idea of the family cycle (or life cycle). R. J. Johnston (1969) argues that the life-cycle succession accounts for a substantial part of intra-metropolitan moves, especially, as he implies, when in-migration is negligible, ethnic composition of inhabitants homogeneous, and the overall income level high. In the case of Melbourne, Johnston identified six stages of the family cycle, the passing from one to the next being conducive to a change in location from the centre to an intermediate zone, then to a marginal zone (which becomes an intermediate zone again with city growth) and back to the centre. This scheme is in accordance with the data on spatial mobility of urban dwellers (Simmons 1968), moreover it is consistent with a concentric pattern of family status variables as depicted in numerous factor ecological studies (see Salins 1971).

In a cross-cultural perspective data on change of residence within urban areas are rather scarce and, if available, they rarely allow the isolation of the individual migration factors (some authors prefer to speak of bundles of factors). Therefore, it is possible to propose other general hypotheses relating intra-urban migrations to social and economic mobility or to changes in the urban environment. Those are, in fact, indirect factors acting through housing market systems and various institutional factors. Their differentiation in the contemporary world restricts considerably the range of generalizations on intra-metropolitan migrations.

The geographical concept of sequent occupance, closely corresponding to a broad interpretation of ecological succession, was recently rediscovered and developed in the Whittlesey tradition by several authors who, however, do not formally claim this inheritance. Reference is made particularly to a study by E. H. Hoover and R. Vernon (1959) who formulated a model of the morphological succession cycle for a large metropolitan area.

Individual concentric zones, according to Hoover and Vernon, tend to experience the following sequence of occupance: (1) development of single-family houses (corresponding to primary succession), (2) spread of multi-family housing, (3) influx of minority ethnic groups and/or economically disadvantaged groups, and conversion of residential buildings, (4) deterioration of housing and decline of overall densities, (5) redevelopment. This model has been empirically confirmed by B. Duncan, G. Sabagh and M. D. Van Arsdol (1962) who received a regular pattern of population growth (with rapid growth in the first stage, slower — in the second stage and negative increase in the remaining stages) for cohorts of census tracts developed during subsequent time periods; and by D. L. Birch (1971) who found a statistical relationship between the population growth rate of an individual zone and its position on the time axis of the morphological cycle, as well as between the age of buildings and a phase of that cycle.

An apparent limitation of the above scheme is that it disregards uses other than residential, in other words, it pertains to about 40–45% of developed urban land. As is implied by L. S. Bourne (1971), a sequent occupance model is far too simple to be applicable in an analysis of succession between individual types of land use. The latter is found to be discontinuous over time and hardly yielding to a deterministic interpretation. A regular pattern of change, according to Bourne, can be seen only at a high level of data aggregation and after smoothing local variations and spatial discontinuities. One of the factors that make for a lack of regularity are the differences between functional succession and a succession of spatial forms, the one characterized by higher inertia, and referred to by M. Colean (1953) in his concept of the renewal cycle. To separate the two processes, however, is not a simple task, as the former takes place within the latter. This was in fact implied in the Hoover–Vernon scheme, where redevelopment stages were separated by stages of adaptation. W. K. D. Davies (1968) described this relationship more precisely by defining individual development stages according to dominant functions and by dividing each stage into the phases of (1) occupance, (2) occupance and conversion, and (3) creation of a new spatial form.

The cyclical nature of urban development is also portrayed in numerous historical studies on the evolution of urban landscape. M. R. G. Conzen (1962), in an important study, on the plan analysis of a city centre introduced the concept of a succession cycle for individual burgages, consisting of four main phases: institutive, repletive, climax, and recessive. The last stage is followed by a period of "urban fallow", and a subsequent redevelopment.

(III) THEORIES OF URBAN LAND MARKET AND LAND USE

Theories conceiving urban structure as the reflection of spatial patterns of transport costs and urban land rent, share with ecological theories a number of basic notions, such as competition in space and dominance. The relationships are genetic since R. Hurd (1903) is recognized as the forerunner of both groups of theoretical concepts. In addition, the "neo-classical" ecological theory (Hawley 1950) assumes a given pattern of rent values as a framework for the development of social patterns over urban space. This was already noted by D. W. G. Timms (1971, p. 94) with respect to the classical urban-ecological models.

There are fundamental differences between the role of land rent in the contemporary Western city and in the socialist city. The "Western" theory of the capitalist city explicitly links land use and urban rent patterns and looks upon allocation as a process based on competition where individual decision-making units aim at occupying locations with the highest utility level (Ratcliff 1949) or, in the case of the residential sector, the highest level of satisfaction (Alonso 1964). B. J. Garner (1967, pp. 335–336) gives a simplified account of this process: "Each activity has an ability to derive utility from every site in the urban area; the utility of a site is measured by the rent the activity is willing to pay for the use of the site... In the long run, competition in the urban land market ...results in the occupation of each site by the "highest and best" use, which is the use able to derive the greatest utility from the site and which is, therefore, willing to pay most to occupy it. As an outgrowth of the occupation of sites by the "highest and best" uses, an orderly pattern of land uses results in which rents throughout the systems are maximized and all activities are optimally located".

In the theory and planning of the socialist city the nation of land rent is not formally applied. Instead, urban land development is based upon principles of investment effectiveness, the productivity of agricultural land (if urban territorial expansion is involved), technical and economic requirements of different uses, and the functional organization of urban space. During recent years a number of concepts of determining urban land prices were proposed. It has been suggested that a uniform and theoretically based system of land prices should aid planning and land development policies. The major concepts, put forth by Polish authors, are comparatively presented and analysed by J. Kolipiński (1970). In most of the methods proposed accessibility is treated implicitly through such variables as the value of investments and the productivity per unit of land. H. Chołaj (1966) considers explicitly site rent as a component of the hypothetical price of urban land, relating its pattern to the actual structure of transportation costs.

R. M. Haig (1927), who laid the foundations of the economic theory of urban spatial structure, considered site rent and transportation costs as complementary values, summing up to total costs of friction. He identified site rent with the value of savings resulting from the accessibility of a given site to the remaining sites within an urban area. Locations with low aggregate transportation costs, because of their limited supply, are becoming occupied by those activities which are able to take advantage of high accessibility most effectively. "The layout of a metropolis — the assignment of activities to areas — tends to be determined by a principle which may be termed the minimizing of the costs of friction" (Haig 1927, p. 39). The resulting land use pattern, with assumptions simplifying the structure of transportation costs, corresponds, in fact, to that portrayed by von Thünen. Haig based his analysis mainly on the location of business and manufacturing oriented on the local market, but he extended his assumptions concerning the role of aggregate transport costs to all kinds of activities, the residential sector included. He considered locations (sites) in terms of points rather than portions of territory.

The latter ommission was erased by W. Alonso (1964) whose theory emphasizes the substitution of transport inputs and lot size. He assumes that in the land development process individual decision makers, be it firms or households, aim at minimizing rent and transportation costs and maximizing the area occupied. A crucial concept in Alonso's theory — the bid rent curve is defined as a set of combinations of rent and transport inputs which represent an equal satisfaction level for an individual. Alonso proves that users with steep bid rent curve gradients tend to occupy locations close to the city centre. This allows him to interprete the classical urban-ecological models. The Burgess-type pattern, in terms of Alonso's theory, is the case when for higher income groups the rate of decrease of the marginal rate of substitution is less than the rate of the increase of land (Alonso 1964, p. 108). The reverse pattern represents a different scale of preferences; factors other than income, influencing preferences are, however, beyond the scope of Alonso's theory. Other concepts of urban land use, which also apply the principle of substituting space for transport inputs, and which explicitly pertain to a competitive, free-market economy, were developed by L. Wingo (1961) and R. Muth (1961, 1969).

The theory of urban land market has been criticized for its static-equilibrium form and the assumptions drastically simplifying the reality, such as the location of all service and employment opportunities at a single city centre, a symmetric pattern of transport costs and the condition of perfect competition. S. Angel and G. M. Hyman (1972) have demonstrated that the substitution of rent and transport costs leads eventually, at a certain distance from the city centre, to negative rent values. The assumption of a constant rent may be a solution but then the condition of a single centre (or a single trip destination point) becomes impossible. The same assumption is questioned by D. E. Boyce (1965) and J. O. Wheeler (1970) who show a poor correspondence between empirical data on directions and length of trips for various socio-economic groups and the pattern postulated by the urban land market theory.[3] Another simplification in the theory which deserves criticism and further study is that it disregards all components of differential rent, other than site rent. Numerous studies (Brigham 1965, Czamański 1966) have empirically established the importance of such factors as topography, directions of urban growth, environmental quality, and historical factors. E. Ullman (1962) defines those factors collectively as site attractiveness, while P. H. Rees (in Berry and Horton, 1970, p. 301) proposes a term amenity rent. Empirical studies (for example Yeates, 1965) demonstrate a growing share of the amenity rent in relation to other rent components.

In the present context the question of generality of the urban land market and land use theory over time and space is of primary importance. J. E. Vance, Jr. (1971) came to the conclusion that the theory can be applied, with some limitations, to an analysis of the spatial structure of the capitalist city, but he denies its relevancy in the case of both the precapitalist (or feudal) and the post-capitalist city. To the latter categories one should add the socialist city. Vance notes that in the feudal city land utilization was of a functional rather than an economic character and the land assignment basis related to the social status of the users rather than the rent-paying ability. In the capitalist city, to continue Vance's arguments, the principle of land inputs — transportation inputs substitution mainly pertains to man's economic activity, while the spatial structure of the housing market is shaped mostly by social factors, such as the class structure of the society, status models and symbols, discrimination (i.e. ethnic ghettos) and others.

In the socialist city, as represented by a large Polish city, the substitution principle is actually of marginal importance.[4] New housing is characterized by practically uniform standards not only in terms of materials and equipment but also in density. Since the overwhelming majority of new dwelling units are in multi-family buildings, net residential densities in housing estates, located in city peripherial zones, do not greatly differ from those in centrally situated residential districts. Although central locations are generally preferred, this does not usually produce a growth in population concentration within the inner

[3] Attempts of W. Alonso (1964, pp. 134–141) as well as G. Papageorgiou and E. Casetti (1971) to construct polycentric urban rent models should be noted at this point.

[4] Recently P. Lentz (1973) applied Muth's model to an analysis of land use and population densities in Warsaw. She found some general similarities between the spatial patterns of Warsaw and of America cities, although, as she admitted, those patterns were produced by different processes.

city districts as, because of the communal ownership of land and the centralization of investment rights, the land development process is largely of a non-competitive character. At the same time, diseconomies associated with both urban sprawl and congestion are recognized by the planners and decision makers. With the spread of single-family housing the substitution rule may grow in importance, although the terms opposite to transportation costs would be the existing structure of housing and amenity factors (particularly relating to air quality and noise disturbances) rather than land rent.

This leads to another critical question which is the implicit time dimension of the urban land market and land use theory. Of particular interest is the effect of changes in the site rent gradient upon land use patterns. If technological progress in transportation and the related decrease in the friction of space lead to the equalization of accessibility values (or aggregate transport costs) throughout the city, their consequence is definitely a flattening site rent gradient. Not every improvement in transportation technology brings however such effects; while some improvements may promote spatial concentration (in a punctual or a linear form), others aid deconcentration. If one assumes, as one finds some empirical evidence [5], that those trends follow one another, then changes in accessibility and in site rent gradients may be of a cyclical character. Two reservations are necessary. First, the progress in transportation technology and the growth of spatial mobility result in an expanding spatial scale of a city. Second, a usual consequence of technological improvements is an increase of the selectivity in the choice of location. Locational decisions are becoming more and more dependent on the amenity factor, representing all the characteristics of a site other than accessibility, and largely disregarded by the classical rent theory.

The evolution of the spatial structure of accessibility and amenity values involve the concentration and deconcentration processes in urban land use patterns. A majority of factors listed by Ch. Colby (1933) in his seminal paper on centrifugal and centripetal forces in urban geography can be, in fact, reduced to the transportation cost and the site attractiveness. The latter is to some degree a function of the former, although its role also depends on the society's level of affluence, the class structure, the income distribution and on the prevailing values and laws. When spatial variations in accessibility (or site rent) tend to be levelled-off, the amenity factors begin to play a more significant part. So the proportions, in the total rent values, of the transportation costs and the site attractiveness components largely determine the balance of the centrifugal and the centripetal forces within an urban area. A change in the balance in favour of either of those forces means, in terms of Alonso's theory, a shift of the tangency point between a bid rent curve and a curve corresponding to the real cost structure, from the city centre outwards. In consequence of diminishing relative variations in the accessibility patterns the individual bid rent curves tend to become less steep. Deconcentration processes in land use patterns occur when, with site amenity values being a positive function of distance from the centre, the aggregate transportation costs gradient fall over time, followed by the flattening of individual bid rent curves.

---

[5] During the rail transportation era the accessibility of the central districts increased faster than that of the remaining city areas. Motor transportation brought a more uniform growth in accessibility, with the most substantial additions in the marginal zones. A subsequent atrophy of the transportation system calls for the introduction of limited access roads and new means of rapid transportation, the elements that destroy the former (a relatively uniform) accessibility map.

(IV) URBAN POPULATION DENSITY MODELS

The Clark's (1951) rule [6] describing an intra-urban population density pattern in terms of a negative exponential function of distance from the city centre, has often been recognized as the most universal statistical generalization in urban geography. Although tested by numerous authors, including B. J. L. Berry *et al.* (1963), W. J. Romashkin (1967), J. E. Brush (1968), and P. Lentz (1974) for cities of varied size, origin, socio-political and cultural milieu, the rule has never been invalidated. In addition, B. J. L. Berry (1964) has found that for most of the cases investigated the density gradient declines with the city size. On this basis he postulated a relation between the density rule and the rank-size rule.

Empirical material suggests that the density gradient (the b parameter) and the central density (d_0) substantially vary between cities (of comparable class categories) located in different countries and different parts of the world, as well as between cities of different age, shape and functions. Those observations prompted many authors to build formulae describing individual types of density patterns more precisely than does the Clark's formula. Thus G. G. Sherrat (1960) put forth a quadratic exponential function:

$$d_x = d_0 e^{-bx^2}$$

while Yu. V. Medvedkov (1968) proposed a function with the b values differing — depending on the direction from the city centre:

$$d_x = d_0 e^{-b(\varphi)x}$$

B. L. Gurevich and J. G. Saushkin (1966) described several hypothetical density patterns, including one with the gradient value constant in all directions but varying with distance from the centre. B. Kostrubiec (1970) identified that pattern with the generalized density profile constructed by S. Korzybski (1952), whose work was parallel to, and independent of, Clark's contributions. The pattern is in fact similar to the one derived by R. Ajo (1965) for London, in which densities in subsequent concentric zones are described by different formulae. Attempts have also been made to build formulae taking into account the central density crater, and to build density models for polycentric patterns, both for a single city (Gurevich 1967, Dacey 1968) and a regional urban system (Papageorgiou and Casetti 1971). In this latter approach the density rule is treated as a general model of population distribution.

E. Casetti (1969) carried out a comparative analysis of the existing empirical and hypothetical density models. While testing the models against data for several large cities he found that the formulae describing the density crater do not give a much better statistical fit than those disregarding depressions at the centre. This feature of the density profile was therefore ruled as secondary. The analysis also confirmed the hypothesis put forth earlier by R. Ajo; according to which density patterns within individual concentric zones of a city are most precisely described by different functions. In general, however, all the models tested by Casetti explained a very high percentage of density variations within the cities investigated.

Although the urban population density concept appeared originally as an "inductive generalization in search of a theory" (Berry, 1964), its explicative layer is now not much thinner than that of other concepts pertaining to the intra-urban structure and growth. Most frequently the density profile is explain-

---

[6] An earlier formulation is by H. Bleicher (1892).

ed in terms of transportation costs and site rent structure; however, E. Casetti (1967), G. J. Papageorgiou (1971) and V. I. Sevostyanov (1972) have attempted more general interpretations, avoiding the notion of land rent. A decline in the population density with distance from the city centre logically follows from the Alonso (1964) theory. When accessibility values decline from the centre the amount of space occupied by individual households increases, and the residential density decreases in a similar manner, with a correction made for variations in family size. The negative exponential decline of density with the distance does not follow directly from the theory, but J. H. Niedecorn (1971) and E. Casetti (1971) derived that function using such assumptions as a concave transportation costs curve and uniform household preferences.

In an earlier study E. Casetti (1967) treated population distribution within a city as an outcome of two opposite locational factors: the centrality preference and the preference for noncongested sites. He demonstrated that the equilibrium density pattern generated by a competitive process is higher than the optimum density. This difference is supposed by Casetti to form a basis for the so-called "deglomeration" policies which have been followed in the socialist countries. Despite certain misconceptions (Brown 1968) Casetti's study was an interesting attempt to look at land use and density patterns from a perspective broader than that of a particular socio-economic system.

G. J. Papageorgiou's (1971) approach is in many respects similar to that of Casetti (1967, 1969). Since the same population distributions are found throughout the world, he proposes to attribute the density gradient phenomenon to "basic locational preferences of the household, preferences strong enough to dominate within entirely different economic systems" (*ibid*, p. 26). Those preferences, according to him, are a general desire for space abundance and accessibility. Unlike Casetti (1969), however, Papageorgiou is less concerned with the empirical performance of individual density functions than with their theoretical consistency. The basic criterion adopted by him is that the function associated with the least restrictive assumptions should be regarded as the best description of the density gradient. Those assumptions are primarily related to transportation costs functions. The least restrictive function, i.e. a constant rate of increase in costs with distance from the city centre is implied by the equations of C. Clark (1951) and E. Casetti (1969), and those models are labelled by Papageorgiou as the most general. The functions accounting for the density crater are associated with estimation difficulties and, in general, are based on what he suggests to be an unrealistic structure of transportation costs. The last conclusion does not seem to be firmly established, since the density crater is a realworld phenomenon and its existence is related to the varying proportions of residential land within individual concentric zones. The Papageorgiou utility model, as well as the earlier Casetti (1967) model, pertain to net population densities, while the empirical functions which they analyse were built to describe gross density profiles. The functions with extrapolated d_0, including the Clark formula, are probably a better approximation of the net than the gross density pattern, but this is a hypothesis which remains to be ascertained.

Equivocal differences between various density models, in their empirical and theoretical performance, suggest that this problem may only be solved within a dynamic framework. Here, theoretical analyses are much less developed than in the case of the equilibrium density patterns. The available data series for the b values show, in a very consistent manner, the progressive flattening of density profiles over time. This may imply that: (1) with the passage of time and with the growth in size, cities are subject to a deconcentration, and/or (2) the deconcentration processes, which are tied to certain technological, economic, and cultural factors, are a contemporary feature of cities (of all size categories) in the developed countries. The former hypothesis can be ascribed to H. H. Winsborough (1963), B. J. L. Berry (1964), and B. E. Newling (1966, 1969), while the latter was exposed mainly by D. J. Bogue (1950), A. H. Hawley (1956), and L. F. Schnore (1957, 1959).

Newling has found that the rate of density increase is a positive, exponential function of distance from the city centre and a function of the overall density level. At any given point within the city the density increases until a maximum is reached and then it falls gradually. The points of maximum density move over time from the city centre outwards. The evolution of density at the centre is portrayed as a quadratic exponential function of time. The maximum, or the critical density is explained by Newling as an optimum value (under given conditions), the overpassing of which involves social costs.

Although the model is quite general, it is descriptive rather than explanatory. However, its expanded interpretation is possible within the framework of the so-called suburbanization theory. The process of suburbanization was defined by D. J. Bogue (1950) and A. H. Hawley (1956) as a rapid increase of the population growth rate in the peripheral zones and a net outflow of people from the city inner areas. In the terminology of the population density concept this means the progressive decline of d_0 and b values over time, which was postulated by Newling and by other authors. H. H. Winsborough (1963) has shown that the population growth of a city may proceed through: (1) an increase in the values of both d_0 (or the concentration level), and b (degree of congestion), (2) an increase of d_0 and a decline, or stabilization, of b values, (3) a decline of both the central density and the gradient value. The above conditions do not take account of the density crater, since its presence would allow a growth of total number of inhabitants with falling d_0 and a constant (the case described by Newling), or a flattening density profile outside of the city centre. The debate between the proponents of the suburbanization theory and its opponents (R. E. Schmitt, H. Blumenfeld) pertained to those conditions exactly. If the former referred to data showing a definite shift, at a given point on the time axis, from the types of change (1) and (2) into the type (3), the latter supported the principle of historical continuity of the development of population density patterns, and treated falling central densities and an accelerated population growth in outer zones as attributes of big cities in general. To solve this dilemma it would be necessary, according to O. D. Duncan (1959), to analyse comparatively a number of cities of a similar size, but for varied historical periods, and to look for analogies and differences in the evolution of density patterns. Since changes of density patterns are of a secular nature, the available data are not likely to allow such a comparative study to be carried out, and the problem itself will for some time remain in the domain of loose hypotheses.

However, contemporary changes in urban density patterns are relatively well documented and on this basis some fragmentary generalizations, on a cross-cultural scale, were attempted by several authors. It has been already asserted that the existing studies pertaining to West European, North American, and Australian cities, consistently show, over time, the parallel decrease of density gradients and of central (as well as maximum) densities. Studies for socialist cities contain findings basically analogous to those noted above. Thus, G. A. Golc (1972) indicated that in the case of Moscow both the maximum density value and the b value have substantially declined during the last decades. In addition,

W. I. Romashkin (1967) found a further decline in population densities within the central districts of Moscow over the 1963–65 period, and a substantial growth of densities in its marginal zones, with inter-sectoral differences becoming more pronounced. The density profiles for Cracow, compiled by K. Bromek (1964) for the 1880–1950 period show the emergence of the central crater about 1920 and a progressive flattening of the density curve since 1930. Over the whole 70-years-long period the maximum density zone moved from the distance of about 0.4 km to 1.5 km from the historical centre of Cracow. Trends in density patterns are often more complicated when a city is subject to a large-scale redevelopment, as has been the case of Warsaw after its nearly-total destruction during the Second World War. Large portions of the central zone in Warsaw were put to residential construction in the 1960s, after the surrounding areas had been already developed. New, centrally situated housing is often in the form of residential towers, a fact which adds to the recent increase of net population densities within the central part of Warsaw. This also points to the broader problem of redevelopment cycles and their effects upon population densities, the problem to be returned to in the concluding paragraph of the present section.

Although the density patterns in cities of the Third World countries remain weakly recognized, B. J. L. Berry *et al.* (1963) ventured a generalization according to which the patterns in question, when seen over time, are characterized by growing d_0 values and a stability of b values. This also suggests a general increase of central densities with growing city size. Those rules were basically supported by J. E. Brush (1968) who investigated in more detail the cities of India. He noted, however, that in the case of some large and relatively young cities, such as Bombay, the increase of density within outlying zones was much more rapid than the growth of the central densities. A comparative study by P. W. Amato (1970) for four capital cities in Latin America, supplies similar findings, i.e. the general growth of population densities and a decrease of b values. Thus, Berry's graphical model found a confirmation in more recent empirical analyses, although they do not answer the fundamental question of whether the recent flattening of density profiles and a parallel increase of maximum densities in cities of Asia and Lating America, follow the principle of a gradual density decline with the growing urban scale or whether those trends are symptomatic of the early stages of deconcentration processes.

The above discussion supports a hypothesis of the suburbanization theory, according to which the deconcentration process represents a certain stage in the development of urban density patterns, and needs not be necessarily associated with the growth of city size. The principal factor of deconcentration can be undoubtedly identified with changes in transportation costs and their effects upon income budgets and the time budgets of individual households. At the same time, the pace and range of deconcentration processes is related to the prevailing socio-economic systems, cultural and ethnic patterns, general intensity of space occupancy, rates of economic growth, rates of overall population increase, and so on. A good illustration is provided by the differences in the nature of deconcentration between North American and West European cities (see Pokshishevsky and Gokhman, 1961).

The conclusion that under given conditions city growth is accompanied by city deconcentration calls for determining the limits of the process. B. Duncan *et al.* (1962, p. 428; also quoted by H. Winsborough, 1962) sees the limits in the following way: „... growth tends to take the form of outward expansion until the spread of the city begins to present a barrier to internal transportation and communication, at which juncture growth takes the form of upward expansion near the metropolitan core". If therefore, one may assume that urban density patterns pass through subsequent stages of concentration and deconcentration, then it is possible to describe their transformations as a cyclical process, connected with the cycles of occupancy and succession, and related to changes in accessibility patterns, which were discussed in the previous section.

(V) MODELS OF INTRA-URBAN FUNCTIONAL PATTERNS

The theoretical background of models and research approaches labelled here "functional" are concepts of social physics (Wilson 1969A), more precisely the concept of human interaction in space. These concepts, introduced to spatial analysis during the 1940s (J. Q. Stewart 1947, G. K. Zipf 1949) have found numerous applications in studies of migrations, settlement network and intra-urban structure. Since models of functional patterns were often developed by planners in connection with particular studies of individual cities or metropolitan areas, they are predominantly of operational character, and in this respect they differ from the models discussed so far. This feature bears on the theoretical content of the functional models. One should agree with R. Colenutt (1970, p. 120) that "the more a model is committed to theoretical concepts and simulation of behaviour of decision makers in the urban system the more difficult it is to operationalise and to use for forecasting"; while, as he notes after H. M. Blalock, "it would be misleading to confuse causal notions with those of prediction". However, it is rather easy to prove that models of functional patterns are based upon a specific understanding of urban spatial structure, which cannot be fully accomodated within any other group of concepts under discussion.

Models of intra-urban functional patterns describe spatial relations between individual spheres of man's activity from the point of view of location of individuals and firms, or the differentiation and succession in land use patterns. Sometimes both approaches are used concomitantly. A "typical" model (see L. J. King 1972, p. 4) contains the so-called exogenous variables, such as size and distribution of basic employment, which serve to generate the total employment figure, the economic structure and the population total of a city, or a metropolitan area. These endogenous variables are distributed over space using particular allocation functions. The structure of land use within individual subareas is derived by the use of land-requirement functions for the different allocated variables.

The allocation functions are based upon accessibility formula the notion of spatial accessibility being here of no less importance than in the case of urban land market theories and the population density concepts. While in the latter approaches this notion pertains to relations between all the structural units of a city, in models of functional patterns it refers only to those elements which are, or may become, functionally interconnected; in other words, it refers to potential interactions. The role of spatial accessibility concepts stems also from the fact that models of functional patterns originated from metropolitan transportation models, and are often called the second generation transportation models. In metropolitan transportation models, for example in the well-known CATS study, the land use pattern was assumed as given and as a basis for the estimation of spatial variations in the demand for transportation. In contrast, models of functional patterns assume the location of certain land use elements, but the derived variations in predicted trip patterns serve to generate the remaining components of land use[7]. The accessibility formulae used are, generally speaking, common to both types of models, as they originate from the classical formulations by W. G. Hansen (1959) and M. Schneider (1959). To these one should add an early potential model by M. O. Chauke (1960), briefly described in P. Korcelli (1973).

It would not be purposeful to describe individual models of intra-urban functional patterns, as they are extremely well covered in the literature (see, for example, Lowry 1968). One may only note that those models range from quite pragmatic constructs, such as the EMPIRIC model, to very intricate formulations like the Herbert–Stevens model and its later developments. Most of the recent work stems from the Lowry (1964) model. Lowry was the first to combine the concepts of spatial accessibility and urban economic structure. Instead, it seems worthwhile to turn now to the question of the cross-cultural validity of the models, or, more exactly, of the principal theoretical assumptions which they apply. Strikingly enough, already a cursory examination of the problem allows one to state that the theoretical assumptions in question — those which relate to principles of human interaction in space and the mechanisms of urban economic base — are generally the same throughout the world, if a given level of economic development is referred to. Those approaches have received considerable attention both in Western, and in Soviet, as well as Polish geographical and planning literature (see for example: Chojnicki 1966, Dziewoński 1967).

Thus, the tendency towards minimizing the distance (measured in time, cost, or trouble of overcoming it) between places of residence and places of work, shopping, and recreation, as depicted by the gravity, potential, and intervening-opportunity models, is found theoretically sound and is empirically supported on a universal scale. In addition, this tendency is interpreted both as a spontaneous phenomenon and as a planning objective. It was defined by G. A. Golc (1972) as a rule of spatial statistical self-organization of population distribution in relation to the patterns of costs of overcoming the distance. In an analogous manner, a specific balance between the exogenous and endogenous activities is treated in terms of both a statistical rule (with strong causal connotations) and a planning goal.

It is therefore possible to conclude that models of intra-urban functional patterns represent a research approach of a very high level of universality, although most of the actual studies which have been carried out pertain to Western cities (among them British cities are certainly overrepresented). However, some attempts to apply trip distribution models and spatial allocation models of the Lowry-type to an analysis of socialist cities have been made, and more are in progress. T. Zipser (1973) for example developed a version of the intervening opportunity model to predict changes in the spatial structure of several Polish cities, including Wrocław, Cracow, and Gliwice. He argues that values of the selectivity parameter in the model are of a universal character and are constant for particular types of trips, independent of the area under study.

---

[7] The factor of accessibility was also extensively treated in the so-called areal association studies — association of individual structural components of the city, or, interdependences between the rate of the population growth and particular elements and features in intra-urban structure. These studies represent another theoretical approach, related genetically to the models of intra-urban functional patterns. See, for example: W. B. Hansen (1961), E. N. Thomas (1960), F. S. Chapin, Jr. and S. F. Weiss (1962).

During recent years the concepts of intra-metropolitan functional patterns have been considerably developed, and, at the same time, have received a good deal of criticism. These theoretical developments are mainly connected with the contributions by A. G. Wilson (1969B, 1970), whose aim has been to construct dynamic models and to consider the relations from the supply, rather than the demand side, which would facilitate a more direct application of models to urban and regional planning. The development and disaggregation of spatial allocation models is performed by Wilson on the basis of his entropy-maximizing approach to model building.

Two general characteristics of the allocation models have been subject to numerous critical evaluations. One of the characteristics are methods of measuring spatial accessibility and spatial interaction, the other is the way the temporal dimension is usually handled. For example, M. A. Stegman (1969) strongly questions the concept of locational choice exclusively based on the spatial accessibility factor. To support his argument he submits data which attest dependencies of intra-metropolitan migrations on variations in the quality of the urban environment, the social differentiation of urban communities, and the structure of housing markets, rather than on location of workplaces, shopping centres or other focal points of social interaction. It should be admitted that similar data are supplied by many ecological studies. On the other hand, analysis of existing distributions consistently shows a very high regularity in spatial relations between patterns of residence, work, shopping, recreation etc. So, even if the statistical interdependence may not apply to individual behaviour, it is of a critical importance on an aggregate level. It has also been noted earlier that when transportation costs decline accessibility tends to yield to other location factors, such as amenity factors. A gradual diminution of its rate is portrayed by M. M. Webber (1964) and, in another context, by L. P. Bucklin (1971) who indicated that it may be deceptive to estimate the degree of interaction on the basis of distributions and *vice versa,* espacially when relatively short distances are involved.

As far as the problem of time dimension is concerned, it is frequently noted that even the so-called quasi-dynamic models do not take account of changes in the nature of interactions, the evolution of the attractiveness criteria and the development of transportation means. It is possible that studies on individual activity patterns, both spatial and temporal, such as those advanced by F. S. Chapin (1968) and T. Hägerstrand (1970) would produce new inputs for the allocation models meeting the requirements of higher theoretical, rather than technical, sophistication.

(VI) SETTLEMENT NETWORKS THEORIES

At least three different theoretical approaches originally developed on the scale of regional or inter-regional settlement networks have been subsequently extended to an interpretation of intra-urban growth and structure. These include: the central place theory, the concepts of linear settlement patterns, and the theory of metropolitan area. All three approaches have been mainly concerned with the principles of spatial distribution of functions and with the hierarchical organization of space.

The possibility of transferring the central place theory onto an intra-urban scale was suggested in the late 1950s by W. L. Garrison and B. J. L. Berry, while the first explicit, detailed analysis along this line was the study by H. Carol (1960) who found three hierarchical levels in the structure of the retail trade in Zurich and identified those levels with three out of the sevel orders in Christaller's hierarchy. Later, B. J. L. Berry (1963) proposed a comprehensive classification of shopping centres within a metropolitan area connecting the morphological and functional approaches. In addition to five hierarchical levels of centres he described two types of non-hierarchical retail clusters, namely commercial ribbons and specialized clusters. Berry found the following explanations of the hierarchical arrangement of business within a city: (1) different commercial functions have different conditions of entry (threshold) and thus demand minimum trade areas measured in size of population for their support, (2) goods and services differ in terms of the demand they generate and in the terms of frequency with which they are bought. In consequence, high threshold functions that meet sporadic needs, or are demanded by a small portion of the population, are found in higher level nucleation serving larger trade areas. These centres also perform lower order functions (as various kinds of goods may be purchased on one shopping trip) and for their lower order goods serve larger areas (population) than centres which are exclusively of that lower level. Thus, according to Berry the hierarchical structure of business centres within cities is generated by the same factors which are responsible for the emergence of central place systems on the regional and the inter-regional scale. Differences between intra-urban and regional patterns are explained by differences in population density; at higher densities the number of the population served by a centre of a particular order is bigger — the phenomenon for which Berry (1967) coined the term phase-shift.

It can easily be seen that the mechanisms of the central place system, as described by Berry, disregard commercial ribbons and specialized retail clusters. For the explanation of the full commercial pattern more realistic assumptions would be required, particularly those relating to spatial accessibility, organization and ownership of trade and services, and the behaviour of different types of shoppers. A tentative model accounting for such disaggregation levels was built by L. Curry (1962) but it was fitted to hypothetical patterns only.

In the studies quoted, the most fundamental assumption of the central place theory — the principle of selecting the nearest centre — was also transfered onto the intra-urban scale. This principle, however, was more recently questioned by G. Rushton, R. G. Golledge and W. A. V. Clark (1967) who instead proposed the principle of selecting the most attractive centre. The measures of attractiveness, from the point of view of individual shoppers, included, in addition to distance, also the size of the centre, as well as personal factors, which explained the existence of indifference zones.

Those theoretical developments have hardly brought closer the possibility of predicting the actual patterns of business, and their change, a fact noted by R. Colenutt (1970) and others. This particularly applies to dynamic analysis. It is a rather symptomatic fact that B. J. L. Berry's early optimism as regards the validity of the central place theory on an intra-urban scale was substantially contracted in his later work. While in 1960 he treated the evolution of spatial patterns of business and services as an outcome of a general concentration process occurring in the central place systems, in 1967 he emphasized the differences between the traditional central place patterns and the contemporary intra-metropolitan patterns, the latter characterized by a functional specialization of centres which are satisfied with attracting a fraction of the total demand generated by the common service area.

It follows from the above that when incomes increase and transportation costs fall, the explanatory power of the central place theory gradually diminishes, at least on an intra-urban scale. Nevertheless, some of its elements retain validity. They include the non-hierarchical Lösch (1944) model of market areas for individual goods, the areas which according to G. Olsson (1967) may also be treated as changeable over time. Another argument in favour of the theory is that the territorial expansion of urban areas and the decline in overall population densities may counterbalance the decrease in transportation costs, hence the role of physical distance in hindering spatial mobility even on the intra-metropolitan scale. With the trend towards an increase in size of firms, units and centres, this may lead to the emergence of a new, relatively regular pattern of business centres, although lacking hierarchy and differing in spatial scale from the traditional patterns.

In contrast to the central place models which apply to one of the components of urban spatial structure, i.e., to commercial and service patterns, models of linear patterns are usually intended to be comprehensive in scope. Their origin is sometimes traced to Lösch (1944)[8] who postulated that central place forces produce what is called city-rich and city-poor sectors. W. Isard and T. Reiner (1962) described the mechanism of expansion within such a pattern. In an early phase growth of the main centre occurs in those portions of its peripheral zone which offer the advantages of low transportation costs and scale economies. Such areas, of course, happen to be located in the city-rich sectors. With a further expansion coming from the centre, consecutive rings of former central places are engulfed by the emerging metropolis. In city-poor sectors growth is relatively retarded; thus the shape of the metropolitan area tends to repeat the original pattern of settlement within the region. A sectoral differentiation of settlement patterns on the regional scale is well illustrated by D. J. Bogue (1949) who analyzed the distribution of population and economic activity within 67 metropolitan regions, each divided into twelve sectors, classified as: (1) inter-metropolitan, (2) subdominant, and (3) local. Bogue proved that subdominant sectors were characterized by higher densities and greater economic potential than the remaining types of sectors; subdominant sectors may be therefore identified with city-rich sectors.

The sectoral pattern of economic landscape, described by Lösch and by Bogue have been interpreted within a framework of the central place theory as well as the theory of metropolitan dominance. On the other hand, the so-called corridors concept was developed by C. F. J. Whebell (1969) as an alternative to earlier theories of settlement networks. Whebell builds his theory upon postulates relating to the differentiation of natural conditions, the linear spread of innovations and the inertia of fixed assets and the socio-economic structure. From those postulates he derives the principle of the linear growth of settlement in each of the following development stages: primary occupance, commercial agriculture, rail transportation, and the stage of metropolization. However, it is not difficult to demonstrate that the initial postulates could produce different spatial patterns depending on prevailing economic and technological conditions; furthermore, it seems that, assuming present conditions, they are not likely to generate a linear structure of settlements. The differentiation of natural conditions may result in irregular patterns of occupancy while the factor of inertia may be conducive to a spatial concentration of human activity. Finally, diffusion of innovations often assumes hierarchical and non-contiguous forms, as a consequence of modern transportation and communication systems.

[8] Normative models of linear settlement patterns date already from the 19th century.

The corridors concept was expanded and modified by B. Malisz (1971) (in fact, his early formulations date from the 1950s) who used it as a framework for his graphic model of the settlement structure of Poland, called the band-nodal model. Malisz analyzed the role of "bundles of communication lines", locational trends in individual sectors of the economy, as well as objectives relating to the quality of life and socio-economic development. He also assumed the pre-existence of central places of three hierarchical levels which determine the subsequent location of corridors. The band-nodal model, according to its author, reflects the mechanism of urban and regional growth and may be used for projections, but it also meets basic socio-economic goals.

Description and evaluation of spatial patterns cannot, and probably should not be completely separated one from another. However, the concern of this paper is the question to what extent the theoretical concepts of linear settlement structure explain existing spatial patterns on both the regional and intra-urban scales. If we assume, as it is implied by the theory, that corridors, or bands, are becoming progressively urbanized in the course of their development and that the nodes, situated at the ends of the corridors, or on their crossings, are gradually expanding, then we cannot predict the future proportions between the linear and concentric elements, and cannot therefore anticipate the viability of a "readable" linear pattern. It should be borne in mind that systems of technical infrastructure, especially in densely populated areas, have to expand in various directions (not particularly along the established corridors), following complex human interaction patterns. In addition, corridors themselves are loosing their original, rigid forms with increasing demand for land claimed by all kinds of economic activity, evolution of time budgets and the growing role of recreation. Taking this into account one could hypothesize that the viability of a linear settlement pattern depends on the presence of physical barriers or zones of high attractiveness, while it can not be justified by the inertia factors alone.

The last of the three groups of concepts discussed in this section, namely, the theory of metropolitan area, can be traced back to the theory of metropolitan dominance (Blanchard 1911, Gras 1922) and the theory of urban regions (Geddes 1915). The functional approach of the former and the morphological approach of the latter concept were integrated in the theory of methopolitan community (McKenzie 1933) and in the related theory of metropolitan region (Bogue 1949, Duncan *et al.* 1960).

R. Blanchard and N. S. B. Gras described the metropolis from the viewpoint of its ties with its surrounding territories, while P. Geddes investigated the spatial expansion of urban areas. He noted that urban growth results not only in economic dominance of smaller towns by metropolitan centres, but also in the physical integration of large urban areas and the formation of conurbations. It was McKenzie's contribution to demonstrate the interrelations between the growth of a city as an economic organism and its territorial expansion and internal rearrangement of functions, leading to the state in which a metropolis becomes identified with the former hinterland of its principal centre and the population occupying the whole territory becomes functionally integrated to form one metropolitan community. According to McKenzie, the rapid development of transportation during the first decades of the 20th century and the resulting growth of spatial mobility of people, followed by the mobility of institutions, were the principal factors of metropolitan development. Similar interpretations prevail in the contemporary literature. For example, G. M. Lappo (1969) proposed a verbal model of the sequence of urban growth starting from a concentrated pattern, typical to the railway era, and ending up in a polycentric city-region pattern. Here again the dominant means of transportation determine to a large extent the major features and parameters of the urban structure.

Several authors observed that monocentric and polycentric agglomerations evolve towards analogous spatial and functional forms. Thus, D. J. Bogorad (1968) distinguished two routes for the development of an urban agglomeration. In the first case a single large city, in the course of its expansion, encroaches upon surrounding settlements which become transformed in terms of both functions and physiognomy. New settlements, linked with the main city, are also established, and the rural-urban fringe continuously expands. In the second case the development of an industrial settlement complex proceeds via concentration and a parallel growth of its constituent parts. Individual mining and industrial towns gradually coalesce to form larger cities of dispersed, and later on, a more concentrated structure. An urban cluster is converted into a continuous urbanized territory, and, subsequently, into an urban agglomeration with a well-developed core area and morphologically diversified outer zones. K. Dziewoński (1973) adds another stage to the growth sequence, namely the stage of modernization and change of the economic base of a conurbation. Its expanding potential attracts new functions, particularly tertiary and quaternary activities which replace the original functions, eliminated by new technological requirements and/or the exhaustion of local resources.

Coming from a regional to an interregional scale one finds two concepts falling, respectively, into the morphological and the functional traditions mentioned at the beginning of the present section. Reference is made to the megalopolis concept (Gottmann 1961; see also Leszczycki, Eberhardt and Herman, 1971, in the case of Poland) and the urban field concept (Friedmann and Miller, 1965). Although based upon the same principles of interplay between concentration on the interregional and deconcentration on the regional (or urban) scale, they produce somewhat contrasting spatial patterns of settlement. It seems, however, that their ranges of applicability somewhat differ. Thus, the megalopolis concept may be relevant in case of a territory with a high overall population density and a historically developed division of functions among the major cities. On the other hand, the urban field concept relates to areas with relatively low average densities and a relatively regular network of settlement in the initial stages of development. Similarly, a high differentiation of environmental conditions and the existence of physical barriers to movement or high attractiveness zones are prerequisits in the concept of linear settlement forms.

Another conclusion which can be reached is that the basic concepts and notions stemming from the settlement network theories, and transposed onto the intra-metropolitan scale, show a cross-cultural range of applicability, a fact which is reflected in the literature quoted above.

(VII) MODELS OF SPATIAL DIFFUSION ON THE INTRA-URBAN SCALE

Many of the models and concepts discussed so far in the paper emphasized spatial patterns rather than the processes responsible for their formation. Models of spatial diffusion pertain first of all to the process itself while treating existing or simulated patterns as possible outcomes of the development process. In turn, spatial diffusion is recognized as a basic process shaping locational patterns, consisting in a spread of phenomena on a given territory over time (Brown 1968B).

Models of spatial diffusion can be traced back to the concepts of the spread of cultures and plants, the sequent occupancy, and the movement of colonization frontiers. Later, they were usually identified with innovation diffusion models (Hagerstrand 1952). From there, via models of migrations (Hagerstrand 1957) and of the development of local settlement networks (Morrill 1965A) stem those approaches which treat urban spatial growth and the change of urban structure as specific types of diffusion processes. Despite dissimilarities between the nature of innovation diffusion on the one hand, and urban growth on the other, the two kinds of models share many common notions. These include the notion of the so-called neighbourhood effect, the notion of the resistance of the environment, and, first of all, the basic concepts of mean information field and physical barriers to diffusion. L. A. Brown (1968A) pointed out market organization and the structure of contact networks as factors largely determining the course of both innovation diffusion and the change in settlement networks. R. L. Morril (1965A) in his early empirical study on the province of Smäland attempted to demonstrate that the existing settlement system was a product of the complex historical process which, evolving irregularly over time and over space, took place in physically restricted and differentiated environments, whose mechanism included random factors, representing locational decisions based on criteria other than economic-man criteria, insufficient information, birth and decay, and technological change. In the case of Smäland those factors explained an apparent irregularity in the spatial and hierarchical arrangement of places on a regional level, but they could also account for differences between the observed patterns on intra-urban scale and those postulated by economic and ecological theories of the city.

To place properly the process of urban expansion in relation to other diffusion processes it is instructive to quote M. J. Beckmann (1970, p. 109) who classified objects of diffusion into states (e.g. information) and physical objects. While changes in state may be reversible or irreversible, physical objects are subject to a law of conservation, or continuity. According to this criteria technical innovations and urban growth belong to the same category of diffusion phenomena. A different classification was proposed by L. A. Brown (1968B) who identified a relocation-type and an expansion-type of spatial diffusion. In the case of relocation a phenomenon or elements of population actually move over space between t and $t+1$. Expansion, on the other hand, consists in a growth of population over the given time span, and the spread of the territory occupied. An interpretation of the expansion — type of diffusion may not be unequivocal if it is associated not only with states but also with physical objects, as Brown himselfs suggests, since, in the latter case a real transfer over space is taking place. The problem can probably be solved using the principle of the hierarchy of spatial systems. Thus, territorial expansion of a city would possess characteristics of both types of diffusion and an increase of the population number through immigration could be interpreted as an inflow from places situated outside the local system.

A specific regularity in the change of intensity forms one of the most distinctive features of diffusion processes and it supports hypotheses according to which which some fundamental features are common to all types of diffusion. Thus, the innovation diffusion cycle, described by T. Hägerstrand (1952) consisted of the following stages: (1) the primary stage, when new centres of innovation acceptance appeared, (2) the proper diffusion stage, when the existing centres became consolidated and the innovation began to spread out from the centres into the surrounding territory, (3) the condensing stage, when the phenomenon in question was commonly known and the number of new accepting units diminished. The condensing stage was followed by the saturation stage (4) when diffusion reached a certain ceiling and could not proceed further under the existing conditions.

A rural settlement diffusion cycle, presented by J. C. Hudson (1969) in the plant-ecological jargon possessed similar characteristics. In both cases an S-shaped curve represents a cumulative increase of population, while a closed unimodal curve describes the change of intensity of the process over time. Its shape, in fact, prompted T. Hagerstrand to introduce the notion of innovation waves, for which R. L. Morrill (1968, 1970) substituted later the more general term of diffusion waves. Morrill observed, however, that only certain diffusion processes have wave-like properties, as some (e.g. immigration into a region) may consist in a gradual intensification of a relatively constant pattern over time. On the other hand, urban spatial growth was found to possess many characteristics of a wave-like process (Korcelli 1970, 1972; Korcelli and Kostrubiec 1973).

