

STUDIES

**IN geographical
methods**

Geographia Polonica 18

PWN – Polish Scientific Publishers

The First Anglo-Polish Geographical Seminar was held at Nieborów, Poland from September 15—18, 1959. Proceedings of the Seminar were published as a separate volume, „Problems of Applied Geography”, Geographical Studies 25, Warszawa 1961.

The Second Anglo-Polish Geographical Seminar was organized at Keele, Great Britain, from September 9—20, 1962. The proceedings were published as Vol. 3 of *Geographia Polonica* (Warszawa 1964).

The present volume contains 18 papers read at the Third Anglo-Polish Seminar held at Baranów Sandomierski, Poland, from September 1—10, 1967. The reports deal with the problems of communication in geography, theoretical problems in economic geography, population studies, urban and rural land use, mathematical and cartographical techniques in geography, methods of geomorphological investigations, the application of satellites in geographical research.

Geographia Polonica 18

Studies in Geographical Methods

Geographia Polonica

18

INSTITUTE OF GEOGRAPHY
POLISH ACADEMY OF SCIENCES

Studies in Geographical Methods

*Proceedings of the 3rd Anglo-Polish Geographical Seminar
Baranów Sandomierski, September 1–10, 1967*

*PWN – Polish Scientific Publishers
Warszawa 1970*

EDITORIAL BOARD

STANISŁAW LESZCZYCKI (Editor-in-Chief)
JERZY KONDRACKI, JERZY KOSTROWICKI, JANUSZ PASZYŃSKI
PIOTR KORCELLI (Secretary)

EDITORS OF THE VOLUME

RICHARD H. OSBORNE, ANDRZEJ WRÓBEL

Address of Editorial Board

Krakowskie Przedmieście 30, Warszawa 64
POLAND

PAŃSTWOWE WYDAWNICTWO NAUKOWE
PWN — POLISH SCIENTIFIC PUBLISHERS
WARSZAWA 1970

Zakłady Graficzne PZWS w Łodzi

CONTENTS

List of Participants in the Third Anglo-Polish Seminar	7
Report of the Meeting	9
<i>Keith M. Clayton</i> : Publication and communication in geography	13
<i>Kazimierz Dziewoński</i> : Geographical scientific information in Poland and its difficulties	21
<i>David W. Harvey</i> : Behavioural postulates and the construction of theory in human geography	27
<i>Zbyszko Chojnicki</i> : Some theoretical aspects of the application of mathematical models in economic geography	47
<i>James H. Johnson</i> : Recent British contributions to the study of population geography ...	57
<i>Leszek Kosiński</i> : The internal migration of population in Poland, 1961—1965	75
<i>Juliitta Rakowicz-Grocholska</i> : Investigations of urban land use in Polish geographical studies	85
<i>Jerzy Kostrowicki</i> : Some methods of determining land use and agricultural „orientation” as used in Polish land utilization and typological studies	93
<i>Christopher Board</i> : The quantitative analysis of land use patterns with special reference to land use maps: shape analysis, with an application	121
<i>H. Brian Rodgers</i> : Random sampling techniques in social geography	139
<i>Derek Thompson</i> : Some comments on the relevance of multivariate analysis to geography: a methodological review	157
<i>Lech Ratajski</i> : The point of change for methods in thematic cartography	177
<i>Christine S. Sunley</i> : Precision in terminology: an example from fluvial morphology	183
<i>Bolesław Dumanowski</i> : Examples of geomorphological investigations in Poland based on quantitative methods	199
<i>Zdzisław Czeppe</i> : Some methods of morphodynamic research used in Cracow	207
<i>V. Bruce Proudfoot</i> : Some recent field and laboratory experiments in geomorphology	213
<i>Marjorie M. Sweeting</i> : Recent developments and techniques in the study of karst landforms in the British Isles	227
<i>Eric C. Barrett</i> : Satellites in geographical research	243