Among empirical models which apply notions and techniques typical of the spatial diffusion concepts, two different types prevail. Type one are models predicting the peripheral expansion of urban areas (see for example R. L. Morrill, 1965C), or the extention of the urban "frontier"; they can be called colonization models. Type two are models portraying the change in one of the components of structure; they in turn can be labelled succession models. Examples are models of expansion of Negro ghettos (see for example, Ch. R. Hansell and W. A. V. Clark 1970; R. L. Morrill 1965B). Both types represent partial models and they hardly meet the theoretical standards put forth by W. Garrison (1960) some fifteen years ago. According to Garrison a comprehensive simultation model of urban growth should be composed of a number of submodels accounting for individual components of the urban structure. He distinguished three component categories. The first category are grid elements governed by the principle of proximity, or contiguity. The second group are "system-wide" elements (mostly networks, but also point lattices, e.g., local government centres) with operating characteristics of their own. Interrelations between grid elements and system-wide elements are of a functional nature (as in the case of trip patterns) and they also pertain to the capacity of the networks. Finally, the third category are external factor for example, external migrations, national and regional rates of economic growth, and intersectoral relations among metropolitan areas.

It can easily be seen that only some of the above research requirements are fulfilled by the existing empirical models. Totally, however, the models cast much light on the mechanisms of intra-urban systems and allow the tracing of their actual behaviour within the theoretical framework outlined by W. Garrison. Other problems related to the application of spatial diffusion models on intra-urban scale include: the measurement of the probability field and its change over time; the evaluation of simulated patterns, and, finally, the determination of the range of applicability of the models in different socio-economic and cultural contexts. The last question will be briefly considered below.

The degree of universality of a simultation model depends largely upon the type of phenomenon studied as well as on the spatial and thematic disaggregation level of input. For example, models of the expansion of Negro ghettos are more restricted in their applicability range than models of urban fringe development. All the studies quoted [9] pertain to contemporary Western, mostly Anglo-American cities, and use assumptions reflecting specific features of those cities, particularly assumptions relating to the operation of the land market and the housing market. Generally, probabilistic simulation models should be most effective in the case of uncontrolled processes involving great numbers of units. Nevertheless, the decision making (in respect to locational selection) on the scale of a large city is an extremely intricate process irrespective of the level of decision centralization and the relationships between plan making and plan implementation. It is therefore possible to think of random variables as components in explanatory models of the development of the contemporary city in general.

(VIII) BASIC THEORETICAL STATEMENTS AND POSTULATES

It has been demonstrated on the preceding pages that the six major theoretical approaches in the field of urban growth and urban structure, although developing rather independently of each other, share a number of common notions and may be regarded as complementary rather than substitutive. It should therefore be possible to arrange the basic findings about the structure and growth of the city in such a manner that contributions of different approaches are taken account of, and much of the overlap is avoided. Such a tentative systematization is presented below. Its logic is a sequence from the description of static patterns to the indication of trends, and the outlining of development cycles.

(1) The intensity of the occupancy within an urban area is a negative function of distance from its centre.

(1.1) The location of individual land users: households, firms, institutions (their distance from the city centre) is determined by the relative importance of the factors of accessibility and space occupied, or, the shape and steepness of locational bid rent curves.

(1.2) With growing distance from the city centre the space occupied by individual land users increases as a result of transportation inputs and land inputs being complementary.

(1.3) The land use pattern assumes the form of specialized concentric zones.

(1.4) Population density is a negative exponential function of distance from the city centre.

(1.5.) Density gradient value is related to the position of a city in the system of cities, and it shows a decrease with the increasing city size.

(2) The intra-urban pattern is composed of concentric and sectoral elements.

(2.1) Differentiation of socio-economic characteristics is sectoral in form.

(2.2) Differentiation of socio-demographic characteristics is concentric in form.

(2.3) Intensity of spatial differentiation is a function of the socio-political. economic, and cultural system.

(3) The spatial functional structure of an urban area is composed of a set of patterns, superimposed one upon another, and related to the main spheres of human activity.

---

[9] Recently K. Dramowicz (1975) applied a simulation technique to predict changes in the settlement pattern within a section of a rural-urban fringe of the city of Płock, in north-central Poland.

(3.1) Functional patterns consist of sets of trip origins and trip destinations — of periodical (mostly daily) and non-periodical trips.

(3.2) Intensity of interaction between elements of different functional patterns is a negative function of distance separating those elements.

(3.3) Spatial relations between trip origins and trip destinations are specific to each type of activity.

(4) The spatial structure of an urban area is built of hierarchical and specialized, as well as non-hierarchical (and non-specialized) elements.

(4.1) Hierarchical and specialized elements assume punctual and linear forms; they build interconnected patterns on the scale of an urban area.

(4.2) The non-hierarchical (and non-specialized) elements assume grid patterns.

(4.3) Urban spatial form is made up of concentric and sectoral elements.

(5) The urban structure is subject to a gradual deconcentration proceeding with economic and technological development.

(5.1) The curve representing aggregate transportation costs in relation to distance from the city centre gradually levels-off.

(5.2) The role of amenities as locational factors increases.

(5.3) Marginal zones take over the functions originally concentrated within the inner zones of an urban area.

(5.4) The population density gradient and the maximum density value gradually decline.

(5.5) The maximum density zone moves outwards from the city centre.

(5.6) The location of groups with highest mobility (or, least constrained in the choice of location) shifts from the centre to marginal zones.

(5.7) Functional patterns related to the main spheres of human activity are subject to a gradual differentiation and dissociation in space.

(5.8) Hierarchical patterns are superseded by patterns characterized by internal specialization of functions.

(6) The process of urban spatial growth is composed of alternate stages of concentration and deconcentration.

(6.1) Technological progress in transportation produces cyclical change in the spatial accessibility gradient.

(6.2) At an advanced stage of urban deconcentration a new increase of density in central zones tends to occur.

(6.3) At an advanced stage of deconcentration and redevelopment of an urban area the groups with least constrained locational choice tend to move towards the city centre again.

(7) The process of urban growth and the change of its spatial scale involves alternate stages of expansion and consolidation.

(7.1) Urban growth leads to the integration and identification of a metropolis with its service area.

(7.2) The spatial structure of monocentric and polycentric urban areas tend to converge.

(8) The process of urban spatial growth is composed of a set of spatial diffusion-succession cycles, including:

(a) — peripheral growth of an urban area (colonization-type of diffusion),

(b) — succession of urban landscape forms (or, sequent occupancy),

(c) — land use succession,

(d) — redevelopment cycle,

(e) — succession of housing occupancy (the filtering-down process),

(f) — succession of socio-economic groups,

(g) — life-cycle succession,

(h) — succession of ethnic groups.

(8.1) Spatial diffusion-succession cycles proceed from the city centre towards marginal zones.

(8.2) Spatial diffusion-succession cycles generally consist in the superseding of less intensive forms of occupancy by more intensive forms.

(8.3) The succession cycles differ in pace and momentum; the succession of functions precedes the succession of physical forms.

(8.4) The succession cycles are mutually interlinked and follow each other.

(8.5) The pace of the succession cycle may vary over time as a result of external factors (migrations, technological change, economic development, etc.).

(8.6) The succession cycles consist of several distinctive phases. The change of their intensity over time is described by a closed unimodal curve (or, a wave-like form), while the change in the degree of their accomplishment is represented by an S-shaped curve.

It should be admitted that the statements listed above very in respect of the generalization level, the fact which is explained by an irregular and uneven development of the theory. While some of the statements, such as the population density rule or the principle of the coexistence of hierarchical and specialized components, are universal on the scale of the contemporary city, others, for example those relating to the succession of ethnic groups, or the life-cycle succession, are more restricted in their temporal and spatial validity range. Since much of the theory of intra-urban structure has been developed in the United States, it explains the growth and structure of the American city to a much larger extent than the structure of other cities, particularly non-Western cities. This is also reflected by general postulates, formulated either explicitely, or implied, from which individual statements follow (see B. Garner 1967; Rogers 1967). The basic traditional postulates are:

(1) Locational decisions are taken so as to minimize cost or maximize profit, or, to maximize utility. More specifically, individuals, households, firms, and institutions aim to:

(a) maximize accessibility within daily activity space,

(b) maximize the area occupied,

(c) maximize social, cultural, natural, esthetic, and salubrious values of the environment.

It was found more recently that the utility of a given site changes as a household moves through subsequent stages of the life cycle. In addition, other postulates have been proposed, partly opposing the economic-man assumptions. They include the following rules:

(2) Locational decisions and the resulting patterns are a function of historical and cultural values (tradition, symbolism etc.)

(3) Locational decisions include a random factor, representing lack of information, inertia, differences in preferences, etc.

It is apparent that all the above postulates may have different validity in different socio-economic and cultural systems. However, the principal, albeit interrelated issues in the cross-cultural perspective are:

(1) Whether the utility to be maximized is (a) individual, or (b) collective utility.

(2) Whether the process of allocation in space is to be (a) competitive, the aim of the competition being the position of dominance, or (b) governed by the principle of mutual cooperation in the pursuit of common social goals.

(IX) CONCLUSIONS

There are two questions that need to be introduced in the closing remarks. First, there is a need to venture a concise, concluding statement on cross-cultural validity of the theoretical approaches discussed. Second, it is conventional to speculate briefly on research prospects and research priorities.

It seems that theories and concepts which represent first of all what may be called the aggregate approach to the study of urban structure and growth, are largely universal in scope. This category includes settlement network theories, population density concepts, and the concepts of functional patterns. On the other hand, theories which emphasize the disaggregation of the metropolitan community into status, income, or ethnic groups, and analyse individual decision making processes, are usually specific to different socio-economic and cultural systems. To the latter group belong in general the ecological concepts, the urban land market theories and, to some extent, the diffusion concepts of urban growth.

As far as research frontiers are concerned, three major developments, already conspicuous, are likely to grow further in importance. First, there is a trend towards an expansion of the functional patterns concepts, mostly in the direction of activity systems and contact networks analysis. This is hardly accidental since the second major development is the growing emphasis on relevancy of research and its applicability to urban and regional planning. Finally, one may perceive a growing share of theoretical studies pertaining to urban structure in the developing countries and in the socialist countries.

Institute of Geography
Polish Academy of Sciences, Warsaw

BIBLIOGRAPHY

Abu-Lughod, J., 1969, Testing the theory of social area analysis: the ecology of Cairo, Egypt, *Amer. Soc. Rev.*, 34, pp. 198–212.

Ajo, R., 1965, On the structure of population density in London's field, *Acta Geogr.*, 18, 4.

Alonso, W., 1964, *Location and land use. Toward a general theory of land rent*, Harvard University Press, Cambridge, Mass.

Amato, P. W., 1970, A comparison: population densities, land values, and socioeconomic class in four Latin American cities, *Land Econ.*, 46, pp. 447–455.

Angel, S., Hyman, G. M., 1972, Urban transport expenditures, *Papers, Reg. Sci. Ass.* 29, pp. 105–123.

Beckman, M. J., 1970, The analysis of spatial diffusion processes, *Papers, Reg. Sci. Ass.*, 25, pp. 109–118.

Berry, B. J. L., 1963, *Commercial structure and commercial blight*, University of Chicago, Department of Geography, Research Paper 85.

Berry, B. J. L., 1964, Cities as systems within systems of cities, *Papers, Reg. Sci. Ass.*, 13, pp. 147–164.

Berry, B. J. L., 1967, *Geography of market centers and retail distribution*. Foundations of Economic Geography Series, Prentice-Hall, Englewood Cliffs N. J.

Berry, B. J. L., 1971, The logic and limitations of comparative factorial ecology, *Econ. Geogr.*, 47, 2 (Suppl.), pp. 209–219.

Berry, B. J. L., Horton, F. E., eds., 1970, *Geographic perspectives on urban systems*, Prentice-Hall, Englewood Cliffs N. J.

Berry, B. J. L., Simmons, J. W., Tennant, R. J., 1963, Urban population densities: structure and change, *Geogr. Rev.*, 53, pp. 389–405.

Birch, D. L., 1971, Toward a stage theory of urban growth, *J. Amer. Inst. Planners*, 37, pp. 78–87.

Blanchard, R., 1911, *Grenoble: étude de géographie urbaine*, Grenoble.

Bleicher, H., 1892, *Statistische Beschreibung der Stadt Frankfurt am Main und ihrer Bevölkerung*, Frankfurt am Main.

Bogorad, D. J., 1968, Zadachi i metody regulirovanya rosta gorodskikh aglomeratsii, in: *Gradostroitelstvo. Rayonnaya planirovka; gorodskiye aglomeratsii*, Budivelnik, Kiev, pp. 10–19.

Bogue, D. J., 1949, *The structure of the metropolitan community. A study of dominance and subdominance*, Contributions of the Institute for Human Adjustment; Social Science Research Project. University of Michigan, Ann Arbor.

Bogue, D. J., 1950, *Metropolitan decentralization: A study of differential growth*, Scripps Foundation Series in Population Distribution, 2, Miami University.

Bourne, L. S., 1971, Physical adjustment processes and land use succession: a conceptual review and central city examples, *Econ. Geogr.*, 47, pp. 1–15.

Boyce, D. E., 1965, The effect of direction and length of person trips on urban travel patterns, *J. Reg. Sci.*, 6, pp. 65–80.

Brigham, E. F., 1965, The determinants of residential land values, *Land Econ.*, 41, pp. 325–334.

Bromek, K., 1964, *Rozwój demograficzny regionu Krakowa w okresie od 1869 do 1950* (Sum.: The development of the demographic region of Cracow), Zesz. Nauk. U. J., 85, Prace Geogr., 9, Prace Inst. Geogr. 31, Kraków.

Brown, K. M., 1968, A note on: Urban population density patterns: an alternative explanation, *Canad. Geogr.*, 12, pp. 203–205.

Brown, L. A., 1968A: *Diffusion dynamics. A review and revision of the quantitative theory of the spatial diffusion of innovation*, Lund Studies in Geography, Ser. B, Human Geography, 29, C. W. K. Gleerup, Lund.

Brown, L. A., 1968B: *Diffusion processes and location. A conceptual framework and bibliography*, Regional Science Research Institute, Bibliography Series Number Four.

Brush, J. E., 1968; Spatial patterns of population in Indian Cities, *Geogr. Rev.*, 58, pp. 362–391.

Bucklin, L. P., 1971, Retail gravity models and consumer choice: A theoretical and empirical critique, *Econ. Geogr.*, 47, pp. 489–497.

Carol, H., 1960, The hierarchy of central functions within a city, *Annals, Ass. Amer. Geogr.*, 50, pp. 211–237.

Casetti, E., 1967, Urban population density patterns: an alternative explanation, *Canad. Geogr.*, 11, pp. 96–100.

Casetti, E., 1969, Alternative urban population density models: an analytical comparison of their validity range, in: *Studies in Regional Science*, 1, A. J. Scott, ed., Pion, London.

Casetti, E., 1971, Equilibrium land values and population densities in an urban setting, *Econ. Geogr.*, 47, pp. 16–20.

Chapin, F. S., Jr., 1968, Activity systems as a source of inputs for land use models, in: *Urban development models*, Proceedings of a Conference held at Dartmouth College, New Hampshire (June 1967), Special Report 97, Highway Research Board, Washington D. C.

Chapin, F. S., Jr., Weiss, S. F., 1962, *Factors influencing land development*, Institute for Research in Social Sciences, University of North Carolina, Urban Studies Research Monograph.

Chapin F. S., Jr., Weiss, S. F., 1968, A probabilistic model for residential growth, *Transp. Research*, 2, pp. 375–390.

Chauke, M. O., 1960, K voprosu izucheniya zakonomernosti vnutrigorodskogo i prigorodskogo rasseleniya, *Probl. Sov. Gradostroit.*, 8.

Chojnicki, Z., 1966, *Zastosowanie modeli grawitacji i potencjału w badaniach przestrzenno-ekonomicznych* (Sum.: The application of gravity and potential models in spatial economic research), Studia KPZK PAN, 15, Warszawa.

Chotaj, M., 1966, *Cena ziemi w rachunku ekonomicznym* (Land price in economic calculation), Warszawa, PWE.

Clark, C., 1951, Urban population densities, *J. Roy. Statist. Soc.*, 64, Ser. A, pp. 490–496.

Colby, Ch., 1933, Centrifugal and centripetal forces in urban geography, *Annals, Ass. Amer. Geogr.*, 23, pp. 1–20.

Colean, M. L., 1953, *Renewing our cities*, The Twentieth Century Fund, New York.

Colenutt, R. J., 1970, Building models of urban growth and spatial structure, in: *Progress is Geography*, 2, Arnold, London, pp. 109–152.

Conzen, M. R. G., 1962, *The plan analysis of an English city centre*, Lund Studies in Geography, Ser. B, 24, Proceedings of the IGU Symposium in Urban Geography, Lund 1960, pp. 383–414.

Cressey, P. F., 1938, Population succession in Chicago, 1898–1930, *Amer. J. Soc.*, 44, pp. 59–69.

Curry, L., 1962, The geography of service centers within towns: the elements of an operational approach, in: *Proceedings of the IGU Symposium in urban geography*, Lund Series in Geography, Ser. B, Human Geography, 24, Lund, pp. 31–53.

Czamański, S., 1966, Effects of public investments on urban land values, *J. Amer. Inst. Planners*, 32, pp. 204–217.

Dacey, M. F., 1968, A model for the areal distribution of population in a city with multiple population centers, *Tijd. Econ. Soc. Geogr.*, 59, pp. 232–236.

Davies, W. K. D., 1968, The morphology of central places, a case study, *Annals, Ass. Amer. Geogr.*, 58, pp. 91–110.

Domański, R., 1967, Konstruowanie teorii w geografii ekonomicznej (Sum.: Construction of theories in economic geography), *Przegl. Geogr.*, 39, pp. 85–102.

Dramowicz, K., 1975, Symulacja cyfrowa i analiza systemowa w badaniach nad procesami urbanizacji wsi (Sum.: Computer simulation and systems approach in research on the urbanization processes of villages (model provided by Biała Stara commune, Płock district), Prace Geogr. IG PAN, 112, Warszawa.

Duncan, B., Sabagh, G., Van Arsdol, M. D., 1962, Patterns of city growth, *Amer. J. Soc.*, 67, pp. 418–429.

Duncan, O. D., 1959, Human ecology and population studies, in: P. M. Hauser, O. D. Duncan, eds., *The study of population*, University of Chicago Press.

Duncan, O. D., Duncan, B., 1957, *The Negro population of Chicago. A study of residential succession*, University of Chicago Press.

Duncan, O. D., Scott, W. R., Lieberson, S., Duncan, B., Winsborough, H. H., 1960, *Metropolis and region*, Resources for the Future, Johns Hopkins University Press, Baltimore.

Dziewoński, K., 1967, *Baza ekonomiczna i struktura funkcjonalna miast. Studium rozwoju pojęć, metod i ich zastosowań* (Sum.: Urban economic base and functional structure of cities), Prace Geogr., IG PAN, 63 (1st edition), 87 (2nd edition, 1971), Warszawa.

Dziewoński, K., 1973, The geographical differentiation of contemporary urbanization, *Geogr. Pol.*, 27, pp. 31–41.

Friedmann, J., Miller, J., 1965, The urban field, *J. Amer. Inst. Planners*, 31, pp. 312–319.

Garner, B. J., 1967, Models of urban geography and settlement location, in: R. J. Chorley, P. Haggett, eds., *Models in geography*, Methuen, London, pp. 303–360.

Garrison, W., 1962, Toward a simulation model of urban growth and development, in: *Proceedings of the IGU Symposion in Urban Geography*, Lund Studies in Geography, Ser. B., 24, Lund 1960, pp. 91–108.

Geddes, P., 1915, *Cities in evolution*, London.

Golc, G. A., 1972, Vliyanye transporta na prostranstvennoe razvitye gorodov i aglomeratsii, in: *Problemy sovremennoy urbanizatsii*, Statistika, Moskva, pp. 159–190.

Gottmann, J., 1961, *Megalopolis, The urbanized northeastern seaboard of the United States*, The Twentieth Century Fund, New York.

Gras, N. S. B., 1922, *An introduction to economic history*, Harper, New York.

Greer-Wootten, B., 1972, *Metropolitan regional analysis*, Progress in Geography, 4, Edward Arnold, London.

Gurevich, B. L., 1967, Plotnost naseleniya goroda i plotnost veroyatnosti sluchaynoy velichiny, *Vestn. Mosk. Univ., Ser. Geogr.*, 1, pp. 15–21.

Gurevich, B. L., Saushkin, J. G., 1966, Matematicheskii metod v geografii, *Vestn. Mosk. Univ., Ser. Geogr.*, 1, pp. 3–28.

Hägerstrand, T., 1952, *The propagation of innovation waves*, Lund Studies in Geography, Ser. B., Human Geography, 4.

Hägerstrand, T., 1957, Migration and area, in: D. Hannerberg, T. Hagerstarand, B., Odeving, eds., *Migration in Sweden: A Symposium*, Lund Studies in Geography, Ser. B, 13.

Hagerstrand, T., 1969, Methods and new techniques of urban and regional research, in: *Urban and regional research*, Economic Commission for Europe, 2, United Nations, New York, pp. 61–72.

Hagerstrand, T., 1970, What about people in regional science, *Papers, Reg. Sci. Ass.*, 24, pp. 7–21.

Haig, R. M., 1927, *Major economic factors in metropolitan growth and arrangement*. Regional Plan of New York and its Environs, 1, New York.

Hansell, Ch. R., Clark, W. A. V., 1970, The expansion of the Negro ghetto in Milwaukee: a description and simulation model, *Tijd. Econ. Soc. Geogr.*, 61, pp. 267–277.

Hansen, W. G., 1959, How accessibility shapes land use, *J. Amer. Inst. Planners*, 25, pp. 73–76.

Hansen, W. B., 1961, An approach to the analysis of metropolitan residential extension, *J. Reg. Sci.*, 3, pp. 37–55.

Harvey, D. W., 1970, Social processes and spatial form: an analysis of the conceptual problems of urban planning, *Papers, Reg. Sci. Ass.*, 25, pp. 47–70.

Hawley, A. H., 1950, *Human ecology. A theory of community structure*, Ronald Press Co., New York.

Hawley, A. H., 1956, *The changing shape of metropolitan America: deconcentration since 1920*, The Free Press, Glencoe, Ill.

Hoover, E. M., Vernon, R., 1959, *Anatomy of a metropolis. The changing distribution of people and jobs within the New York Metropolitan Region*, Harvard University Press, Cambridge, Mass.

Hudson, J. C., 1969, A location theory for rural settlement, *Annals, Ass. Amer. Geogr.*, 59, pp. 365–381.

Hurd, R. H., 1903, *Principles of city land values*, New York, The Record and Guide.

Isard, W., Reiner, T. A., 1962, Regional science and planning, *Papers, Reg. Sci. Ass.*, 8, pp. 1–36.

Janson, C. G., 1971, A preliminary report on Swedish urban spatial structure, *Econ. Geogr.*, 47, No. 2 (Suppl.), pp. 249–257.

Johnston, R. J., 1969, Some tests of a model of intra-urban population mobility: Melbourne, Australia, *Urban Studies*, 6, No. 1.

Johnston, R. J., 1972, *Towards a general model of intra-urban residential patterns: some cross-cultural observations*, Progress in Geography, 4, Edward Arnold, London.

King, L. J., 1972, Models of urban land-use development, in: *Models of urban structure*. D. C. Sweet, ed., Lexington, Mass., pp. 3–26.

Kolipiński, J., 1970, *Metody rachunku efektywności w miejscowym planowaniu przestrzennym w Polsce* (*Studium krytyczne*), (Sum.: Methods of the efficiency calculation in spatial planning — applied in Poland), Studia KPZK PAN, 32, PWN, Warszawa.

Korcelli, P., 1970, A wave-like model of metropolitan spatial growth, *Papers, Reg. Sci. Ass.*, 24, pp. 127–138.

Korcelli, P., 1972, Urban spatial growth: a wave-like approach, *Geogr. Pol.*, 24, pp. 45–55.

Korcelli, P., 1973, Urban growth: some models and generalizations, *Geogr. Pol.*, 27, pp. 133–141.

Korcelli, P., Kostrubiec, B., 1973, Harmonic analysis of urban spatial growth, *Geogr. Pol.*, 25, pp. 93–101.

Korcelli, P., 1974, *Teoria rozwoju struktury przestrzennej miast* (Sum.: Theory of intra-urban spatial structure), Studia KPZK PAN, 45, PWN, Waszawa.

Korzybski, S., 1952, Le peuplement des grandes agglomérations urbaines, Londres et Paris aux XIXe et XXe siècles, *Population*, 3, pp. 485–520.

Kostrubiec, B., 1970, Badania rozwoju przestrzennego aglomeracji miejskiej metodą profilów (Sum.: Spatial expansion of the urban agglomeration investigated by means of the method of profiles), *Przegl. Geogr.*, 42, pp. 235–248.

Lentz, P., 1973, *The influence of a socialist economic environment on intra-urban residential spatial structure*, *Clark Univ.*, Worcester, Mass. Ph.D. thesis (typescript).

Leszczycki, S., Eberhardt, P., Herman, S., 1971, Aglomeracje miejsko-przemysłowe w Polsce, 1966–2000, (Sum.: Urban-industrial agglomerations in Poland, 1966–2000), *Biuletyn KPZK PAN*, 67, Warszawa.

Leven, Ch. L., 1969, Determinants of the size and spatial form of urban areas, *Papers, Reg. Sci. Ass.*, 22, pp. 7–28.

Linday, S., 1971, The ecology of Islamic cities: the case for ethnocity, *Econ. Geogr.*, 47, No. 2 (Suppl.), pp. 303–313.

Lösch, A., 1944, *Die räumliche Ordnung der Wirtschaft*, Gotha.

Lowry, I. S., 1964, *A model of metropolis*, The Rand Corporation, Memorandum RM-4035-RC, Santa Monica, California.

Lowry, I. S., 1968, Seven models of urban development: a structural comparison, in: *Urban development models*, Proceedings of a conference held at Dartmouth College, New Hampshire (June 1967) Special Report 97, Highway Research Board, Washington D.C.

Lappo, G. M., 1969, *Geografia gorodov s osnovami gradostroitelstva*, Izd. Mosk. Univ., Moskva.

Malisz, B., 1971, Prognoza zmian sieci osadniczej w Polsce (A prognosis of changes in the settlement network of Poland) in: *Prognozy rozwoju sieci osadniczej*, Komitet Badań i Prognoz, Polska 2000 PAN, 2, Warszawa.

McElrath, D. C., 1962, The social areas of Rome: a comparative analysis, *Amer. Soc. Rev.*, 27, pp. 376–391.

McKenzie, R. D., 1933, The metropolitan community, McGraw-Hill, New York.

Medvedkov, Yu. V., 1965, *Ekonomicheskaya izuchennost rayonow kapitalisticheskogo mira*, Vyp. 2, Prilozhenya matematiki v ekonomicheskoy geografii, VINITI, Moskva.

Medvedkov, Yu. V., 1971, Internal structure of a city: an ecological assessment, *Papers, Reg. Sci. Ass.*, 27, pp. 95–117.

Morrill, R. L., 1965A, *Migration and the spread and growth of urban settlement*, Lund Studies in Geography, Ser. B, Human Geography, 26.

Morrill, R. L., 1965B, The Negro Ghetto: problems and alternatives, *Geogr. Rev.*, 55, pp. 339–361.

Morrill, R. L., 1965C, Expansion of the urban fringe: a simulation experiment, *Papers, Reg. Sci. Ass.*, 15, pp. 185–202.

Morrill, R. L., 1968, Waves of spatial diffusion, *J. Reg. Sci.*, 8, pp. 1–18.

Morill, R. L., 1970, The shape of diffusion in space and time, *Econ. Geogr.*, 46 (Suppl.), pp. 259–268.

Morris, F. B., Pyle, G. F., 1971, The social environment of Rio de Janeiro in 1960, *Econ. Geogr.*, 47, No. 2 (Suppl.), pp. 286–299.

Murdie, R. A., 1969, *Factorial ecology of metropolitan Toronto, 1951–1961.* The University of Chicago, Department of Geography, Research Paper 116, Chicago.

Muth, R. F., 1961, The spatial structure of the housing market, *Papers and Proceedings, Reg. Sci. Ass.*, 7, pp. 207–220.

Muth, R., 1969, *Cities and housing*, Chicago.

Newling, B. E., 1966, Urban growth and spatial structure: mathematical models and empirical evidence, *Geogr. Rev.*, 56, pp. 213–225.

Newling, B. E., 1969, The spatial variation of urban population densities, *Geogr. Rev.*, 59, pp. 242–252.

Niedecorn, J. H., 1971, A negative exponential model of urban land use densities and its implications for metropolitan development, *J. Reg. Sci. Ass.*, 11, pp. 317–326.

Olsson, G., 1967, Central place systems, spatial interaction, and stochastic processes, *Papers, Reg. Sci. Ass.*, 18, pp. 13–45.

Pahl, R. E., 1967, Sociological models in geography, in: R. J. Chorley, P. Haggett, eds., *Models in geography*, Methuen, London.

Papageorgiou, G. J., 1971, A theoretical evaluation of the existing population density gradient functions, Econ. Geogr., 47, pp. 21–26.

Papageorgiou, G. J., Casetti, E., 1971, Spatial equilibrium residential land values in a multicenter setting, *J. Reg. Sci.*, 11, pp. 385–389.

Pederson, P. O., 1967, *An empirical model of urban population structure. A factor analytical study of the population structure in Copenhagen.* Proceedings of the First Scandinavian-Polish Regional Science Seminar, Studia KPZK PAN, 17, pp. 193–216.

Pióro, Z., 1962, *Ekologia społeczna w urbanistyce* (Social ecology in urban planning), Arkady, Warszawa.

Pokshishevski, V. V., 1973, Urbanization and ethnographic processes, *Geogr. Pol.*, 27, pp. 79–85.

Pokshishevski, V. V., Gokhman, V. M., 1967, Problema giperurbanizacii v razvitykh kapitalisticheskikh stranakh i yego geograficheskiye aspekty, in: *Nauchnye problemy geografii naseleniya*, Izd. Mosk. Univ., Moskva, pp. 151–163.

Ratcliff, R. V., 1949, *Urban land economics*, McGraw-Hill Publishing Co., New York.

Rees, P. H., 1970, The factorial ecology of Metropolitan Chicago. Chapter 10, in: B. J. L. Berry, F. E., Horton, eds., *Geographic perspectives on urban systems,* Prentice-Hall, Englewood Cliffs, N. J.

Robson, B. T., 1969, *Urban analysis: a study of city structure with special reference to Sunderland,* Cambridge University Press, Cambridge.

Rogers, A., 1967, Theories of intra-urban spatial structure: a dissenting view, *Land Econ.,* 43, pp. 108–112.

Romashkin, V. J., 1967, Matematicheskaya model raspredeleniya plotnosti naseleniya Moskvy, *Vestn. Mosk. Univ., Ser. Geogr.,* 3, pp. 54–58.

Rushton, G., Colledge, R. C., Clark, W. A. V., 1967, Formulation and test of a normative model for the spatial allocation of grocery expenditures by a dispersed population, *Annals, Ass. Amer. Geogr.,* 57, pp. 389–400.

Salins, P. D., 1971, Household location patterns in American metropolitan areas, *Econ. Geogr.,* 47, No. 2 (Suppl.), pp. 234–248.

Schmitt, R. E., 1956, Suburbanization: statistical fallacy?, *Land Econ.,* 32, pp. 85–87.

Schneider, M., 1959, Gravity models and trip distribution theory, *Papers and Proceedings, Reg. Sci. Ass.,* 5, pp. 51–56.

Schnore, L. F., 1957, Metropolitan growth and decentralization, *Amer. J. Soc.,* 63, pp. 171–180.

Schnore, L. F., 1965, On the spatial structure of cities in the two Americas, in: P. H. Hauser, L. F. Schnore, eds. *The study of urbanization,* Wiley and Sons, New York, pp. 347–398.

Senior M. L., 1973, Approaches to residential location modelling; Urban ecological and spatial interaction models (A review), *Environment and Planning,* 5, pp. 165–197.

Sevostyanov, V. I., 1972, Plotnost naseleniya kak kharakteristika protsessa razvitiya goroda, *Vestn. Mosk. Univ.,* 2, *Ser. Geogr.,* pp. 40–46.

Sherratt, G. G., 1960.. A model for general urban growth, in: C. W. Churchman, A. Verhulst, eds., *Management sciences, models and techniques,* New York, Pergamon Press, pp. 147—159.

Simmons, J. Wm. 1968, Changing residence in the city: a review of intra-urban mobility, *Geogr. Rev.,* 58, pp. 622–651.

Stegman, M. A., 1969, Accessibility models and residential location, *J. Amer. Inst. Planners,* 35, pp. 22–29.

Stewart, J. Q., 1947, Empirical mathematical rules concerning the distribution and equilibrium of population, *Geogr. Rev.,* 37, pp. 461–548.

Tanabe, K. 1970, Intra-regional structure of Japanese cities, in: *Japanese cities: a geographical approach,* Special Publication 2, The Association of Japanese Geographers, Tokyo.

Thomas, E. N., 1960, Areal association between population growth and selected factors in the Chicago Urbanized Area, *Econ. Geogr.,* 36, pp. 158–170.

Timms, D. W. G., 1971, *The urban mosaic. Towards a theory of residential differentiation,* The Cambridge University Press, Cambridge 1970.

Ullman, E. R., 1962, The nature of cities reconsidered, *Papers and Proceedings, Reg. Sci. Ass.,* 9, pp. 7–23.

Vance, J. E.-Jr., 1969, Urban structure and the geography of choice, *Annals, Ass. Amer. Geogr.,* 59 (abstract of a paper prepared for 64th Annual Meeting AAG).

Vance, J. E., 1971, Land assignment in the precapitalist, capitalist, and postcapitalist city, *Econ. Geogr.,* 47, pp. 101–120.

Webber, M. M., 1964, The urban place and the non-place urban realm, in: *Explorations into urban structure,* University of Pennsylvania Press, Philadelphia.

Werwicki, A., 1973, *Struktura przestrzenna średnich miast ośrodków wojewódzkich w Polsce*, (Sum.: Internal structure of Polish medium size towns, voivodship capitals), Prace Geogr. IG PAN, 101, Warszawa.

Węcławowicz, G., 1975, *Struktura przestrzeni społeczno-gospodarczej Warszawy 1931 i 1970 w świetle analizy czynnikowej* (Sum.: The structure of socio-economic space of Warsaw of 1931 and 1970 in the light of factor analysis), Prace Geogr. IG PAN, 116, Warszawa.

Whebell, C. F. J., 1969, Corridors: a theory of urban systems, *Annals, Ass. Amer. Geogr.*, 59, pp. 1–26.

Wheeler, J. O., 1970, Transport inputs and residential rent theory: an empirical analysis, *Geogr. Analysis*, 2, pp. 43–54.

Whittlesey, D., 1929, Sequent occupance, *Annals, Ass. Amer. Geogr.*, 19, pp. 162–165.

Wilson, A. G., 1969A, Notes on some concepts in social physics, *Papers, Reg. Sci. Ass.*, 22, pp. 159–193.

Wilson, A. G. 1969B, Developments of some elementary residential location models, *J. Reg. Sci.*, 9, pp. 373–385.

Wilson, A. G., 1970, Disaggregating elementary residential location models, *Papers, Reg. Sci. Ass.*, 24, pp. 103–125.

Wingo, L., Jr., 1961, *Transportation and urban land*, Resources for the Future, Washington D.C.

Winsborough, H. H., 1962, City growth and city structure, *J. Reg. Sci.*, 4, pp. 35–50.

Winsborough, H. H., 1963, The ecological approach to the theory of suburbanization, *Amer. J. Soc.*, 68, pp. 565–570.

Yeates, M. H., 1965, Some factors affecting the spatial distribution of Chicago land values, 1910–1960, *Econ. Geogr.*, 41, pp. 57–70.

Zipf, G. K., 1949, *Human behavior and the principle of least effort*, Addison-Wesley, Reading.

Zipser, T., 1973, A simulation model of urban growth based on the model of the opportunity selection process, *Geogr. Pol.*, 27, pp. 119–132.

GEOGRAPHIA POLONICA 31, 1975

THE GRAVITY MODEL OF SPATIAL INTERACTION: AN APPRAISAL

Ali Mekky

1. INTRODUCTION

The hypothesis of interaction in social sciences has been recognized since 1858.[1] Groups of people interact more as they become larger, nearer, faster, with integrable activities. To put it in concrete mathematical form, it was noticed that the dimensions of group interaction parallel the dimensions of physical gravitation. Therefore, Newton's gravitation law has been borrowed. Using suitable definitions and interpretations for masses and distance, the model has been developed and extensively used to evaluate and explain spatial interactions which manifest themselves as flows of people, goods, money or information. The model has found wide applications especially after the great contributions made by Prof. Wilson at the University of Leeds. The model has been used in urban and regional planning as a transport flow[2] model, a location model and a retail model.

This paper is concerned with the evaluation of the Gravity Model (GM concept), and its applications from both philosophical and operational points of view. The main concern will be the model as a simulator for a trip distribution phase of the transportation planning process, although some of the criticisms apply to the other applications of the model as well.

In section 2 the GM will be presented on the basis of an analogy with Newtonian mechanics, statistical mechanics and information theory. The generalised form of the GM and its modal split implications which have been developed in the frame-work of Entropy Maximising (EM) methodology will be described briefly. In section 3 the GM will be evaluated from the philosophical, performance, and operational points of view. EM approach will be considered throughout the work as a methodology which leads to a special case of GM which we shall call Entropy Maximising Gravity Model (EMGM). A derivation of GM without total cost constraint will be developed. Summary and conclusions will be given in section 4.

2. THE GRAVITY MODEL THEORY

The GM is a synthetic model proposed to simulate trip distribution between zones of the study area, that is the process by which calculated trip ends (i.e. generations and attractions) are distributed between zones of the study area.

---

[1] It was pioneered by H. C. Carey.
[2] Flows of goods and persons.

Like any trip distribution model, the GM gives an estimate t_{ij} for the number of trips per unit time to be made from zone i to zone j as a function of the number of trips attracted to zone j d_j, the number of trips generated at zone i, o_i and a measure of the spatial separation of zone i from zone j c_{ij} (expressed in terms of distance, time, cost or combinations of them).

$$t_{ij} = F\{o_i, d_j, c_{ij}\}^3 \tag{1}$$

The three variables o_i, d_j, c_{ij} could, of course, themselves be functions of other independent variables.

The GM as a result of three types of analogy will be presented as well as the generalized formula of the model.

2.1. ANALOGY WITH NEWTON'S GRAVITATION LAW

The traditional way of presenting the model is simply as a trial to parallel the gravitation concept proposed by Newton in 1686.

The gravitational force F_{ij} between two masses m_i and m_j separated by a distance d_{ij} is

$$F_{ij} = \gamma \frac{m_i m_j}{d_{ij}^2} \tag{2}$$

where γ is a constant (universal one). The analogous transport model is then

$$t_{ij} = K \frac{o_i d_j}{c_{ij}^2} \tag{3}$$

where K is a constant and the travel cost c_{ij} is interpreted as distance.

Equation (3) has at least one obvious defect; if o_i and d_j are doubled, t_{ij} would quadruple instead of doubling. To make up for this criticism, the following constraints should be added

$$\sum_j t_{ij} = o_i \tag{4}$$

$$\sum_i t_{ij} = d_j \tag{5}$$

These constraints can be satisfied if sets of constants a_i and b_j associated with generation zones and attraction zones respectively, are introduced. They are sometimes called balancing factors (or normalizing factors).

After some experience in applying the model it came out that there was no reason for assuming that the distance plays its part in the transport model (3) as it does in the world of Newtonian mechanics (Wilson, 1970a). Therefore, a general function of distance $f(c_{ij})$ was introduced as a measure of the average area-wide effect of spatial separation c_{ij} on the amount of travel from i to j, t_{ij}.

The modified formula becomes [4]

$$t_{ij} = a_i b_j o_i d_j f(c_{ij}) \tag{6}$$

---

[3] A more general version can be written if we take into consideration the socio-economic linkages between zones

$$t_{ij} = F\{o_i; d_j, c_{ij}, k_{ij}\}$$

[4] Notice that the summation signs will always be written as $\sum_j$ instead of $\sum_{j=1}$ for simplicity.

where:

$$a_i = l/\sum_j b_j d_j f(c_{ij}) \tag{7}$$

$$b_j = l/\sum_i a_i o_i f(c_{ij}) \tag{8}$$

Equations (7) and (8) can be solved iteratively.

It is important to notice that the amount of travel between i and j is directly proportional to $f\ (c_{ij})$, which is a decreasing function [5] of the separation that may be expressed as distance, time, or cost.[6]

2.2. ANALOGY WITH STATISTICAL MECHANICS TECHNIQUES

Wilson (1970a) writes: "Given total number of trip origins o_i" and destinations "d_j for each zone for a homogenous person trip purpose category; given the cost of travelling between each pair of zones c_{ij}, given that there is some fixed amount of expenditure on transport C in the region at the given point in time, then there is a most probable distribution of trips between zones and this distribution is the same as one normally recognized as the GM, though with the negative exponential function appearing as the preferred form of attenuation function."

$$t_{ij} = a_i b_j o_i d_j \exp\left(-\beta c_{ij}\right) \tag{9}$$

where:

$$a_i = l/\sum_j b_j d_j \exp\left(-\beta c_{ij}\right) \tag{10}$$

$$b_j = l/\sum_i a_i o_i \exp\left(-\beta c_{ij}\right) \tag{11}$$

β is constant.

These results were achieved while seeking the distribution $[t_{ij}]$ with whom maximum number of states w $[t_{ij}]$ (or S) is connected. It can be proved that

$$S = \frac{T!}{\prod_{i,j} t_{ij}!} \tag{12}$$

where T is the total number of trips.

Therefore the problem is mainly maximizing S or In S subject to the constraints (4), (5) and

$$\sum_i \sum_j t_{ij} c_{ij} = C \tag{13}$$

This procedure which is analogous to the microcanonnical ensemble technique in statistical mechanics, is called EM technique, because a quantity analogous to In S is defined to be "entropy" in statistical mechanics.

2.3. ANALOGY WITH INFORMATION THEORY TECHNIQUES

Jaynes (1957) writes: "... The great advance provided by information theory lies in the discovery that there is a unique, unambiguous criterion for the amount of uncertainty represented by a discrete probability distribution, which

---

[5] We shall refer to this function always as either the cost function or the distribution function; other names can be found in the literature as deterrence fn., resistance fn., travel time fn. ... etc.

[6] See section 2.4.4. for the definition of the generalized cost.

agrees with our intuitive notions that a broad distribution represents more uncertainty than does a sharply peaked one and satisfies all other conditions which make it reasonable."

This measure of uncertainty is defined to be the "entropy" of the probability distribution p_{ij}

$$H = -\sum_{i,j} p_{ij} \ln p_{ij} \tag{14}$$

where:

$$p_{ij} = \frac{t_{ij}}{T} \tag{15}$$

Therefore, to estimate a probability distribution on the basis of partial information, we have to choose the probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased estimate which represents our state of knowledge about the system. Therefore, in our problem we have to maximize (14) subject to constraints (4), (5) and (13). The result is exactly as (6). This is because H and ln S are linearly related.[7]

$$\ln S = (\ln T! - T\ln + T) + TH \tag{16}$$

2.4. GENERALIZED FORM OF THE GRAVITY MODEL

As a result of the extensive use of GM in transportation planning, it was realized that for obtaining better model representation, it is necessary to subdivide the groups of travellers into more homogeneous groups, because different kinds of trips have different characteristics (trip length frequency characteristics and cost function shape) and different end opportunities.

Disaggregation of the model has been suggested to be according to:

(1) Trip purpose. Trip purpose categories can be as little as one and as many as nine or more (BPR 1965).

(2) Mode of travel. Usually two categories are used, namely, private drivers and public transport.

(3) Person type. Traditionally two categories are used: car owners and non car owners.

(4) Household income.

(5) Household structure and size.

(6) Socio-economic occupation. Some studies have differentiated between "blue" and "white collar" workers. This is due to the concentration of work opportunities for each class in particular groups of zones.

(7) Trip length. Some studies have used different distribution functions for each part of the study area which has the same average trip length, e.g., Toronto Study (TARMS) used three sectors, namely, CBD, metropolitan area, and the region outside the metropolitan area.

For the same reason different models are generally calibrated for internal trips and external trips.

(8) Time of day. Usually two categories are used, namely, average daily and peak hour.

The generalized GM forms the way the behaviour of these subgroups can be estimated, linked and aggregated together.

The basic generalization approach happened in the framework of EM methodology. As an example, Wilson's model which recognizes the modes of travel

---

[7] Assuming that the Stirling approximation is valid i.e.

$$\ln N! \cong N \ln N - N$$

and the person types will be presented together with its aggregation forms and model split implication.

2.4.1. THE MODEL

The exponential GM expressed by equations (9)–(11) can be generalized by introducing several person types n and several transport modes k (within a single trip purpose). Notice that:

(1) Usually the production ends are characterized by household characteristics (here person types), while the attraction ends are characterized by land use type, therefore, we expect o_i to be disaggregated into o_i^n but not d_j. To put this another way we might assume that different types of people at zone i compete for the same attractions at zone j.

(2) t_{ij}^{kn} is the number of trips generated in zone i — by person of type n using the mode k — and attracted to zone j.

Therefore, the end constraints for each o_i^n, d_j are

$$\sum_j \sum_{k \in M(n)} t_{ij}^{kn} = o_i^n \tag{17}$$

$$\sum_i \sum_n \sum_{k \in M(n)} t_{ij}^{kn} = d_j \tag{18}$$

the cost constraint for each person type n is

$$\sum_i \sum_j \sum_{k \in M(n)} t_{ij}^{kn} c_{ij}^k = C^n \tag{19}$$

Now maximize

$$ln\left[\frac{T!}{\prod_{ijk} t_{ij}^{kn}!}\right]$$
$$k \in M(n)$$

subject to (17)–(19), we obtain

$$t_{ij}^{kn} = a_i^n b_j o_i^n d_j \exp\left(-\overset{n}{\beta} c_{ij}^k\right) \tag{20}$$

where:

$$a_i^n = \left[\sum_j \sum_{k \in M(n)} b_j d_j \exp\left(-\overset{n}{\beta} c_{ij}^k\right)\right]^{-1} \tag{21}$$

$$b_j = \left[\sum_i \sum_n \sum_{k \in M(n)} a_i^n o_i^n \exp\left(-\overset{n}{\beta} c_{ij}^k\right)\right]^{-1} \tag{22}$$

Equation (20) is a linked set of GM's for each k–n category. The linking is through the b_j's

2.4.2. AGGREGATION[8] OF THE MODEL OVER PERSON TYPES n

$$t_{ij}^{k\cdot} = a_i' b_j o_i' d_j \exp(-\beta' c_{ij}^k) \tag{23}$$

$$a_i = \left[\sum_j \sum_k b_j d_j \exp(-\beta' c_{ij}^k)\right]^{-1} \tag{24}$$

$$b_j = \left[\sum_i \sum_k a_i' o_i' \exp(-\beta' c_{ij}^k)\right]^{-1} \tag{25}$$

[8] Notice that when an index is replaced by a point then this denotes summation over the index.

2.4.3. AGGREGATION OF THE MODEL OVER MODES K

It is easy to prove that the models formulas are

$$t_{ij}^{n} = a_i^{n} b_j o_i^{n} d_j \sum_{k \in M(n)} \exp\left(-\frac{n}{\beta} c_{ij}^{k}\right) \tag{26}$$

$$a_i^{n} = \left[\sum_j b_j d_j \sum_{k \in M(n)} \exp\left(-\frac{n}{\beta} c_{ij}^{k}\right)\right]^{-1} \tag{27}$$

$$b_j = \left[\sum_i \sum_n a_i^{n} o_i^{n} \sum_{k \in M(n)} \exp\left(-\frac{n}{\beta} c_{ij}^{k}\right)\right]^{-1} \tag{28}$$

2.4.4. MODAL SPLIT IMPLICATION OF THE MODEL

$$\frac{t_{ij}^{k\cdot}}{t_{ij}^{\cdot\cdot}} = \frac{\exp(-\beta' c_{ij}^{k})}{\sum_{k'} \exp(-\beta' c_{ij}^{k'})} \tag{29}$$

Notice that c_{ij}^{k} is the generalized cost expressed as

$$c_{ij}^{k} = \sum_r a_r x_r(i, j, k) \tag{30}$$

Where $x_r's$ are variables like fares, travel time, excess time, and distance; $a_r's$ are wights.

In practical applications of the last formula of modal split (29) it was found necessary to use β' for modal split other than that for trip distribution.

3. GRAVITY MODEL EVALUATION

3.1. PHILOSOPHICAL CONCEPT

The different bases of analogy between the trip distribution on the one hand, and Newtonian mechanics, statistical mechanics and information theory on the other hand will be discussed.

It is evident that the supposed basis of analogy with Newtonian mechanics is confined to the "gravitation formula", while the other analogies apply to the methodology of entropy maximization and not to special concrete formula.

3.1.1. THE ANALOGY WITH NEWTONIAN MECHANICS

Is there any real kinship between the Newtonian concept of gravitation and the trip generation system, or between the two formulas? The answer is no, for the following reasons:

(1) Newtonian gravity is an energy-force field characterising the motions of particles, not their intentions. It does not deal in statistical units. A cluster of mass points under the influence of circumambient masses does not break into flights aimed at the separate attractive poles, but each member of the cluster is content to be governed by the vector resultant obtaining at its point in space and this vector varies from point to point only in a mathematically continuous fashion. "Lines of force" is a reasonable visualization, but this is not at all the same as lines of movement (Schneider, 1959).

(2) There is a basic difference between Newton's formula and the so called GM for trip distribution. That is the feed back, or iterative nature of the second formula (between L. H. S. and R. H. S.), which does not exist in the first one.

$$F_{ij} = \frac{m_i m_j}{d_{ij}^2} \tag{31}$$

$$t_{ij} = a_i b_j \sum_i t_{ij} \sum_j t_{ij} f_{ij} \tag{32}$$

The masses in Newton's formula have nothing to do with the "force", while in (32) the "masses" are algebraic sums of the "forces"!!

(3) There is no evidence for the universality of the structural parameters in the GM for trip distribution. The more so as the stability over time is highly questionable.

3.1.2. THE ANALOGY WITH STATISTICAL MECHANICS AND INFORMATION THEORY

The essence of the methodology is "maximizing" the entropy, subject to linear constraints (end constraints+total cost constraint), i.e., it is simply a mathematical programming technique where the objective function is the entropy of the system or the distribution.

The following discussion will be concerned with the role of the entropy expression and the constraints in shaping the model and defining its predictability.

(1) The role of the entropy expression. As discussed before, two entropy expressions are in use, namely

$$\ln S = \ln T! - \sum_i \sum_j \ln t_{ij}! \tag{33}$$

$$H = -\sum_i \sum_j \frac{t_{ij}}{T} \ln \frac{t_{ij}}{T} \tag{34}$$

Both of these expressions lead to the same result after maximization because they are linearly related.

Examining near the maximum of the entropy distribution ln S, Wilson (1970a) has proved that

$$d[\ln S] = -\frac{1}{2} \sum_i \sum_j p^2 t_{ij} \tag{35}$$

where p is the percentage change in each t_{ij} away from the most probable distribution.

It is easy to prove that, if we maximize the entropy subject to no constraints, we may get the distribution

$$t_{ij} = \frac{T}{k} \tag{36}$$

where k is constant.

The distribution (36) has a very sharp peak — exactly the same peak as the distribution (9) — yet it is highly impractical. The addition of any set of linear constraints does not affect the sharpness of the distribution but it affects the final form of the model and its ability to explain and predict the phenomenon. This conclusion contradicts the opinion of Jaynes (1957) as he writes: "Consider now the case when the theory makes definite predictions and they are not borne out by experiments. This situation cannot be explained by concluding that the initial information was not sufficient to lead to the correct prediction; if that were the case the theory would not have given a sharp distribution at all.

The most reasonable conclusion in that case is that the enumeration of the different possible states (i.e. the part of the theory which involves our knowledge of the laws of physics) was not correctly given. This experimental proof that a definite prediction is incorrect gives evidence of the existence of new laws of physics."

Therefore, as it was argued before, the situation which may arise from having the broadest possible probability distributions over the microstates with an enormously sharp peak over the macrostates, yet not standing up to experimental tests, this situation may be explained by lack of information. This conclusion shows the incorrectness of Jaynes' reasoning, cited above.

Therefore, the main role of the entropy expression is defining the sharpness of the trip distribution expression and its maximum whatever the linear constraints are.

(2) The role of the end constraints.

$$\sum_j t_{ij} = o_i \tag{4}$$

$$\sum_i t_{ij} = d_j \tag{5}$$

It is clear that the presence of the end constraints (4) and (5) give rise to the balancing factors a_i and b_j, yet it is merely a necessary condition for this presence — but not a sufficient one. It may seem odd to say that the presence of the total cost constraint does affect the existence of the balancing factors. It is easy to prove it simply by developing the model without cost constraint. In this case the balancing factors would not be needed (see equation 37).

(3) The role of the cost constraint.

$$\sum_i \sum_j t_{ij} c_{ij} = C \tag{13}$$

The above total cost constraint developed by Wilson (1970a) is probably introduced for one or two reasons:

(a) Maximizing the entropy subject to end constraints only gives

$$t_{ij} = \frac{o_i d_j}{T} \tag{37}$$

This result, in spite of having a very sharp maximum, fails the experimental test. Adding the total cost constraint (13), the resultant model is similar to the familiar well tested GM, though with exponential distribution function. Wilson (1970b) writes: "The negative exponential function arises because of the nature of the constraint (13). In this equation and in the model equation it has been tacitly assumed so far that travellers perceive costs (subjectively) in the way in which they are (objectively) measured. This is a hypothesis about behaviour. Thus if this is correct the entropy maximizing method suggests that the negative exponential function is the correct form of impedence function".

This author believes that the EM model in its essence is rather far from representing any behavioural approach as the total cost constraint contains no behavioural content of any sort.

(b) In an analogy with statistical mechanics which uses total energy constraint. Jaynes writes: "In conventional statistical mechanics, the energy plays a preferred role among all dynamical quantities, because it is conserved both in the time development of isolated systems and in the interaction of different systems".

It is clear that the total cost of travel for an urban system is neither conserved in time nor in the interaction with other urban systems. Therefore while there is some reason to introduce a third constraint beside the two end constraints — to make the resultant GM formula an operational one — there is no reason to make such constraint similar to the total energy contraint, since energy and cost have different roles and characteristics in physics and social physics respectively. On the other hand, if the formula (13) could be assumed, then it may be equally probable to assume other formulas which may lead to different forms of the GM. (see Appendix 1). Also different assumptions can lead to the exponential function without using total cost constraint. This fact invalidates the reasoning that the total cost constraint "generates" operational GM. Actually the "generation" can be done without any cost assumption as (13) (see Appendix 2).

Therefore, it may not be very sound to assign any theoretical basis to the exponential distribution function (which is an admittance rather than impedence function) in the GM or the EMGM.

In the author's opinion, if we can assume anything without good reasons (e.g. the total cost constraint) it may just as well be the final GM formula.[9] There is no apparent conceptual content in the model. It is not a behavioural description but a formula which empirically fits observed facts, simply some kind of curve fitting or data summarizing.

3.2. OPERATION AND PERFORMANCE

The GM has appeared to give rather satisfactory results if properly calibrated and tested.

Bouchard and Pyers (1965) write: "The level of accuracy obtained by forecasting trip distribution patterns (using GM) was comparable to the level of model accuracy for the base year".

Heanue and Pyers (1966) write: "The overall accuracy of the GM (using k_{ij} factors) in the base year simulation and in forecasting ability has proved to be slightly better than the accuracy of the intervening opportunity model."

Lanson and Dearinger (1967) compared and tested four models of trip distribution for industrial work trips, including the GM, the electrostatic model, and the multiple regression model. They concluded that the GM which included socio-economic adjustment factors gave the most accurate distribution pattern by RMS for the theoretical models.

Connected with the use of the GM, there are advantages, disadvantages and problems in application. These items will be treated briefly in the following two subsections.

3.2.1. ADVANTAGES OF THE GM

(1) Simple to understand and apply.

(2) It is the first to recognize the importance of trip purpose and socio-economic status.

(3) It accounts for the competition which exists between different land uses.

(4) It considers the effect of transport network changes on travel patterns.

(5) The model in principle can evaluate the influence that social and economic factors have on travel patterns.

---

[9] The EM derivation of the GM "resting" heavily on the arbitrary assumption of the total cost constraint, reminds me of the Hindu's old view about the world resting upon a big elephant, and the elephant resting on a tortoise. But what about the tortoise?!

(6) The GM as a synthetic model, not only can be used to predict future trips, but can also synthesize the base year flows without having to survey every individual call in the trip matrix. It has been proved that predictions made when information is missing are consistent with that obtained using all information[10] (Kirby, 1972). This leads to a substantial decrease in the cost of data collection, as small samples are sufficient.

(7) Also, being a synthetic model, errors in the original data are not magnified in the forecasting process as is the case with the previous techniques (the growth factor methods). Errors, instead are smoothed out.

(8) Well documented.

3.2.2. SHORTCOMINGS AND PROBLEMS CONNECTED WITH THE GM APPLICATIONS

(1) Study-area boundaries. Assuming elementary GM, Schneider (1960) has shown theoretically that the GM formulation is not generally valid over an unlimited or an undefined range of the distance variable, but can be entertained only within a region between some stated minimum and maximum distance. When these limits are given, however, the formula becomes a function of them, and no change in the formula, even change in the exponent can yield the same results if the arbitrary boundaries of the region are moved.

(2) Zoning problems. Many transportation studies have found that the choice of zoning system affects the values of parameters. If, for example, 90% of the movements are within the zones, then we will not have a good model.

Broadbent (1969) has proposed a set of criteria which might be useful in choosing the zone size:

Small zones are needed to optimize:

(a) Locating the zone centroide.

(b) Defining spatial distributions and attractions.

(c) Forecasting of interzonal movement.

Large zones are needed to optimize:

(a) Conditions necessary for the application of statistical models where validity depends upon having large numbers.

(b) Accuracy of measurement and forecasting of zonal variables.

(c) Data collection time and cost.

(d) Accuracy of measurement of intrazonal distance.

A compromise must be sough for the optimum size of zones which take into consideration the following:

(a) Compact zones are preferred.

(b) Size of the zone may vary inversely with the density of "masses".

(c) Availability of required statistical data for the size chosen.

The dependence of the value structural parameters upon the number of zones, can be explained by the fact that the parameters are an outcome of the distribution function assumed, and the regression analysis used. Therefore, the result of such treatment will be eventually dependent on the number of points in the regression (which equals the number of zones having non-zero observations).

(3) The distance between zones. The changing nature of (time) distance between zones with time of day makes questionable the use of simple value of travel time, regardless of what transportation facilities are available. This effect may be reduced by using the generalized cost function.[11]

---

[10] This is called "the separability property".

[11] There is a difficulty in the solution of using generalized cost, mainly in forecasting these costs.

(4) Iteration difficulties. The basic operational difficulty with the GM application is that it requires a considerable amount of adjustment and manipulation to achieve satisfactory results. This difficulty may be reduced by using the non-iterative methods of calibration developed recently, e.g., Kirby's method.

(5) The form of the cost function. Different forms of the cost function, $f(c_{ij})$ have been suggested by different investigators e.g.:

(a) Power function

$$f(c_{ij}) = l/c_{ij}^{\alpha}$$

suggested by Lill (1891), many investigators have used this function with α equals 1, 2 or 3 early in the development of the model especially for the intercity movements which exercised the minds of that time.

(b) Exponential functions

$$f(c_{ij}) = \exp(-\beta c_{ij}) \text{ suggested by Barbe (1963).}$$

$$f(c_{ij}) = \exp(-\beta c_{ij}^2) \text{ (Angenot, 1961).}$$

$$f(c_{ij}) = \exp(-\beta c_{ij})\cdot c_{ij} \text{ (Dieter, 1962).}$$

The first form is the most popular one especially after Wilson's derivation of the EMGM.

(c) Gamma function

$$f(c_{ij}) = \exp(-\beta c_{ij})\cdot c_{ij}^{-\alpha} \text{ (Tanner, 1962).}$$

(d) Pearson function type I and type III. suggested by Ashford and Covault (1969) in an attempt to correlate the curves parameters with the area-wide characteristics.

(e) Intervening opportunity function

$$f(c_{ij}) = \exp(-\beta v_{(i)j}) \text{ suggested by Wilson (1970a).}$$

Where j is the destination zone in the rank order away from (i). $V_{(i)j}$ is the number of opportunities passed, up to and including j.

(f) No functional form.

B.P.R. (1965) suggested using empirically derived travel time factors. A criticism against this approach is that since the cost function is not absolutely analytical, it is more susceptible to the effects of chance variations in calibration data due to sampling and actual errors, than models which utilize completely analytical expressions.

Therefore, there is no agreement between investigators about the form of cost function.

(6) The cost function dimensions. As we mentioned earlier in section 2.4., the model — and accordingly the cost function — has to be disaggregated according to many factors e.g. trip purpose, mode, etc. Isard (1960) writes: "...disaggregation is desirable when additional information and precision is obtainable and when such disaggregation does not destroy to any great degree the inherent meaning and internal structural unity of the mass or population. Under these circumstances it will be fruitful to employ distinguishing exponents, weights, etc., ...But if the inherent meaning and internal structural unity of the mass tend to be significantly destroyed... disaggregation does not yield productive results".

(7) The structural parameters stability overtime

Distance exponent has been found to range from 0.5 to 3.5 for all journey modes, with the lower values indicating higher levels of mobility or accessibility as their effect on the distance measurement is more minimal. Usually, short distance trips are much more sensitive to alternative exponent values, while the longer trips allow for more constancy. This is due not just to greater variety of travel patterns over small distances, but also because of the greater propensity to move in any event over small distances. The result of this situation is that in GM calibration while we may get good fits for total volumes of interaction, the distribution of component flows be over- or under-estimated in various different zones.

Very often however, the more special effects on mobility and accessibility can be better controlled by manipulations of the mass variables rather than the distance, for clearly the two sets of variables are strongly interlinked. It has been noticed that as the spatial units increase in size, become more compact (not irregular in shape) and more numerous in occurrence, then the relative importance of the weights and mass components become minimized and the distance exponent tends to crystallize around 2.0 (Davies, 1969).

It was found that the structural parameters vary with many factors (see section 2.4). This problem can be treated by further disaggregation of the GM.

(8) The time stability of the cost functions. One of the weakest assumptions made in the use of the GM is that the form and parameters of the cost function are constant in time. No agreement between research workers about that stability, has been reached. Most of the opinions agree that the function may vary over time (i.g. Whitemore, 1965). Solutions for this problem were sought:

(a) Ashford and Covault (1969) tried to correlate the parameters of a suggested parametric curve — for the cost function — with the area-wide parameters which are known to vary with time.

(b) Alan Vorhees and Ass. (1966) tried to fit a gamma distribution for the cost function (calculated for every zone). It was found out that shape parameter is highly correlated with the mean opportunity trip duration.

(c) Kirby (1971) has shown that a part of the instability of the cost function may be an apparent one and he concluded that using the generalized cost may decrease such apparent instability.

(d) Ottesen (1971) using data from Copenhagen, noticed stability over time for the relation between the mean trip lengths for residential zones and the distance from the city centre which can be used as a basis for prediction.

(9) The time stability of the k factors. The need for k factors may result due to

(a) Insufficient trip purpose stratification.

(b) The use of area-wide travel time factors.

(c) The existence of special links between pairs of zones. It has been shown by Hansen (1962) that the stability over time of the k factors or its relation with income is highly questionable. The problem can be reduced by:

(a) No need for the k factors is small urban areas (for population less than 50,000).

(b) For more complex urban areas the problem can be reduced by finer stratification of work and shopping trips (Bouchard and Pyers 1965).

(c) If there is a great need for k factors, they must be related to the characteristics of the zones so that they may be forecast as a function of socio-economic conditions estimated for the future land use plan.

4. SUMMARY AND CONCLUSIONS

The Gravity Model of spatial interaction has found wide applications. One of these important applications is in the trip distribution phase of the transportation planning process, which is the main concern of this paper. The GM theory is presented and the model is evaluated from the philosophical, operational and performance points of view, with the following results:

(1) The main advantage of the GM are simplicity, ability to evaluate the socio-economic influence on travel, having separability property and being well documented.

(2) There is no kinship between Newton's gravitation formula and the GM for trip distribution.

(3) EMGM, practically, has a sharp maximum which does not depend upon linear constraints, but mainly upon the definition of entropy.

(4) The non-validity of some of the EM models, in spite of having sharp peaks, may happen due to lack of information (constraints); as linear constraints affect the form of the model but not the sharpness of its peak.

(5) The balancing factors appear not only because of the nature of the end constraints, but also due to the presence of the total cost constraints.

(6) The total cost constraint is artificial and has no rigid basis.

(7) There is no "preferred form of cost function" derived by the EM approach, as many functions can be assumed with equal probability.

(8) The cost function has many dimensions of variability and is rather unstable over space and over time, besides the basic disagreement between researchers about its form.

(9) There is large correlation between instability in time and in space for both the cost function and the k factors.

(10) There is no apparent conceptual content in the GM.

(11) The well known GM can be derived from other more practical assumptions which do not include constant total cost constraint.

(12) If we can assume anything without good reasons (as the assumption of the total cost constraint), it may just as well be the final GM formula itself.

Therefore, the GM should not be considered more than an efficient technique utilizing some empirical regularities to which weak theoretical bases were attributed. We need to direct our research towards a more logical consistent theory of trip distribution.

APPENDIX

1. DERIVATION OF DIFFERENT FORMS OF THE DISTRIBUTION FUNCTION

It may be equally probable to assume many forms of the cost constraint, *ceteris paribus*, we get different forms of the distribution function, e.g.

(a) Using general function of the cost $\phi(c_{ij})$ then the cost constraint becomes

$$\sum_i \sum_j t_{ij}\, \phi(c_{ij}) = c$$

the model formula becomes

$$t_{ij} = a_j b_j o_i d_j \exp(-\beta\phi(c_{jj}))$$

It is clear that this is due to the well known special cases of distribution function namely, the power, the exponential and the gamma functions.

(b) Adding the constraint

$$\sum_i \sum_j t_{ij} c_{ij}^2 = \overline{c^2}$$

the resulting distribution function is truncated Gaussian.

(c) Adding the constraints

$$\sum_i \sum_j t_{ij} \ln c_{ij} = \overline{\ln c_{ij}}$$

$$\sum_i \sum_j t_{ij} \ln (l - c_{ij}) = \overline{\ln(l - c_{ij})}$$

the resulting distribution function is "Beta" function.

2. DERIVATION OF GM WITHOUT TOTAL COST CONSTRAINT

It was argued that the addition of the total cost constraint (13) is rather artificial and has no rigid basis. If we seek the most probable distribution which takes into account the tendency of the individual towards minimizing their costs of travel, then this idea may be expressed as maximization of the function

$$F = \ln S - \beta \sum_i \sum_j t_{ij} c_{ij} \tag{38}$$

subject to the end constraints

$$\sum_j t_{ij} = o_i$$

$$\sum_i t_{ij} = d_j$$

the solution is

$$t_{ij} = a_i b_j o_i d_j \exp(-\beta c_{ij})$$

which is the same as equation (9).

Here, we recognize two special cases:

(a) If β is very small, therefore, t_{ij} tends to

$$\frac{o_i d_j}{T}.$$

(b) If β is very large, therefore, the solution is the same as that of the Linear Programming transportation problem.

The assumptions above, seem to be more practical than those which include total cost constraint and ignore the tendency of travellers towards minimizing their travel costs.

Department of Traffic Research
Alexandria Region General Transport Administration, and
Economic Academy, Cracow

BIBLIOGRAPHY

Angenot, L. H. J., 1961, De afstandsfactor der verkeerspolen teorie, *De Ingenieur*, 23, 15, pp. 11–13.

Ashford, H., Covault, D. O., 1969, The mathematical form of travel time factors, *HRR*, 283.

Barbe, H. B., 1963, *The transportation balance for Zurich metropolitan area.*
Bentfeld, G., 1964, *Concerning attempts to predict flows of motor traffic on the basis of population and other structural data.* Seventh International Study Week in Traffic Engineering, London.
Bouchard, R. J., Pyers, C. E., 1965, Use of gravity model for describing urban travel: an analysis and critique, *HRR,* 88, pp. 1–43.
B.P.R., 1965, *Calibrating and testing a gravity model for any size urban area,* US Dept. of Transport, Bureau of Public Roads, Washington, DC.
Broadbent, T. A., 1967, *An unpublished paper given at an exposition on interurban models at the Ministry of Transport.*
Broadbent, T. A., 1969, *Zone and spatial interaction: a preliminary discussion,* CES Draft Working Note 106.
Carey, H. C., 1958, *Principles of social science,* J. Lippincott, Philadelphia.
Devies, R. L., 1969, *Problems of variable selections and measurement,* Lanchester Polytechnic.
Dieter, K. H., 1962, Distribution of work trips in Toronto, *J. City Plann. Div.,* 88, No. CP1, pp. 9–28.
Hansen, W. G., 1962, Evaluation of gravity model trip distribution procedures, *HRB Bull.,* 347, pp. 67–76.
Heanue, K. E., Pyers, C. E., 1966, A comparative evaluation of trip distribution procedures, *HRR,* 114.
Isard, W., 1960, *Methods of regional analysis: an introduction to regional science,* John Wiley and Sons, London.
Jaynes, E. T., 1957, Information theory and statistical mechanics, *Phil. Rev.,* 106, p. 620–630.
Kirby, H. R., 1971, *Interim review of the result of gravity model calibration,* PTRC symposium, London, pp. 97–110.
Kirby, H. R., 1972, *Theoretical requirements for calibrating gravity model,* University College London (in print).
Lanson, H., Dearinger, J., 1967, A comparison of four work journey distribution models, *Proc. Am. Soc. Civ. Eng., Highway Div., 93,* No. HW2.
Lill, E., 1891, *Das Reisegesetz und seine Anredung auf den Eisenbahnverkehr,* Wien.
Ottesen, K. A., 1971, *The use of a gravity model to analyse the relation between homes and workplaces in a metropolitan area,* PTRC, London.
Schneider, M., 1959, Gravity models and trip distribution theory, *Papers, Reg. Sci. Ass.,* 5. pp. 51–56.
Schneider, M., 1960, Panel discussion on inter-area travel formulas, *HRB Bull.,* 253, pp. 128–138.
Tanner, 1961, Factors affecting the amount of travel, *RRL Tech. Pap.,* 51, London.
Tarms, 1968, *Toronto Area Region Model Study.*
Voorhees, A. M., Bellomo, S. J., Schofer, J. L., Cleveland, D. E., 1966, Factors in work trip length, *HRR,* 141, pp. 24–46.
Whitemore, R. E., 1965. *Graphical and mathematical investigation of the differences in travel time factors for the gravity model trip distribution formula in several specific urban areas,* University Tennessee.
Wilson, A. G., 1970a, *Entropy in urban and regional modelling,* Pion, London.
Wilson, A. G., 1970b, Advances and problems in distribution modelling, *Transp. Res.,* 4, pp. 1–18.

ON THE ORGANIZATION OF POLITICAL SPACE *

W. A. Douglas Jackson and Edward J. F. Bergman

POLITICAL SPACE

In their study of political space, geographers have traditionally focused their attention on the large territorial-administrative unit or the State. This interest has been essentially chorological; that is, its purpose has been to describe the characteristics and features of the State and to attempt some type of regional synthesis. Many of such studies have been little more than regional inventories; this has been particularly true where an attempt has been made to assess the natural resources that contribute to the power of the State.

Political scientists have shared with geographers this interest in the State. The former, however, have been more inclined to deal with man in society, organized to achieve political goals and to establish the machinery of government. Political scientists above all have been principally concerned with power and the formal political institutions of the State. The focus of the geographer, on the other hand, has been more on the territorial aspect: what territory, how organized, and, to some extent, why. Seldom, however, has the geographer examined political organization, or political activities in the broadest sense below the State level. Still it has been a focus on the State region that has made the work of the geographer "political".

After World War II, as geography tended to move away from chorology, an effort was made to inject a more dynamic element into the geographical study of States. A leader in this direction was Richard Hartshorne, who developed what was termed a "functional approach". [1] Accepting the given organized State as the core of political geographical study, Hartshorne proposed that the geographer should seek to determine how effectively the State area or region functions. This objective would be achieved through a careful evaluation and weighing of the "forces" that compete within each State or affect each State: forces that are (a) centrifugal and tend to prohibit or discourage successful integration of people and territory, and (b) centripetal that, on the contrary, promote internal unity and purpose. An assessment of these centrifugal and centripetal forces would then provide some indication of how well the State could govern throughout its legal, sovereign territory.

---

* Editor's note: Political geography is a relatively poorly developed branch of geographical sciences. In the present article, prof. W. A. Douglas Jackson, a reknown specialist in the field and author of the book on: *Politics and Geographical Relationships* (Prentice Hall, Englewood Cliffs 1971) undertakes a revaluation of some of the basic concepts and notions. The article is of a discussion character.

[1] Hartshorne, R., The functional approach in political geography, *Ann. Ass. Amer. Geogr.*, 40 (1950), 95–130.

The political-geographical phenomenon known as the State, however, is only one manifestation of political activity. In an age of strong nationalism the State, of course, remains the focus of our most intense loyalties. Still there are other political activities in which men engage, which lead to the formation of political activity regions but do not necessarily culminate in formal State organization. Among geographers who have urged that a broader view of political activity be taken, may be mentioned Stephen Jones. Indeed, he suggested that attention be focused on political process, that geographers examine the spatial implications of the entire field of politics.

With this in mind, Jones proposed the notion of a theoretical chain linking the formulation of a political idea to its ultimate materialization in a political area, State or otherwise. In an essay entitled *A Unified Field Theory of Political Geography,* he identified the links in the chain as Political Idea, Decision, Movement, Field and Political Area.[2] The process by which Political Idea comes to fruition in Political Area may be traced at any level of organization.

As Jones made clear, one may trace the manner in which political ideas culminate in political areas at any level of human activity. The process may entail nothing other than the desire of an urban neighborhood to establish a local park. The achievement of the goal may require leadership, organization and other political activities.

On the other hand, it is possible to demonstrate, as Jones noted, what steps ought to be taken beginning with the Political Area to promote the development of an activity field within the area and, finally, the formulation of a Political Idea. The purpose is to ensure that the Political Area — whether a State or any other previously delineated region, such as a country — becomes an effectively functioning field of political activity.

The interest of geographers in political processes at whatever level of human activity tends to parallel new approaches that have been introduced in political science. These represent in both geography and political science a desire and an effort to develop a more conceptual approach to political man — to political activity in all its manifestations and at every level of organization. This shift in focus sees man within a vast network of overlapping political organizations, each characterized by its own pattern of human interaction. These networks have been termed "systems", and they come into existence for the expressed purpose of achieving some desired goal. They entail, accordingly, sets of relationships between people, which may be described as exchanges and transactions. One way, therefore, to analyze the political use and organization of the earth's surface is to examine how these systems come into being, how they operate, and the way in which they hope to achieve their objectives. The political scientist, David Easton, has made major contributions to political systems analysis.[3]

Political space should not be accepted as rigidly structured — as any political-administrative map might suggest — but as fields wherein the act of organization, or formation of systems, is an on-going dynamic process. Political processes, moreover, define their own space, their own functioning regions. The application of system analysis to political geography is consistent with trends in other branches of geographical science and affords valuable new insights into

[2] Jones, S. B., A unified field theory of political geography, *Ann. Ass. Amer. Geogr.*, 44 (1954), 111–123.

[3] Easton, D., *The political system* (New York: Knopf, 1953); *A framework for political analysis* (New York: Prentice-Hall, 1965); *A systems analysis of political life* (New York: John Wiley and Sons, Inc., 1965).

the dynamism of political organization and the nature of the geography of politics.

THE DIMENSIONS OF POLITICAL SYSTEMS

Human organization is a political process. The very act of coming together to deal with problems of mutual concern is a political one. This process entails, in effect, a system of behavior and represents or establishes relationships or patterns of interaction that provide the machinery to make decisions and to resolve problems. These processes and patterns of human organization are not everywhere the same. This is because each system is "surrounded by its own physical, biological, social and psychological environments", which influence or modify it. [4] Political systems, therefore, may differ in size, structure, resources, and purpose.

The view of political systems as a means of solving societal problems parallels the use of systems analysis in the study of corporate efficiency. The criteria of success or efficiency in political or governmental institutions cannot be equated entirely with those in business, but there is value in bringing to political systems some ideas that have emerged from the study of organizational efficiency in modern corporations. In business management a system has been defined as "a problem-solving vehicle, scheme, process, structure or methodology designed to regularize business decision-making in a given area or context".[5] Systems analysis is an attempt to find the most feasible, suitable, and acceptable means for accomplishing a given purpose, taking all of the environmental factors into account.

Business organizations are frameworks for the production and distribution of goods and services to achieve profit. Political systems are frameworks for the management of political power and for the regulation of society to achieve public goals. The actors within the political system are those individuals and groups which determine and allocate the distribution of power and authority. Specifically, then, a political system is a pattern of interaction wherein power and authority are generated and allocated to regulate the behavior of constituents to resolve problems and achieve desired goals.

THE GOALS OF POLITICAL SYSTEMS

Every political system must have goals or objectives. The goals that are sought or the problems to be solved determine the area and the constituency of the system created to achieve those objectives.

An awareness of objectives, particularly with respect to the organization of States, has been an important part of the intellectual equipment of political geographers since the late nineteenth century. Writing in the 1890's the German geographer Friedrich Ratzel noted, that a State comprises a section of land and a section of humanity organized according to a particular distinctive idea. This distinctive state-idea has been termed the *raison d'etre* of the State. Richard Hartshorne has referred to it as "some concept or idea justifying the existence

---

[4] Easton D., *A systems analysis of political life*, p. 18.

[5] Minnich, Ch. and Nelson, O., *Systems management for greater profits and growth* (Englewood Cliffs: Prentice-Hall, 1966), p. 1.

of this particular state incorporating these particular regions"[6]. He noted that the "*raison d'etre* must be based upon desires or values of first importance to the populations of the regions included in the State. Desires for freedom from war's devastation or for economic advantages are common to all regions, but which of all possible state-areas may offer any particular region the best advantages in these respects will depend in part on its particular geographic conditions."[7]

Sometimes the *raison d'etre* may be the sense of commitment commonly held by a number of people to a certain way of doing things, to membership in a political community; sometimes it may be an idea in the mind of a conqueror, such as that held by Napoleon, or a vision of empire such as Cecil Rhodes entertained. Both the identity and the stability of the reason or idea are important aspects of the study of political systems at whatever level of organization, for not all reasons are of equal stability. People may change their political commitments or allegiance, and states and empires have disappeared at the death of a strong leader or ruler. Still, the fact that the idea may not long hold the support of the participants does not lessen its importance in the formation, development, and the study of political systems.

The State or the Nation-State remains, for the most part, the largest and most fully integrated political system in the modern world. The period since World War II has seen the appearance of many new sovereign States, ranging in size from large and populous countries like India and Nigeria to others as tiny as the Maldive Islands and Singapore. Often composed of diverse cultural and racial groups, such new States often originally lack articulated national goals generally accepted by the people. The official granting of independence has legitimized the successor political structure, but the task of achieving internal unity based on some sense of community or a national identity in the new State and its goals often remains to be realized. Many newly independent States have particular difficulty in obtaining the allegiance of all their people because their international boundaries still follow old colonial administrative delineations. These lines on maps drawn in London or Berlin in the 19th century may have indiscriminately separated groups sharing some common precolonial allegiance, or they may have associated groups having traditional animosities. The Ewe people, for instance, were divided between Togo and Ghana, while Arabs and Blask Africans found themselves together in the Sudanese State. Such conditions make difficult the establishment of political unity and can create unstable boundary problems with serious implications. As Julius Nyerere, President of Tanzania, has noted in reference to African States: "These new countries are artificial units, geographical expressions carved on the map by European imperialists. These are the units we have tried to turn into nations."[8] These new States, then, strive to define their reasons for existing, their *raison d'etre*, and to achieve a common commitment of their peoples.

Political systems, as has been noted above, however, should by no means be identified exclusively with the sovereign State. At the other end of the scale of human organization, the League of Women Voters, or even a faction of a Parents' and Teachers' Association advocating the addition of sex education to the curriculum of the local school, qualifies by definition as a political system. Provinces, countries, and townships, not to mention special purpose govern-

[6] Hartshorne, R., *op. cit.*, p. 110.

[7] Hartshorne, R., The concepts of *Raison d'etre* and 'Maturity' of states; Illustrated form the Mid-Danube Area, abstract in: *Ann. Ass. Amer. Geogr.*, 30 (1940), 59–60.

[8] *Time*, vol. 97, No. 5 (February 1, 1971), p. 39.

mental units, are systems as well, though established by administrative means within the framework of the sovereign State. The objectives of provinces, countries, and other general purpose systems are to carry out the general goals of the State system at the local level. The objectives of special purpose units, such as sewage or education districts, are for the most part specifically limited. However, whether the organization is the League of Women Voters or a unit of government such as a county, both have a legitimacy of their own, even though they are subordinate to the sovereign State and their goals may be limited.

Every political system (or subsystem) must be able to regulate its membership in order to achieve the goals commonly sought. For this it must possess the machinery or institutions. In time these may come to embody and characterize the system, tending to formalize and to stabilize it. If, for example, the issue of sex education is brought up at only one P.T.A. meeting, and the members advocating sex education lack the means to create an effective formal organization to lobby for their point of view, then that system will die a-borning. If, on the other hand, the members organize and regulate themselves to pursue their goal, a system takes shape and so, concomitantly, does the essential framework or machinery. They may create an organization called MASE, Mothers Advocating Sex Education, and set to work in the school district to achieve their goals.

TERRITORIAL SOVEREIGNTY AND POLITICAL STRUCTURES

In the world today only the State claims exclusive and exhaustive territorial sovereignty. Within the State authority is ordinarily divided and subdivided down through a hierarchy of subunits. The education district (a special-purpose subdivision) in which the MASE ladies form a pressure group is just a tiny fraction of sovereignty within a small space or territory. The district is strictly delimited in both area of space and field of competence or function.

A complex of political regionalization or structuring exists in two dimensions: horizontal and vertical. On a horizontal level structures are coordinate (Fig. 1); that is, they are exclusive in territory (countries or provinces do not overlap) or are differentiated by role. In the latter case they may include the

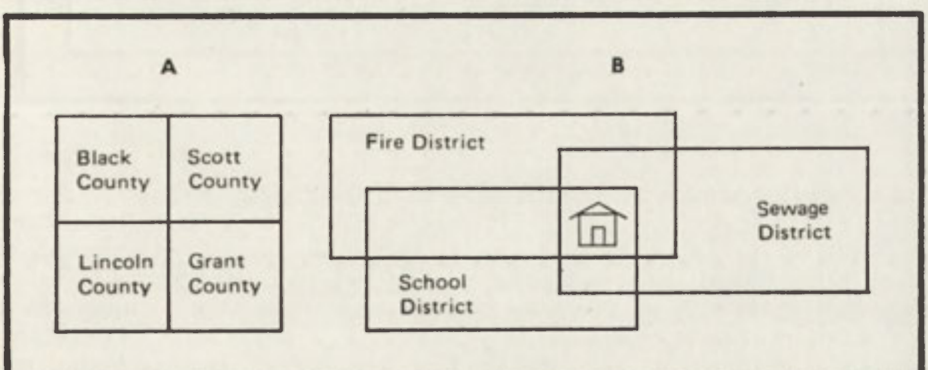


Fig. 1. Horizontal structures

A — no overlap
B — overlap in area but not in function

same territory, but possess different functions (sewage districts and education districts). The fact that general structures of a horizontal level are coordinate does not mean that their boundaries always coincide with political activity fields. If a political system is not coterminous with its allotted space, boundary difficulties may result. On the political map of the world each State is shown as having exclusive control over a territory, while this may not be an accurate reflection of reality. In his functional analysis of States, Hartshorne always investigated whether political control was suffused throughout the entire State area to achieve political integration. Control and integration require transportation and communication facilities and, perhaps even as a final resort, semi-military occupance of remote or rebellious regions. The Brazilian political system, for example, has not been able to integrate effectively the huge Amazon Basin; it remains a vaste territory with a sparse population, both difficult to organize and to control. Similarly, it will require more than an act of Parliament to ensure Canadian sovereignty and effective control over the Arctic Islands astride the Northwest Passage.

Insurgent systems within the political area may challenge a State's ability to integrate and to control its allotted space. This was the case with the Confederacy in America and, more recently, the abortive formation of Biafra in Eastern Nigeria.

On the other hand, a system can overspread its legal space, spilling beyond its boundaries to embrace territory belonging to a neighbour, or territory that has not been effectively integrated. The numerous groups making territorial claims on the Austro-Hungarian Empire before the political map of the area was redrawn at Versailles afford a clear illustration of this problem.

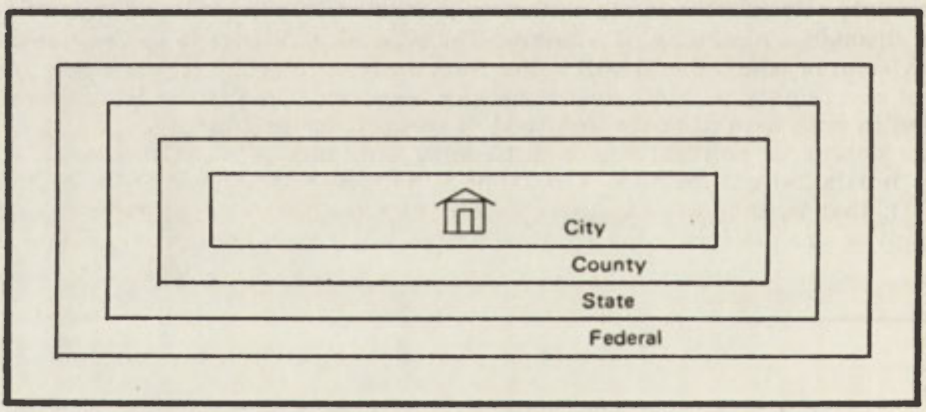


Fig. 2. Vertical telescoping

Such a spillover may take forms other than the strictly territorial. Penetration of the state-area of one political system by another may come about through irredentism or spread of an ideology, or through economic investment. One problem facing Canadians today is the extent to which United States investment in Canadian industry and resource development affects the formulation and implementation of policies designed in Ottawa to promote purely Canadian interests. In a metropolitan area the field of the central large city may cover numerous suburban communities.

Political structures, as noted above, exist not only on horizontal levels, but also telescope in a vertical direction (Fig. 2). An industrial site may be in a city, in a county, in a state or province and in a Nation-State. This constitutes a nested hierarchy, each level of which subsumes lower levels. Most of the inhabitants of this planet belong to a multiplicity of structures both horizontal (counties, sewage districts, airport districts, soil conservation districts) and vertical (cities, counties, states, and so forth). Similarly, any square mile of territory may be incorporated within the bounds of a variety of these political units and structures.

A basic problem inherent in any political system that has a multiplicity of substructures is that the political subunits cannot always properly or effectively cooperate to solve common problems. Contractual cooperation among the states of the United States is expressly forbidden by the United States Constitution (Article One, Section 10, paragraph 3). The collapse of the West Indian Federation was due principally to the inability, for sound historic and geographic reasons, of the constituent states, formed primarily of a number of islands extending over approximately a thousand miles of sea space, to overcome their parochialism and to function as a united whole. That parochialism was not simply a reflection of distance so much as it was an indictment of the old colonial system in which historic associations were principally between each island and mother country and not among the islands.

The complexity of contemporary societal problems requires that attention be given to the spatial patterns of political organization. If political systems are to function effectively to achieve the desired goals (assuming, of course that goals are not in conflict), it would seem that some model might be established. For theoretical purposes at least, the following criteria of territorial organization, at least at the substate level, need to be reconciled:

(1) The territorial jurisdiction of a government should cover and unite both the activity field and the area of any specific problem to be solved. A political structure can function best when all interested parties and the entire area of the problem are included. This is why, on the local level, different governmental structures are created to deal with different problems. The best pattern of sewage districts, based upon topography, watersheds, and type of land use does not always coincide with the best pattern for school districts based primarily upon population and population characteristics. As new problems arise or political systems arise, the pattern of jurisdiction districts should be reorganized to conform to the new activity areas. On the level of States lack of territorial coincidence between areas of jurisdiction and areas of political activity often results in boundary problems.

(2) The territorial jurisdiction should be large enough to have a sufficient resource base to meet societal needs. For some service of local government, for instance, a large area is a requirement for meeting minimum standards of economy and performance; smaller service areas result in high unit costs.

(3) The patterns of jurisdiction of overlapping governments on a horizontal level should be drawn so that the various governments can cooperate in planning or organizing the use of space. Different patterns are best for the provision of different services, but proliferation of overlapping and noncoterminous governments makes coordination in any area impossible. The best sewage and school districts may not be coincident, but a school district should not plan a new school at a site at which adequate sewage facilities cannot be provided, or which the county government wants for an airstrip. The service of public needs is thus thwarted by the confusion which results from superabundant overlapping jurisdiction.

(4) The guarantee of effective popular control over government should be assured. The multiplicity of governmental structures over any area or citizen often endangers the achievement of this goal. In a broader sense what is essential is a sense on the part of the citizenry of participation in the system. The lack of this sense of participation seriously weakens the effective functioning of the system or region.

While these four criteria are individually desirable, it is obvious that when considered collectively, simultaneous and complete achievement of each particular goal may not be possible. The second criterion of an adequate resource base, for instance, is often sacrificed to the idea of local autonomy, which represents, in effect a belief in the primacy of the local political system. This takes place at all levels of government. Local governments in the United States are heavily subsidized by higher levels of government, and many of the new micro-states in the United Nations remain economically viable only with substantial foreign aid. In the development of any governmental framework, therefore, some balance must be reached, some compromise attained among the criteria. That balance must be that which best serves the particular problem, society, and place.

THE LOCUS OF POWER IN THE POLITICAL SYSTEM

Political authority is usually represented as being housed in the capital city, the head office, or in some centralized place. This image is so generally accepted that newspapers often report, "Moscow's reaction to the State Department announcement was..." Most students of international politics would recognize the significance of references to the Kremlin, the Quai d'Orsay (the French Foreign Office), Number 10 Downing Street (the official residence of the British Prime Minister), and a host of other very specific place names.

Each State has a different division of power between the central government and its subunits. A unitary form of government allows considerable concentration of political authority at the centre. The government of England, for instance, is concentrated in London. All subunits of government in England are creatures of and subordinate to the central government. Not all of the political authority for the United Kingdom of Great Britain and Northern Ireland, however, is resident in London. Authority within Northern Ireland until recently was shared by London with Belfast. In the formation of the United States, specific powers were granted to the federal government, while all remaining authority was retained by the states. The states, then, have distinct legal rights within the federal system.

The division of power among subunits can be studied even within a large city. The political framework of the city of Chicago is what political scientists call a "weak mayor" form. The mayor presides over the City Council, but the fifty aldermen, by means of executive committees paralleling the mayor's departments, can intervene in the affairs of the mayor's office. The aldermen are so powerful in their own wards that they are known as Chicago's "fifty little mayors". Obviously, if the mayor is a dominating personality, he can effectively centralize power and control the city.

While the legal division of power among subunits and the location of power can be determined for any level of political organization, the Chicago example well illustrates that the power to influence and to make decisions ultimately lies in the hands of men. Thus the legal forms do not always tell us where and in what manner decisions are really made. While Washington, D. C., remains the official residence of the President of the United States, President Nixon has moved the decision-making around the country to a far greater degree than any previous president by establishing alternate residences in the South and in the West. In the case of a large city, many politically powerful men such as wealthy industrialists who contribute heavily to campaign funds, may live outside the city limits in suburban areas. Important decisions can thus be made not in city hall, but in suburban country club.

SYSTEM. SUPPORTS

While political systems differ in size and complexity, as well as in their goals, they must, in order to survive, have a number of basic supports. These include territory (having a particular size, shape, and location if the system exercises dominium), population (numbers, density, distribution, skills), resources (type, quantity, quality, location, accessibility), and a degree of cohesiveness of society expressed in a common acceptance of institutions, effective communications, and so forth.

Traditionally the political geographer, assessing State power, has been most concerned with compiling an inventory of physical resources; this is of value. but measuring the sheer size of the resources does not provide sufficient insight into the ability of a political system to maintain itself. It is necessary to know whether the system is functionally effective and cohesive, and how it mobilizes its resources. One State's population explosion, for example, may be another's human resource base. Similarly, an inventory of a State's mineral resources does not tell us whether that State has the technology or economic infra-structure to extract and put those resources to work.

The enhancement of the quantity or quality of resources may in itself be a goal of a political system. This may be manifested in territorial aggression and aggrandizement to obtain additional resources, which occurs frequently when one State invades another. But, on a local level, the desire to acquire a larger support base to enhance tax income may be behind the move of a city to incorporate its surrounding suburbs.

The continuing dedication of the population to common goals is one of the firmest and most basic supports of a political system. Without that commitment no system can survive. President Lincoln most eloquently analyzed the Civil War in terms of continuing commitment to the goals of the United States in two famous sentences:

"Fourscore and seven years ago our fathers brought forth on this continent a new nation, conceived in liberty and dedicated to the proposition that all men are created equal.