LIST OF PARTICIPANTS IN THE THIRD ANGLO-POLISH SEMINAR

GREAT BRITAIN

1. Prof. R. H. Osborne, University of Nottingham
2. Prof. S. H. Beaver, University of Keele
3. Prof. K. M. Clayton, University of East Anglia
4. Prof. K. C. Edwards, University of Nottingham
5. Dr. J. H. Johnson, University College, London
6. Mr. H. B. Rodgers, University of Keele
7. Dr. E. C. Barrett, University of Bristol
8. Dr. C. Board, London School of Economics
9. Dr. D. W. Harvey, University of Bristol
10. Dr. V. B. Proudfoot, University of Durham
11. Miss Ch. S. Sunley, Queen Mary College, London
12. Dr. M. M. Sweeting, University of Oxford
13. Dr. D. Thompson, University of Sheffield

POLAND

1. Prof. S. Leszczycki, Institute of Geography PAN, Warsaw
2. Prof. M. Dobrowolska, Teachers' Training College, Cracow
3. Prof. K. Dziewoński, Institute of Geography PAN, Warsaw
4. Prof. R. Galon, University of Toruń
5. Prof. M. Kielczewska-Zaleska, Institute of Geography PAN, Warsaw
6. Prof. J. Kostrowicki, Institute of Geography PAN, Warsaw
7. Prof. A. Wrzosek, Jagellonian University, Cracow
8. Doc. J. Braun, Research Centre PAN, Zabrze
9. Doc. Z. Chojnicki, Adam Mickiewicz University, Poznań
10. Doc. Z. Czeppe, Jagellonian University, Cracow
11. Doc. B. Dumanowski, University of Warsaw
12. Doc. T. Lijewski, Institute of Geography PAN, Warsaw
13. Doc. L. Starkel, Institute of Geography PAN, Cracow

14. Doc. L. Ratajski, Institute of Geography PAN, Warsaw
15. Doc. A. Wróbel, Institute of Geography PAN, Warsaw
16. Dr. L. Barwińska, Maria Skłodowska-Curie University, Lublin
17. Dr. E. Dowgiałło, University of Warsaw
18. Dr. S. Gilewska, Institute of Geography PAN, Cracow
19. Dr. J. Januszewski, University of Wrocław
20. Dr. A. Kęsik, Maria Skłodowska-Curie University, Lublin
21. Dr. M. Koter, University of Łódź
22. Dr. A. Lisiecki, Teachers' Training College, Gdańsk
23. Dr. A. Werwicki, Institute of Geography PAN, Warsaw
24. Mrs. J. Rakowicz-Grocholska, Institute of Geography PAN, Warsaw
25. Mrs. E. Iwanicka-Lyra, Institute of Geography PAN, Warsaw
26. Mr. B. Czyż, Institute of Geography PAN, Warsaw
27. Mr. M. Stalski, Institute of Geography PAN, Warsaw

REPORT OF THE MEETING

The Third Anglo-Polish Geographical Seminar, organized by the Institute of Geography, Polish Academy of Sciences, with the co-operation of the Institute of British Geographers, was held in Baranów Sandomierski. The British group, consisting of twelve persons, from ten university Departments, was led by its Organizing Secretary, Professor R. H. Osborne, of the University of Nottingham (owing to the illness of Professor S. H. Beaver). The Polish participants, under the leadership of Professor S. Leszczycki, represented both the Institute of Geography of the Polish Academy of Sciences and several university centres.

On September 1st the members of the Seminar gathered in Warsaw, where they visited the Geographical Institutes of the Academy and the University and became acquainted with the research work carried on in the individual sections and departments. The next day was occupied by the journey from Warsaw to Baranów Sandomierski. During a stop in Lublin the participants visited the Geographical Institute of Maria Skłodowska-Curie University and also toured the town.

The three-day discussions in Baranów were held in a beautifully restored Renaissance castle, made available to the Seminar by the Ministry of the Chemical Industry. All the members were accommodated in a near-by hotel. The whole meeting, divided into six sessions, was devoted to the application of precise methods to geographical investigation. Professor S. Leszczycki delivered the introductory address.

At the first session, under the chairmanship of Professor R. Galon, four geomorphological papers were read and discussed. These included Dr. V. B. Proudfoot's work on field and laboratory experiments and Dr. M. M. Sweeting's paper on recent developments and techniques in the study of karst landforms in the British Isles, as well as papers by Docent B. Dumanowski on geomorphological investigations in Poland based on quantitative methods and by Docent Z. Czeppe, reporting on methods of morphodynamic research carried on in Kraków.