Now we are engaged in a great civil war, testing whether that nation or any nation so conceived and so dedicated can long endure."

The concept of belonging together shared by all the people is another powerful support of a political system. The political scientist, Karl Deutsch, has sought to identify the processes that promote political and social integration.[9] Among the factors he has identified are the development of exchange economies and of communication grids, the awakening of ethnic awareness, and the acceptance of shared symbols. In the maintenance of a sense of community geo-

[9] See, for instance, The growth of nations: *Some recurrent patterns of political and social integration, World Politics*, 5 (October 1952–July 1953), 168–195; and *Nationalism and social communication* (Cambridge: The M.I.T. Press, 1966).

graphic symbols may play a role that is often overlooked. Some ethologists have suggested that territoriality is as basic an instinct in man as in a number of other animals. Scientists recognize that we each feel and possess personal space around us, but there is yet little concrete evidence to support the contention that groups of men, or whole nations, have a biological urge to possess and defend territory.[10] If we cannot prove that group values toward a shared space are instinctual, we do, however, have ample evidence that such values are propagated by governments or in the folk culture. The school system of each State concentrates on national geography, and a distinct map-image of the State is imprinted in every schoolchild's mind. Geography in the history class can be influential in creating a national image. Surely such geography lessons have political undertones. They may contribute to the political climate for potential territorial claims or boundary disputes. Folk culture supporting a geographic national image includes everything from the song "This Land is My Land" to legends of Johnny Appleseed and Paul Bunyan.

Some commonly held ideas of geography form a part of what the French geographer, Jean Gottmann, calls the "iconography", a shared set of symbols, of a political system.[11] Coming to share in this iconography is part of the political socialization of the citizen, and thus it furthers the integration of the political system.

The supports of any political system are very difficult to measure effectively. Coal and iron resources can be quantified, but it is more complicated to measure the potential of a nation to mobilize such resources and to put them to work. A population can be counted easily, but how can the people's "patriotism," or the real depth of their commitment, be measured? If satisfactory measures of systems' supports (or resources, in the broadest sense of the term) are lacking, this does not necessarily preclude their consideration and appraisal. Inventories tell us little about how resources are brought into the system, but the way in which a political system uses its resources may indicate what is essential to its purpose, or whether the political system may be using those resources as productively as possible.

THE POLITICAL SYSTEM AND ITS ENVIRONMENT

Political systems do not exist in a void, but in a total environment which includes not only the physical setting, but the social and cultural milieu. Political geographers have usually concentrated on the physical environment. As we have seen, the physical resource base does provide some of the material supports of the political system. Obviously, the system must have adequate resources to achieve the goals it seeks and to satisfy the wants of its population.

---

[10] For studies of animal behavior and territory see: Eibl-Eibesfeldt, I., *Ethology: The biology of behaviour* (New York: Holt, Rinehart and Winston, 1970); Howell, F. C. and Bourliere, F. (eds.), *African ecology and human evolution* (Chicago: Aldine Press, 1963); Klopfer, P. H., *Habitats and territories: A study of the use of space by animals* (New York: Basic Books, 1969); Schaller, G. B., *The Mountain Gorilla: Ecology and behaviour* (Chicago: University of Chicago Press, 1963). Man and personal space is discussed in Hall, E. T., *The hidden dimension* (Garden City: Doubleday and Co., 1966) and Sommer, R., *Personal space* (Englewood Cliffs: Prentice-Hall, 1969). Robert Ardrey, however, in his popular *The territorial imperative* (New York: Atheneum Press, 1966) stands almost alone in extrapolating the findings of scholars to suggest territoriality as a biological drive in man.

[11] See, The political partitioning of our world: An attempt at analysis, *World Politics*, 4 (1952), 512–519.

Some historians have gone so far as to argue that certain aspects, features, or elements of the natural environment determine both what man will attempt and what he will achieve through political activity. Such a point of view has been termed environmental determinism. The ancient Greeks thought that civilization could develop only in a mild climate. Arnold Toynbee has reversed this idea with his popular "challenge-response" theory, in which he argues that an unfavorable environment challenges men and thereby calls forth a better and more successful effort.[12] Karl-August Wittfogel has stated that "institutional conditions being equal, it is the difference in the natural setting that suggests and permits — or precludes — the development of new forms of technology, subsistence, and social control." [13] He thus tries to explain the rise of "oriental despotism" in a society based on irrigation agriculture where the climate is dry.

In spite of such arguments, which have also been advanced at times by geographers, it cannot be proven that the physical environment determines what man will undertake or what his course of action will be. An alternate viewpoint recognizes that while the endowment of the physical setting may limit what is possible for a political system to achieve, the natural environment cannot determine what a political system may attempt. Geographers call this point of view environmental possibilism. For instance, without the necessary resources Ghana may nevertheless struggle to build an iron and steel potential, but it probably will not be achieved without heavy imports and substantial cost.

On the other hand how man perceives his environment may certainly have a decided influence on what he attempts to do. Political systems themselves cannot perceive, and so we must speak of the perception or understanding of the decision-makers in any political system. Walter Prescott Webb argued convincingly that one of the greatest difficulties in settling the Great Plains was the formulation of a suitable legal-institutional framework. Since most of the national lawmakers came from the relatively well-watered East, they did not understand that the settlement of a drier area necessitates a different body of laws regarding homestead size and riparian rights. The slow struggle over the formulation of a body of laws suitable to the environmental possibilities offered in the west well illustrates the role of environmental perception in the structuring of a legal system. The settlement of the plains proceeded with the development of a new code of laws.[14]

The total environment of decision-makers, of course, is not only a physical, but a social and a cultural environment as well. The domestic political and social environment in the limited states includes for example, student campus disorders, black unrest, and built into the political system, the scheduling of elections by the calendar every two years rather than on issues, as in a parliamentary system.

The situation of any political system relative to other political territories is another important aspect of a system's environment. A significant factor in Panamanian politics, for instance, is the fact that Panama is an isthmus through which the United States built and operates a canal as the shortest water route "across" the United States. The importance of the situation may be demonstrated in other ways as well. Throughout history a number of political units, such as Switzerland, have been created or allowed to exist in peace as "buffer sta-

[12] Toynbee, A., *A study of history* (London: Oxford University Press, 1954), vol. VIII, pp. 466 ff; 532–533 (1961), vol. XII, pp. 254–263.
[13] Wittfogel, K., *Oriental despotism* (New Haven: Yale University Press, 1957), p. 11.
[14] Webb, W. P., *The Great Plains* (Boston: Ginn and Co., 1931).

tes." A buffer is clearly a result of its situation. Location is always relative, and this relative location is an important part of the environment of any political system.

THE CHALLENGE OF CHANGE

Political reality is an on-going process; political systems, structures, and institutions are created and dissolved. Politics, then, should be viewed as a process with a continuous shifting of power among groups and between institutions and structures as new problems emerge. Governmental institutions are society's response in any particular time and place to the social problems of that time and place. Thus, as the nature of the problems changes, the institutions as a response to those problems ought to change too. Perhaps this is what Thomas Jefferson, a government-builder himself, meant when he suggested that each generation must have its own revolution. The Constitution of the United States remains fundamentally the same as it was fifty years ago, and yet no one would argue today that the division of power between the state and federal governments or even among the branches of the federal government has remained the same. These shifts of power among levels and branches of government have left a variety of political structures with uncertain purpose or viability. The Oklahoma Legislature, in an attempt to increase efficiency, effectively eliminated one whole level of government by withdrawing the right of townships to levy taxes for local purposes.

Transformations of institutions and structures are essential if political organization is to remain responsive to the needs of the people and the time. David Easton, in his book, *A Systems Analysis of Political Life*, esks: "How can any political system ever persist whether the world be one of stability or of change?" He notes in response that political systems have the ability to "regulate their own behavior, transform their internal structure, or even go so far as to remodel their fundamental goals." [15] This last point, the remodeling of fundamental goals, may be viewed as the determination of a new *raison d'etre*.

The institutions and structures established to achieve goals tend to become more rigid than the systems they supposedly represent. Hartshorne has argued that "inevitably a structure that has lost its original *raison d'etre* without evolving a new one, cannot hope to stand..." [16] This would seem to make sense, but it is not always true. The San Ramon Water District in Contra Costa County, California, a local special-purpose government, was formed in 1956 to pass a bond issue. The issue passed, and the water district, levying taxes but no longer serving any corporate functions, continued to exist for another fourteen years. Only in 1970 did a perceptive housewife, studying local government, force the dissolution of the obsolete political structure.

On any level of human organization the formation of a new political system (or subsystem) may be suppressed by another. Thus, in metropolitan areas central city interests often oppose the independent incorporation of suburban areas, preferring themselves to annex these areas to gain their tax revenues. The efforts of colonial peoples to achieve independence is another example of such conflict. In the American Revolution the thirteen colonies felt they could best pursue their own interests by dissolving the bonds that linked them to Great

[15] Easton, D., *A systems analysis of political life*, pp. 15, 19.
[16] Hartshorne, R., *The functional approach in political geography*, p. 111.

Britain. John Adams remarked that the Revolution was decided in the minds of Americans long before it was decided on the battlefield.

New political systems and structures often can arise only on the ruins of old ones. The Austro-Hungarian Empire is an archetype of an old political structure that proved incapable of dealing with the emergence of new political systems. As the various groups within the empire were stirred by nationalist aspirations during the nineteenth century, the empire did not satisfactorily, in Easton's phrase, regulate its behavior, transform its internal structure, or remodel its goals in order for it to survive. The Austro-Hungarian case was aggravated by the fact that many of the emerging nations, such as the Polish, Romanian, Czech and South Slav, could not be satisfied within the territorial framework of the empire. Since large groups of these peoples lived outside the imperial system, or astride its borders, the new national communities that took shape in Eastern Europe did not conform to the existing political map. The satisfaction of national aspirations could be achieved, therefore, only through an international solution, and thus the Habsburg empire was torn apart. The Ottoman Empire met a similar fate for many of the same reasons, and the Nation-State idea with all of its attendant problems, emerged triumphant in Europe from World War I.

As new problems arise, new structures with new jurisdictions must be created to meet public needs. Contemporary political systems and structures on all levels of organization are now confronted with the problem of the deterioration of the human habitat. Existing political jurisdiction seem inadequate to deal with pollution, whether the concern be with a contaminated lake bounded by a number of municipalities or with the pollution of the whole earth's atmosphere by industrial combustion, to say nothing of the threats of nuclear war and fall-out. The spread of urbanization across city, county, and even state lines is another problem demanding the emergence of broader political systems and new jurisdictional areas to conform to changing land uses and regional functions. Such new problems severely challenge existing and traditional administrative and governmental structures. Clearly, the legal and territorial frameworks of some of our cities, states, and even perhaps much-revered Nation-State have outworn and outlived their usefulness. In many instances, they have become obstacles in the way of serving society and the needs of humanity.

University of Washington,
Herbert H. Lehman College

GEOGRAPHIA POLONICA 31, 1975

THE METHODS OF CONSTRUCTING CLIMATIC MAPS OF VARIOUS SCALES FOR MOUNTAINOUS AND UPLAND TERRITORIES, EXEMPLIFIED BY THE MAPS PREPARED FOR SOUTHERN POLAND

MIECZYSŁAW HESS, TADEUSZ NIEDŹWIEDŹ AND BARBARA OBRĘBSKA-STARKEL

INTRODUCTION

The aim of this paper is to present the methods of constructing climatic maps of various scales for mountainous and upland territories worked upon in the last ten years in the Department of Climatology of the Geographical Institute of the Jagellonian University based on investigations carried out in Southern Poland. This paper forms a contribution to the discussion on drawing general and detailed climatic maps useful for estimating the qualities of the natural environment, and solving numerous economic problems, especially those of spatial planning.

In the construction of climatic maps, various guiding criteria are applied, depending on the aim the map is to serve, and on its scale. On macroclimatic maps representing the differentiation of a particular component of the environment over large areas, the mountains are either treated separately by way of distinguishing a different type of climate, but without more detailed precision of its pecularities (e.g. E. Romer, 1949; W. Gorczyński, 1945), or the changing climatic phenomena are treated zonally depending on the altitude above sea level corresponding to climatic zones in the lowlands (e.g. W. Koppen, 1936; C. W. Thornthwaite, 1931), and ascribing to them the same characters of temperature and humidity.

In the macroscale, the maps of particular climatic elements are most often encountered; they are constructed mostly by way of interpolation between the values recorded by the climatological stations. The great differentiation of the climatic conditions in the mountains and the loose network of measurement stations is often the reason why these maps do not show precisely the spatial distribution of these elements of climate in mountainous areas.

In the mesoclimatic scale, there is a great variety of methods in constructing maps, depending on the accepted definition of the mesoclimate, as well as on the scope of the measurement material possessed. With respect to the manner of mapping in the field, these methods may be divided in general into two groups, those of physical measurement and those of biological observation (F. Schnelle, 1963). To the former belong the following methods:

— The instrumental method, the one most commonly used, which consists of determining the differentiation of mesoclimatic conditions based on a dense network of profile measurements, both horizontal and vertical, of the temperature and humidity of air and soil in some representative areas (R. Geiger, 1961; A. Mäde, 1956, A. Baumgartner, 1961; E. Quitt, 1965; I. Gugiuman, 1971 F. Schnelle, 1972).

— The method of thermal balance, elaborated in recent years and assuming the differentiation of its particular components to be the basis of the processes of heat exchange and local air circulation within the particular forms of the territory (J. Paszyński, 1964; B. A. Ayzenshtat, 1971; J. Skoczek, 1971).

— The method of estimating the deviations of the particular elements of the mesoclimate in various forms of the macroclimatic background (I. A. Goltsberg, 1962; A. V. Shakhovich, A. A. Wilkens, 1963; Z. A. Mishchenko, 1962). This method, used by Soviet research workers, is partly derived from instrumental methods and adapts the results of measurements in some representative areas to the morphometric and morphographic characterizations of particular territories.

Most of the biological observational methods are formed by visual observation and information on the differentiation of some phenomena of a local character and detrimental from the point of view of the vegetation, e.g., damage by frost, wind throw, radiation fogs (S. Ziobrowski, 1933; S. Uhlig, 1954; F. Schnelle, 1963; W. Weischet, 1963). The data collected in this way often complement those obtained by means of the physical measurement methods. The mapping of the disappearance of the snow cover and of the therms of some chosen phenological appearances (A. Baumgartner et al., 1956; J. Klein, 1967; B. Obrębska-Starklowa, 1968; H. Dickel. 1972) forms a sort of bridge between these two groups of methods.

In laboratory work, when constructing mesoclimatic maps, cartometric methods are generally used; they consist for example in the elaboration of mathematical models of the distribution of the energy of radiation depending on the exposure and inclination of slopes (W. Stružka, 1954; H. Turner, 1966) or, on the formation of the wind field in territories of a determined type of relief (W. Nägeli, 1971).

The methods of microclimatic mapping are the least developed ones. This is connected with the divergence of opinion on the very essence of the microclimate, and with the necessity of taking into account a considerable number of factors influencing the complexity of climatic elements in the air layer above the ground. Field investigations in this scale are usually carried out during short periods of time, and the results represent the microclimatic differentiation of some types of weather only, which often have been chosen at random. Therefore, any comparison of the material collected in areas of different qualities of the geographical environment is rendered impossible.

It is generally known that mountainous and upland territories are characterized by a very great differentiation of all elements of the geographic environment and thus also of climatic conditions. The differentiation is in the macro-, meso-, and microscale. These regions require a different methodical approach than those used in low-lying areas, both in typology and in regionalization of the climatic conditions occurring in the mountains and uplands.

MACROCLIMATIC MAPS

To learn the macroclimatic differentiation of the Carpathian Mountains the authors used a method (Hess 1965) which, according to them, enables the determination of values for the particular elements and indices of the climate in every place of the vertical profile of the mountains. It consists in the *determination of the interdependency between the altitude above* sea level and the values for the various elements and indices of climate at a simultaneous quantitative determination of the correlations among them. There exist particularly close relationships between the mean annual temperature and the annual values for some other elements of the climate (Fig. 1), and between that temperature and the mean temperatures for the particular months (Fig. 2). Similar close interdependence has been established between the mean monthly temperature and the monthly values for other elements of the climate (M. Hess, 1966). Thus, it has been proved, among others, that the mean annual and the mean monthly temperature are no abstractions, but real, complex climatic phenomena closely dependent on a range of other elements and factors of the climate. Thus, it appears that on the basis of the given value of one climatic element we are able to determine a whole range of other components of the climate. The interdependences established may be given an algebraic form; by means of simple equations we are able to calculate real values for the given climatic elements e.g. on the basis of the mean annual temperature.

Table 1 presents several equations which make possible the calculation of real values for numerous elements and indices of the climate, including mean monthly temperatures, on the basis of the mean annual temperature. Table 2 lists the values for several elements of the climate calculated in this way; they

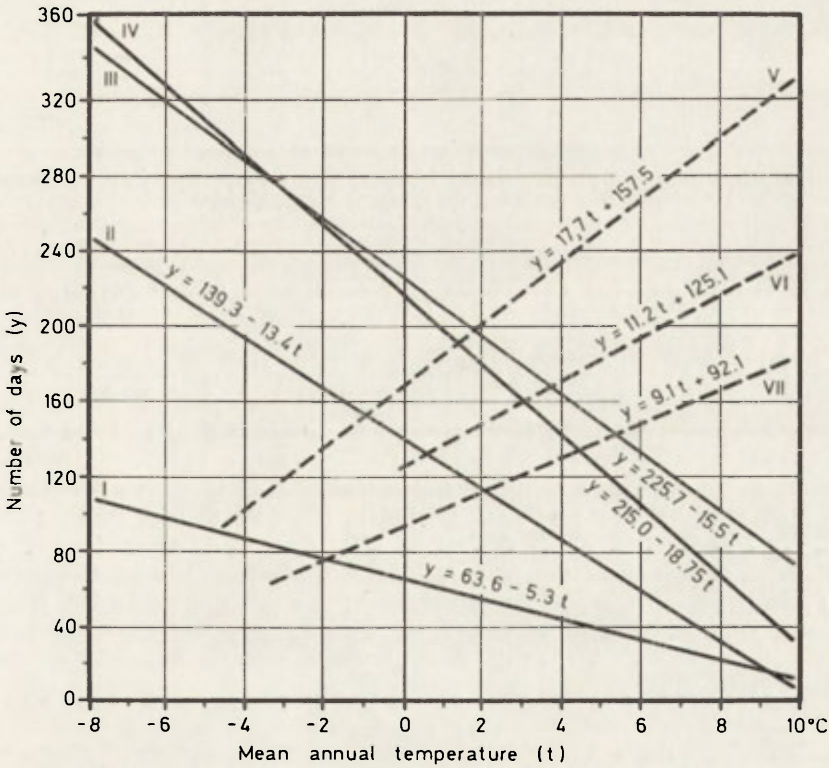


Fig. 1. Correlation between the mean (1952–1961) annual temperature (t) and the number of days (y) with: temperature minimum $< -10°$ (I), temperature maximum $< 0°$ (II), temperature minimum $< 0°$ (III), snow cover (IV), mean diurnal temperature $> 0°$ (V), mean diurnal temperature $> 5°$(VI), temperature minimum $> 0°$ (VII) in the Western Carpathians

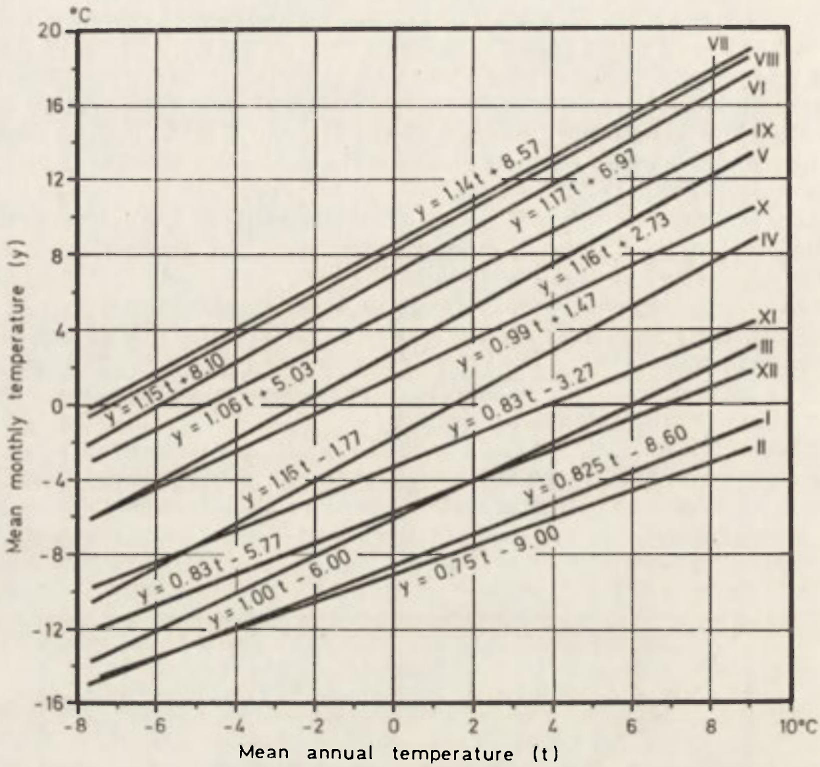


Fig. 2. Correlation between mean annual temperature (t) and monthly temperature (y) in the Western Carpathians (1952–1961)

correspond to the two-grade intervals of the mean annual temperature. We may read in it the values for these elements of climate which correspond to the given mean annual temperature. It is most essential that the dependences presented do not lose their value even at a very great range of the humidity of the climate, since in the vertical profile of the West Carpathians the coefficient of humidity (the precipitation-evaporation ratio) changes ten times and these dependences concern both the foothills and the highest peaks of the Carpathian Mts. This enables the determination of climatic zones in the Carpathians (M. Hess, 1965) on the basis of a complex index linked closely with a whole range of climatic elements. The role of such an index is taken by the mean annual air temperature which has become the skeleton of the whole classification scheme. Based upon the analysis of the interdependence between the altitude above sea level (h) and the mean annual temperature (t) it was established (M. Hess, 1965) that in the vertical profile of the West Carpathians the mean annual temperature changes from +8°C at the foot of the mountains to −4°C on the highest summit of the Tatras. This correlation is linear

$$t = 9{,}231 - 0{,}00498h$$

which gives an average gradient of temperature 0.5°C per 100 m. It is of essential importance that the limits of the particular zones of vegetation are connected with the two-grade intervals of the mean annual temperature. The upper limit of the zone of alpine meadows lies at the annual isotherm of 0–2°C, the upper limit of the dwarf pine at the isotherm of 0°C the upper forest limit

TABLE 1. Straight-line equations $y = ax+b$ and correlation coefficients (r), defining the interdependences between the mean (1952–1961) annual temperature (t) and the other climatic elements and indices (y) in the Western Carpathians

Climatic elements and indices (y)	Equation	r
Absolute annual maximum temperature	$y = 1.46t+24.2$	0.99
Mean annual maximum temperature	$y = 1.10t+3.85$	0.95
Mean annual minimum temperature	$y = 0.84t-3.22$	0.99
Number of days with min. temperature $< -10°$	$y = 63.6-5.3t$	−0.95
Number of days with max. temperature $< 0°$	$y = 139.3-13.4t$	−0.98
Number of days with min. temperature $< 0°$	$y = 225.7-15.5t$	−0.99
Number of days with max. temperature $> 25°$	$y = 5.7t-16.7$	0.95
Dates of last frost change days[1]	$y = 78.7-6.14t$	−0.99
Dates of first frost change days[2]	$y = 3.33t+17.0$	0.96
Duration of period without frost change days[3]	$y = 9.12t+92.1$	0.96
Duration of period with mean diurnal temperature $< -10°$	$y = -26.6t-43.6$	−0.98
Duration of period with mean diurnal temperature $< -5°$	$y = 95.1-13.8t$	−0.99
Duration of period with mean diurnal temperature $< 0°$	$y = 207.5-17.7t$	−0.99
Duration of period with mean diurnal temperature $> 0°$	$y = 17.7t+157.5$	0.99
Duration of period with mean diurnal temperature $> 5°$[4]	$y = 11.23t+125.05$	0.97
Duration of period with mean diurnal temperature $> 10°$[5]	$y = 16.14t+39.3$	0.98
Duration of period with mean diurnal temperature $> 15°$[6]	$y = 31.07t-143.7$	0.96
Number of days with snowfall (in %% of total number of days with precipitation)	$y = 61.9-5.9t$	−0.98
Sum of snowfalls (in %% of total precipitation)	$y = 55.3-5.7t$	−0.98
Number of days with snow cover	$y = 215.0-18.75t$	−0.98
Mean temperature of January	$y = 0.825t-8.60$	0.99
Mean temperature of February	$y = 0.75t-9.00$	0.99
Mean temperature of March	$y = 1.00t-6.00$	0.99
Mean temperature of April	$y = 1.16t-1.77$	0.99
Mean temperature of May	$y = 1.16t+2.73$	0.99
Mean temperature of June	$y = 1.17t+6.97$	0.99
Mean temperature of July	$y = 1.14t+8.57$	0.99
Mean temperature of August	$y = 1.15t+8.10$	0.99
Mean temperature of September	$y = 1.06t+5.03$	0.99
Mean temperature of October	$y = 0.99t+1.47$	0.99
Mean temperature of November	$y = 0.83t-3.27$	0.99
Mean temperature of December	$y = 0.83t-5.77$	0.99

[1] Correlation apply for annual temperature $> 0°$ (counting from April 1 on).
[2] Correlation apply for annual temperature $> 0°$ (counting from September 1 on).
[3] Correlation apply for annual temperature $> 0°$,
[4] Correlation apply for annual temperature $> -2°$,
[5] Correlation apply for annual temperature $> +2$.
[6] Correlation apply for annual temperature $> +5°$.

TABLE 2. Interdependence between mean annual temperature and other climatic elements and indices in the Western Carpathians

Climatic elements and indices	Mean annual temperature (in °C)						
	−4	−2	0	2	4	6	8
Absolute annual maximum temperature	18.4	21.3	24.2	27.1	30.0	33.0	35.9
Mean annual maximum temperature	−0.6	1.6	3.0	6.0	8.2	10.4	12.6
Mean annual minimum temperature	−6.6	−4.9	−3.2	−1.5	0.1	1.8	3.5
Number of days with minimum temperature < −10°	85	74	64	53	42	32	21
Number of days with maximum temperature < 0°	193	166	139	112	86	59	32
Number of days with minimum temperature < 0°	288	257	226	195	164	133	102
Number of days with maximum temperature > 25°	·	·	·	·	6	18	29
Dates of last frost change days	·	·	18 VI	5 VI	24 V	12 V	30 IV
Dates of first frost change days	·	·	17 IX	24 IX	30 IX	7 X	14 X
Duration of period without frost change days	·	·	92	110	129	147	165
Duration of period with mean diurnal temp. < −10°	63	10	·	·	·	·	·
Duration of period with mean diurnal temp. < −5°	150	123	95	68	40	12	·
Duration of period with mean diurnal temp. < 0°	278	243	208	172	137	101	66
Duration of period with mean diurnal temp. > 0°	87	122	157	193	228	264	299
Duration of period with mean diurnal temp. > 5°	·	·	125	148	170	192	215
Duration of period with mean diurnal temp. > 10°	·	·	·	·	104	136	168
Duration of period with mean diurnal temp. > 15°	·	·	·	·	·	43	105
Number of days with snowfall (in %% of total number of days with precipitation)	85	74	62	50	38	27	15
Sum of snowfalls (in %% of total precipitation)	78	67	55	44	33	21	10
Number of days with snow cover	290	252	215	178	140	102	65
Mean temperature of January	−11.9	−10.2	−8.6	−7.0	−5.3	−3.6	−2.0
Mean temperature of February	−12.0	−10.5	−9.0	−7.5	−6.0	−4.5	−3.0
Mean temperature of March	−10.0	−8.0	−6.0	−4.0	−2.0	0.0	2.0
Mean temperature of April	−6.4	−4.1	−1.9	0.5	2.8	5.1	7.5
Mean temperature of May	−1.9	0.5	2.7	5.1	7.4	9.7	12.0
Mean temperature of June	2.3	4.7	7.0	9.3	11.6	14.0	16.3
Mean temperature of July	4.1	6.4	8.6	10.9	13.2	15.5	17.7
Mean temperature of August	3.5	5.8	8.1	10.3	12.7	15.0	17.3
Mean temperature of September	0.8	2.9	5.0	7.1	9.3	11.3	13.5
Mean temperature of October	−2.5	−0.5	1.5	3.5	5.5	7.4	9.4
Mean temperature of November	−6.6	−5.0	−3.3	−1.6	0.0	1.7	3.4
Mean temperature of December	−9.1	−7.4	−5.8	−4.1	−2.4	−0.8	0.9

is in line with the annual isotherm +2°C, the upper limit of the lower montane forest zone — with the annual isotherm of +4°C, and the upper limit of the submontane zone with the isotherm of +6°C. This enabled the authors to distinguish in the vertical profile of the West Carpathian Mts. six climatic zones with which the vegetational zones are linked (Fig. 3). The climatic conditions of each zone may be learned from the educed equations which determine the interdependences between the mean annual temperature and other elements of the climate (Tables 1 and 2).

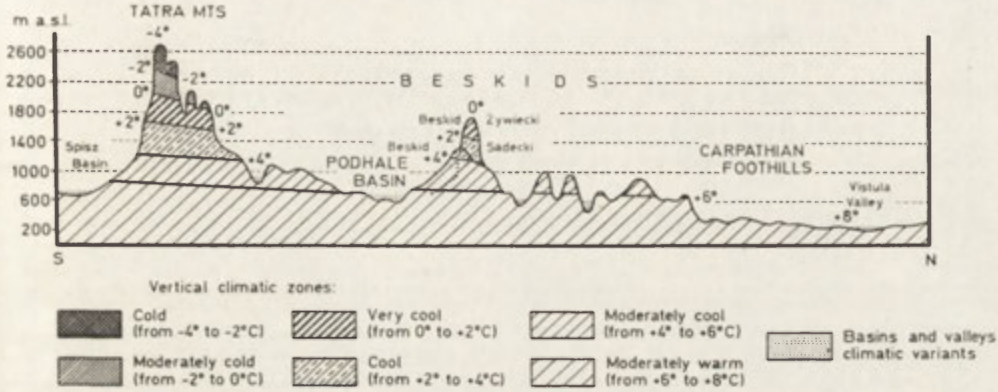


Fig. 3. Vertical climatic zones in the Polish Carpathians

Thus, the method elaborated enables the determination of the climatic relationships in different places and regions on the basis of data easily available and concerning one of the components characteristics of the thermal régime. This method was also used in the determination and characterization of the climatic zones in the East Alps, Sudetes and South Carpathian Mts. (M. Hess, 1968b, 1971).

The described method of characterizing the climatic conditions which prevail in mountains, enables the construction of macroclimatic maps in an easy and objective way. To this purpose it is enough to draw proper isolines on a topographic map, the most suitable being the annual isotherms, and in this way, based upon the knowledge of the quantitative links of the annual temperature with the other elements of the climate (Tables 1 and 2) to characterize in detail the climatic units thus distinguished.

MESOCLIMATIC MAPS

The characterization of the mesoclimatic differentiation conditioned by the particular elements of the configuration of the territory is regarded by the present authors to be the next stage in the process of studying the climate of mountains.

To learn the climatic differentiation of the main elements of the relief of mountains (valleys, basins, ridges, etc.) the authors applied the method of quantitative interdependences between the altitude above sea level and the mean annual temperature, and between the latter and other elements of the climate (M. Hess, 1968a, 1968b). The results of measurements obtained at the climatological stations, which form a dense network in the belt between the Vistula river valley and the Tatras, situated in different forms of the territory and at various altitudes above sea level, enabled the authors to establish that between the altitude above sea level and numerous elements and indices of the climate there exist strict dependences on the convex and concave forms of the territory and on the slopes of different exposure. The strict and linear dependence is shown by the mean annual temperature. Thus, on the basis of equations (Table 3) the mean annual temperature can be determined at every point of the vertical profile of the mountains in any of the elements of the configuration of the territory. It is worth noting that there are different dependences represented for every type of territory, which proves their great impact on the climatic relationships.

TABLE 3. Straight-line equations $y = ax+b$ and correlation coefficients (r), defining the relation between the altitude (h) in metres a.s.l. and the mean (1952–1961) annual temperature (t) on convex and in concave land forms and on slopes of northern and southern exposure in the Polish Western Carpathians (with h varying from 200 to 1700 m a.s.l.)

Land forms	Equation	r
Convex land forms	$t = 8.82-0.00433h$	−0.986
Concave land forms	$t = 9.02-0.00552h$	−0.977
Slopes of northern exposure	$t = 9.24-0.00496h$	−0.994
Slopes of southern exposure	$t = 9.27-0.00441h$	−0.991

Consequently, the authors were able to state that between the mean annual temperature and the annual values of numerous elements of the climate there also exist linear dependences (Table 4) in the convex and concave forms of the relief as well as on the slopes of northern and southern exposure. It thus appears that the climate of the main elements of the configuration of the territory in the whole vertical profile of the Carpathian Mts. can be characterized on the basis of the interdependences established, and that the comparability of the climate prevailing at similar altitudes and at every altitude above sea level is guaranteed, which is of essential importance.

At the same time it has been stated that the climatic differentiation between the particular forms of the configuration of the territory is not the same in the whole vertical profile of the mountains but varies with the rising altitude above sea level. For example, the difference in the number of days with strong frost in the concave forms and on south-exposed slopes at a mean annual temperature of +2°C (i.e. in the alpine part of the Carpathians) amounts to 23 days, while at the annual temperature of +8°C (i.e. the foothills) it falls to 4 days, that means it is almost six times. A reverse tendency exists in the dependence of the altitude above sea level on the length of the period without slight frost in the various forms of the configuration of the territory: at the temperature of +2°C this difference amounts to 11 days if concave forms are compared with south-facing slopes, and at the temperature of +8°C it rises to 31 days, being thus nearly three times greater.

This means that the climatic differentiation occurring among the particular elements of mountainous relief is strictly depedent on the macroclimatic relationships. It seems therefore that by this method the authors have gained a key

TABLE 4. Straight-line equations $y = ax + b$ and correlation coefficients (r) defining the interdependence between the mean (1952–1961) annual temperature (t) and the values of a number of other climatic elements and indices (y) on convex and in concave land forms and on slopes of northern and southern exposure in the Western Carpathians

Climatic elements and indices (y)	Convex land forms		Concave land forms		Slopes with northern exposure		Slopes with southern exposure	
	Equation	r	Equation	r	Equation	r	Equation	r
Mean annual minimum temperature	$y = 1.04t - 4.1$	0.99	$y = 0.96t - 4.7$	0.92	$y = 0.65t - 2.0$	0.94	$y = 0.72t - 1.9$	0.94
Mean annual maximum temperature	$y = 1.12t + 3.4$	0.99	$y = 1.14t + 4.5$	0.97	$y = 1.13t + 3.2$	0.98	$y = 1.0t + 4.5$	0.98
Absolute annual maximum temperature	$y = 1.58t + 23.2$	0.99	$y = 1.64t + 24.6$	0.98	$y = 1.2t + 24.7$	0.97	$y = 1.32t + 24.9$	0.97
Number of days with min. temperature $< -10°$	$y = 60.0 - 5.0t$	−0.95	$y = 81.6 - 7.2t$	−0.90	$y = 53.7 - 3.83t$	−0.83	$y = 52.0 - 4.0t$	−0.84
Number of days with max. temperature $< 0°$	$y = 121.4 - 10.8t$	−0.99	$y = 102.0 - 8.5t$	−0.99	$y = 134.0 - 12.5t$	−0.95	$y - 114.3 - 10.67t$	−0.96
Number of days with min. temperature $> 0°$	$y = 219.0 - 14.0t$	−0.98	$y = 221.8 - 12.6t$	−0.87	$y = 223.3 - 12.54t$	−0.98	$y = 223.4 - 14.2t$	−0.99
Mean dates of last frost change days[1]	$y = 81.6 - 7.8t$	−0.99	$y = 79.0 - 5.0t$	−0.79	$y = 79.7 - 5.83t$	−0.94	$y = 78.3 - 6.17t$	−0,94
Mean dates of first frost change days[2]	$y = 4.3t + 15.4$	0.99	$y = 2.8t + 12.4$	0.92	$y = 2.83t + 19.3$	0.80	$y = 3.0t + 23.0$	0.80
Mean duration period without frost change days	$y = 12.0t + 87.0$	0.96	$y = 6.2t + 91.4$	0.81	$y = 8.67t + 91.7$	0.98	$y = 9.67t + 94.7$	0.98
Duration of period with mean diurnal temperature $< -5°$[3]	$y = 97.0 - 15.0t$	−0.89	$y = 108.5 - 14.75t$	−0.81	$y = 94.5 - 13.25t$	−0.86	$y = 82.0 - 13.0t$	−0.80
Duration of period with mean diurnal temperature $< 0°$	$y = 206.6 - 17.2t$	−0.99	$y = 206.6 - 17.2t$	−0.99	$y = 208.0 - 18.0t$	−0.99	$y = 208.0 - 18.0t$	−0.99
Duration of period with mean diurnal temperature $> 0°$	$y = 17.23t + 158.2$	0.99	$y = 17.23t + 157.9$	0.99	$y = 18.0t + 157.0$	0.99	$y = 18.0t + 157.0$	0.99
Duration of period with mean diurnal temperature $> 5°$	$y = 11.4t + 123.8$	0.98	$y = 11.4t + 123.8$	0.98	$y = 11.2t + 122.6$	0.99	$y = 10.5t + 132.0$	0.99
Duration of period with mean diurnal temperature $> 10°$	$y = 17.3t + 27.4$	0.98	$y = 16.7t + 35.6$	0.98	$y = 14.8t + 47.4$	0.99	$y = 15.3t + 48.4$	0.99
Duration of period with mean diurnal temperature $> 15°$[4]	$y = 30.33t - 134.6$	0.92	$y = 30.33t - 134.6$	0.92	$y = 31.0t - 147.5$	0.98	$y = 31.0t - 147.5$	0.90
Number of days with snow cover	$y = 230.3 - 20.66t$	−0.98	$y = 237.7 - 20.83t$	−0.96	$y = 233.6 - 20.69t$	−0.98	$y = 223.2 - 21.53t$	−0.96

[1] counting from April 1 on [2] counting from September 1 on [3] correlation apply for annual temperature $< +6.5°$ [4] correlation apply for annual temperature $> 5.0°$

to learn the quantitative mesoclimatic differentiation of the elements of the mountainous relief and in connection with the macroclimatic relationships, which change with the growing altitude above sea level.

It is obvious that the interdependences stated make possible the calculation of an *average* mesoclimatic differentiation only, and do not take into account the deviations from the average, conditioned by the differentiated morphometric characters of the investigated forms of the relief. To learn in detail the meso- and microclimatic relationships in some chosen forms of the relief characteristic of the given part of the vertical profile of the Carpathian Mts., the Department of Climatology of the Jagellonian University carried out detailed field investigations.

These investigations which were performed in representative territories corresponding to the fundamental types of the relief in the Carpathian Foothills (T. Niedźwiedź, 1967, 1968, 1973), and in the Holy Cross Mts. (T. Niedźwiedź, B. Obrębska-Starkowa and Z. Olecki, 1973) made it possible to learn the regularities of the mesoclimatic differentiation in stream valleys. It is a generally known phenomenon that in territories distinguished by a variegated relief there occurs an inversional distribution of air temperature at night in the depressions. Therefore, the differentiation of thermal and humidity relationships depends on the relative elevation above the bottom of the valley or basin, which offer conditions (flat or slightly inclined surface of bottom) suitable for the stagnation of cool and moist air.

Thus, while the altitude above sea level is a factor deciding upon the macroclimatic differentiation in the mountains, it is the relative elevation above the bottom of the valley or basin which influences the mesoclimatic differentiation. This fact gives authority to the conclusion that the most important guiding indices which should be taken into account in the work of constructing mesoclimatic maps are as follows: minimum air temperature or mean temperature at night, day-and-night amplitude of air temperature, frequency of occurrence of slight frosts, and air humidity at night. The impact of the relief upon the other indices of temperature and humidity is smaller as the latter change under the influence of the exposure of the territory and the altitude above sea level.

Based upon the investigations carried out on the slopes of valleys it has been stated that up to an elevation of 150–200 m above the bottom of concave forms the distribution of the minimum temperature is inversional, and the greatest gradients occur in the lower part of the profile. This is well illustrated by the data obtained in the area of the foothills of the Raba river valley which is distinguished by denivelations of the order of 100 m. In these particular elevations, the deviations of the mean annual minimum temperature (in the years 1967 and 1968) from the values at the valley bottom were as follows:

Relative elevation, m	0	5	10	20	50	100	150
Deviation of minimum temperature (centigrades) (°C)	0.0	+0.6	+1.0	+1.5	+1.9	+2.5	+2.8

A similar distribution of the index discussed in the altitudinal profile of valleys was found in other mountainous areas (H. Aulitzky 1967, 1968; B. Obrębska-Starklowa, 1969, 1970), and on upland (R. Geiger, 1961; H. G. Koch,

1961; F. Schnelle, 1963; T. Niedźwiedź, B. Obrębska-Starklowa and Z. Olecki, 1973). This differentiation of thermal relations occurs on the background of the macroclimatic differentiation conditioned by the change in the altitude above sea level. The alterations in the mean annual minimum temperature connected with the altitude above sea level in the West Carpathians are illustrated by the straight line equations determined for the forms showing extreme morphological conditions (M. Hess, 1966, 1967, 1968a, 1969):
for convex forms (above the limit of inversion)

$$t_{\min} = 5.10 - 0.0045h$$

$$r = -0.99$$

for the bottoms of concave forms

$$t_{\min} = 3.96 - 0.0053h$$

$$r = -0.99$$

These equations form the basis for the construction of a nomograph which enables the calculation of the mean annual minimum temperature for any place in the Carpathian Mts. depending on its situation in relation to the bottom of the nearest concave form of the configuration of the territory which has conditions suitable for the stagnation of cool air at night (Fig. 4). The use of that

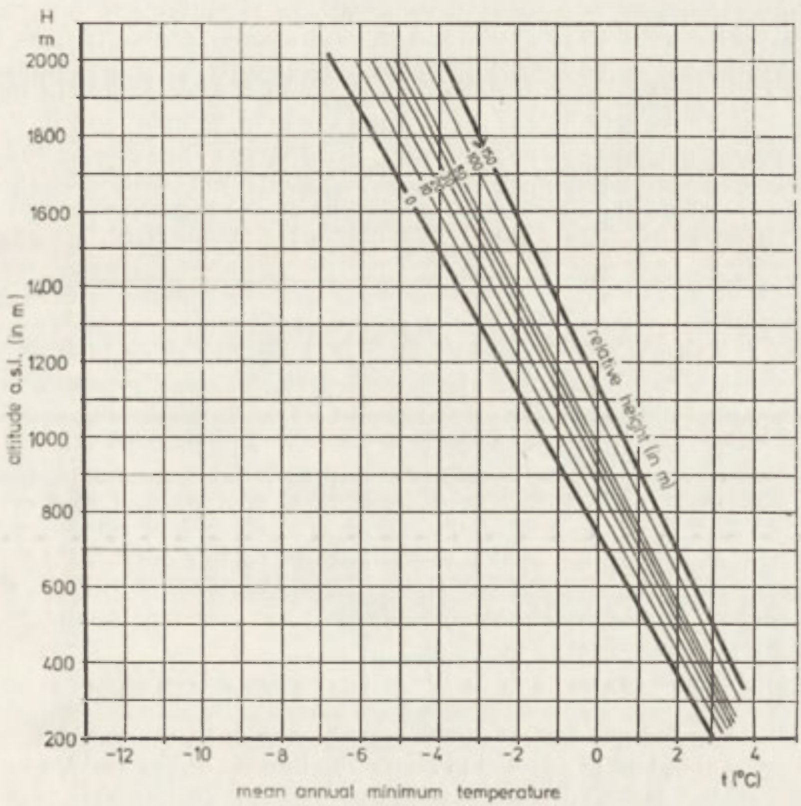


Fig. 4. Nomograph for determining the mean annual minimum of temperature in relation to the altitude a.s.l. and the relative height above the valley bottom (according to formulas given by M. Hess)

nomograph enables the drawing of detailed maps of the isotherms of minimum air temperatures. Such a differentiation of the minimum temperature is reflected in the spatial distribution of the frequency of both the occurrence of slight frosts and the length of the frostless period. For example, the differences in the frequency of occurrence of slight frosts in the Raba river valley exceed 30 days yearly, and those in the length of the frostless period amount to 60 days.

The changes in air temperature described above as connected with the relative elevation, are also linked with the changes in its humidity. The latter show at night the influence of the relief, because the bottoms of valleys and basins are the places of the greatest air humidity manifested as the radiation fogs which fill this kind of territory. The mean annual differences in the area of the Carpathian foothills at 9 p.m. attain 11% of the relative humidity, and 2 mb in the air saturation-deficit.

The studies on the stratigraphy of temperature and humidity on the slopes of the Carpathian valleys enabled the authors to distinguish the typological mesoclimatic units as follows:

I — The lowerings in valleys with the greatest gradients of temperature and humidity. The authors propose to determine the conventional limit of the areal ranges of that type of mesoclimate to be equal to 2/3 of the value for the intensity of inversion of minimum temperature (i.e. the difference between the minimum temperature at the valley bottom and the temperature at the upper limit of inversion). In the area of the Carpathian Foothills and low mountains the range varies from 40 to 60 m, and in medium mountains it rises to about 80 m above the valley bottom. The mean annual difference in minimum temperatures amounts in that layer to about 2°C, which corresponds to the differentiation of the days with slight frost to the order of 25 days, and in the length of the frostless period — to the order of 30–50 days. This also is the area in which radiation fogs occur most frequently. The differentiation being so great it is necessary to distinguish several sub-types of the mesoclimate.

Ia — Valley bottoms of a relative elevation up to 5 m with a relief typical of the foothills, and about 20 m in that of low and medium mountains. These are bottoms of reservoirs of cool air with the lowest minimum temperature, the highest day-and-night amplitude of temperature, and the greatest frequency of occurrence of slight frost and considerable air humidity at night.

Ib — Terraces of the central and lower parts of slopes, typical of the foothills and uplands, up to a relative elevation of 5 to 20 m, having milder mesoclimatic conditions than those of the valley bottoms in subtype Ia.

Ic — Valley sides, pronouncedly warmer and dryer at night, situated at elevations from 20 to 40–60 m in the foothills, uplands and low mountains, and 20 to 80 m in medium mountains, showing deviations of minimum temperature from that prevailing at the valley bottom to the order of 1.5–2°C in a yearly scale. If compared with that lasting at the bottoms of valleys, the frostless period is prolonged by about 15–30 days in the lower part, and 30–50 days in the upper parts of these slopes.

II — Slopes and lower ridges with optimum mesoclimatic qualities (corresponding to the "thermal belt on slopes") situated at altitudes exceeding 40–60 m above the bottoms of the valleys in the foothills, on uplands, and in low mountains, and from about 80 to 300–400 m above the bottom of the valleys in medium mountains. In the areas within the range of that typological unit of the mesoclimate there occurs most often the upper limit of the inversion of minimum temperature. Thus, this is the area in which the day-and-night oscillations of temperature are the least and the frostless period the longest.

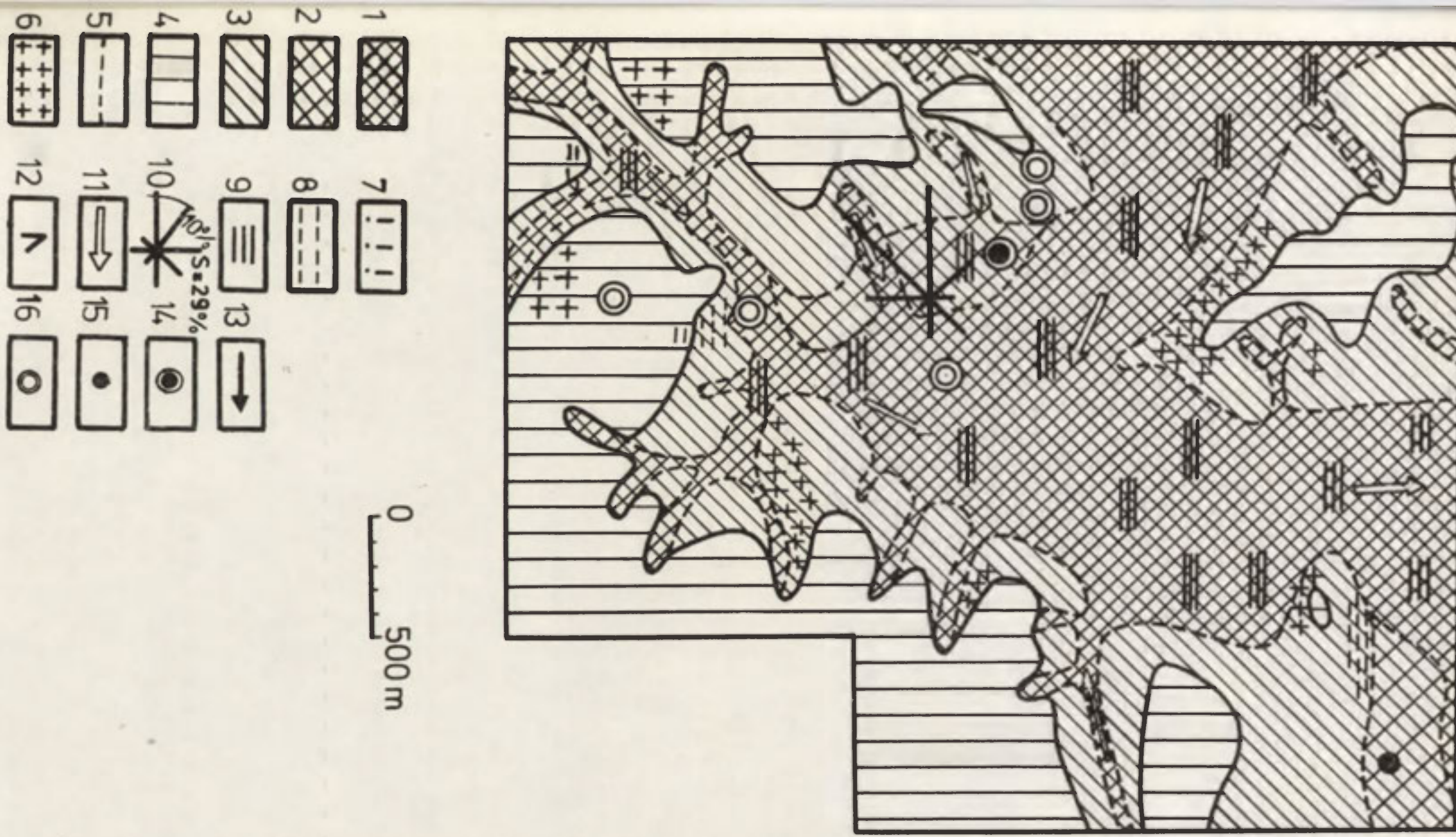


Fig. 5. Mesoclimatic map of the Raba river valley in the Carpathian Foothills

1 — Mesoclimate of valley bottoms (Ia), 2 — Mesoclimate of the middle and lower parts of valley sides (Ib), 3 — Mesoclimate of higher parts of valley sides (Ic), 4 — Mesoclimate of slopes and lower ridges characterized by the most favourable qualities of air temperature and humidity (thermal belt on slopes II), 5 — Limits of the sub-types of mesoclimate, 6 — Areas of great sums of relative insolation (over 110% in relation to horizontal surface), 7 — Areas with small sums of relative insolation (below 90% in relation to horizontal surface), 8 — Stated damage by frost, 9 — Areas with the longest duration of radiation fogs, 10 — Wind roses, 11 — Directions of the cold air flows during radiation nights, 12 — Wind throws, 13 — Flag-trees (as index of the direction of the prevailing strong winds), 14 — Climatological stations, 15 — Pluviometer stations, 16 — Points of mesoclimatic measurements.

III — Ridges, summits and slopes in low and medium mountains situated at elevations exceeding 300–400 m, above the bottoms of valleys, cooler and situated beyond the influence of the local circulation in the valley. The values of air temperature change here depending on the altitude above sea level. Although the mean minimum temperatures may have similar, and at high altitudes above sea level even lower values than those in unit I, the essential difference in respect of the latter lies in the existence of normal (and not inversional) gradients of temperature, and in the much lesser day-and-night oscillations in air temperatures.

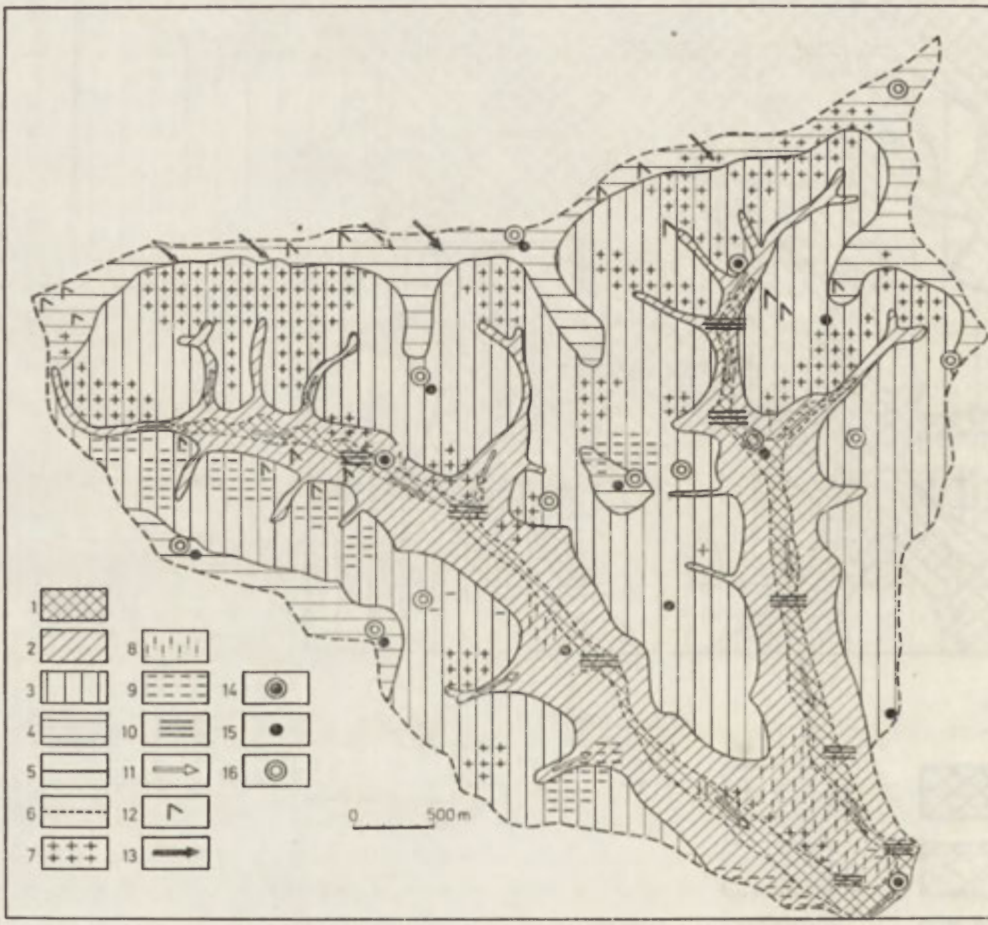


Fig. 6. Mesoclimatic map of the Jamne i Jaszcze valleys in the Gorce range (medium mountains)

1 — Mesoclimate of valley bottoms (Ia), 2 — Mesoclimate of higher parts of valley sides (Ic), 3 — Mesoclimate of slopes and lower ridges characterized by the most favourable qualities of air temperature and humidity (thermal belt on slopes II), 4 — Mesoclimate of higher and colder ridges, slopes and summits beyond the influence of the local air circulation within valleys (I ridges, slopes and summits beyond the influence of the local air circulation within valleys (III), 5 — Limits of the types of mesoclimate, 6 — Limits of the sub-types of mesoclimate, 7 — Areas of great sums of relative insolation (over 110% in relation to horizontal surface), 8 — Areas with small sums of relative insolation (below 90% in relation to horizontal surface), 9 — Stated damages by frost, 10 — Areas with the longest duration of radiation fogs, 11 — Directions of the cold air flows during radiation nights, 12 — Wind throws, 13 — Flag-trees (as index of the direction of the prevailing strong winds), 14 — Climatological Stations, 15 — Pluviometer stations, 16 — Points of mesoclimatic measurements.

The areal distribution of the units mentioned above form the background of the mesoclimatic map (Figs. 5 and 6). This map may be complemented with additional information obtained either on the basis of an analysis of a topographic map (relative insolation calculated by means of the method applied by V. Struzka 1954, H. Turner 1966), or on the basis of observations carried out in the field (scope of damage by frost and rediation fogs, trends of the cold air flow, occurrence of windthrown and flag trees, etc.). On the mesoclimatic maps of industrialized territories there should be submitted the data concerning the range of transgression of the admissible norms of air pollution. It also is advisable to mark the network of climatological stations in which the mesoclimatic measurements were executed. At the climatological stations wind roses should be marked.

The map described above, which should be prepared to the scale of 1 : 25,000 and may be published to those of 1 : 50,000 or 1 : 100,000, presents most important data concerning the thermal conditions of the air. Such elements as slight frosts, radiation fogs, insolation of the territory, and wind are taken into account in the map and make it useful for the purposes of country planning, both urban, agricultural and sylvicultural, in mountainous territories. A map of this type may also help in planning and organizing detailed work and mesoclimatic and microclimatic field investigations. It also forms a basis for the preparation of general maps to the scales of 1 : 200,000 and 1 : 500,000.

MICROCLIMATIC MAPS

The microclimatic relationships, i.e. all of the climatic factors in the air layer lying closely above the ground in a small territory, are influenced by the character of the substratum, i.e. the distribution of the plant cover, soils, surface waters, and microforms of the relief.

The microclimatic differentiation in the mountains depends not only on the season of the year and weather conditions but also proves to be considerably influenced by the altitude above sea level, and this influence varies according to different weather conditions and seasons of the year. Moreover, the microclimatic invsetigations carried out simultaneously in the whole vertical profile of the West Carpathian Mts. showed that in the particular climatic zones the differentiation of the gradients of temperature in the air layer above ground and in soil varies (M. Hess, 1969, 1970).

A cartographic representation of the microclimates on detailed maps to the scales of 1 : 10,000 and 1 : 5,000 requires the execution of a series of field measurements in clear and windless weather, both at the stationary and "en route" measurement points forming a dense network. The situation of the measurement localities should take into account the most characteristic differences in the distribution of the fundamental plant communities (forest associations included), microforms of relief, types of soil, and humidity of the substratum. Therefore, it is rational to carry them out in the areas for which phytosociological, edaphic, and land utilization maps have already been prepared. In the preliminary survey of the territory, before undertaking these measurements, the disappearance of the snow cover, and eventually, the terms of phenological appearances should be mapped.

In the typology of the microclimatic relationships in the various parts of the Carpathians representing the types of medium mountains (Gorce range), low mountains (Beskid Niski range), and uplands (Carpathian Foothills) the characters of the day-and-night course of air temperature have also been taken into consideration (B. Obrębska-Starklowa, 1969, 1974). The author discussed the areal distribution of the values for the differences in the mean temperatures of day and night and of the amplitudes of day-and-night temperatures. The latter index is especially suitable for forming a common basis of determining the types of both meso- and micro-climates (B. Obrębska-Starklowa, 1974) because its values are influenced in an equal measure by the maxima of temperature depending for example on the exposure of slopes, width of the valley, type of the plant community, and by the day-and-night minima connected with the stratification of the air at night in a territory with a variegated relief. On the other hand, such indices as the length of the frostless period, or the mean minimum temperature, fail here.

The forest communities typical of the particular climatic zones exercise a varied influence upon the thermal relationships of the air layer above ground, mostly by the composition of the stand, and within it by the changing percent age of coniferous and broadleaved trees, as well as by the horizontal and vertical structure of these communities.

On the slopes overgrown with a forest uniform as regards its structural character there develops an inversional distribution of air temperature (B. Obrębska-Starklowa 1969), much the same as on the slopes devoid of forests, and the prolongation of the frostless period, passing from the valley bottom along the slopes towards the ridges, shows a similar tendency as in an open space. However, at the same altitude of the measurement points above the valley bottoms, the duration of that period may be shorter in the forest than on the meadow (transition formation from *Piceetum abietetosum* to *Fagetum carpaticum, Tilio--Carpinetum* association, B. Obrębska-Starklowa, 1969, 1970), or it may be similar in the two communities (*Dentario glandulosae-Fagetum*; B. Obrębska--Starklowa, 1974). This depends on the structure and composition of the given stand and on the origination of the first and last slight frosts. On the whole, in all types of forest associations there follows a retardation of the date of the first slight frost by 1–3 weeks. In spring, in the stands with broadleaved species prevailing or with a weakly developed ground flora in which air passage is rendered possible, the slight frosts end earlier. In coniferous forests with a well developed layer of shrubs the air stagnates; therefore, the last date of slight frosts may be retarded in relation to those in the open spaces. As is seen from the above, the length of the frostless period depends on several factors, and for that reason it cannot serve as a uniform guiding criterion in microclimatic typology.

A comparison of the data provided by the stations in forests and of those in open spaces show during the whole year the modifying influence of the forest upon the value of the day-and-night amplitude of temperature in the above--ground layer, that is in the bottom part of the stands there occur less day-and--night thermal oscillations. The screening influence of the crowns of trees is marked in all forest communities by lower maxima of temperature, which is particularly emphasized by the data on *Dentarioglandulosae-Fagetum*, an association distinguished by a dense canopy and a developed ground flora, or on *Alnetum incanae*. An even course of day-and-night temperature during the year is characteristic of the communities in which the minima of temperature exceed those in the open space (transition patch from *Picęetum abietetosum* to *Fagetum carpaticum*), in which the tree crowns counteract the radiation from the secondary active surface, and there is no possibility of air stagnation at the bottom of the forest. Such conditions are also offered in summer by the *Tilio-Carpinetum* and *Querco-Carpinetum* associations. On the other hand, in winter, these two associations show equal or slightly higher minima (up to 0.5°C) than those recorded on the adjacent areas devoid of forests. The higher day-and-night amplitude in April, just before the development of leaves, is very characteristic of the lime-hornbeam and oak-hornbeam stands, which is evident in the rise of the maxima in relation to the open space, in consequence of the heating of a greater mass of trees.

The microforms of the territory are a further differentiating factor influencing the thermal relationships in the above-ground air layer. The widenings in the valley bottom are distinguished by strong heating in the daytime, unlike the narrow bottoms of the shadowy gorges. However, the lowest temperatures in the 24-hour-period are approximate. The differences in the compared values of the day-and-night amplitude of temperature recorded at the localities enable the distinction of two microclimates: (1) one prevails in the widening of the valley bottoms which are strongly heated during the day, while at night they form reservoirs of the cold, thus showing the greatest thermal contrasts in the scale of the 24-hour-period; (2) the other occurs in the narrowing of valley bottoms, which are shadowy and less heated; at night, there accumulates in them

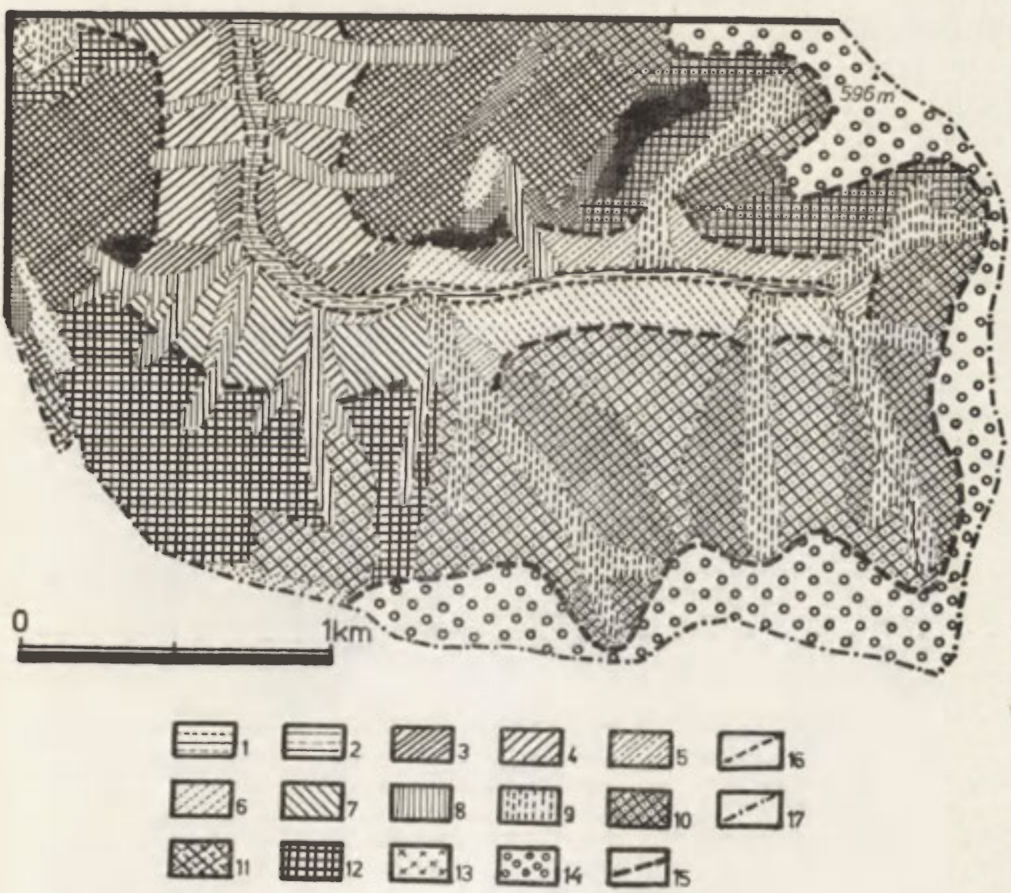


Fig. 7. Example of differentiation of meso- and microclimatic conditions on the border between the Carpathian Foothills zone and the lower mountain (Beskid Niski range)
Explanation of signs in Table 5

TABLE 5. Typology of meso- and microclimatic conditions at Szymbark (some selected examples of units taken from the table of explanations)

Type of mesoclimate	Subtype of mesoclimate	Type of microclimate	Subtype of microclimate	Signature on map
I. Depressions in valleys with greatest day-and-night oscillations of air temperature and humidity e.g. of a most contrasting character of mesoclimatic conditions. Average frostless period 169–189 days. Mean daily range of air temperature 13–17° Mean daily range of air saturation-deficit 15 mb	A. Valley bottoms and flood terraces up to 20m high, with greatest temperature and humidity contrasts. Frostless period 169–179 days. Mean daily range of temperature 15–17°. Mean daily range of air saturation-deficit 16.5 mb		a_1. Cool air reservoirs in open widenings of valley bottoms, with the greatest daily oscillations of air temperature; mean daily range of temperature 16.0–17.2°	1
			a_2. Cool air reservoirs in closed valley bottoms widenings, with impeded flow of cooled air at night, strongly heated during the day, mean daily range of temperature 15.8–16.5°	2
	B. Lower parts of slopes, up to 40 m high with decreasing daily oscillations of temperature and air saturation-deficit, starting from valley bottom upwards. Frostless period 180–190 days. Mean daily range of air temperature 13–15°. Mean daily range of air saturation-deficit 15–16.5 mb	b. Warmer slopes with favourable (sums exceeding 108%) and fairly favourable (sums 100–108%) relative insolation conditions. Mean daily range of air temperature 14.5–15.5°	b_1. The warmest frostless south-exposed slopes inclined at 6° and east and west-exposed slopes inclined at more than 11°; mean daily range of air temperature 13.5–15.0°	3
			b_2. Warm forestless south-exposed slopes, inclined less than 6° and east- and west-exposed slopes, inclined less than 11°; mean daily range of temperature air 14.7–15.0°	4

	c. Cooler and more humid slopes with not too favourable relative insolation conditions (sums below 100%). Mean daily range of air temperature 13–15°	c_1. Afforested south-, west- and east-exposed slopes, with milder daily air temperature oscillations; mean daily range of air temperature 14–15°	5
		c_2. Wooded patches on south-, west- and east-exposed slopes with lowered air temperature maximum values as compared with open spaces; mean daily range of air temperature 14.5–15.0°	6
		c_3. Forestless slopes with northern exposure and inclination below 17°, cooler; mean range of air temperature 13.5–14.0°	7
	d. Cool air flow channels in valley and landslide headwalls on slopes	d_1. Forestless air flow channels with greater daily oscillations of temperature and air humidity; mean daily range of temperature above 17.5°, mean daily range of saturation-deficit ca 15 mb	8
		d_2. Afforested flow channels with softened daily ranges of temperature and air humidity; mean daily range of temperature 14.5°, mean daily range of saturation-deficit about 14 mb	9
II. Warm and dry slopes and lower ridges of the Carpathian Foothills with optimum temperature and air humidity conditions. Average frostless period 190–199 days. Mean daily range of air temperature on open slopes 6–13.5° on ridges up to 14.5°. Mea daily range of air saturation-deficit 7–15 mb, in forests 4.5–-7 mb	c. Warm east- and west-exposed slopes with favourable (sums above 108%) and fairly favourable (sums 100–108%) insolation conditions. Mean daily range of air temperature 8–13°; mean daily range of air saturation-deficit 10–14 mb	c_4. Compact forest complexes on slopes with eastern and western exposure with softened daily oscillations of air temperature and humidity; mean daily range of air temperature 8–9°; mean daily range of air saturation-deficit 5 mb	10
		c_5. Wooded patches on east- and west-exposed slopes with softened daily maximum values of air temperature; mean daily range of air temperature 9.5–-10.5°	11

Table 5 Cont.

	d. Moderately warm north-exposed slopes with sums of insolation below 100%. Mean daily range of air temperature 11.5°. Mean daily range of air saturation-deficit 7–10 mb	d_1. Moderately warm forestless north-exposed slopes, inclined less than 17°; mean daily range of air temperature 11–11.5°	12
		d_2. Forest glades on north-exposed slopes with greater daily oscillations of air temperature; mean daily range of air temperature 12–13°	13
III. Upper parts of slopes on insular mountains of the Beskid Niski range, within the range of advection. Frostless period less than 200 days. Mean daily range of air temperature 6-11°. Mean daily range of air saturation-deficit about 7 mb.			14

Signature 15 denotes the range of mesoclimatic types.
Signature 16 denotes the range of mesoclimatic subtypes.
Signature 17 denotes the limit of the part of the area investigated at Szymbark.

Note: The values given in the table above are only guiding character; the differentiation of frostless period is cited according to 1969; daily ranges of air temperature and saturation deficit concern mainly the results of microclimatic investigations at radiation weather type in July 1968.

equally cool air, but considering the lower maxima, the narrowing is distinguished by a "milder" course of temperature during the 24-hour-period. The slopes of landslides represent a highly varied type of microclimatic conditions because there occur on them microforms of different size, both convex (hummocks) and concave (depressions), and also cuttings.

It is beyond any doubt that while considering the results of a shorter series of observations, not multiannual or annual, the usefulness of the day-and-night amplitude of air temperature as a uniform and guiding criterion of mesoclimates and microclimates acquires still greater importance. An example of its utilization for these purposes is shown in the maps of meso- and microclimatic conditions in medium-height mountains and at the boundary of low mountains and foothills (Fig. 7, Table 5 B. Obrębska-Starklowa, 1969, 1974). They are constructed on the basis of the several-day-series of microclimatic observations during sunny and windless weather. Besides the day-and-night amplitude, the typological characterization comprises the mean day-and-night temperature, mean temperatures of the day and night, or the length of the frostless period.

The selected areas of detailed investigations represent mainly the climatic conditions of one vartical zone: a moderately warm one (Szymbark), and a moderately cool one (Gorce). With this background the authors tried to work out a uniform meso- and microclimatic typology, in which they utilized all the indices proposed by them to delimitate the structural units of the climate.

FINAL CONCLUSIONS

The methods of typology and construction of climatic maps in the various scales elaborated in the Department of Climatology of the Geographical Institute of the Jagellonian University and discussed in the present paper are based mainly on thermal indices which show a close relationship with other elements of the climate.

It is the mean annual air temperature which has proved especially suitable for macroclimatic studies on the vertical zones of mesoclimatic conditions in the mountains because it changes with the altitude above sea level according to a straight-linear function. The whole complex of the climatic elements and indices which are closely linked with the mean tend to be altered accordingly.

While passing to meso- and microclimatic elaborations, which take into account the influence of the relief and the varied substratum, we must look for more "susceptible" indices. This role is best played by the characterization of the day-and-night course of air temperature. In the work on the typology of mesoclimates, this characterization includes the indices of the stratification of air temperature at night in the altitudinal profiles of large concave forms of the territory, i.e. the minimum air temperature, mean temperature at night, number of days with frost and slight frost, dates of the beginnings and ends of slight frosts, and the length of the frostless period. Air humidity shows a lesser differentiation during the 24-hour-period but also is pronouncedly manifested during the night hours. One may pass from the determined macro- to mesoclimatic influences by seeking the dependences of these indices on the altitude above sea level and, simultaneously, on the relative elevation above the bottom of the valley.

In microclimatic typology, the role of the guiding index may be played by the day-and-night amplitude of temperature in the above-ground layer. It reflects the influence of the heating of various substrata and microforms of the territory during the day as well as the impact of the cooling of the surface by radiation on the distribution of temperature in the above-ground air layer. If the typology of the meso- and microclimates is elaborated simultaneously, the day-and-night amplitude can become the guiding index which enables the creation of a division of the climatic conditions uniform to its structure.

Macroclimatic maps should be based upon data collected during many (at least 10) years by the greatest possible number of stations distributed in the altitudinal profile of the whole mountainous territory investigated. The map of the areal distribution of the mean annual temperature showing the limits of the particular climatic vertical zones should be complemented by a table representing the data on the values of the elements and indices of climate corresponding to the determined values of the typological guiding index. Such maps may be constructed starting with a 1 : 500,000 scale.

Mesoclimatic maps should also be based on multiannual data of temperature and humidity of air provided by a dense network of observation points distributed in the transversal profiles of some selected valleys with consideration of the particular elements of the relief. An analysis of results, performed separately for the stations situated at the bottom of the concave forms of the territory and for those lying on convex forms of the relief, allows the determination of the scope of the mesoclimatic differentiation in the whole altitudinal profile of the given massif of mountains.

Next, there is the stage of searching for the regularities in the changes of the minimum temperature depending on the elevation above the bottom of valleys and basins. To that purpose, use can be made of both the data provided by the network of existing stations and of those obtained by a shorter series of measurements executed on the sides of some chosen valleys. The construction of adequate nomographs enables the preparation of mesoclimatic maps based on topographic maps.

The preparation of microclimatic maps should be preceded by field investigations of the differentiation of the character of the substratum in the sections typical of the given territory, and by the establishment of a network of points in which day-and-night gradient and *en route* measurements would be executed in the above-ground air leyer. The results concerning the temperature, air humidity and wind should next be compared with those obtained at the stationary mesoclimatic measurement stations. The analysis of the areal differentiation of the conditions of temperature and humidity, and the determination of the areal ranges of the particular typological microclimatic units should proceed with reference to the phytosociological, edaphic and land utilization maps. A transposition of the results of measurements to adjacent areas is possible if there is a suitable selection of stations.

The contents of all the maps discussed above can be utilized in spatial planning; the macroclimatic maps can serve as guides in the scale of large regions, e.g. they enable the determination of areas having suitable snow-cover conditions for winter sports. The meso- and microclimatic maps find application in the various fields of economy, especially in building, urban, agricultural and sylvicultural planning. Based upon these maps, one can construct other climatic maps to help the calculation of the yield.

Jagellonian University

BIBIOGRAPHY

Ayzenshtat, B. A., 1971, Metody i niekotorye rezultaty izucheniya teplogo balansa gornykh rayonov, *Zesz. Nauk. UJ*, 244, *Prace Geogr.*, 26, pp. 365–376.

Aulitzky, H., 1967, Lage und Ausmass der "warmen Hangzone" in einem Quertal der Innenalpen, *Ann. Meteor., N. F.*, 3, pp. 159–165.

Aulitzky, H., 1968, Die Lufttemperaturverhältnisse einer Zentralpinen Hanglage, *Arch. Meteor., Geoph., Biokl.*, ser. B, 16, pp. 18–69.

Baumgartner, A., 1961, Die Lufttemperatur als Standortfaktor am Gr. Falkenstein (Bayer Wald), *Forstwiss. Zbl.*, 81, pp. 18–47.

Baumgartner, A., Kleinlein, G., Waldman, G., 1956, Forstlich-phänologische Beobachtungen und Experimente am Gr. Falkenstein, *Forstwiss. Cbl.*, 75, pp. 290–303.

Dickel, H., 1972, Methodologische Aspekte praktischer Galändephänologie, in: *Mesoklima v komplexu zivotniho prostředi*, Brno, pp. 7–23, *Studia Geogr.*, 26.

Geiger, R., 1961, *Das Klima der bodennahen Luftschicht*, Braunschweig.

Goltsberg, I. A. (Ed.), 1962, *Mikroklimat kholmistogo relyefa i yego vliyaniye na selskokhoziaystvennye kultury, Leningrad.*

Gorczyński, W., 1945, *Comparison of climate of the United States and Europe with special attention to Poland and her Baltic coast*, Polish Institute, 7 p.

Gugiuman, I., *et al.*, 1971, Profils topoclimatiques dans les Carpates Orientales, *Zesz. Nauk. UJ*, 244, *Prace Geogr.*, 26, pp. 279—287.

Hess, M., 1965, Piętra klimatyczne w polskich Karpatach Zachodnich (Sum.: Vertical climatic zones in the Polish Western Carpathians), *Zesz. Nauk. UJ*, 115, *Prace Geogr.*, 11.

Hess, M., 1966, Znaczenie średniej temperatury roku dla poznania warunków klimatycznych (Sum.: Significance of mean annual temperature in recognizing climatic conditions), *Przegl. Geogr.*, 38, 1, pp. 17–40.

Hess, M., 1967, Versuch der Unterscheidung und Charakteristik der klimatischen Höhenstufen am Beispiel der polnischen Westkarpaten, *Petermanns Geogr. Mitt.*, 111, 1, pp. 1–12.

Hess, M., 1968a, A new method of determining climatic conditions in mountain regions (with the Western Carpathians as example), *Geogr. Pol.*, 13, pp. 57–77.

Hess, M., 1968b, A method of distinguishing and specifying vertical climatic zones in temperate zone mountains (with the Western Carpathians and the Eastern Alps as examples), *Geogr. Pol.*, 14, pp. 133–140.

Hess, M., 1969, Główne problemy klimatologiczne Karpat (Sum.: Principal climatological problems of the Carpathians), *Zesz. Nauk. UJ, Prace Geogr.*, 25, pp. 7–45.

Hess, M., 1970, Zróżnicowanie stosunków mikroklimatycznych w profilu pionowym Karpat (Sum.: Differentiation of the microclimatic conditions in the vertical profile of the Carpathians), *Fol. Geogr., ser. geogr.-phys.*, 4, pp. 43–61.

Hess, M., 1971a, Piętra klimatyczne w Karpatach Północnych i Południowych i ich charakterystyka termiczna (Sum.: Vertical climatic zones in the North and South Carpathians and their thermic characteristics), *Fol. Geogr., ser. geogr.-phys.*, 5, pp. 15–23.

Hess, M., 1971b, Studien über die quantitative Differenzierung der makro-, meso- und mikroklimatischen Verhältnisse der Gebirge — als Grundlage zur Konstruierung detaillierten Klimakarten, *Zesz. Nauk. UJ*, 244, *Prace Geogr.*, 26, pp. 265–277.

Hess, M., 1972, Über die Grundlage der detailliierten Klimakarten für Gebirgsregionen, in: *Mesoklima v komplexu zivotniho prostředi*, Brno, pp. 67–84, *Studia Geogr.*, 26.

Hess, M., 1969 (1973), Methode pour determiner l'influence des formes de terrain sur le climat dans les montagnes, *Meteorologie*, 10–11, pp. 75–85.

Klein, J., 1967, Powiązanie między pokrywą śnieżną a zbiorowiskami roślinnymi w Ojcowskim Parku Narodowym (Sum.: Correlation between the snow cover and plant communities in the Ojców National Park), *Fragm. flor. geobot.*, 13, 1, pp. 77–100.

Koch, H. G., 1961, Die warme Hangzone. Neue Anschauungen zur nächtlichen Kaltluftschichtung in Tälern und an Hängen, Z. *Meteor.*, 15, 1–6 pp. 151–171.

Koppen, W., 1936, *Das geographische System der Klimate*, I, C, Berlin.

Mäde, A., 1956, Über die Methodik der meteorologischen Geländevermessung, Dtsch. Akad. Land. Wiss., 5, pp. 1–25.

Mishchenko, Z. A., 1962, *Sutochnyi khod temperatury vozdukha i ego agroklimaticheskoe znacheniye*, Leningrad.

Nägli, W., 1971, Der Wind als Standortsfaktor bei Aufforstungen in der subalpinen Stufe (Stillbergalp im Dischmatal, Kanton Graubünden), *Mitt. Schweiz. Anst. Forstl. Versuchswesen*, 47, 2, pp. 39–147.

Niedźwiedź, T., 1967, Mezoklimat doliny Raby w okolicy Gaika Brzezowej (Mesoclimate of the valley of the Raba river in the vicinity of Gaik Brzezowa), *Zesz. Nauk. UJ, Prace Geogr.*, 16, pp. 57–80.

Niedźwiedź, T., 1968, Przymrozki w pogórskim odcinku doliny Raby (Sum.: The occurrence of frost in the upland of the valley of Raba river), *Zesz. Nauk. UJ*, 179, *Prace Geogr.*, 18, pp. 53–69.

Niedźwiedź, T., 1973, Temperatura i wilgotność powietrza w warunkach rzeźby pogórskiej Karpat (na przykładzie doliny Raby koło Gaika Brzezowej) (Sum.: Temperature and air humidity of a hilly landscape (as exemplified by the Raba valley near Gaik Brzezowa)), *Zesz. Nauk. UJ*, 327, *Prace Geogr.*, 32, pp. 7–88.

Niedźwiedź, T., Obrębska-Starkel, B., 1972, Beispiel der Anfertigung einer mesoklimatischen Karte für die Karpaten, in: *Mesoklima v komplexu zivotniho prostředi*, Brno, pp. 85–96, *Studia Geogr.*, 26.

Niedźwiedź, T., Obrębska-Starklowa, B., Olecki, Z., 1973, Stosunki termiczno-wilgotnościowe wybranych zbiorowisk roślinnych w zachodniej części Swiętokrzyskiego Parku Narodowego (Sum.: The hygrothermic conditions of chosen vegetational associations in western part of the Świętokrzyski National Park (Holy Cross Mts.)), *Fol Geogr., ser. geogr.-phys.*, 7, pp. 27–75.

Obrębska-Starklowa, B., 1968, Pokrywa śnieżna we wschodniej części Gorców (Sum.: The snow cover in the eastern part of the Gorce Mts.), *Zesz. Nauk. UJ*, 179, *Prace Geogr.*, 18, pp. 27–49.

Obrębska-Starklowa, B., 1969, Stosunki mikroklimatyczne na pograniczu pięter leśnych i pól uprawnych w Gorcach (Sum.: Microclimatic conditions on the border of forest belts and arable land in the Gorce Mts.), *Zesz. Nauk. UJ*, 214, *Prace Geogr.*, 23.

Obrębska-Starklowa, B., 1970, Mezoklimat zlewni potoków Jaszcze i Jamne (The mesoclimate of the Jaszcze and Jamne drainage areas), *Studia Naturae*, ser. A., 3, pp. 7–100.

Obrębska-Starklowa, B., 1974, Mezo- i mikroklimat gromady Szymbark (Sum.: Meso- and microclimatic conditions at Szymbark), *Dok. Geogr.*, 5.

Paszyński, J., 1964, Topoclimatological investigations on heat balance, *Geogr. Pol.*, 2, pp. 69–77.

Quitt, E., 1965, Metody konstrukce mezoklimatických map, *Sborn. Čsl. Společ. Zemep.*, 70, 3, pp. 232–250.

Romer, E., 1949, Regiony klimatyczne Polski, *Prace Wrocł. Tow. Nauk.*, ser. B, 16, pp. 5–27.

Schnelle, F., 1963, *Frostschutz im Pflanzenbau. Quellenwerk über den Nachtfrost, seine Entstehung und Abwehr*, München–Basel–Wien.

Schnelle, F., 1972, Lokalklimatische Studien im Odenwald, *Ber. Dtsch. Wetterdienstes*, 17, pp. 3–38.

Skoczek, J., 1971, Vliyaniye mestnykh uslovii gornoy doliny na strukturu teplovogo balansa, *Zesz. Nauk. UJ*, 244, *Prace Geogr.*, 26, pp. 377–386.
Stružka, V., 1954, *Metody bioklimatickych pruzkumu. Praktikum fytocenologie, ekologie, klimatologie a pudoznastvi*, Praha, pp. 249–267.
Shakhnovich, A. V., Wilkens, A. A., 1963, Mikroklimaticheskaya ocenka i kartografirovaniye otdelnogo khoziaystva bez nabliyudenii, *Trudy Vsesoyuz. Nauch. Meteor. Soveshch.*, 8, ser. Agrometeorologia, pp. 190–202.
Thornthwaite, C. W., 1931, The climates of North America according to a new classification, *Geogr. Rev.*, 21.
Turner, H., 1966, Die globale Hangbestrahlung als Standortsfaktor bei Aufforstungen in der subalpinen Stufe, *Mitt. Schweiz. Anst. Forst. Versuchswesen*, 42, 3, pp. 109–168.
Uhlig, S., 1954, Beispiel einer kleinklimatologischen Galändeuntersuchung, *Z. Meteor.*, 8, 2/3, pp. 66–75.
Weischet, W., 1963, Grundvoraussetzung, Bestimmungsmerkmale und klimatologische Aussagemöglichkeit von Baumkronendeformation. Studien zum Problem der Deformation von Baumkronen durch Wind, *Freiburger Geogr. Hefte*, 1.
Ziobrowski, S., 1933, Wpływ ciężkiej zimy 1928/29 na roślinność drzewiastą w dolinie Raby (The influence of hard winter 1928/29 on the arborescent flora in the Raba valley), *Acta Soc. Bot. Pol.*, 10.

GEOGRAPHIA POLONICA 31, 1975

GEOGRAPHICAL GRADIENTS OF AIR TEMPERATURE IN POLAND

MARIA STOPA-BORYCZKA

PREFACE

The principal aim of the present study is to evaluate numerically the relations occurring in the area of Poland between air temperature and geographical and meteorological parameters. Hence, to start with, the author investigated how far some geographical and meteorological parameters jointly affect the formation of air temperature; afterwards she established the impact exerted by each of these parameters. In the course of her further reflections the author dealt separately with the particular geographical co-ordinates (geographical latitude φ and longitude λ, and altitude a.s.l. H), investigating their comprehensive and their partial impact upon air temperature. Next she proceeded in a similar way with the meteorological parameters, but dividing these into two groups of variables. To one of these groups the author assigned the mean values of the basic meteorological elements (total radiation J and absorbed radiation J_p, atmospheric pressure p and wind velocity v, relative humidity f and undersaturation of air humidity Δ, clouding N and sum of precipitation O). To the second group she assigned days in some way characteristic (number of days with precipitation $L_{\bigcirc}$, number of fair weather days $L_{\odot}$ and of cloudy days $L_{\bullet}$. number of calm days L_c and days with strong winds L_v and, finally, days with fog L_m). In these calculations the author treated as dependent variables the indices of air temperature (mean diurnal T, maximum T, minimum T, and diurnal amplitude A).

In solving her problem the author took into consideration the data gained from observations made at 120 meteorological stations distributed all over Poland and recorded for the 10-year period from 1951–1960. She limited her research to what she called the cool half-year or season (October through March) and the warm season (April through September). In a multidimensional aspect the author investigated the interdependence between the meteorological parameters by means of well known methods of mathematical statistics.

The dependent variable was marked by y, the independent variables by the symbols

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$$

* The study represents a part of a habilitation thesis on "The thermic features of Poland's climate" (Stopa-Boryczka, 1973) prepared at the Department of Climatology, Geographical Institute, Warsaw University.

Assuming between them a linear dependence in the form of

$$y = a_1x_1 + a_2x_2 + \ldots + a_kx_k + a_o \tag{1}$$

and making use of the method of least squares

$$\sum_{i=1}^{n}(y_i - a_1x_{1i} - a_2x_{2i} - \ldots - a_kx_{ki} - a_o)^2 = \min(n = 120)$$

one obtains a system of normal equations

$$\begin{aligned} s_{11}^2 a_1 + s_{12}a_2 + \ldots + s_{1k}a_k &= s_1 \\ s_{21}a_1 + s_{22}^2 a_2 + \ldots + s_{2k}a_k &= s_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ s_{k1}a_1 + s_{k2}a_2 + \ldots + s_{kk}^2 a_k &= s_k \end{aligned} \tag{2}$$

where: $s_{11}, s_{22} \ldots s_{kk}$ — are the standard deviations of the variables $x_1, x_2, \ldots x_k$, and $s_{12}, s_{13}, \ldots s_{1k}$ the covariants of the variables $(x_1, x_2) \ldots (x_1, x_k)$ etc. By $s_1, s_2 \ldots s_k$ are marked the covariants of variable y with regard to $x_1, x_2 \ldots x_k$.

Written in their matrix form these equations appear in the form of:

$$SA = S_k \tag{3}$$

where:

$$S = \begin{bmatrix} s_{11}^2 & s_{12} & s_{1k} \\ s_{21} & s_{22}^2 & s_{2k} \\ \ldots & \ldots & \ldots \\ s_k^1 & s_k^2 & s_{kk}^2 \end{bmatrix}, \quad S_k = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{bmatrix} \tag{4}$$

is the variance-covariance matrix, and

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \tag{5}$$

The coefficients of multiple regression $a_1, a_2 \ldots a_k$ were determined from the equation

$$A = S^{-1}S_k \tag{6}$$

The gauge of the correctness of fitting the hyperplanes to the empiric points

$$X_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ki} \end{bmatrix}$$

is the residual variance

$$W^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - a_1x_{1i} - a_2x_{2i} - \ldots a_kx_{ki} - a_o)^2 \tag{7}$$

As index of the interdependence between the variable y and $x_1, x_2 \ldots x_k$ the author adopted the coefficients of multiple correlation R which represents the gauge for the joint impact of the variables X upon y $(0 \leqslant R \leqslant 1)$.

AIR TEMPERATURE AND GEOGRAPHICAL AND METEOROLOGICAL PARAMETERS

The gauges indicating the comprehensive impact of the geographical (φ, λ, H) and the meteorological (J, J_p, p, v, f, Δ, N, O) parameters upon selected thermic indices (T, T_{max}, T_{min}, A) are the coefficients of multiple correlation (R) which assume different values in different seasons (Table 1).

TABLE 1. Coefficients of the multiple correlation R for T, T_{max}, T_{min}, and A with regard to meteorological and geographical parameters in Poland's territory

	X–III	IV–IX
T	0.94	0.99
T_{max}	0.80	0.98
T_{min}	0.90	0.63
A	0.85	0.81

From this schedule it appears, that in most cases the coefficients of multiple correlation R approaches the value one, and it can be concluded that air temperature develops foremost in correlation with the parameters taken into consideration, that is, the geographical parameters φ, λ, H and the meteorological parameters J, J_p, p, v, f, Δ, N, O).

The author limited her investigations of the impact of the particular geographical and meteorological parameters upon air temperature to the determination of the coefficients of partial regression. The hyperplanes of regression have the following equations:

For the cool season (X–III):

$$\begin{aligned} T = & -0.464\varphi - 0.227\lambda - 0.006H + 0.004J - 0.092J_p - 0{,}0010p + \\ & + 0.057v - 0.002f + 1.305\Delta + 0.691N + 0.004O - 0.014L_{\bigcirc} + \\ & + 0.006L_{\odot} - 0.012L_{\bullet} - 0.001L_c + 0.015L_v - 0.005L_m + 309.100. \\ T_{min} = & -0.142\varphi - 0.186\lambda - 0.007H + 0.012J + 0.021J_p - 0.015p - \\ & - 0.013v - 0.038f + 0.705\Delta + 0.186N + 0.005O - 0{,}010L_{\bigcirc} - \\ & - 0.017L_{\odot} - 0.011L_{\bullet} - 0.003L_c + 0.014L_v + 0.003L_m + 299.200. \\ A = & -0.255\varphi - 0.028\lambda + 0.000H - 0.012J + 0.051J_p - 0.006p + \\ & + 0.003v - 0.003f - 0.017\Delta - 0.144N - 0.003O + 0.011L_{\bigcirc} + \\ & + 0.014L_{\odot} + 0.006L_{\bullet} + 0.004L_c - 0.012L_v - 0.007L_m + 26.260. \end{aligned} \tag{8}$$

For the warm season (IV–IX):

$$\begin{aligned} T = & -0.362\varphi + 0.035\lambda - 0.004H - 0.015J - 0.004J_p + 0.004p - \\ & - 0.004v + 0.018f + 0.543\Delta + 0.140N - 0.000O - 0.014L_{\bigcirc} + \\ & + 0.009L_{\odot} + 0.003L_{\bullet} - 0.001L_c - 0.004L_v - 0.004L_m + 299.100. \\ T_{max} = & -0.416\varphi + 0.061\lambda - 0.004H - 0.048J + 0.038J_p + 0.007p - \\ & - 0.006v + 0.001f + 0.760\Delta + 0.176N - 0.001O - 0.008L_{\bigcirc} + \\ & + 0.018L_{\odot} + 0.007L_{\bullet} - 0.000L_c - 0.022L_v - 0.003L_m + 305.100. \end{aligned} \tag{9}$$

$$T_{min} = -0.421\varphi - 0.078\lambda - 0.005H + 0.137J - 0.039J_p + 0.010p + \tag{9}$$
$$+ 0.005v + 0.074f + 0.644\Delta + 0.876N + 0.001O + 0.010L_{\bigcirc} +$$
$$+ 0.037L_{\odot} - 0.000L_{\bullet} - 0.002L_c - 0.005L_v + 0.009L_m + 273.600.$$
$$A = -0.125\varphi + 0.062\lambda + 0.000H - 0.076J + 0.086J_p + 0.001p -$$
$$- 0.005v - 0.024f + 0.570\Delta - 0.000N - 0.002O + 0.015L_{\bigcirc} +$$
$$+ 0.007L_{\odot} + 0.005L_{\bullet} + 0.002L_c - 0.031L_v + 0.002L_m + 13.780.$$

The coefficients of partial regression define how much the particular geographical and meteorological parameters affect air temperature. They express an increase in air temperature (ΔT) per unit of a geographical parameter ($[\varphi] = 1°$, $[\lambda] = 1°$, $[H] = 1$ m) and per unit of a meteorological parameter ($[J, J_p] = 1$ kcal/cm², $[p] = 1$ mb, $[v] = 1$ m/s, $[f] = 1\%$, $[\Delta] = 1$ mb, $[N] = 1$, $[O] = 1$ mm, $[L_{\bigcirc}, L_{\odot}, L_{\bullet}, L_c, L_v, L_m] = 1$ mean day) — all these values holding good for Poland's territory. These small units explain, why in many instances the values obtained are so very small. Hence, in analyzing these coefficients it is advisable to calculate changes of the thermic indices for larger differences of parameters, especially concerning precipitation and the number of characteristic days. This same advice also applies to altitude a.s.l. It happens, that in Poland's territory the parameters mentioned are subject to exceptionally large fortuitous disparities; this is why at times a change of but one unit fails to reveal noticeable temperature changes in the near-ground air layer. The higher value of the coefficient of partial regression, the stronger is the impact of the given element upon air temperature. In addition, the signs added to the figures indicate the kind of impact: whether an increase of a given parameter denotes an increase or a decrease in air temperature.

Among geographical agencies, *geographical latitude* stands out with regard to a_j and to maintaining the minus sign during both seasons. Assuming the remaining parameters to be stable, a 1° change (increase) in latitude is always associated with a decrease of air temperature T amounting to 0.5–0.4°C, of T_{max} to 0.5–0.2°C. of T_{min} to 0.4–0.1°C, and of A to 0.3–0.1°C, depending on the season. Negative increase of $\frac{\Delta T}{\Delta \varphi}$ signify additionally, that air temperature drops with growing latitude — in other words, that the horizontal gradient is always directed towards the south. This type of interdependence reflects foremost the strike angle of solar radiation and duration of insolation which, decreasing with growing φ, causes a reduction in the inflow of radiation upon the surface of the Earth and, consequently, a drop in air temperature (Equations (8), (9)).

Regarding *geographical longitude* the horizontal gradients of air temperature $\frac{\Delta T}{\Delta \lambda}$ average a level nearly twice lower than $\frac{\Delta T}{\Delta \varphi}$, and their signs change depending on the season. Minus values during the cool season signify, that the increased distance from the Atlantic Ocean is accompanied by a decrease in all thermic features (T, T_{min}, A). On the other hand, the plus values observed in the warm season are evidence of a change in the direction of the temperature gradient, from west to east, that is, in a direction opposite to the influence of the Atlantic Ocean. A smaller temperature increase per 1°λ during the cool season, compared with the warm season, indicates a decrease in the warming effect of the ocean due to the impact of the continent of Asia. It is well known, that this period brings a peak of activity to the nearest baric configurations (the Iceland low and the Siberian high) and that Poland is under their impact.

An exception is T_{min} which in summer keeps up its sign — probably due to local conditions.

In keeping with a commonly known regularity, to every increase in *altitude* a.s.l there is a corresponding decrease in air temperature. Calculating the value of $\frac{\Delta T}{\Delta H}$ for 100 m, the average vertical thermic gradients oscillate for T within the range of 0.4–0.6°C, for T_{max} between 0.4–0.5°C, and for T_{min} between 0.5–0.6°C. With increasing altiudes H a.s.l. the decrease of extreme temperatures causes the diurnal amplitudes to be reduced to 0°C.

With regard to *total radiation* $\left(\frac{\Delta T}{\Delta J} = \frac{1°C}{kcal/cm^2}\right)$, the values of the coefficients of partial regression for air temperature are rather low, compared with absorbed radiation $\left(\frac{\Delta T}{\Delta J_p} = \frac{1°C}{kcal/cm^2}\right)$. This indicates a higher effect of the last parameter; yet not always does this increase in J and J_p correspond to a plus increase in ΔT. The reason for this type of anomaly should be looked for mainly in the radiation balance which in our geographical latitudes is negative. Since Poland's climate is not subject to stable cooling, these heat losses during radiation must have their compensation in other processes, for the most part in the general atmospheric circulation, which to some degree equalizes the thermic disparities occurring on the surface of the Earth. Moreover, upon the quantity of radiating energy arriving from the Sun and upon the degree of its absorption by the surface of the Earth depends the intensity of a further process — the circulation of water in nature. Increased evaporation leads to a rise in air humidity, and this in turn brings an increase in clouding and precipitation and, in consequence, a decrease in air temperature. Ultimately J and J_p do not always cause a temperature increase; in many instances they may result in its decrease, especially when accompanied with increased clouding and precipitation.

The effect of *atmospheric pressure* upon air temperature depends upon the season of the year. In the warm season, especially during summer, a pressure increase in the 1 mb level causes an increase in air temperature averaging 0.001°C, whereas in winter in most cases it causes a decrease. After eliminating the influence of the remaining meteorological and geographical parameters, especially of altitude a.s.l., the role played by baric configurations and the types of weather connected with them is conclusive. In the warm season a temperature increase per pressure unit $\frac{\Delta T}{\Delta p}$ is connected with improving weather conditions, usually marked by slight clouding. During daytime this causes intensive heat to be transmitted to the air from the surface of the Earth; during the night the Earth is unable to emit all of the heat energy absorbed in the daytime, because of the short duration of the night, and ultimately this leads in Poland's territory to an increase in all thermic indices (T, T_{max}, T_{min}, A). In turn, during the cool season the occurrence of a high may cause a temperature increase in the noon hours. On the other hand, at night a lack of clouding fails to protect the surface of the Earth from excessive radiation, and this is reflected mainly in the values of minimum temperatures, mean diurnal temperatures, and amplitudes. Additionally propitious is the existence of a snow cover.

An increase in temperature per unit of *wind velocity* $\left(\frac{\Delta T}{\Delta v} = \frac{1°C}{m/s}\right)$ is definitely negative in the warm and positive in the cool season. Not only the resultant horizontal and vertical exchange of air masses (v) is conductive here, but also the main direction of inflow of these air masses. In Poland's territory the predominance of winds from the western sector cause them to exert a different influence upon air temperature, a cooling influence in summer and a warming one in winter. This regularity does not apply to minimum temperatures because, just before they are about to set in (temperature inversion, low wind velocities), local conditions become decisive in view of the reduced horizontal and vertical air exchange. In a similar way the coefficients of regression between air temperature and the number of days with strong winds develop while this does not apply to the number of calm days. In the latter case, of marked significance are the near surroundings of the meteorological stations.

The presence of *water vapour* in the air protects the surface of the Earth from excessive heating during the daytime and in summer, and from excessive emission of radiation during the night and in the winter. These facts further an increase in air temperature. The close relation between air temperature and air humidity during the warm season results for the most part from the physical character of these parameters. The intensity of evaporation and the contrasting thermic features between arriving air masses and the ground surface are also highly important being the main causes of water storage in the soil and of water vapour in the atmosphere.

Calculations show that all indices of air temperature depend on the degree of *clouding*. No matter what season, an increase of general clouding goes hand in hand with an increase of mean diurnal air temperature (T) and of minimum air temperature (T_{min}), as well as with a decrease in diurnal amplitude (A). Maximum air temperatures are also lower, but only during the cool season. In the warm season an increase in air temperature is connected with increased clouding — undoubtedly caused by the predominance of convective-type clouding in the warmer season which usually develops in the daytime reaching its maximum in the afternoon hours. That the coefficients of partial regression of the temperature vary less with regard to the number of fair-weather days and cloudy days in the different seasons must be ascribed to the rather small variations they undergo in Poland's territory. As far as fogs are concerned, they are accompanied by a temperature decrease, especially of the maximum and the mean diurnal temperature, and also of the amplitude. On the other hand, on days with fog the temperature minima are higher, caused by the rate of effective radiation.

On the whole, larger sums of *precipitation* are connected with lower air temperatures. In most cases the coefficients of regression $\frac{\Delta T}{\Delta O}$ show a minus sign. Admittedly, a 1 mm increase in the sum of precipitation and a 1 day increase in cloudy days cause a very small decrease in air temperature in Poland — merely in the order of 0.001 to 0.01°C; however, the coefficients of partial regression are practically always negative. This kind of dependence of air temperature upon the sums and the number of days with precipitation can be explained mostly by increased clouding which lessens the daytime inflow of solar energy. Moreover, any kind of precipitation, especially when cool, absorbs part of the heat from the atmosphere whilst falling. And, finally, part of the heat energy is also consumed by the very process of water evaporation from the surface of the Earth.

AIR TEMPERATURE AND GEOGRAPHICAL PARAMETERS

The next step in these reflections is the dissociation of the impact of geographical and of meteorological agencies upon air temperature.

The joint impact of the group of geographical variables (φ, λ, H) upon air temperature (T, T_{max}, T_{min}, A) is expressed by the coefficients of multiple correlation which undergo variations in the annual profile (Table 2).

TABLE 2. Coefficients of the multiple correlation R for T, T_{max}, T_{min}, A with regard to geographical parameters in Poland's territory

	X–III	IV–IX
T	0.88	0.97
T_{max}	0.89	0.94
T_{min}	0.85	0.59
A	0.75	0.70

Extreme values of the coefficient R appear in the warm season: maxima for mean diurnal air temperature ($R = 0.97$), while minima coincide with minimum temperature ($R = 0.59$). It can be seen that among the variables T, T_{max}, T_{min}, A it is the minimum air temperature which least depends upon the geographical agencies φ, λ, H. On the whole it might be said, that the degree of how much the thermic indices depend upon geographical parameters is fairly high.

The share of particular components of the geographical agencies in their joint impact upon air temperature is illustrated in regression planes (Equations 10 and 11).

For the cool season (X–III):

$$\begin{aligned} T &= -0.376\varphi - 0.246\lambda - 0.004H + 299.2 \\ T_{min} &= -0.124\varphi - 0.233\lambda - 0.003H + 281.9 \\ A &= -0.502\varphi - 0.026\lambda - 0.001H + 33.29 \end{aligned} \tag{10}$$

For the warm season (IV–IX):

$$\begin{aligned} T &= -0.514\varphi + 0.071\lambda - 0.006H + 313.2 \\ T_{max} &= -0.771\varphi + 0.118\lambda - 0.008H + 331.3 \\ T_{min} &= -0.439\varphi - 0.052\lambda - 0.005H + 306.5 \\ A &= -0.460\varphi + 0.108\lambda - 0.003H + 32.9 \end{aligned} \tag{11}$$

After taking into account the influence of geographical latitude and longitude and of altitude a.s.l., the particular features of air temperature (T, T_{max}, T_{min}, A) show that to the highest degree they depend upon φ, that is, the height of the Sun and the duration of solar radiation. The meridional gradients of air temperature vary, contingent upon the area under investigation, as to maximum temperature (T_{max}) at the rate of some 0.8°C, to mean diurnal temperature (T) of 0.4–0.5°C, to minimum temperature (T_{min}) of 0.1–0.4°C, and as to diurnal amplitude of approximately 0.5°C.

Assuming geographical latitude and absolute altitude to be constant, one can define the relation between air temperature and geographical longitude, which is, in the case of Poland, reflected by the extent and magnitude of the influence of the Atlantic Ocean and the Asiatic continent upon Poland's thermic conditions. In the cool season a 1° increase in geographical longitude is associated with a descrease of all thermic features, and this effect appears most strikingly in the maximum and minimum values of the mean diurnal air temperatures $\frac{\Delta T}{\Delta\lambda} \approx 0.2°C/1°\lambda$, and least clearly in the diurnal temperature amplitudes ($-0.03°C/1°\lambda$). In contrast, the warm season causes the signs of the coefficients of partial regression to change into minus signs, meaning, that in Poland checking from west to east an average increase of some $0.1°C/1°\lambda$ sets in, in the maximum temperature (T_{max}), diurnal amplitude (A) and mean diurnal temperature (T). An exception is minimum temperature which changes in the opposite direction, that is, fails to increase and rather decreases about $0.05°C/1°\lambda$. This pattern of increases in air temperature per unit of horizontal distance ($1°\lambda$) must be mainly ascribed to Poland's situation with regard to the most active centres of atmospheric activities, the impact of which changes during the year.

As to altitude a.s.l., this parameter always affects all features of air temperature (T, T_{max}, T_{min}, A), in the same way, irrespective of seasons. In other words, an 100 m increase in altitude causes maximum temperatures to drop 1°C, mean diurnal temperatures some 0.6–0.4°C, minimum temperatures 0.5–0.3°C, and diurnal amplitudes 0.3–0.1°C. This also clearly shows, that the vertical gradients of air temperature are markedly larger in summer than in winter — mainly due to a more intensive heat exchange between the surface of the Earth and the atmosphere in summer than in winter.

AIR TEMPERATURE AND OTHER METEOROLOGICAL PARAMETERS

Here again the author deals separately with meteorological conditions, investigating their comprehensive and partial impact upon air temperature in Poland. She treats the particular temperature features (T, T_{max}, T_{min}, A) as dependent variables and the remaining meteorological elements (J, J_p, p, Δ, O) as independent variables, but eliminates the days with special features (L_o, $L_{\circledcirc}$, $L_{\bullet}$, L_m).

The author now determines how much the coefficients of multiple correlation R of the above mentioned thermic indices change with regard to meteorological parameters (Table 3).

TABLE 3. Coefficients of the multiple correlation of T, T_{max}, T_{min}, and A with regard to other meteorological parameters in Poland's territory

	X–III	IV–IX
T	0.84	0.93
T_{max}	0.88	0.90
T_{min}	0.77	0.56
A	0.78	0.72

In conformity with their definition, the coefficients of multiple correlation which are the yardstick of the impact of meteorological conditions upon air temperature can oscillate between zero and one. The upper limit reflects the maximum and the very close dependence of a given element upon the remaining elements, while the lover limit denotes a lack of mutual dependence. From Table 3 it appears, that the particular meteorological elements are mutually correlated and that they jointly bear upon air temperature in a definite manner. The figures obtained reveal, that air temperature is one of the outstanding meteorological elements, illustrating or even conditioning the values of the remaining elements and their components. The highest degree of interdependence with meteorological conditions show mean diurnal values ($R = 0.84$–0.93) and minimum temperature ($R \approx 0.90$), whereas the lowest interdependence is the degree for minimum temperatures ($R = 0.56$–0.77). This implies that during minimum temperatures meteorological conditions are far divergent from the average. In part this belief seems justified, because during minimum temperatures the intensity of turbulent heat exchange wanes considerably. This is why another factor usually not fully taken into account, plays an important role — local conditions, i.e. the type of ground relief.

Further information about the effect of meteorological conditions upon air temperature is gained from the coefficients of partial regression, determined for T, T_{max}, T_{min}, and A with regard to the following eight meteorological parameters: J, J_p, p, v, f, Δ, N, O. The hyperplanes of regression for thermic features with regard to the remaining meteorological parameters are defined in Equations (12) and (13).

For the cool season (X–III):

$$\begin{aligned} T &= -0.002J + 0.288J_p + 0.043p - 0.041v - 0.006f + \\ &\quad + 1.984\Delta + 0.211N + 0.001O + 223.900. \\ T_{min} &= -0.008J + 0.119J_p + 0.045p - 0.027v - 0.017f + \\ &\quad + 1.160\Delta - 0.217N + 0.003O + 225.900. \\ A &= -0.015J + 0.283J_p - 0.013p - 0.226v - 0.031f + \\ &\quad + 0.144\Delta - 0.018N - 0.004O + 20.410. \end{aligned} \tag{12}$$

For the warm season (IV–IX):

$$\begin{aligned} T &= 0.073J - 0.142J_p + 0.025p - 0.050v - 0.016f + \\ &\quad + 1.266\Delta - 0.204N + 0.001O + 260.700. \\ T_{max} &= 0.096J - 0.179J_p + 0.022p - 0.079v - 0.059f + \\ &\quad + 1.172\Delta - 0.286N + 0.001O + 271.300. \\ T_{min} &= 0.120J - 0.129J_p + 0.034p - 0.044v + 0.045f + \\ &\quad + 1.079\Delta + 0.400N + 0.003O + 235.100. \\ A &= 0.036J - 0.068J_p - 0.009p - 0.050v - 0.073f + \\ &\quad + 0.948\Delta - 0.161N - 0.000O + 22.530. \end{aligned} \tag{13}$$

As can readily be seen, elimination of the geographical elements from the impact upon air temperature leads to a considerable increase of the effect of meteorological agencies; previously this came to light in the coefficients of multiple correlation R, while now it appears in the coefficients of partial regression. Apart from the investigated features of temperature and season, the highest degree of interdependence between air temperature and undersaturation of humidity (for T, $a_j = 0.90–1.98$, for T_{max}, $a_j = 1.15–2.57$, for T_{min}, $a_j = 0.60–1.16$, and for A, $a_j - 0.14–0.95$), especially during the warm season. An increase of air dryness at the rate of 1 mb (with values of other elements assumed constant) is mostly associated with an almost identical increase in air temperature, that is, 1°C $\left(\frac{\Delta T}{\Delta\Delta} \approx 1\right)$. This indicates that these two elements are closely correlated, showing identical changes in periodicity, diurnal and annual. This has its effect upon their mean values and upon their spatial disparity (between particular stations) in Poland's territory. The air temperature continues to be the least correlated with the sums of atmospheric precipitation, and the reason for this should be seen in the wide variation of the sums of precipitation in both time and space. Moreover, the very fact of precipitation does not affect the course of air temperature as essentially as does the increase in clouding associated with precipitation — the more so since precipitation may affect air temperature in different ways, depending on the time of day and the year, as seen in the pattern of the coefficients of partial regression (Equations (12) and (13)). More in evidence in this aspect of the problem is the part played by total and absorbed radiation, especially during the warm season. Any increase in total radiation (J) by 1 kcal/cm² between stations is associated with an increase in air temperature of the order of 0.1°C. For absorbed radiation (J_p) things are the opposite. Two further elements: atmospheric pressure and wind velocity, which both contribute much to shaping air temperature are practically the same. Usually a rise in atmospheric pressure by 1 mb is connected with a rise air temperature (0.01–0.4°C), whereas an increase in wind velocity by 1 m/s usually causes a 0.02–0.08°C decrease in air temperature.

AIR TEMPERATURE AND NUMBER OF CHARACTERISTIC DAYS

The last group of variables for which the author investigated the impact upon air temperature are, what are called "characteristic" days; most often they are separated out in climatological studies, comprising: days with precipitation marked $L_{\bigcirc}$, fair-weather days $L_{\circledcirc}$, cloudy days $L_{\bullet}$, days with strong winds L_v and calm days L_c, and finally, days with fog L_m. Table 4 presents the coefficients of multiple correlation indicating air temperature for such days.

TABLE 4. Coefficients of the multiple correlation for T, T_{max}, T_{min}, and A with regard to the number of characteristic days in Poland's territory

	X–III	IV–IX
T	0.58	0.92
T_{max}	0.62	0.90
T_{min}	0.54	0.50
A	0.71	0.71

This table shows, that here also a high degree of correlation exists between the thermic indices and the number of characteristic days, especially regarding mean diurnal and maximum values; this may partly be explained by the fact, that they mostly refer to diurnal and seasonal periods. Generally speaking, the coefficients of multiple correlation R are higher in the warm season than in the cool season.

The way that, in terms of a concise relation of temperature to the number of characteristic days, air temperature depends upon particular days, appears from the coefficients of partial regression. The hyperplanes of regression are formulated in the following equations:
For the cool season:

$$\begin{aligned} T &= -0.010L_{\bigcirc}-0.023L_{\odot}-0.041L_{\bullet}-0.003L_c-0.006L_v- \\ &\quad -0.026L_m+280.20. \\ T_{\min} &= -0.005L_{\bigcirc}-0.038L_{\odot}-0.038L_{\bullet}-0.007L_c-0.006L_v- \\ &\quad -0.022L_m+276.800. \\ A &= 0.011L_{\bigcirc}+0.028L_{\odot}-0.021L_{\bullet}+0.009L_c-0.016L_v+ \\ &\quad +0.003L_m+6.748. \end{aligned} \tag{14}$$

For the warm season:

$$\begin{aligned} T &= -0.087L_{\bigcirc}+0.013L_{\odot}+0.016L_{\bullet}+0.001L_c-0.021L_v- \\ &\quad -0.027L_m+292.800. \\ T_{\max} &= -0.101L_{\bigcirc}+0.026L_{\odot}+0.034L_{\bullet}+0.002L_c-0.044L_v- \\ &\quad -0.026L_m+298.100. \\ T_{\min} &= -0.057L_{\bigcirc}+0.023L_{\odot}+0.024L_{\bullet}-0.001L_c-0.017L_v- \\ &\quad -0.021L_m+284.900. \\ A &= -0.039L_{\bigcirc}+0.013L_{\odot}+0.029L_{\bullet}+0.003L_c-0.040L_v- \\ &\quad -0.001L_m+11.860. \end{aligned} \tag{15}$$

In these equations the days, selected from the viewpoint of frequency of precipitation ($L_{\bigcirc}$), of extreme values of wind velocity (L_c, L_v), and of the degree of clouding $L_{\odot}$, $L_{\bullet}$), as well as the days with fog (closely related to cloudy days) (L_m), show the least association with air temperature compared with the groups of variables previously discussed. The reason seems to be, that days of this kind constitute usually a minor percentage compared with the total number of days of the season under examination. Proportions most favourable in this respect occur for days with precipitation (their frequency is 22% in the cool and 42% in the warm season), for days with clouding (frequencies 50% and 28%) and for calm days (frequencies 40% and 48%). On the other hand, the lowest percentages were observed for fair-weather days (frequencies 10% in the cool and 12% in the warm season), for days with strong winds (frequencies 12% and 7%), and for days with fog (18% and 9%).

Taking into account the relation of the number of days of a definite type to the total number of days in a season, one observes that the greatest decreases in air temperature occur during an increase in the number of strong winds (some 0.01°C per day in the cool and 0.2–0.4°C in the warm season). Negative increases in air temperature, irrespective of the season, are also caused by days with precipitation (some 0.01°C per day in the cool and some 0.1°C in the warm season) and by days with fog (0.02–0.1 and 0.2°C). A distinct change of the signs of the coefficients of regression a_i during the year appears in the relation of T to the number of fair-weather days. Negative values of a_i during winter indicate, that hand-in-hand with their increase goes a decrease in air temperature amounting to 0.04–0.1°C per day; the reverse applies in the summer (some 0.01–0.02°C per day). The thermic conditions in Poland's territory for cloudy days are similar: in winter the particular features of T, T_{max}, T_{min} and A decrease with an increase in the number of cloudy days (0.02–0.04°C per day), while in summer they increase (0.02–0.03°C per day). This latter anomaly is presumably associated with a predominance of convection-type clouding which usually is accompanied by fairly high air temperatures. Among all days, the least degree of interdependence occurs between air temperature and the number of calm days, even regarding minimum temperature which promotes this phenomenon (due to conditions of a stable equilibrium). Nonetheless, the effect of calm days differs depending on the season: in winter an increase in the number of calm days is in most cases reflected by a decrease in air temperature, while in summer the opposite is the rule. As far as daily temperature amplitudes are concerned, on the average they are higher for localities where calm days prevail.

DISTRIBUTION OF COEFFICIENTS OF REGRESSION OF AIR TEMPERATURE WITH REGARD TO OTHER METEOROLOGICAL PARAMETERS IN POLAND'S TERRITORY

Lines of identical correlation were for the first time compiled for Poland's territory in 1917 by W. Gorczyński (1917, 1918). The domain of interest to this author was, amongst other factors, the variability of temperature and atmospheric pressure in both time and space, and for determining them he applied the correlation calculus. Choosing Warsaw as the point of reference, he investigated the interdependence between temperature changes (1886–1910) and atmo-

Fig. 1. Distribution of meteorological stations in Poland

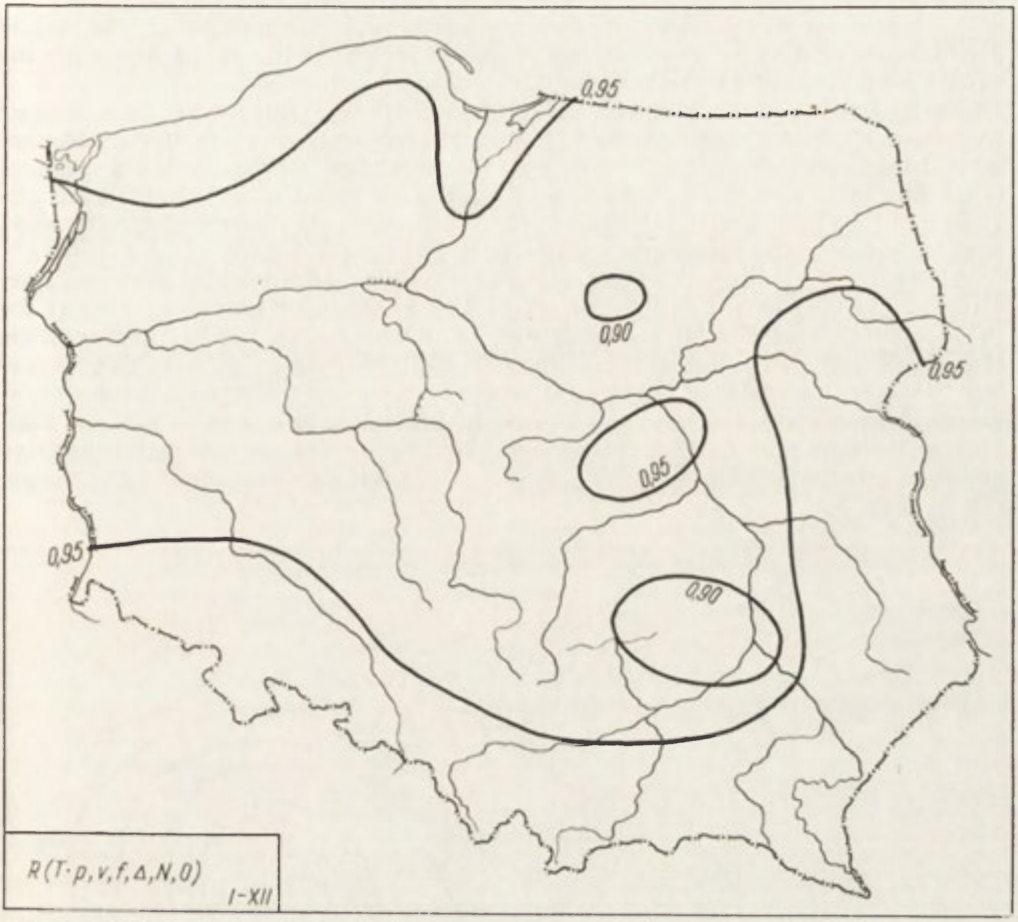


Fig. 2. Multiple isocorrelates of air temperature and other meteorological parameters: R ($T \cdot p$, v, f, Δ, N, O) for one year

spheric pressure (1851–1900) for a variety of localities. The lines, which on his map join localities of identical values of the coefficient of correlation regarding Warsaw, he called "equicorrelates". It is important in climatological investigations not only to know the interdependence of changes of a given element occurring between neighbouring stations, but also to understand the interdependence between various meteorological elements at given points or at different localities situated in the region under investigation (Merecki, 1914; Michna, 1972). This is why the principal aim of this chapter is the study of the spatial distribution of the coefficients of regression in a multidimensonal aspect, between the basic feature of air temperature (the mean diurnal T) and other meteorological parameters such as pressure (p), wind velocity (v), air humidity (f, Δ), clouding (N) and precipitation (O) in a time profile. In this case the author took into account annual values from records for the 1951–1960 period; here it should be mentioned that in this study she limited herself to 60 synoptic stations distributed over Poland's territory. The distribution of these 60 stations is illustrated in Fig. 1. The comprehensive and the partial impact of meteorological parameters upon mean diurnal air temperature all over the country has been pictured graphically by maps presenting spatial patterns of the coefficients of multiple correlation and of partial regression. The gauges of the interdependence applied by the author have been indicated in accordance with the equations presented in the preface.

As the gauge of the comprehensive impact of meteorological elements upon the mean diurnal air temperature the author continues to use the coefficient of multiple correlation R which in Poland's territory shows a value approaching 1 (Fig. 2). From the distribution of coefficient R it appears that, in accordance with the meteorological values taken into account, T assumes for the most part its value in the south and the north of Poland ($R > 0.95$).

The author started her investigation of the effect of particular meteorological parameters upon air temperature from an analysis of isoline a_j = const for the coefficients of partial regression of T and p. On an annual average, the rise of atmospheric pressure causes the greatest temperature drop in the Masovian Plain ($a_j = -0.34$ at Ostrołęka) and on the Lesser Poland Plateau ($a_j = -0.37$ at Kielce). On the other hand, in the Carpathians ($a_j = 0.55$ on Kasprowy Wierch) and in the Karkonosze Mts. ($a_j = 0.41$ on Śnieżnik) positive pressure increases tally not with a decrease but with an increase in air temperature (Fig. 3).

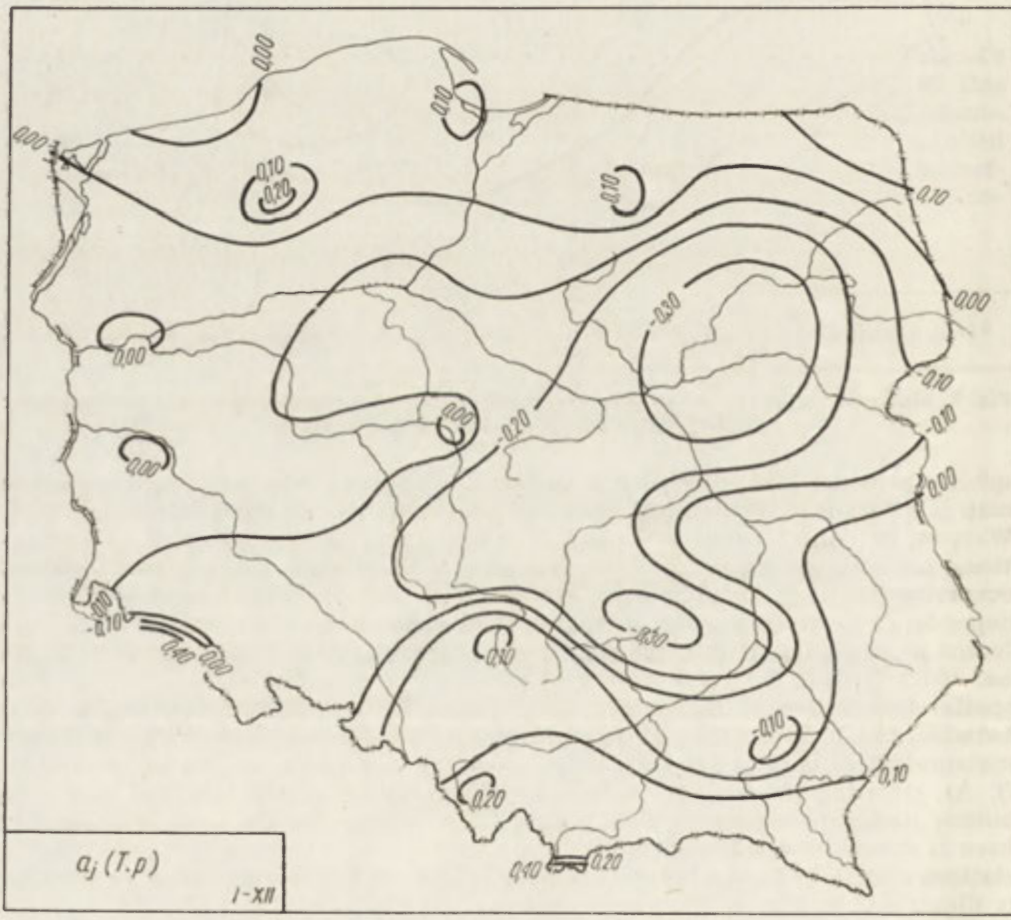


Fig. 3. Isolines of partial regression of air temperature with regards to atmospheric pressure; $a_j (T \cdot p)$ for one year

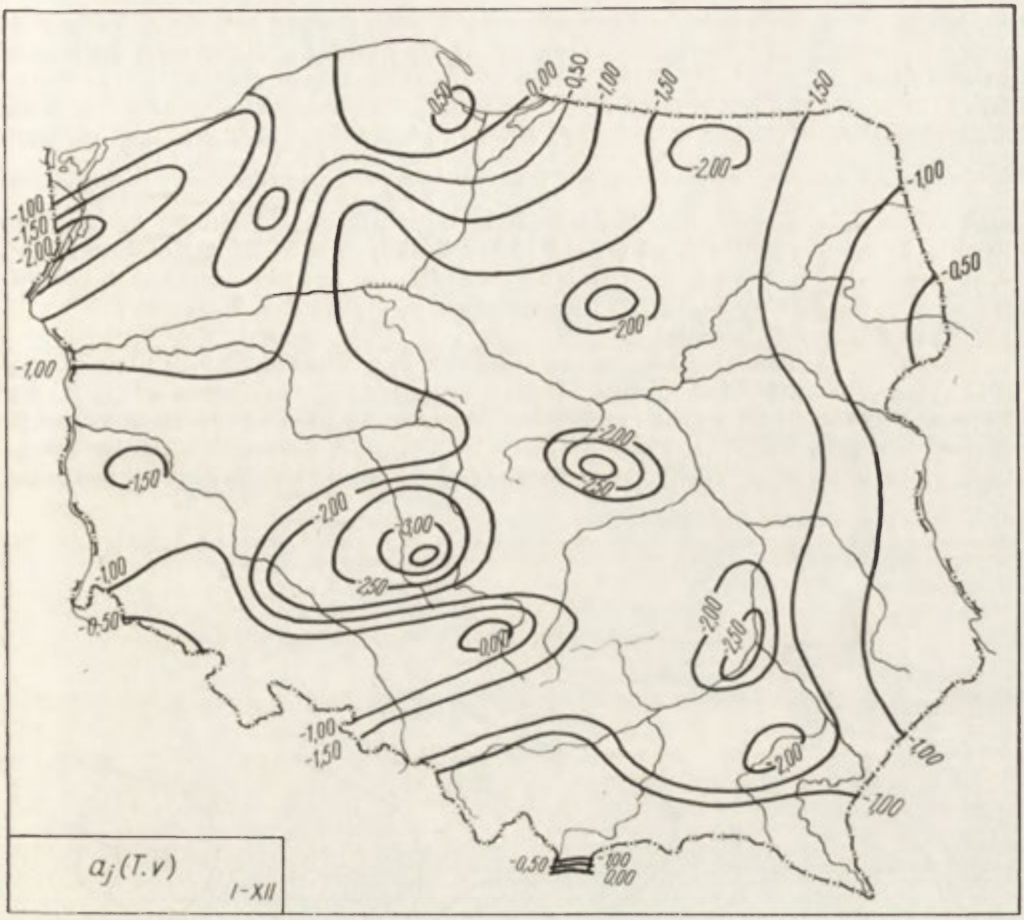


Fig. 4. Isolines of partial regression of air temperature with regard to wind velocity: $a_j(T \cdot v)$ for one year

On the whole the increase of T per unit of *wind velocity* is negative (Fig. 4), especially on the NW and E side of the Lesser Poland Plateau ($a_j = -3.54$ at Wieluń and $a_j = -2.75$ at Sandomierz). Particularly remarkable is the run of the isarithm $a_j = -1.50$ which separates the eastern, southern, western and north-western part of Poland where the coefficients of partial regression lie above -1.50, from the central part of Poland, including the centre of the Mazurian Lake District and the Sandomierz Basin, where a_j is much higher than -1.50. An exception is the eastern part of the Pomeranian Lake District including its surrounding shore belt, where a positive increase of T per unit of v_j can be observed ($a_j = 0.61$ at Gdańsk, 0.43 at Hel, 0.36 at Szczecinek and 0.22 at Wałcz).

In the relation of T to the *relative air humidity* the coefficients a_j are always above zero. Greatest are the changes in $\frac{\Delta T}{\Delta f}$ locally at Ustka ($a_j = 0.78$) and at Świnoujście ($a_j = 0.72$). Smaller by half are these coefficients in western Poland and in its north-west. Least connected with f is the air temperature in a central belt marked by isoline $a_j = 0.20$ (Fig. 5). In the eastern part of Poland a_j values of the 0.20–0.30 range predominate.

Much the same as the figures of a_j for T and f, the coefficients of partial regression for T and for *undersaturation of air humidity* differ only in regard to absolute values. On the whole, the changes of T per unit of Δ are fairly large. The isolines of regression $a_j = \text{const } T$ with regard to Δ (Fig. 6) show a pattern similar to that of T and f. Where the increases of T per unit of f are lowest, we nowadays observe rather low increases of $\frac{\Delta T}{\Delta \Delta}$ (the central meridional belt of Poland with $\lambda = 20$–$22°$). Highest values of coefficient a_j continue to rule along the Baltic coast ($a_j = 6.33$ for Ustka, $a_j = 5.56$ for Kołobrzeg, and $a_j = 4.94$ for Świnoujście). Approaching these figures are the values of a_j recorded in the mountains ($a_j = 6.22$ for Śnieżka, $a_j = 5.23$ for Kasprowy Wierch, and $a_j = 5.22$ for Zakopane).

The way that *clouding* affects air temperature is illustrated in a further map (Fig. 7), showing the distribution of the coefficients of regression a_j (T, N) for Poland's territory. The areas situated in the east, separated by isolines -1.00 and -1.50 (the Masovian Plain, the Siedlce Plateau), deserve particular attention; from these areas fairly large values (< -1.00) or the largest temperature

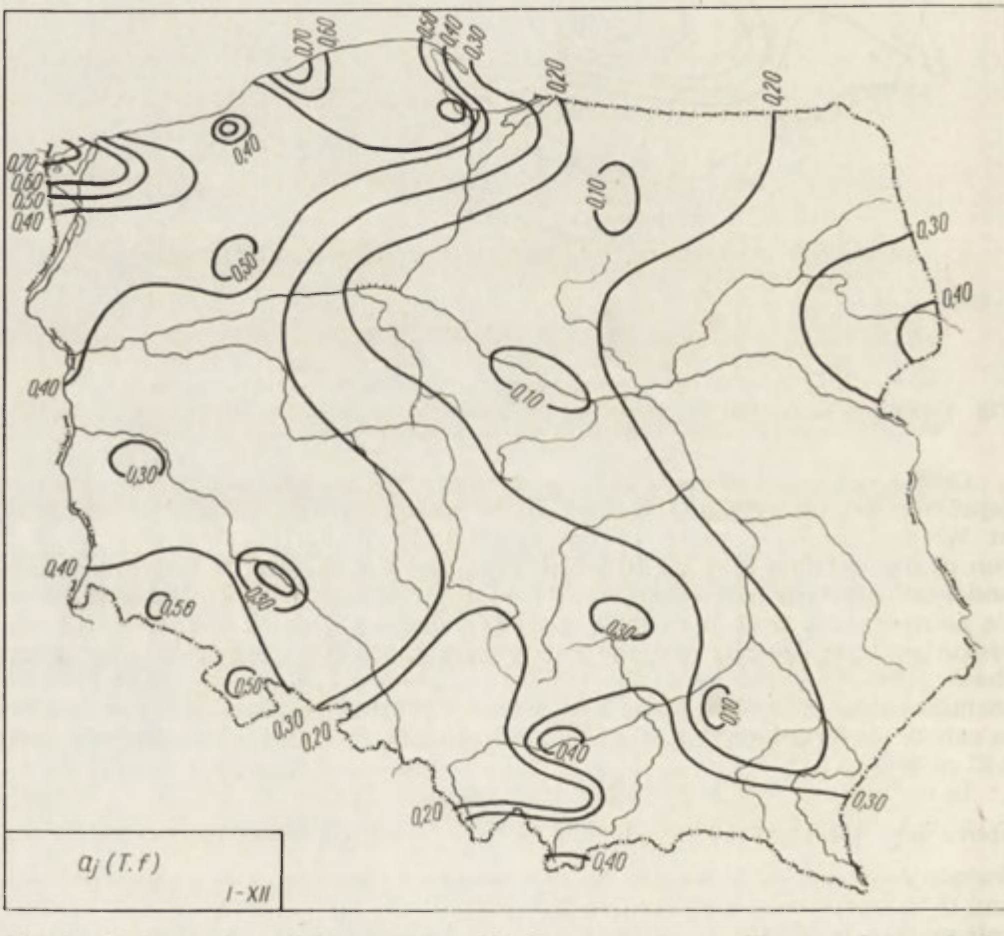


Fig. 5. Isolines of partial regression of air temperature with regard to relative humidity: $a_j (T \cdot f)$ for one year

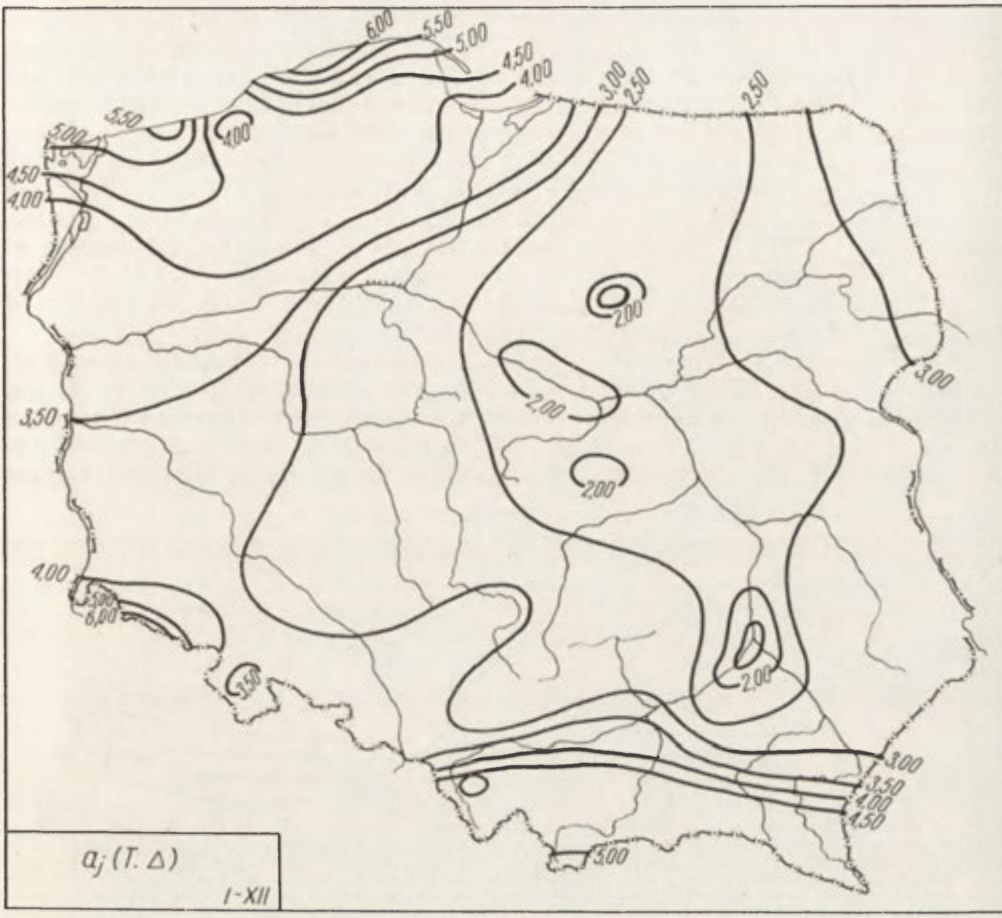


Fig. 6. Isolines of partial regression of temperature with regard to undersaturation of air humidity: $a_j(T \cdot \Delta)$ for one year

decreases per degree of clouding ($a_j < -2.00$) are recorded. Negative increases of $\frac{T}{N}$ prevail in all of Poland, but as a rule they are lower in the west than in the east of Poland. An exception are the Carpathians, the Karkonosze Mts. and a large part of the Pomeranian Lake District, where an increase in clouding is associated not with a decrease but an increase of T.

The distribution of the coefficients of regression for T with regard to the sums of *atmospheric precipitation* in Poland's territory, as it appears in Fig. 8, reveals rather low values of a_j — an obvious fact. A 1 mm increase in precipitation cannot be reflected in a temperature increase higher than that of the order of one hundredth parts of one °C — the more so since here we are dealing with annual values. Here the run of isoline $a_j = 0.04$ is most remarkable. It clearly separates the eastern part of Poland, where a temperature decrease twice as large per unit of precipitation prevails ($a_j = 0.04$–0.06 and $a_j > 0.06$), from the western part of Poland where a_j oscillates between 0.02 and 0.04.

Our analysis of the distribution of the coefficients of regression for Poland's territory indicates, that the links of correlation are closest between T and Δ

($a_j = 2.50$–3.50) and between T and v where the coefficients of partial regression lie between $a_j = -1.50$ and -1.00. Next it appeared, that the coefficients of regression show the lowest values for T and O ($a_j = 0.02$–0.04 and 0.04–0.06). In the majority of cases the coefficients of regression differ when passing from the eastern to the western part of Poland, whereby for T_p, T_v, TN and TO these values decrease considerably while for Tf and $T\Delta$ they increase. In the majority of maps (with TN and TO omitted) it can be seen, how the central belt of Poland stands out, extending more or less meridionally ($\lambda = 19$–$22°$) and featured by the lowest values for coefficients of regression for the relation Tf and $T\Delta$, and highest values for T_p and T_v. This central belt does not include the Carpathian range, along which the isolines of regression run for the most part in the direction of the parallels of latitude.

Our analysis of the coefficients of multiple correlation of mean diurnal air temperature with regard to four meteorological parameters (p, v, Δ, N) has revealed, that for these agencies the values are lower than those in which the six meteorological parameters (p, v, f, Δ, N, O) where taken into account. On the other hand, the coefficients of regression increased at the rate highest

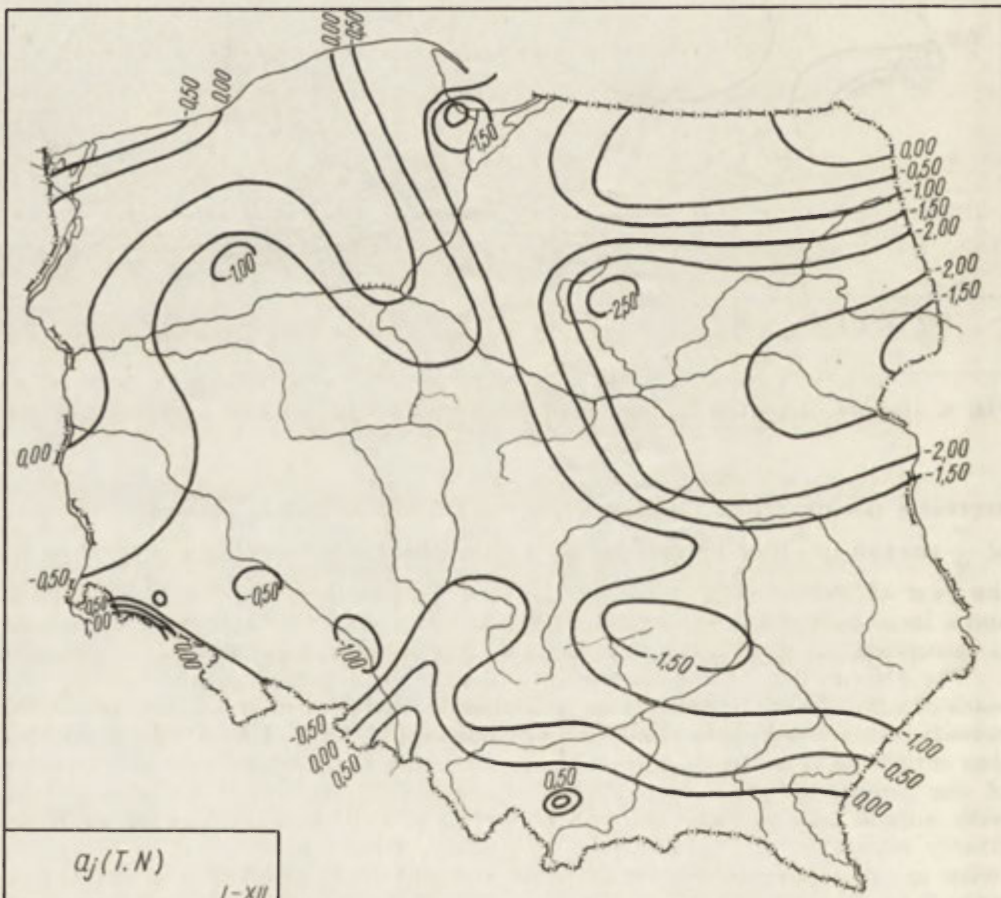


Fig. 7. Isolines of partial regression of air temperature with regard to clouding: $a_j(T \cdot N)$ for one year

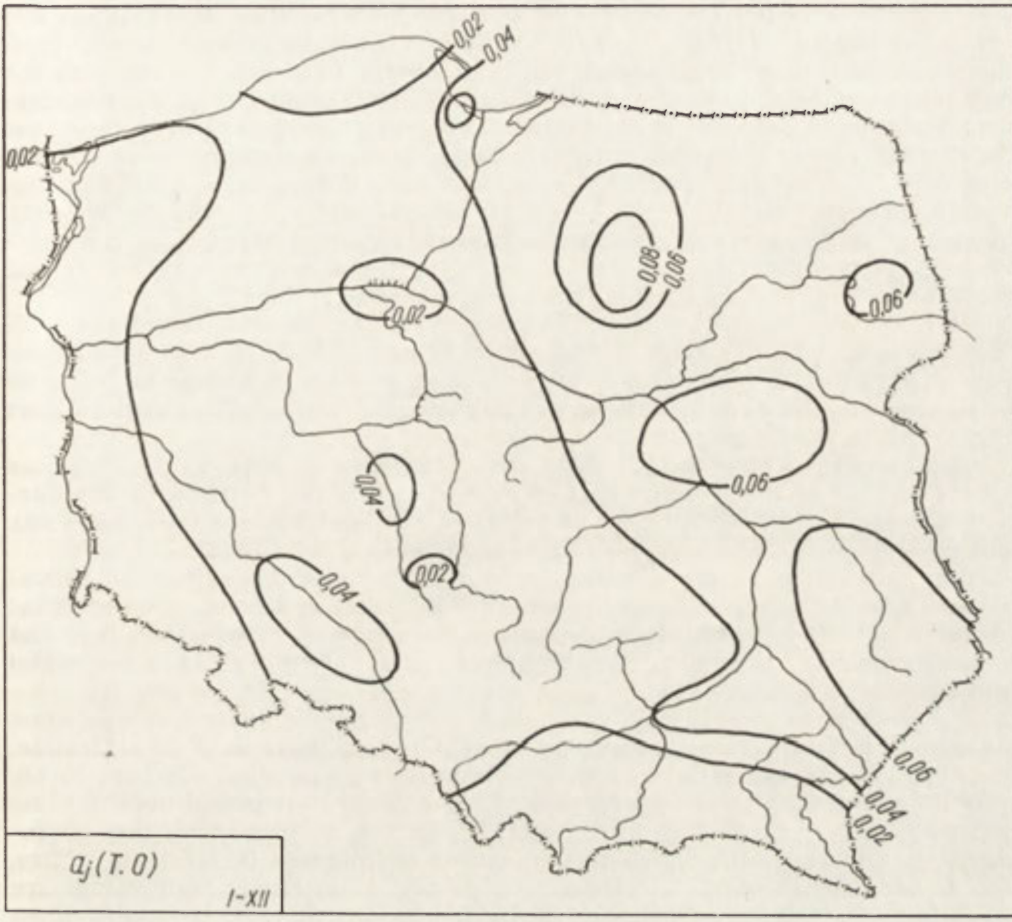


Fig. 8. Isolines of partial regression of air temperature with regard to atmospheric precipitation: $a_j\,(T \cdot O)$ for one year

for the components of water circulation in nature ($T\Delta$ and TN), and lowest for Tp. No significant changes appeared with regard to the run of the lowered and raised values of isarithms, nor to the lowest and highest isarithms with regard to the mean values for Poland.

After eliminating the effect of clouding upon air temperature (T and p, v, Δ), the coefficients of multiple correlation showed practically no changes with regard to the value of R which had taken clouding also into account (T and p, v, Δ, N). However, with regard to the former group of variables (T and p, v, f, Δ, N, O) the values decreased at an average of 0.04. As to the coefficients of partial regression for Tp and Tv, these resemble their corresponding variable in the former group and are slightly lower than their corresponding variables in the latter group. For the last pair of variables $T\Delta$, the a_j coefficients maintain the values they had in the latter group — thus they showed an increase with regard to those of the former group.

The author decided to change diametrically the composition of the latter group of meteorological parameters, retaining wind velocity and adding, from the components of water circulation in nature, clouding and precipitation. Considering this changed composition of the group of variables the greatest decrease, averaging 0.10–0.20, compared with previously mentioned values, show the coefficients of multiple correlation. This observation does not apply to the coefficients of partial regression. Investigating the dependence of air temperature upon wind velocity, a certain deviation was determined here from the regularities so far observed: a_j reveals the highest temperature decreases per unit of wind velocity, especially in north-eastern Poland ($a_j = -5.97$ for Suwałki) and at the margin of the Lesser Poland Plateau ($a_j = -6.79$ for Wieluń). The isoline $a_j = -2.50$ embracing areas of fairly high regression, runs in a nearly meridional direction; it separates the eastern and western part of Poland from the central belt in which a_j shows values below -2.00.

The coefficients of partial regression for TN can only be correlated with the first and second groups of variables. In this aspect an abrupt increase appeared in the a_j values which vary from < -3.00 in the west to > -3.00 in the east, in the eastern part of the Pomeranian Coast, in the western part of the Lublin Plateau, and in the Carpathians.

The highest increase in a_j values shows the relation of T to O. They rise almost tenfold in the western part of Poland ($a_j > 0.13$), fourfold in the Carpathians and in the central part of the Lesser Poland Plateau ($a_j = 0.06$–0.08), and twofold in the east and the centre of Poland ($a_j = 0.10$–0.12).

Our calculation of the coefficients of multiple correlation and of partial regression of air temperature with regard to different groups of meteorological parameters makes it possible to determine the range of changes which R and a_j are undergoing, depending on the number and the type of variables taken into account.

The coefficients of multiple correlation revealed their highest values when all meteorological parameters (p, v, f, Δ, N, O) were taken into consideration, and their lowest values after atmospheric pressure was eliminated (v, N, O).

With reference to the hyperplanes of regression, most conspicuous are the coefficients of regression a_j for $T\Delta$ and Tv, on the understanding that an increase in undersaturation of humidity causes an increase in air temperature, and an increase in wind velocity a temperature drop. Fairly high values are also reached by the coefficients of partial regression for TN, especially when pressure and relative humidity are eliminated. On the other hand, lowest a_j values show air temperature and precipitation. Regarding the spatial pattern of the investigated rates of interdependence between R and a_j, the greatest are the diversities appearing in the coefficients of partial regression for T and v; this can probably be ascribed to local conditions which in the highest degree affect the horizontal component of air motion (v). Generally speaking, the coefficients a_j show the greatest density of isolines between the eastern (partly also the central) and western part of Poland.

FINAL COMMENT

In the present study, bearing a methodical and cognitive character, the author expressed by empiric equations both the spatial pattern of air temperature (T, T_{max}, T_{min}, A) in Poland's territory and the interdependence between air temperature and other meteorological parameters.

The determined equations of the hyperplanes of regression and the coefficients of multiple correlation furnish a comprehensive picture of the impact of geographical and climatic conditions upon air temperature in Poland. Calculating the partial coefficients of regression, the author eliminated the impact upon

T, T_{max}, T_{min}, and A of all other meteorological parameters correlated both with themselves and with air temperature. In this way the author expressed by equations the thermic features of Poland's climate, based on fortuitous variables as they were recorded by 120 meteorological stations.

The equations of the hyperplanes of regression which approximately define groups of meteorological parameters in Poland's territory, can be used for supplementary interpolations of the lacking parameters of T, T_{max}, T_{min}, and A.

By reading from the map the geographical co-ordinates of a given point (φ, λ, H), its geographical latitude φ, its geographical longitude λ and its altitude H a.s.l., one can fairly accurately define for that point the values of the unknown parameters T, T_{max}, T_{min} and A.

Table 5 indicates the order of magnitude of the differences between parameters determined from equations of regression and from measurements.

TABLE 5. Differences of air temperature and of its diurnal amplitude between ten-year values (1951–1960), and values determined from equations of planes of regression — as per equations (10) and (11)

	φ	λ	H	ΔT		ΔA	
				X–III	IV–IX	X–III	IV–IX
Koszalin	54.1	16.1	33	+0.3	−0.5	+0.1	−0.6
Suwałki	54.1	22.6	165	−0.4	+0.2	+0.5	+0.4
Gorzów	52.4	15.2	65	−0.3	+0.2	−0.3	−0.3
Siedlce	52.1	22.2	146	−0.2	−0.1	+0.2	+0.4
Lublin	51.1	22.3	171	+0.1	−0.1	−0.2	0.0
Kielce	50.5	20.4	261	−0.3	−0.1	+0.3	+0.5
Cracow	50.0	19.6	206	−0.1	+0.1	−0.5	+0.8
Krynica	49.3	20.6	613	−0.5	−0.6	+0.3	+0.3
Śnieżka	50.4	15.4	1603	+0.4	−0.8	−0.9	+1.1
Kasprowy Wierch	49.1	19.6	1991	−0.3	−0.6	−0.7	+0.9

By determining for 60 control stations in Poland the coefficients of multiple correlation and of partial regression for the mean diurnal air temperature with regard to a variety of groups of meteorological parameters, the author's aim was to define the range of changes these coefficients are undergoing, dependent upon the number and the kind of meteorological parameters taken into account.

The coefficients of multiple correlation attained their highest values when all variables were taken into consideration (p, v, f, Δ, N, O), and their lowest values after elimination of atmospheric pressure (v, N, O). This shows, that air temperature is strongly correlated with pressure which considerably restricts the intensity of heat exchange and water circulation in nature — and in this way pressure clearly bears upon the behaviour of particular meteorological elements. Any increase in wind velocity and in the degree of clouding is accompanied by a drop of air temperature.

The impact of atmospheric pressure (baric configurations) upon air temperature appears also in the coefficients of multiple correlation; but here this impact is less evident due to the opposite action of some meteorological parameters (f, Δ, O).

As to the spatial disparity of the investigated values for the interrelation between R and a_j, this disparity is greatest between the eastern (partly also the central) and the western part of Poland.

The transitional nature of Poland's climate, between maritime in the west and continental in the east, appears most strikingly in the coefficients of partial regression of air temperature regarding relative humidity and undersaturation of humidity (Tf and $T\Delta$), and regarding wind velocity (Tv), thus regarding parameters which usually are omitted in studies of Poland's climate. The width of the transitional zone differs depending upon the meteorological parameters investigated; mostly it is bordered by the 19–21° meridians.

The fact that in the annual values of interdependences the transitional zone stands out, indicates the important part played by the cool season in which a meridional pattern of isarithms and isocorrelates predominates (Atlas klimatyczny..., 1971; Gorczyński, 1918; Kaczorowska, 1959; Michna 1972; Okołowicz, 1962, 1967; Paszyński 1966; Romer, 1949; Schmuck, 1959; Stopa, 1962, 1968, 1873, 1974; Warakomski, 1964; Wiszniewski et al., 1949).

Warsaw University

BIBLIOGRAPHY

Atlas klimatyczny Polski. Temperatura powietrza. Część tabelaryczna (Climatic atlas of Poland. Air temperature. Part containing tables), 1971, Warszawa.

Boryczka, J., 1974, *Empiryczne równania klimatu Polski* (Empiric equations for Poland's climate), Rozpr. Uniw. Warsz. (in print).

Gorczyński, W., 1917, Linie jednakowej korelacji ciśnienia i temperatury w stosunku do Warszawy (Res.: Équicorrelates de la pression et de la temperature de l'air par rapport à Varsovie), *Spraw. Tow. Nauk. Warsz., Wydz. Nauk. Mat.-Przyr.*, 10, 2. Warszawa.

Gorczyński, W., 1918, *Nowe izotermy Polski, Europy i kuli ziemskiej* (Res.: Nouvelles isothermes de la Pologne, de l'Europe et du globe terrestre), Warszawa.

Gumiński, R., 1952, Materiały do poznania genezy i struktury klimatu Polski (Sum.: Origin and structure of Poland's climate), *Przegl, Geogr.*, 24, 3.

Hess, M., 1965, Piętra klimatyczne w polskich Karpatach Zachodnich (Sum.: Vertical climatic zones in the Polish Western Carapathians), *Zesz. Nauk. UJ*, 115, *Prace Geogr.*, 11.

Kaczmarek, Z., 1970, *Metody statystyczne w hydrologii i meteorologii* (Statistical methods in hydrology and meteorology), Warszawa.

Kaczorowska, Z., 1959, *Cechy charakterystyczne klimatu Polski* (Characteristic features of Poland's climate), *Prace Komit. Gosp. Wodnej*, 1, Warszawa.

Kaczorowska, Z., 1962, *Opady w Polsce w przekroju wieloletnim* (Sum.: Precipitation in long-period averages), Prace Geogr. IG PAN, 35, Warszawa.

Kosiba, A., 1952, *Wstęp do klimatologii Polski* (Introduction to Poland's climatology), Wrocław.

Merecki, R., 1914, *Klimatologia ziem polskich* (The climatology of Polish lands), Warszawa.

Michna, E., 1972, O wilgotności względnej powietrza w Polsce (Sum.: On relative atmospheric humidity in Poland), *Przegl. Geofiz.*, 17, 1.

Okołowicz, W., 1962, *Zachmurzenie w Polsce* (Sum.: Cloudiness in Poland), Prace Geogr. IG PAN, 34, Warszawa.

Okołowicz, W., 1967, Amplitudy temperatury powietrza w Polsce (Sum.: Air temperature amplitudes in Poland), *Przegl. Geofiz.*, 12, 3–4.

Parczewski, W., 1971, Dynamiczne aspekty klimatu Polski (Sum.: Dynamic aspects of Poland's climate), *Przegl. Geogr.*, 43, 4.

Paszyński, J. (Ed.), 1966, Atlas bilansu promieniowania w Polsce (Atlas of radiation balance in Poland), *Dokum. Geogr.*, 4.

Romer, E., 1949, Regiony klimatyczne Polski (Poland's climatic regions), *Prace Wrocł. Tow. Nauk.*, ser. B, 16.

Schmuck, A., 1959, *Zarys klimatologii Polski* (An outline of Poland's climatology), Warszawa.

Stopa, M., 1962, *Burze w Polsce* (Sum.: Thunderstorms in Poland), Prace Geogr. IG PAN, 34, Warszawa.

Stopa, M., 1968, Temperatura powietrza w Polsce (Air temperature in Poland), 1, *Dokum. Geogr.*, 2.

Stopa-Boryczka, M., 1973, *Cechy termiczne klimatu Polski* (Thermic features of Poland's climate), Rozpr. Uniw. Warsz., 72.

Stopa-Boryczka, M., Boryczka, J., 1974, *Atlas współzależności parametrów meteorologicznych i geograficznych w Polsce* (Atlas showing interdependence between meteorological and geographical parameters in Poland), Warszawa.

Warakomski, W., 1964, Izonefy miesięczne Polski (Sum.: Monthly isonephs in Poland), *Ann. UMCS*, sec. B, 17, Lublin.

Wiszniewski, W., Gumiński, R., Bartnicki, L., 1949, Przyczynki do klimatologii Polski (Sum.: Contribution to the knowledge of the climate of Poland), 2, *Wiad. Służby Hydrol.*, 5.

GEOGRAPHIA POLONICA 31, 1975

PENULTIMATE PERIOD OF DEGLACIATION IN THE GRUDZIĄDZ BASIN, LOWER VISTULA RIVER VALLEY: AN INTERSTADIAL-LIKE INTERVAL OF THE MIDDLE WÜRM

EUGENIUSZ DROZDOWSKI

INTRODUCTION

In the present paper an attempt has been made to reconstruct the processes of down-wasting of the penultimate ice sheet of the Last (Baltic, North Polish, Würm) Glaciation in the area of the Grudziądz Basin based on the analysis of the deposits and fossil landforms observed in three exposed profiles, including first radiocarbon dating. These profiles, selected from several ones which have been studied in the Lower Vistula Valley (E. Drozdowski, 1974), indicate quite well the type and succession of the deglaciation processes; in consequence they present a general picture of palaeogeographical conditions and make it possible to correlate them with definite evolutionary stages of the Last Glaciation observed in other areas.

THE AREA INVESTIGATED, AND ITS LOCATION WITH REGARD TO THE MAIN MARGINAL ZONES OF THE LAST GLACIATION

The Grudziądz Basin is the largest of three basin-like widenings of the Lower Vistula River Valley. It is up to 18 km wide and covers an area of some 240 sq.km (L. Czajkowski, 1969). Because of its specific natural features, the Grudziądz Basin has been separated in Poland's physico-geographical subdivision as a mesoregion within the macroregion of the Lower Vistula Valley (J. Kondracki, J. Ostrowski, 1973).

Of particular significance from the point of view of the stratigraphical and palaeogeographical problems considered, is the location of the area investigated in relation to the main marginal zones of the Last Glaciation. As shown in Fig. 1, the Grudziądz Basin lies south of the marginal forms of the Pomeranian Stage, that is, in the zone of the marginal forms of the Krajna–Wąbrzeźno Phase belonging to the Poznań State (R. Galon 1961, 1968; J. Kondracki 1965; S. Kozarski, 1962; W. Niewiarowski, 1959; L. Roszko 1956, 1968).

A characteristic feature of the Grudziądz Basin distinguishing it from the remaining basin-like widenings of the Lower Vistula Valley, are three "moraine plateau islands" called *kępy* (Fig. 1), elevated up to 60 m above the valley floor. East of the river rises Kępa Forteczna, with its summit at 86.1 m a.s.l., south of its lies Kępa Strzemięcińska 79.3 m a.s.l. high and, west of the river, separated from the channel by a flood plain some 4 km wide, lies Kępa Górnej Grupy, with its summit at 77.3 a.s.l. In a section perpendicular to the valley axis all these three hills show asymmetrical profiles; the river-facing slopes are steeper than those inclined towards the opposite side, where well developed terraces occur locally.

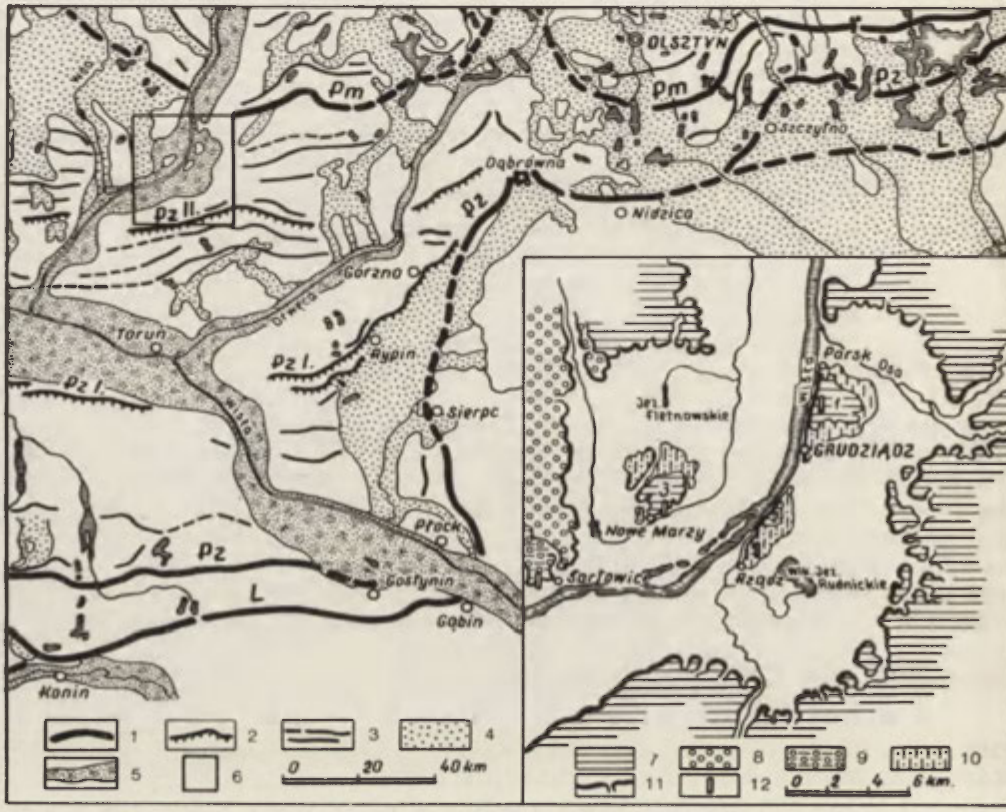


Fig. 1. Geomorphological sketch of the Grudziądz Basin, shown against the background of the lines of inland ice stoppage of the Baltic Glaciation as compared by L. Roszko (1968)

1 — stages: L — Leszno, Pz — Poznań, Pm — Pomeranian; 2 — phases: PzI — Cujavian, PzII — Krajna-Wąbrzeźno; 3 — positions of recessive glacier-margin; 4 — outwash plains; 5 — valleys and ice marginal channels (*pradolinas*); 6 — outline of geomorphological sketch of Grudziądz Basin; 7 — moraine plateau; 8 — valley outwash; 9 — erosive surface of meltwater flow; 10 — higher fluvial terraces surrounding three moraine plateau islands called: 1. Kępa Forteczna, 2. Kępa Strzemięcińska, 3. Kępa Górnej Grupy; 11 — rim of moraine plateau; 12 — documented exposures

LITHO-STRATIGRAPHICAL PROFILE OF THE LAST GLACIATION

A typical developed profile of deposits of the Last Glaciation in the Grudziądz Basin consists of three till strata and glacifluvial deposits separating them (Fig. 2). The lower boundary of this profile has recently been determined on the basis of palaeobotanical evidences in newly discovered sites of the Eemian Interglacial at Grudziądz–Mniszek and Rządz (E. Drozdowski, K. Tobolski 1972) where the Eemian deposits, developed in a facies of lacustrine sediments, underlie stratified fine-grained sands and grey varved clays resting on the locally occurring third (counting from today's surface) till stratum. Thus the entire thickness of sediments deposited during the Last Glaciation in the Grudziądz

Basin is in the order of 70 m; that is, it reaches upwards from some 10 m a.s.l. (the top of the palaeobotanic determined deposits of the Eemian Interglacial) to some 80 m a.s.l. (the mean altitude of the *kępy* summits and of the surface of the neighbouring morainic plateau).

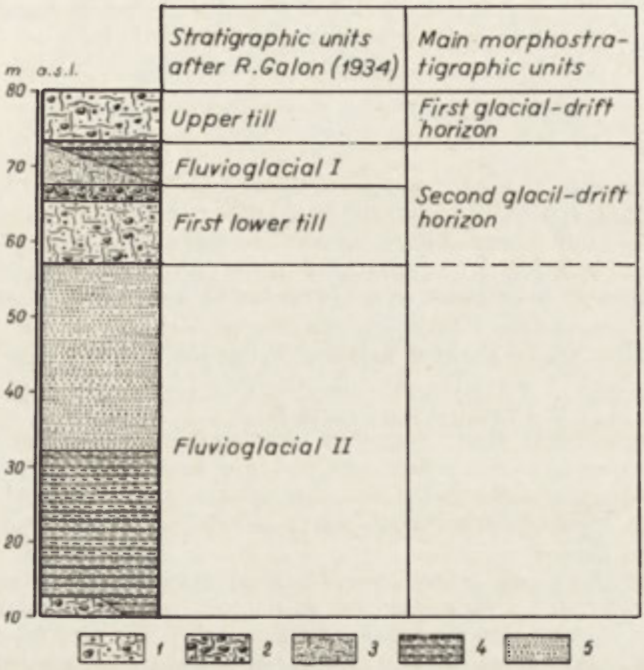


Fig. 2. Litho-stratigraphical profile of deposits of the Baltic (Würm) Glaciation in Grudziądz Basin

1 — lodgment till; 2 — supraglacial subaquatic till; 3 — supraglacial-solifluction till; 4 — varved clays; 5 — stratified fine-grained sands

Worth particular attention is the lower part of the stratigraphical profile, embracing a series of stratified fine-grained sands and, underlying them, grey varved clays (the latter being exploited by brickyards in Grudziądz and Świerkocin). Due to their considerable thickness and broad extension, R. Galon (1934) separated out these deposits as the stratigraphical key-horizon of the Pleistocene in the area of the Lower Vistula valley, calling them Fluvioglacial II Fig. 2. This series overlies the second till stratum which is commonly separated from the first, youngest till stratum by ice-dammed lake sediments (Fluvioglacial I).

The second till stratum often is bipartited. Investigations carried out (E. Drozdowski 1974) show that the lower layer of this stratum, of a brown colour, represents a lodgment till, originated in subglacial and englacial conditions, while the upper layer, usually of a grey colour, is a supraglacial subaquatic till which developed in peculiar conditions due to stagnation or to the obstructed runoff of supraglacial waters. This interpretation is supported principally by the stratigraphical sequence and grain-size composition of the deposit. The supraglacial till shows enrichment by both extreme fractions, i.e. by a fine clayey material and by a very coarse material consisting of pebbles and boulders.

Least elucidated so far are the age and the palaeogeographical conditions of the accumulation of the deposits. This particularly refers to the accumulation of tills and their relationship to the marginal zones. R. Galon (1961, 1968) and J. E. Mojski (1969) believe, that the two youngest till strata should be correlated with the end moraines and deposits of the two main marginal zones of the Baltic Glaciation, the former with the Poznań Stage, the latter with the Leszno Stage. However, this correlation is rather hypothetic, being based only on disputable evidence of a numerical concordance of the till strata in any given geological section, and main marginal zones situated south of this section.

DEFINITION OF THE OBJECT AND METHODS OF STUDY

Rock-stratigraphic classification turned out unpractical for the chronological and palaeogeographical reconstructions, therefore a subdivision into morphostratigraphical units is used here, as was proposed by J. C. Fry and H. B. Willman (1962). This subdivision adopts as a basis of reference the morphological action of the inland ice, linked with different in space and time processes of glaciation and deglaciation which are considered not only from the view-point of climatic changes and a general balance of the ice mass (T. Bartkowski 1969), but also as an effect of other factors, like the relief of the substratum of the ice sheet, movements of the Earth's crust, and the runoff conditions of the meltwater (E. Drozdowski 1973). Such a way of classification has found its expression in complexes of deposits, distinguished in regard to origin and period of occurrence. A unit of higher order, what the author calls a glacial-drift horizon, comprises the glacial and glacifluvial deposits developed during the advance and the decay of one inland ice sheet.

The object of the study is the second glacial-drift horizon, covering the deposits laid down by the penultimate Scandinavian inland ice in the area under discussion. In the synthetic stratigraphical profile of the Pleistocene in the Lower Vistula Valley (R. Galon 1934) this horizon is represented by the first lower till stratum and, superimposed on it, the so-called Fluvioglacial I (Fig. 2). On top of this horizon lies the upper, i.e. the youngest till stratum, while underneath it extend fine-grained sands and grey varved clays, the so-called Fluvioglacial II. It should be mentioned, that so far the origin of these latter sands and clays of Fluvioglacial II has not been interpreted in a convincing way, and therefore their morphostratigraphical position is uncertain. Yet it seems most probable, that genetically and chronologically these deposits should be assigned to the second glacial-drift horizon, and for this reason the lower boundary of the second glacial-drift horizon should be looked upon as preliminary.

The subsurface occurrence of the discussed deposits, as well as the author's effort to arrange these deposits in proper order in accordance with their genetic-facial and chronological interpretation, made him apply diverse and mutually supplementing research methods. Three methods proved of cardinal importance: the stratigraphical method as a basis for establishing the position of the investigated deposits in the total and partial stratigraphical profile of the Last Glaciation, the structural and the palaeomorphological methods which both supplied information about the dynamic and the palaeogeographical conditions under which the deposits and landforms have been developed.

For obtaining additional criteria for interpreting the till genesis and its morphostratigraphical classification, the grain-size and till-fabric analysis were used. In measuring the fabric attention was paid both to the orientation and the dip of the long axes of pebbles (*a*), granting that in a lodgment till they will tend to be dipped mainly on two sides, i.e. according to the ice-movement direction and in the opposite direction (P. W. Harrison 1957; E. Rukhina 1960; R. P. Kirby 1969; G. S. Boulton 1972). Divergence from this pattern was interpreted as indication of deviating conditions of glacial accumulation or redeposition of the material.

The evidences presented here are derived from three large profiles which, in fact, being natural or artificial exposures, were freed of loose material and deepened. These exposures will be described in chronological order of the most important geological and palaeogeographical events recorded, i.e., at Parsk, Nowe Marzy and Rządz.

PARSK

This exposure, facing the river, is located in the western slope of Kępa Forteczna near a triangulation point at 75.0 m a.s.l. (Fig. 1).

Here the top of the fine-grained sands of the intermorainic key-horizon (Fluvioglacial II) is overlain by two complexes of glacial and glacifluvial deposits which were conventionally called the lower and the upper accumulation complex. The lower complex, some 7 m thick, consists of glacifluvial deposits covered by two till layers which in turn are separated by clayey silt and sand (Fig. 3: 4–12), while the upper complex, up to 3 m thick, consists of a greyish-green till, underlain by a lense of fine-grained sand and, higher up, of a till stratum of dark-brown colour (Fig. 3: 2–3a).

The base part of the lower accumulation complex consists of stratified fine-grained sands showing flow structures (Fig. 3: 11–12). The azimuths of the run of the laminae oscillate between 150° and 190°, the runs from 22° to 44° towards the NE. The primary structure of this deposit is disturbed by faults extending in various directions.

The next (counting in an upward direction) sedimentation unit is a series up to 3,5 m thick of poorly washed, non-stratified fine-grained sands with an admixture of coarser sandy and gravelly fractions (Fig. 3: 10), bordered by stratified silty sands (Fig. 3: 9), with the inclination of the top layers from 20° to 25° in ENE direction.

Higher up one observes a bed of grey till (Fig. 3: 8), separated from the successive bed of brown-red clay (Fig. 3: 4) by dark-brown clayey silts and fine-grained sand (Fig. 3: 5–7). In a parallel plane as well as in a plane perpendicular to the run of the bed the borders of the till beds are rough and uneven. Both till beds contain a considerable amount of clayey and boulder-gravel material and therefore they are dense and "stony", resembling supraglacial (subaquatic) till in the typically developed bipartite till profile of the second glacial horizon (E. Drozdowski, 1974). The side wall of this exposure shows that the deposits are inclined eastwards at 22° to 45°, that is, in a direction opposed to the present inclination of the *kępa* slope; moreover, nests and lenses of fine-grained sands (Fig. 3: 6–7) are visible in the silty-clayey deposits also eastward slanting.

The structural features of the glacifluvial deposits underlying the till beds (Fig. 3: 9–12) indicate that they have been laid down under variable hydrodynamic conditions of material transport and a changing inclination of the surface of the accumulated material. This kind of environment is probably created locally by water passing through basins developed in ice crevasses or in ice holes. Initially the process of material transport and accumulation must have proceeded while the accumulation surface was steeply inclined; this seems to be indicated by flow structures observed in the sands of the bottom of the lower complex, and by the steep slant of their laminae which locally exceed the natural angle of repose of sandy deposits. After an erosive cutting of the previously deposited material, there set in a sedimentation of poorly washed sands and gravels (Fig. 3: 10), initiated and terminated by an accumulation of silty sands (Fig. 3: 9).

Two till beds with clayey silt separating them were next deposited. Together they represent supraglacial ablation deposits redeposited by flowages and slidings on the eastward inclining slope of the ice crevasse. Evidence for this interpretation are the results of till fabric analysis. As seen in fabric diagrammes

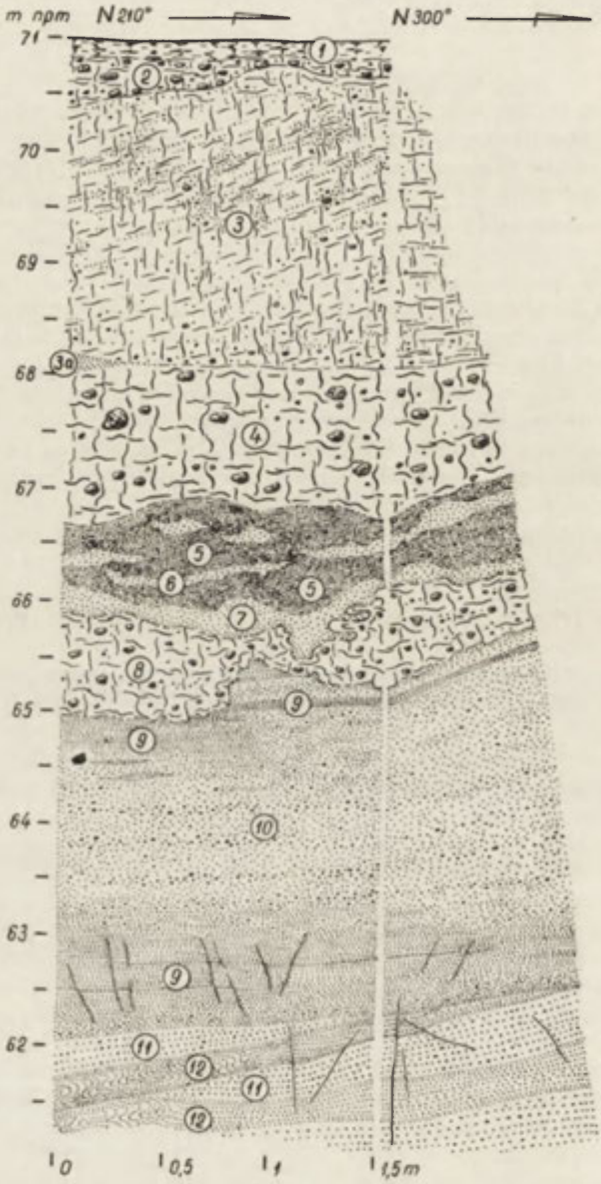


Fig. 3. Parsk. Geological profile of the upper part of the western slope of Kępa Forteczna

1 — humus; 2 — dark-brown till (supraglacial flowtill); 3 — greyish-green sandy till (supraglacial-solifluction till); 4 — brown-red till (supraglacial flowtill); 5 — dark-brown clayey silt; 6 — fine-grained sand; 7 — fine-grained sand with admixture of coarser sand-gravel fractions; 8 — grey till (supraglacial flowtill); 9 — silty sand; 10 — fine-grained sand with admixture of coarser sand-gravel fractions; 11 — fine-grained sand with gravel grains; 12 — fine grained sand

(Fig. 4: b, c), in both till beds the *a*-axes show an E-W orientation, and for the most part they dip in an eastward direction.

The slight disjunctive disturbances of the sandy deposits probably originated from an uneven collapsing of the deposits, caused by the melting of buried winter ice lenses or of minor glacier ice lumps.

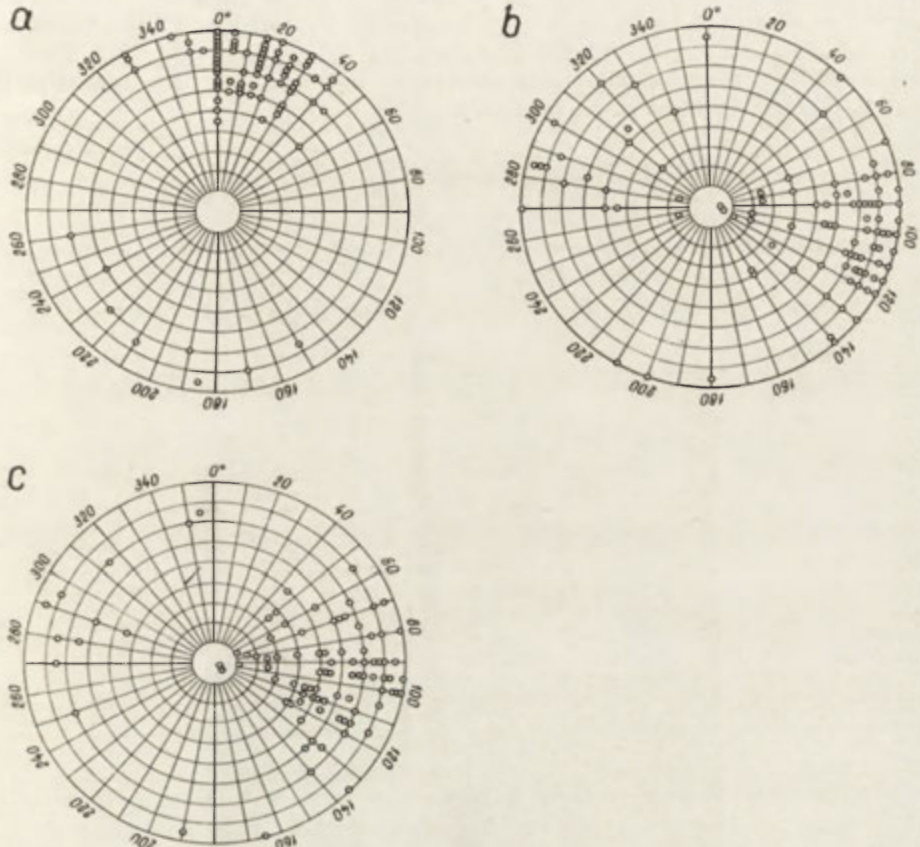


Fig. 4. Parsk. Till fabric
a — in deposit No. 3; b — in deposit No. 4; c — in deposit No.8

The upper accumulation complex registers the next phase of decay of the ice sheet. This phase preceded the deposition of the youngest till strata (not shown in this exposure) which in the summit part of Kępa Forteczna approaches a thickness of 7 m.

The greyish-green till of the upper accumulation complex (which after drying assumes a light-brown tint with clearly noticeable streaks of iron compounds) differs from the subaquatic supraglacial till most of all by its grain-size composition. It consists mainly of a sandy fraction (40–45%) and reveals a paucity of coarsest particles like pebbles and boulders (E. Drozdowski, 1974). The upper and middle parts of the till contains sandy lamellae inclined northwards at 12 —20° (Fig. 5). In the bottom occur sandy inclusions with an admixture of organic substance. A search for sporomorphs proved futile (B. Noryśkiewicz 1971, personal communication).

The *a*-axes of pebbles in the greyish-green till are predominantly arranged in a meridional direction, and in their majority they show a NNE dip (Fig. 4a). Thus the direction in which the long axes are inclined agrees with the dip of the sandy lamellae. On the assumption that the sandy lamellae represent the slip--planes of the deposits, the fabric would be consistent with the stone arrangement in the solifluction deposits (G. Lundqvist, 1949). However the exposure under discussion fails to support this elucidation by the line of the fossil slope and the downward increasing thickness of the redeposited material. Such evidence regarding till of similar characteristics has been gained from the next exposure mentioned, that at Nowe Marzy.

Fig. 5. Parsk. Stratified structure of the supraglacial till displaced by solifluction (Fig. 3 : 3)

The greyish-green till is again overlain by a "clayey-stony" till (Fig. 3: 2) which as to grain size and colour ties in with the redeposited subaquatic supraglacial till of the lower accumulation complex.

NOWE MARZY

This exposure, consisting of two parts called A and B, is located in the upper part of the eastern slope of Kępa Górnej Grupy, north of the triangulation point 77.7 m a.s.l. which is the highest point of the *kępa* (Fig. 1).

Noticeable below the brown till, the youngest in the area investigated, occurs a fragment of a fossil depression filled with sandy greyish-green till known from the previously presented exposure (Fig. 6). Exposure A pictures the succession of the main lithological units in the marginal part of this depression (Fig. 7). Underlying a stratum of the upper till (Fig. 7: 1) of an entire thickness of some 7 m extends a layer of fine-grained sand, several centimeters thick (Fig. 7: 2), next in downward order lies a darkbrown clayey silt, with streaks of sandy silt,

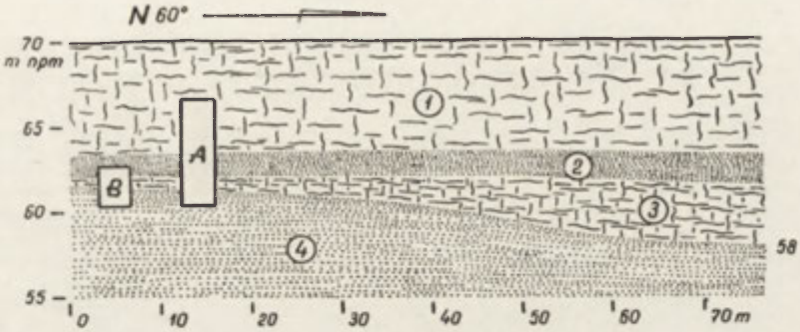


Fig. 6. Nowe Marzy. Geological structure of the upper part of the Kępa Górnej Grupy
1 — brown till; 2 — dark-brown clayey silt; 3 — greyish-green till (supraglacial-solifluction till); 4 — fine-grained sand; A, B — exposures

about 1 m thick (Fig. 7: 3), and beneath this extends the above mentioned greyish-green till, gradually wedging out in the marginal part of the fossil depression (Fig. 7: 4). This latter deposit lies conformably on the sandy sediments of Fluvioglacial II which in their top part consist o ffine- and medium-grained sands with some gravel grains (Fig. 7: 5).

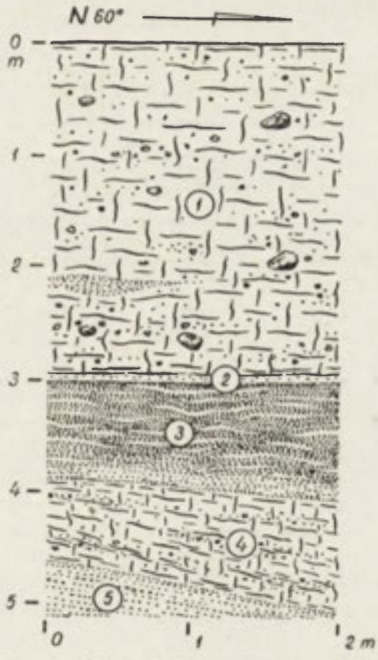


Fig. 7. Nowe Marzy. Exposure A
1 — brown till; 2 — fine-grained sand; 3 — dark-brown clayey silt streaked with sandy silt; 4 — greyish-green till (supraglacial-solifluction till); 5 — fine- and medium-grained sands with gravel grains

How the greyish-green till is deposited can be observed in detail in exposure B, situated 4 m south of exposure A (Fig. 6 and 8). In the front wall of this exposure one observes the characteristic increase in thickness of the greyish-green till bed in a direction parallel to the slope inclination of the fossil depression. In a distance of 1.30 m this thickness grows twofold, from some 30

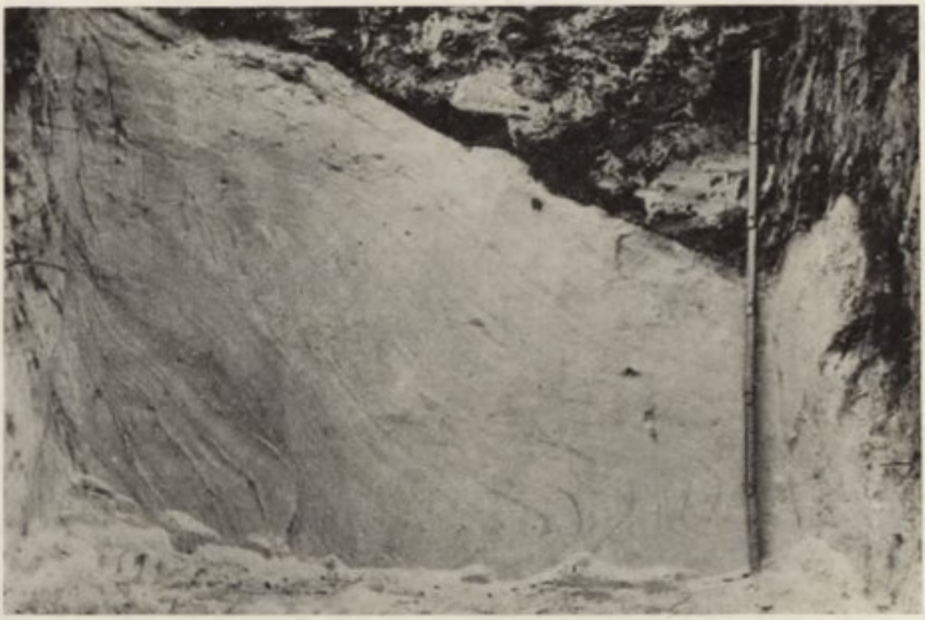


Fig. 8. Nowe Marzy. Exposure B (Fig. 6). Structure of sandy ablation deposits and their contact with supraglacial-solifluction till. Scale: 1 m

Fig. 9. Nowe Marzy. Till fabric in deposit No. 3 (Fig. 6)

to 60 cm. The dip angles of the bottom and top of this bed differ: the bottom drops at a 58°, the top at only 18°. In contrast to its corresponding exposure in the western slope of Kępa Forteczna, no traces of a stratification can be seen here. Till fabric analysis shows, that in most places the long axes are inclined towards the centre of the depression at an angle equal to, or smaller than, the slope of the depression (Fig. 9). However, the till filling the central part of the depression shows variable fabrics. The angles of the long axes inclination are often steeper than those of the slope of the depression (or the till base); some pebbles occur even standing up vertically. These features may be due to changes in the direction of pressures exerted by mass-flow, caused by the collision of mutually opposed lobate till flows in mode similar to that described by A. Jahn (1970) and G. S. Boulton (1972).

All the facts enumerated above, that is, the stratigraphical position of the till under discussion, its palaeomorphological position, its thickness, structure, fabric pattern, as well as the differences in grain-size composition compared with the subaquatic supraglacial till (E. Drozdowski 1974) seem to indicate this to be a supraglacial-solifluction till — in other words, a subaerial ablation moraine which has been displaced by solifluction on the slopes of an already decayed and broken-up ice sheet. This assumption has found its confirmation in the petrographical composition of gravels in the 5–10 mm and 2–5 mm fractions, which revealed, among other features, in the supraglacial-solifluction till a share of quartz grains twice as high as in the subaquatic supraglacial till (E. Drozdowski, 1974). This phenomenon can be explained by frost disintegration of the crystalline rocks.

The structural properties of the sands underlying this till and the nature of the contact zone between these two different types of deposits suggest an attempt of a closer reconstruction of palaeomorphological and palaeoclimatic conditions. The supraglacial-solifluction till overlies sandy deposits, stratified parallel with the inclination of the slopes of the depression (Fig. 8). The steep dip angle of the laminae, locally exceeding the angle of repose of sandy deposits (35–52°), as well as local flow deformations implied that these deposits might be ablation products accumulated *in situ*; their origin might go back to flow of rock material melted out from the ice and washed down onto the slopes of the ice hole.

The supraglacial-solifluction till, signifying a further stage of decay of the ice sheet, has probably been laid down on the surface of autochthonic winter ice which for some time served as protection to the ice hole. This time may have occurred during some cooler oscillation of the period of ice-sheet decay — a period which probably promoted complete freezing of shallow lakes. Afterwards, during renovated climatic amelioration, melting of the winter ice started and continued even after the solifluction processes ended and when on top of the redeposited material a new water basin developed. This interpretation seems to be supported, on the one hand, by the undisturbed original sedimentation structure of the ablation sands and their sharply marked conformable contact with the superimposed supraglacial-solifluction till and, on the other hand, by the fact that both the till and the overlying clayey silts show thicknesses increasing towards the centre of the depression.

All the arguments presented above suggest that the accumulation of the deposits has proceeded here similar to those in the western slope of Kępa Forteczna, in two stages. The first stage has been recorded in the sandy supraglacial deposits, the second in supraglacial-solifluction till and in the sediments of glacial lakes.

The upper glacial till, showing features typical for lodgment till, goes back to the last glacier advance, the youngest in this area.

RZĄDZ

This exposure shows a fragment of the southern wall of a gravel pit situated in the south-eastern slope of Kępa Strzemięcińska, within the area of the fluvial terrace V which here lies at 35–36 m a.s.l. (Fig. 1).

Exposed under the erosional surface of this terrace, can be seen a section across glacial and glacifluvial deposits filling a local depression in the former glacier substratum, some 60 m in diameter and up to 5 m deep. Fig. 10 shows the eastern, most interesting part of this section.

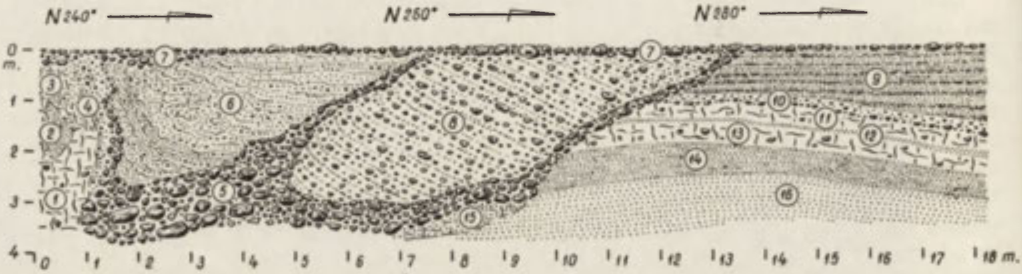


Fig. 10. Rządz. Structure of glacial deposits in the southern wall of the gravel pit
1 — till; 2 — fine-grained sand; 3 — sandy till; 4 — sandy gravel, in sides and in downward direction passing into: 5 — gravel with boulders; 6 — various-grained sands with gravel admixture; 7 — boulder pavement lying on erosive surface of terrace; 8 — rhythmically stratified sandy gravels with pebbles, and sands with admixture of gravels and lumps of clayey silts; 9 — clayey silts in varved structure; 10 — various-grained sands with gravel admixture and boulders in the bottom; 11 — till; 12 — fine-grained sand; 13 — till; 14 — silty sand; 15 — clayey sand; 16 — fine-grained sand

Fig. 11. Rządz. Fragment of the southern gravel pit wall (Fig. 10)
A — series of glacial and glacifluvial deposits lying *in situ* (Fig. 10:9–16); B, C, D, E — series of redeposited sediments filling fossil erosive landform (Fig. 10:1–7)

In general, one distinguishes here five markedly different series of deposits (Fig. 11); in ascending accumulation order they are: Series A — glacial and glacifluvial deposits lying *in situ* (Fig. 10: 14–9), Series B — pebbles and gravels resulted from till washings and moved down by gravity (Fig. 10: 5), Series C — deposits of fluvial, lacustrine and glacial accumulation slid down *en bloc* (Fig. 10: 8), Series D — sandy supraglacial deposits displaced *en bloc* (Fig. 10: 6), and Series E — morainic and sandy-gravelly deposits of the "crevasse-filling" type (Fig. 10: 4–1).

The stratigraphical position of these deposits is indicated foremost by series A which was laid down on the primary depression of the glacier substratum consisting of fine-grained sands (Fig. 10: 16). These sands as well as their underlying varved clays (which also occur in the slope of the *kępa* some 1 km from the exposure) represent the stratigraphical key-horizon of the Pleistocene in the area along the Lower Vistula Valley, called Fluvioglacial II (Fig. 2). The bottom of series A contains two till beds separated by fine-grained sand (Fig. 10: 13–11). The lower till bed is 30–40 cm thick, the upper 30–50 cm, whilst the thickness of the dividing sand layer is some 10 cm. The contacts are indistinct, the tills pass gradually into sands. In their bottom parts both till beds show a weak horizontal stratification parallel with their bottom fissility. The top of the upper till bed shows an even surface and is covered by a gravel-and-sand bed 27 to 52 cm thick (Fig. 10: 10). In the lower contact zone boulders occur of sizes up to 30 cm in diameter, upwards they pass into cross-stratified medium-grained sands with a gravel admixture; this sands shows a westward inclination of 12–18°.

The youngest sedimentation unit of series A are dark-brown clayey silts of varved structure (Fig. 10: 9). Their thickness gradually grows towards the centre of the depression from some 1.0 to 1.5 m.

The remaining series B, C, D and E fill a fossil depression carved by a meltwater flow in the base of the ice crevasse. In contrast with series A, they constitute deposits of an indirect glacial accumulation, reworked and redeposited by water action and gravity movements.

Series B consists of boulder-gravel material of the nature of a boulder pavement (Fig. 10: 5), and of clayey sands laid down in the marginal parts of the depression (Fig. 10: 15). The greatest thickness of the series B deposits appear to be in the centre of the depression below series D and E; the bottom of series B was not reached at the depth of 1.50 m. The form in which the boulder-gravel material occurs and the position of the larger is remarkable. As visible in Fig. 10, all the "layers" separating series A from C, series C from D and series D from E are gradually widening downwards, resembling the shape of debris cones; at the same time, in a downward direction grows the number and the size of the boulders. Beneath series D and E lie the largest boulders, some reaching as much as 50 cm in diameter. This seems to indicate that, after carving a trough in the bottom of the ice crevasse and during the accumulation of series C and D, this material was carried downwards by gravity as loose debris and in dependence on the inclination of the ice-and-mineral slopes, underwent a gravitational segregation. However, the strong abrasion and roundness of the crystalline rocks caused apparently by turbulent flowing water prior to the final depositional process (A. Cailleux, J. Tricart, 1959) is worthy of attention.

Genetically different deposits occur in series C and D. Series C (Fig. 10: 8) constitutes a complex of rhythmically stratified sandy gravels with pebbles (layers 7–15 cm thick) and of sands containing some gravel and clay lumps (layers about 1 cm in thickness). These deposits, separated by a distinct line from the gravels and boulders of series B, show throughout a northward dip of 18°. An upward bend of the layers observed by the borders of the series is noticeable here.

The deposits of series C very much resemble sediments of the marginal lakes (lons) of Efstaffelsvatn and Gjanupsvatn near the margin of the Hoffelsjokull in Iceland, described and photographed by V. Okko (1955). This author had occasion to observe sediments of these lakes directly after they had been drained in July 1949. He describes them as consisting of rhythmically stratified boulders, gravels and sands. According to V. Okko's interpretation they reflect changes in hydrodynamic conditions of material transportation stimulated by oscillations of the glacier margin. It appears that similar conditions might have existed in rivers and lakes temporarily occurring in the ice or on the ice surface during the deglaciation period discussed.

Series D has developed as poorly sorted sandy deposits (Fig. 10: 6). It consists of alternating lamellae of fine-grained sands of a few centimeter thickness, containing an admixture of coarser sand and gravel fractions, and fine-grained sands with a silt admixture of the order of one or a few millimeters. In the dip plane these lamellae appear in the shape of stepped lines (Fig. 12), with their angle of inclination growing in eastward direction. Along the borders of the series especially in its western part, the lamellea show a slight upward bend.

Fig. 12. Rządz. Structure of sandy flow-type ablation deposits (supraglacial or intraglacial in origin) redeposited *en bloc* (Fig. 10 : 6, Fig. 11 : D)

The structural features of the deposits of series D indicate that they have originated as a result of flowage of water-laid sediments in supraglacial or intraglacial conditions.

The facts discussed above suggest that both series C and D are redeposited fragments of larger rock bodies which initially were built up on the top of stagnant ice or — as seems more likely with regard to series C — in tunnels or caves sculptured by water within the ice. When later an intensive decay of the ice sheet set in, these deposits slid down as stiff frozen blocks into newly formed crevasses, and subsequently, shrinking in size and becoming plastic due to thawing, collapsed under their own weight and the weight of overlying rock masses. This interpretation is supported by the sharp border separating the deposits of series C and D from the series B, and the upward bend of lamellae observed at the borders of series C and D.

Fig. 13. Rządz. Structure of "crevasse fillings" deposits (Fig. 10 : 1–6, Fig. 11 : E)

The final stage of filling-in the erosive landform is recorded by series E which consists of till laterally passing into sandy gravels with boulders (Fig. 10: 4–1 and Fig. 13). In the lower part the till is dense and clayey, whilst in the upper part it is considerably coarser or more washed, containing sands with an admixture of disintegrated gravels, sands and silts.

The structural characteristics of the deposits of series E and their position in the section exposed indicate that they are typical "crevasse fillings" (R. F. Flint, 1957). They were redeposited during the final stage of deglaciation due to mass movement down ice slopes. In this material a shell fragments of fossil molluscs were found — presumedly the remnants of animals once living on the top of the buried dead ice (E. Drozdowski, 1974). Dating of some unidentified shell fragments carried out by S. Hakansson has shown the following age: Rządz ... Lu-1071 38, 100 $^{+2900}_{-2100}$ B. P.

The geological, palaeomorphological, and chronological data presented here suggests that the process of deposition within the depression in the glacier substratum took place in two stages. Stage one indicated the accumulation of two till beds in the primary depression. The gradual lithological transitions between these till beds and sand layers seem to prove that the deposition of these deposits proceeded during one continuous period of subglacial accumulation. This assumption is confirmed also by similar results of analysis of petrographical composition of the gravels and by the orientation of the *a*-axes of pebbles in both till beds (E. Drozdowski, 1974). After the till beds had been laid down, a strong action of meltwater flow set in, as shown by the gravel and sand beds, underlain by a boulder pavement (Fig 10: 10).

The first stage of accumulation recorded in the exposure ended with the sedimentation in glacial lake. Because superimposed sediments are not visible today, it is impossible to discern the conditions under which the water basin has developed. However, both the structure and the facial pecularities of the deposits seem to point out subglacial conditions, although it is also possible that these deposits have been laid down in a water basin developed in an ice crevasse or an ice cave open upwards.

The second stage of accumulation visible in the exposure, separated from the first stage by a so far unknown time interval in which material was accumulating on the top and in the stagnant (dead) ice, was initiated by an intensive erosion caused by meltwater flow in the bottom of the ice crevasse. The subglacial deposits were eroded, producing a hollow which subsequently was filled-in by previously deposited intraglacial and supraglacial deposits.

The phenomena described above characterize the final disappearance of the penultimate Scandinavian ice sheet in the Grudziądz Basin. As was determined by investigations (E. Drozdowski, 1974), the erosive forms developed at that time were mostly filled in by glacifluvial material lying *in situ*. The discussed exposure, showing an erosive hollow filled-in by material in the way of mass-movement, represents a fairly rare and exceptional case, because of the well-preserved primary structure of the redeposited sediments. This exposure is apt to supply information of great importance about the mode of the ice sheet decay, likewise the palaeomorphological conditions and the period in which the buried erosive forms have originated.

RECONSTRUCTION OF THE SUCCESSION OF THE DEGLACIATION PROCESSES

For reconstructing the succession of the deglaciation processes the supraglacial tills appear to be of peculiar importance. In the area under discussion they occur in four genetic-facial types: (1) supraglacial subaquatic till, lying *in situ*, (2) its redeposited modification, i.e. supraglacial flowtill, of subaquatic origin, (3) supraglacial-solifluction till, of subaerial origin, and (4) diagenetic flowtill of the "crevasse-filling" type, usually redeposited together with other deposits of supraglacial or intraglacial origin.

An index type of supraglacial deposits is represented by the supraglacial subaquatic till lying *in situ*. This deposit occurs as a continuous bed only in the "normally" developed bipartite till profile on top of the lodgment till. Its thickness in most of examined exposures (Morsk, Dolne Sartowice, Strzemięcin) varies from 0.5 m to 2.0 m. The stratigraphical position of this bed seems to assign it to the first type of ablation moraine distinguished and described by A. Jahn (1952/1953); the interpretation of this author is that this type signifies deglaciation during an abrupt warming of the climate, which caused a rapid melting of the ice.

As mentioned before, the outstanding characteristic of supraglacial subaquatic till is its fairly high content of clayey fraction, which commonly exceeds two times the percentage of this fraction held in the top part of the underlying lodgment till. This feature proves the existence of supraglacial water environments like lakes or areas flooded by stagnated waters, in which during the melting-out the finest material has not been washed away but enriches the supraglacial drift cover.

The existence of water bodies on the surface of the till-covered stagnant ice seems to be indicated by the inclination of the top of the second glacial-drift horizon dipping perpendicular to the axis of the Vistula valley, and by a mantle of varved clays overlying the supraglacial till. These features are probably linked with a large valley depression (E. Rühle, T. Sokołowski, M. Tyska, 1958) existing here in the substratum of the Quaternary, and with a variable glacio-isostatic movements which occurred as the ice sheet was waning. The peripheral parts of the depression, i.e. the relatively mobile area of the Wąbrzeźno Elevation in the east (E. Rühle, 1965) and the elevated Quaternary substratum in the west, on top of which the ice sheet was less thick, reacted to ice lowering by compensating upward movement more quickly than the central part of the depression which was overlain by a thicker ice mantle. In consequence the initially plane surface of the stagnant ice sheet became increasingly concave, assuming gradually the form of an assuaged copy of the depression in the substratum (E. Drozdowski, 1973).

The supraglacial blanket of clayey-stony till up to 2 m thick protected the underlying ice from the thermic action of the atmosphere, so that much of this ice managed to survive for a long time. As deglaciation proceeded, however, differences in the thickness of the supraglacial deposits (in the exposures such differences in till thicknesses are clearly visible) and the growing impact of meltwater and, probably also, of precipitation waters led to increasingly different rates of ice melting; where conditions were favourable, these processes caused a disintegration of the ice sheet and an intensive sculpturing of its surface. This fact is divulged by layers, streaks and lenses of the supraglacial flowtill occurring on both the parent deposits — the supraglacial subaquatic till *in situ*, and on the supraglacial sand and gravel deposits laid down in crevasses and ice holes (Fig. 3).

The supraglacial-solifluction till presents a record of the next, the second stage of ice sheet wasting; during this stage ice melting has been slowing down due to a considerable cooling of the climate. This deposit is seen in the upper part of the geological profiles of all three *kępa*-hills, as well as in the eastern and northern slopes of the Grudziądz Basin. However, this deposit has not been found south of the Grudziądz Basin, in the Vistula Valley reach from Sartowice to Świecie where the second glacial-drift horizon occurs in its typical development as a bipartite till stratum (with the supraglacial subaquatic till lying *in situ*).

The formation of the supraglacial-solifluction till has been ascribed to areas of breaking up and very irregular surface of the dead ice. Evidence for this thesis was provided by the stratigraphical position of the supraglacial-solifluction till in the Parsk exposure (Fig. 3) where the till lies between two layers of supraglacial flowtill. This implies that on the dead ice surface near the glacial-mineral slopes covered by supraglacial-solifluction till water-filled basins must have existed, allowing the primary ablation deposits (supraglacial subaquatic till) to persist up to the final disappearance of the ice.

During the third, final stage of deglaciation the dominating processes were erosive and accumulative action of meltwater streams within ice crevasses and in-between the blocks of dead ice. In the erosive forms developed during this stage, material released from the ice was deposited as well as material washed out and displaced by gravity from the ice surface and from intraglacial voids (Fig. 10).

A supraglacial deposit delimiting the final stage of deglaciation is the diagenized supraglacial flowtill of the "crevasse-filling" type, which originated from material melted out from the ice under varied local conditions, subaerially and subaquatically, and laid down in between dead ice blocks and in ice crevasses. This deposit contains commonly admixtures of lacustrine and fluvial sediments of supraglacial or intraglacial origin and a great deal of pebbles disintegrated as a result of frost action; moreover, often well-preserved remnants of fossil mollusc fauna are found in it. A close typological similarity can be seen between the till discussed above and a modern diagenetic supraglacial till described by M. Klimaszewski (1960) from the marginal zones of Spitsbergen glaciers. Further detailed observations of till of this type have revealed that in the processes of transformation an important part is played by the phenomena of multiple gravitational redeposition (J. Szupryczyński, 1963, 1964), and it seems probable that in the process of diagenesis of the above discussed supraglacial till similar phenomena have also taken part.

What has been presented above suggests two general conclusions:

(1) In the Grudziądz Basin of the Lower Vistula River Valley, the penultimate period of deglaciation during the Middle Würm Interstadial took place by process of large-scale stagnation. In the process of down-wasting of the ice masses by surface ablation, i.e. melting and evaporation, a great part was played by mechanical and thermic erosion of meltwaters and, appearently also, precipitation waters.

(2) The processes of deglaciation proceeded in at least three climatically conditioned stages, in succession from a warmer phase through a cooler one to a renewed warmer phase.

AGE OF THE SECOND GLACIAL-DRIFT HORIZON

Conforming to the geochronological interpretation of the Pleistocene in the Lower Vistula Valley (R. Galon, 1961; J. E. Mojski, 1969), the deposits of the second glacial-drift horizon have been assigned to the younger pleniglacial (or to the younger part of the Main Stage, after E. Rühle's (1965) terminology). According to this interpretation, the second glacial-drift horizon should correspond to the period of the maximum extent of the Würmian ice sheet in Poland's territory, the so-called Leszno (Brandenburg) Stage, and the first horizon to the Poznań (Frankfurt) Stage. Hence the sediments deposited between these two till strata (called Fluvioglacial I in R. Galon's division) should be assigned to the interval separating the Leszno Stage from the Poznań Stage, the period which the majority of research workers (N. Chebotareva, R. Galon, J. Gellert, L. Serebryanny, 1965; R. Galon, 1968; J. E. Mojski, 1969; L. R. Serebryanny, A. V. Raukas, 1970) believe to have been an Interphase of ice recession of the Leszno Stage.

That the deposits of the first glacial-drift horizon were left behind by the inland ice of the Poznań Stage is obvious, because the investigated area of the Grudziądz Basin is located (Fig. 1) between the marginal zones of the Poznań Stage and the Pomeranian Stage (R. Galon, 1961, 1968; L. Roszko 1956, 1968; S. Kozarski 1962; J. Kondracki, 1965). However, the connection of the second glacial-drift horizont with the inland ice of the Leszno Stage seems questionable, as might be concluded from the relationship between the stratigraphical position of this horizon and the number of main ice-marginal zones delineated in the present relief of the Baltic Glaciation south of the Grudziądz Basin. Several facts of wider palaeogeographical significance, together with obtained recently first radiocarbon date indicate that the second glacial-drift horizon is older than the Leszno Stage and should be assigned to the inland ice which in the region under discussion was laid down prior to the Middle-Würm Interstadial and was waning during this Interstadial (Interpleniglacial).

This thesis is supported by the following arguments:

(1) The manner of the occurrence of the deposits of the second glacial-drift horizon, especially of the deposits of the second and third stage of deglaciation which were laid down under conditions of a far-advanced decay of the ice sheets.

In this respect the Nowe Marzy exposure (Fig. 6) is revealing. Here the deposits of the second glacial-drift horizon fill the fossil depression below the horizontally spread youngest till strata. From these stratigraphic-palaeomorphological evidences confirmed in many other geological sections of the area under investigation, it appears that in the period preceding the expansion of the last Scandinavian ice sheet there had been smoothed over all elevations in the ice--and-rock relief connected with the decay of the penultimate ice sheet or, in other words, that all processes of down-wasting of the ice sheet had ended. Based on this conclusion one might assert that the zone of ice stagnation (with the ice mass balance in equilibrium) — which according to common belief in this matter (K. Büllow 1927; R. F. Flint, 1957; T. Bartkowski, 1953, 1969; W. Niewiarowski, 1963; A. Aseyev, 1963; J. Szupryczyński 1963; L. Clayton, 1967) proceded in time and space the final decay and disappearance of the ice sheet (i.e. a negative balance of the ice mass) — had moved considerably northward towards the centre of glaciation. The processes of deglaciation, taking place at such a scale in time and space in lowland conditions (where gravity as the possible source of change in the balance of the ice mass can be disregarded), should be ascribed to a higher-grade general warming-up of the climate. However, this interpretation would be at variance with the course of climatic changes in the extraglacial zone of the Last Glaciation (J. Dylik, 1964, 1966; K. Rotnicki, 1966; J. Malinowski, 1964; H. Maruszczak, 1968; J. E. Mojski, 1965; J. Jersak, 1965; M. Sobolewska, L. Starkel, A. Środoń, 1964; R. G. West, 1968) if this warming-up is attributed to the Pre-Poznań Interphase; in contrast, it would be fully concordant when correlated with a Middle-Würm (Paudorf, Stillfried B.) Interstadial.

(2) The climatic multiplicity of the discussed period of deglaciation recorded in deposits and glacial forms and in their stratigraphic-palaeogeographical position.

This multiplicity of climatic changes within one prolonged interval of the nature of an Interstadial can be correlated with the Interstadial Complex of the Middle Würm, as determined both in Poland's territory (M. Sobolewski, L. Starkel, A. Środoń, 1964) and in other areas of the northern hemisphere (T. Van der Hammen, G. C. Maarleveld, J. C. Vogel, W. H. Zagwijn, 1967; P. Woldstedt, 1966; W. H. Zagwijn, R. Paepe, 1968; M. E. Vigdorchik et. al., 1972; A. Dreimanis, A. Raukas, 1973).

(3) Lithological examinations (E. Drozdowski, 1974) which have revealed differences in the petrographic composition of gravels (5–10 mm and 2–5 mm) and in orientation of the long axes of pebbles between the lodgment tills of the second and the first glacial-drift horizons.

In the second horizon the orientation of the long axes oscillates within the NNW-NNE sector, whilst in the NNE-NE sector in the first horizon. These observations, supplementing each other, are proof of two fully dissimilar inland ices, differing as to the areas of alimentation and direction of movement. This confirms the opinion of some Scandinavian scientists (F. Ljungner, 1949; J. Lundqvist, 1967, 1971; K. Korpela, 1969; E. Olausson, 1971) about a division of the Last (Würm) Glaciation into two main glacial periods separated by a Middle-Würm Interstadial, during which the ice sheet decreased, receding to the area of the Scandinavian mountains. In the area under discussion the climatic conditions ruling during this Interstadial were, even at their optimum, of a periglacial character, because the dead ice — as shown by investigations (E. Drozdowski, 1974) — had survived beneath a blanket of supraglacial deposits up to the advance of the last Scandinavian inland ice due to continuous existence of permafrost.

(4) Radiocarbon dating of shell remnants from the site Rządz which has given the age $38{,}100^{+2900}_{-2100}$ B. P.

CONCLUSIONS

Since the age of the period of deglaciation discussed above can be considered to be of the Middle Würm i.e. older than the maximum extent of the Würmian ice sheet in the Poland's territory (about 25,000 years), the profile of the Würm Glaciation deposits in the area investigated can be safely subdivided in two main parts: a lower part embracing the deposits of the second glacial-drift horizon and, underlying them, fine-grained sands and grey varved clays and, in addition, the third locally appearing till strata (Fig. 2) — and an upper part represented usually by one, the first glacial-drift horizon. The thickness of the lower part of the profile, counting upwards from the top of the palaeobotanically dated sediments of the Eemian Interglacial at Grudziądz–Mniszek (E. Drozdowski, K. Tobolski, 1972), to the top of the deposits of the second glacial-drift horizon, is some 60 m, and this would correspond to the period preceding the Leszno Stage. The upper part, with a thickness varying from 6 to 10 m, corresponds to the period covering apparently both the Leszno and the Poznań Stages.

Institute of Geography,
Polish Academy of Sciences, Toruń

REFERENCES

Aseyev, A. A., 1963, Die geomorphologische Zonalität des alten Eisgebietes der Russischen Ebene, *Report VIth Congr. INQUA*, 3, Łódź, pp. 9–20.

Bartkowski, T., 1953, The role of the buried "dead" ice in the formation of the postglacial landscape of Central Great Poland, *Bull. Soc. Amis Sci. Pozn.*, ser. B, 3, 2, pp. 123–245.

Bartkowski, T., 1962, Deglacjacja strefowa deglacjacją normalną na obszarach niżowych (na wybranych przykładach z Polski zachodniej i północnej) (Zsf.: Zonaler Eisabbau — ein "normaler" Eisabbau des Tieflandsgebietes (erläutert an Beispielen aus West- und Nordpolen)), *Bad. Fizjogr. Pol. Zach.*, 23, pp. 7–33.

Boulton, G. S., 1972, Modern Arctic glaciers as depositional models for former ice sheets, *J. Geol. Soc. London*, 128, pp. 361–393.

Büllow, K., 1927, Die Rolle der Toteisbildung beim letzten Eisrückgang in Norddeutsch land, *Z. Dtsch. Geol. Ges.*, 79, 314, Berlin, pp. 273–283.

Cailleux, A., Tricart, J., 1959, *Initiation à l'étude des sables et galets*, C.D.U., Paris.

Chebotareva, N., Galon, R., Gellert, J., Serebryanny, L., 1965, Morphology of marginal zones and their correlation, in: *Posledniy evropeyskiy lednikoviy pokrov*, Moskva. pp. 178–202.

Clayton, L., 1967, Stagnant-glacier features of the Missouri Cotteau in North Dakota, in: *Glacial geology of the Missouri Cotteau*, Misc. Ser., 30, Grand Forks, pp. 25–46.

Czajkowski, L., 1969, Powierzchnia regionalnych jednostek fizycznogeograficznych Polski (Sum.: Land measures of physico-geographical units of Poland), *Prace i Studia IG UW*, 2, Warszawa, pp. 7–20.

Dreimanis, A., Raukas, A., 1973, Did Middle Wisconsin, Middle Weichselian, and their equivalents represent an interglacial, or an interstadial complex in northern hemisphere? in: *Abstracts IX Congr. INQUA Christchurch*, New Zealand, pp. 84–85.

Drozdowski, E., 1973, Podłoże czwartorzędu i jego wpływ na rozwój procesów glacjalnych w środkowej części dolnego Powiśla (Sum.: The Quaternary substratum, and its effect upon the evolution of glacial processes in the central part of the Lower Vistula River Valley), *Przegl. Geogr.*, 45, 3, pp. 517–533.

Drozdowski, E., 1974, *Geneza Basenu Grudziądzkiego w świetle osadów i form glacjalnych* (Sum.: Genesis of the Grudziądz Basin in the light of its glacial forms and deposits), Prace Geogr. IG PAN, 104.

Drozdowski, E., Tobolski, K., 1972, Stanowiska interglacjału eemskiego w Basenie Grudziądzkim (Sum.: Sites of Eem Interglacial in Grudziądz Basin), *Bad. Fizjogr. Pol. Zach.*, 15, ser. A, pp. 75–91.

Dylik, J., 1964, Sur le changements climatiques pendant la derniére période froide, *Report VIth Congr. INQUA*, 4, Łódź, pp. 55–66.

Dylik, J., 1966, Znaczenie peryglacjalnych elementów w stratygrafii plejstocenu (Rés.: Importance des éléments périglaciaires dans la stratigraphie du Pleistocene), *Czas. Geogr.*, 37, 2, pp. 131–151.

Flint, R. F., 1957, *Glacial and Pleistocene geology*, New York–London, Wiley.

Frye, J. C., Willman, H. B., 1962, Note 27, Morphostratigraphic units in Pleistocene stratigraphy, *Am. Ass. Petrol. Geol. Bull.*, 46, 1, pp. 112–113.

Galon, R., 1934, *Dolina dolnej Wisły, jej kształt i rozwój na tle budowy dolnego Powiśla* (Zsf.: Die Gestalt und Entwicklung des unteren Weichseltales in Beziehung zum geologischen Aufbau des unteren Weichselgebietes), Bad. Geogr. Pol. Pn.-Zach., 12–13.

Galon, R., 1961, General Quaternary problems of North Poland, in: *Guide-book of excursion from the Baltic to the Tatras, part I, North Poland, VIth INQUA Congr.*, Warszawa, pp. 9–53.

Galon, R., 1968, Przebieg deglacjacji na obszarze Peribalticum (Sum.: The course of deglaciation in the Peribalticum), in: *Ostatnie zlodowacenie w Polsce*, Prace Geogr. IG PAN, 74, Warszawa, pp. 201–212.

Hammen Van der, T., Maarleveld, G. C., Vogel, J. C., Zagwijn, W. H., 1967, Stratigraphy, climatic succession and radiocarbon dating of the last glacial in the Netherlands, *Geol. en Mijnbouw*, 46, 3, pp. 79–95.

Harrison, P. W., 1957, A clay-till fabric: its character and origin, *J. Geol.*, 65, 3, pp. 275–308.

Jahn, A., 1952–53. Lodowce typu Baffina i problem moren ablacyjnych (Glaciers of Baffin type and problem of ablation tills), *Czas. Geogr.*, 23/24, pp. 3–14.

Jahn, A., 1970, *Zagadnienia strefy peryglacjalnej* (Problems of the periglacial zone), Warszawa.

Jersak, J., 1965, *Stratygrafia i geneza lessów okolic Kunowa* (Res.: Stratigraphie et genese des loess aux environs de Kunów, Acta Geogr. Lodz., 20.

Kirby, R. P., 1969, Till fabric analyses from the Lothians, Central Scotland, *Geogr. Ann.*, 51 A, 1–2, pp. 48–59.

Klimaszewski, M., 1960, *Studia geomorfologiczne w zachodniej części Spitsbergenu między Kongs-Fjordem a Eidem-Bukta* (Geomorphological studies of the western part of Spitsbergen between Kongsfjord and Eidembukta), Zesz. Nauk. UJ, 32, Prace Geogr., 1.

Kondracki, J., 1965, *Geografia fizyczna Polski* (Physical Geography of Poland), Warszawa.

Kondracki, J., Ostrowski, J., 1973, Regiony fizycznogeograficzne Polski (Physico-geographical regions of Poland), in: *Narodowy Atlas Polski,* IG PAN, Warszawa, pl. 41.

Korpela, K., 1969, *Die Weichsel-Eiszeit und ihr Interstadial in Peräpohjola (nördliches Nordfinnland) im Licht von submoränen Sedimenten,* Ann., Ac. Sci. Fennicae, ser. A, III, 99.

Kozarski, S., 1962, *Recesja ostatniego lądolodu z północnej części Wysoczyzny Gnieźnieńskiej a kształtowanie się pradoliny Noteci–Warty* (Sum.: Recession of the last ice-sheet from the northern part of the Gniezno Pleistocene plateau and its influence upon the formation of the Noteć–Warta ice-marginal valley), Prace Komis. Geogr.-Geol., Pozn. Tow. Przyj. Nauk, 2, 3.

Ljungner, E., 1949, East-west balance of the Quaternary ice caps in Patagonia and Scandinavia, *Bull. Geol. Inst. Upsala,* 33, pp. 11–96.

Lundqvist, G., 1949, The orientation of the block material in certain species of earth. flow, *Geogr. Ann.,* 31, pp. 335–347.

Lundqvist, J., 1967, *Submoräna sediment i Jämtlands län* (Sum.: Submorainic sediments in the county of Jämtland, Central Sweden), Sverig. Geol. Unders., ser. C, 618.

Lundqvist, J., 1971, The Interglacial deposits at the Leveäniemi mine, Svappavaara, Swedish Lapland, *Sverig. Geol. Unders.,* ser. C, 658, pp. 1–35.

Malinowski, J., 1964, *Budowa geologiczna i własności geotechniczne lessów Roztocza i Kotliny Zamojskiej między Szczebrzeszynem i Turobinem* (Sum.: Geological structure and geotechnical properties of loesses in Roztocze and in the Zamość Basin, between Szczebrzeszyn and Turobin), Prace Inst. Geol., 41, Warszawa.

Maruszczak, H., 1968, Przebieg zjawisk w strefie peryglacjalnej w okresie ostatniego zlodowacenia w Polsce (Sum.: The course of phenomena in the Periglacial zone during the Last Glaciation), in: *Ostatnie zlodowacenie skandynawskie w Polsce,* Prace Geogr. IG PAN, 74, Warszawa, pp. 157–200.

Mojski, J. E., 1965, Stratygrafia lessów w dorzeczu dolnej Huczwy na Wyżynie Lubelskiej (Sum.: Loess stratigraphy in the drainage basin of the Lower Huczwa River in the Lublin Upland), *Biul. Inst. Geol.,* 187, Warszawa, pp. 145–216.

Mojski, J. E., 1969, The stratigraphy of the Last Glaciation in the territory of Poland, *Geogr. Pol.,* 17, pp. 72–91.

Niewiarowski, W., 1959, *Formy polodowcowe i typy deglacjacji na Wysoczyźnie Chełmińskiej* (Sum.: Glacial forms and types of deglaciation on the moraine plateau of Chełmno, Bydgoszcz district), Stud. Soc. Sci. Tor., Sec. C, 4, 1.

Niewiarowski, W., 1963, Some problems concerning deglaciation by stagnation and wastage of large portions of the ice-sheet within the area of the last glaciation in Poland, *Report VIth Congr. INQUA,* 3, Łódź, pp. 245–256.

Okko, V., 1955, *Glacial drift in Iceland, its origin and morphology,* Bull. Comm. Geol. Finlande, 170.

Ollausson, E., 1971, Oceanographic aspects of the Pleistocene of Scandinavia, *Geol. Foren. Stockh. Förh.*, 93, pp. 459–475.
Roszkówna, L., 1956, *Zagadnienie zasięgu stadium pomorskiego nad dolną Wisłą* (Sum.: The question of the limits of the Pomeranian Stage on the Lower Vistula), Stud. Soc. Sci. Tor., Sec. C, 3, 1.
Roszko, L., 1968, Recesja ostatniego lądolodu z terenu Polski (Sum.: Recession of the last inland ice from Poland's territory), in: *Ostatnie zlodowacenie skandynawskie w Polsce* (Last Scandinavian Glaciation in Poland), Prace Geogr. IG PAN, 74, Warszawa, pp. 65–100.
Rotnicki, K., 1966, *Rzeźba Wzgórz Ostrzeszowskich jako rezultat rozwoju stoku podczas Würmu* (Sum.: The relief of the Ostrzeszów Hills as result of slope development during the Würm), Prace Komis. Geogr.-Geol., Pozn. Tow., Przyj. Nauk, 5, 2.
Rukhina, E., 1960, *Litologiya morennikh otlozheniy* (Lithology of morainic deposits), Leningrad.
Rühle, E., 1965, Czwartorzęd Polski (The Quaternary of Poland), in: M. Książkiewicz, J. Samsonowicz, E. Rühle, *Zarys geologii Polski,* Warszawa, pp. 299–350.
Rühle, E., Sokołowski, T., Tyska, M., 1958, *Mapa geologiczna Polski 1 : 2,000,000* (Carte géologique de la Pologne, 1 : 2,000,000), Inst. Geol., Warszawa.
Serebryanny, L. R., Raukas, A. V., 1970, Über die eiszeitliche Geschichte der Russischen Ebene im oberen Pleistozän, *Petermanns Geogr. Mitt.*, 114, 3, pp. 161–172.
Sobolewska, M., Starkel, L., Środoń, A., 1964, *Młodoplejstoceńskie osady z florą kopalną w Wadowicach* (Sum.: Late-Pleistocene deposits with fossil flora at Wadowice — West Carpathians), Folia Quatern., 16.
Szupryczyński, J., 1963, *Rzeźba strefy marginalnej i typy deglacjacji lodowców południowego Spitsbergenu* (Sum.: Relief of marginal zone of glaciers and types of deglaciation of Southern Spitsbergen glaciers), Prace Geogr., IG PAN, 39, Warszawa.
Szupryczyński, J., 1965, Zagadnienie genezy krajobrazów wysoczyzny morenowej płaskiej i falistej (Sum.: The problem of the origin of landscapes of a flat and undulant morainic plateau), *Przegl. Geogr.*, 37, 1, pp. 163–169.
Vigdorchik, M. E., Auslender, V. G., Snamenskaya, O. M., Dolukhanov, P. M., 1972, Noviye danniye o rasvitii valdayskogo oledeneniya i yego krayevikh son (New data on the development of Valdai Glaciation and its marginal zones), in: *Krayeviye obrasovaniya materikovikh oledeneniy,* Moskva, pp. 166–176.
West, R. G., 1968, *Pleistocene geology and biology,* Longmans, Green and Co Ltd., London.
Woldstedt, P., 1966, Der Ablauf des Eiszeitalters, *Eisz. Gegenw.*, 17, pp. 153–158.
Zagwijn, W. H., Paepe, R., 1968, Die Stratigraphie der Weichselzeitigen Ablagerungen der Niederlande und Belgien, *Eisz. Gegenw.*, 19, pp. 129–146.

CONTENTS OF VOLUMES

GEOGRAPHIA POLONICA

Vol. 1. 11 papers devoted to the present status of geography in Poland and 3 papers giving the results of research. List of Polish geographers, geographical institutions and geographical periodicals, 262 pp., 20 Figures, 1964 (out-of-print).
Vol. 2. 34 papers prepared by Polish geographers for the XXth International Geographical Congress in London, July 1964, 259 pp., 91 Figures, 1964.
Vol. 3. Problems of Applied Geography II. Proceedings of the Second Anglo-Polish Seminar at Keele—Great Britain, September 9–20 1962, Co-edited by the Institute of British Geographers. 21 papers by British and Polish geographers, 274 pp., 69 Figures, 1964.
Vol. 4. Methods of Economic Regionalization. Materials of the Second General Meeting of the Commission on Methods of Economic Regionalization, International Geographical Union, Jabłonna — Poland, September 9–10, 1963. Reports, communications and discussion, 200 pp., 6 Figures, 1964.
Vol. 5. Land Utilization in East-Central Europe. 17 case studies on land use in Bulgaria, Hungary, Poland and Yugoslavia, 498 pp., 104 Figures, 16 colour maps, 1965.
Vol. 6. 14 papers prepared by Polish geographers for the Seventh World Conference of INQUA in U.S.A., September 1965, 150 pp., 86 Figures, 1965.
Vol. 7. 10 papers on the geography of Poland, mostly dealing, with the economic-geographical problems of Poland, 132 pp., 46 Figures, 1965.
Vol. 8. Aims of Economic Regionalization. Materials of the Third General Meeting of the Commissionon Methods of Economic Regionalization IGU, London, July 23, 1964. Report and 5 papers, 68 pp., 7 Figures, 1965.
Vol. 9. Colloque de Géomorphologie des Carpathes. Materials of the geomorphological symposium held in Cracow and Bratislava, September 17–26, 1963. Report, 7 papers, 2 summaries, 118 pp., 22 Figures, 1965.
Vol. 10. Geomorphological Problems of Carpathians II. Introduction and 6 papers by Rumanian, Soviet, Polish, Hungarian and Czech geographers, 172 pp., 68 Figures, 1966.
Vol. 11. 11 papers prepared by Polish geographers dealing with the history of Polish geography, Polish studies on foreign countries and different economic-geographical questions concerning Poland, 154 pp., 36 Figures, 1967.
Vol. 12. Formation et l'Aménagement du Reseau Urbain. Proceedings of the French-Polish Seminar in urban geography, Teresin, Poland, September 20–30, 1965. Papers by French and Polish geographers, discussion, 298 pp., 51 Figures, 1967.
Vol. 13. 9 papers embracing different fields of both, physical and economic geography, all of which have been devoted to methodological problems and research techniques, 130 pp. 4 Figures, 1968.
Vol. 14. Special issue for the 21st International Geographical Congress in New Delhi, 968, 43 papers pepared by Polish geographers: 24 dealing with physical and 19 with economic and human geography 406 pp., 80 Figures, 1968.
Vol. 15. Economic Regionalization and Numerical Methods. The volume contains the final report on he activities of the IGU Commission on Methods of Economic Regionalization, as well as a collection of 8 papers by American, Canadian, Soviet and Polish authors, 240 pp., 54 Figures, 1968.
Vol. 16. 11 papers dealing with research problems and techniques in both economic and physical geography, 136 pp., 27 Figures, 1969.
Vol. 17. Special issue prepared for the 8th Congress of the International Union for Quaternary Research Paris, 1969, 28 papers by Polish authors, including studies in stratigraphy and neotectonics (6), geomorphology and paleohydrology (10), paleobotany (3), sedimentology (5), archeology (4), 428 pp., 122 Figures, 1969.
Vol. 18. Studies in Geographical Methods. Proceedings of the 3rd Anglo-Polish Geographical Seminar, Baranów Sandomierski, September 1–10, 1967, 260 pp., 54 Figures, 1970.
Vol. 19. Essays on Agricultural Typology and Land Utilization, 20 papers presented at the meeting of the Commission on World Agricultural Typology of the IGU, held 1968 in New Delhi, 290 pp., 97 Figures, 1970.

Vol. 20. 9 papers on various aspects of both physical and economic geography, including urbanization, international trade, changes in rural economy, industrial development, urban physiography and hydrographic mapping, 183 pp., 69 Figures, 1972.

Vol. 21. 10 papers dealing with selected problems of economic growth, transportation, cartographic methods and theory, climatology and geomorphology, 147 pp., 82 Figures, 1972.

Vol. 22. 15 papers prepared for the 22nd International Geographical Congress in Montreal, August 1972, 205 pp., 43 Figures, 1972.

Vol. 23. Present-day Geomorphological Processes. Issue prepared for the 22nd International Geographical Congress by the IGU Commission on Present-day Geographical Processes, 180 pp., 82 Figures, 1972.

Vol. 24. Geographical aspects of urban-rural interaction. Proceedings of the 4th Anglo-Polish Geographical Seminar, Nottingham, September 6—12, 1970, 256 pp., 76 Figures, 1972.

Vol. 25. Perspectives on spatial analysis. 7 papers presented at the meeting of the Commission on Quantitative Methods of the IGU, held 1970 in Poznań, Poland, 140 pp., 51 Figures, 1973.

Vol. 26. Scientific results of the Polish geographical expedition to Vatnajökull (Iceland), 311 pp., 253 Figures, 1973.

Vol. 27. 20 papers presented by the Soviet and Polish Geographers at the First Polish-Soviet Geographical Seminar in Warsaw, Szymbark and Cracow, 22nd May to 1st June, 1971, 189 pp., 25 Figures, 1973.

Vol. 28. 9 papers embracing different fields of geography. 144 pp., 36 Figures, 1974.

Vol. 29. Modernisation des Campagnes. Actes du IV^e Colloqué Franco-Polonais de Geographes. Augustów, Pologne, Septembre, 1973. 444 pp., 145 Figures, 1974.

Vol. 30. Proceedings of the second Polish-GDR Seminar, Szymbark (Poland), April, 1972, 151 pp., 21 Figures, 1975.

Subscription orders for the GEOGRAPHIA POLONICA should be placed with
FOREIGN TRADE ENTERPRISE ARS POLONA — RUCH

Warszawa, Krakowskie Przedmieście 7, Poland
Cables, ARSPOLONA, Warszawa

Cena zł 68.—