The next session, with Professor K. C. Edwards in the chair, was devoted

to geographical terminology and problems of communication. Professor K. M. Clayton read a paper on publication and communication in geography, while Professor K. Dziewoński spoke on geographical scientific information in Poland. The third paper of the evening, by Miss C. S. Sunley, concerned some terminological questions in fluvial morphology.

On September 4th the meetings were under the chairmanship in turn of Professor J. Kostrowicki and Professor K. M. Clayton. The morning session dealt with models in geography. Dr. D. W. Harvey spoke on the construction of theory in human geography from the standpoint of the behavioural sciences and Dr. D. Thompson on the relevance of multivariate analysis to geography, while Docent Z. Chojnicki discussed some aspects of the application of mathematical models to economic geography, and Docent A. Wróbel presented a gravity model for interregional commodity flows. At the evening session Dr. E. C. Barrett spoke on the implications of advancing space technology for geographical research and Docent L. Ratajski discussed the use of some alternative methods in thematic cartography.

On the third day of the Conference, under the chairmanship of Professor R. H. Osborne and Professor K. Dziewoński, the papers lay within the field of land utilization and population geography. Dr. C. Board read a paper on some concepts relating to the quantitative analysis of land-use patterns, Professor J. Kostrowicki spoke on methods of agricultural typology, and Mrs. J. Rakowicz-Grocholska gave an account of Polish geographical studies dealing with urban land use.

The work of the afternoon session was focused on population and urban problems. Dr. J. H. Johnson, reporting on British population studies, was followed by Mr. H. B. Rodgers, who spoke on random sampling techniques in urban geography.

During the Seminar several short trips to near-by places of interest to the geographer were organized. These included an open-pit sulphur working at Piaseczno and a sulphur-processing plant in Machów, as well as new housing developments in the town of Tarnobrzeg. After the Conference the British group, accompanied by the majority of the Polish members of the Seminar, took part in an excursion, passing through Krosno, Biecz, and Gorlice to Krynica. September 7th was occupied by a trip on rafts down the Dunajec river. On the next day the party returned to Warsaw. The programme for September 9th included a tour of the city and an evening performance at the Grand Theatre and Opera. Some of the British guests stayed on for a few more days in Warsaw. Professor R. H. Osborne also visited the Geographical Institutes in Łódź and Wrocław.

At the conclusion of the meetings in Baranów Sandomierski the participants of the Seminar approved the following resolution:

The Third Anglo-Polish Geographical Seminar, meeting from 3rd to 5th September, 1967 in Baranów Sandomierski, Poland, resolves as follows:

1. That the mutual exchange of experience and opinions has been of great value to all concerned;
2. That geographical science has a vital part to play in the more effective analysis and utilization of the territory of the two countries;
3. That reports of the Seminar shall be published in Polish and British geographical journals, as well as in the I.G.U. Newsletter, and that the Proceedings of the meeting shall be published in Poland in *Geographia Polonica*;
4. That the next Seminar shall be held in the United Kingdom after a suitable interval of years and that the date and the theme of the Fourth Seminar shall be decided on after consultation between the Institute of British Geographers and the Institute of Geography of the Polish Academy of Sciences.
5. That a vote of thanks be accorded to the Polish Academy of Sciences for making the necessary financial provision for the success of the Seminar and also to the Ministry of the Chemical Industry for arranging accommodation at Baranów Sandomierski.

PUBLICATION AND COMMUNICATION IN GEOGRAPHY

KEITH M. CLAYTON

As far as I am aware, no one has measured the current annual output of publications in geography, but it probably lies between 10,000 and 20,000. This is, of course, only a fraction of the output in such fields as chemistry or biology, yet it still poses considerable problems. In any one week some 40 United States and 40 Russian papers will be published, with another 30 from the UK and a further 150 or so from all the other countries of the world. Since none of us reads 30 papers a day in our own language, let alone coping with the scores of languages represented, we select from this total output. Ideally we select those papers that are closest to our own research interest, although this is always complicated by the relevance of techniques established in one type of investigation to work in other fields. This means that all active research workers should attempt to survey quite a wide sample of the literature. The need to do this quickly (or they will have time to do nothing else) is one important reason for restricting most of their reading to publications from their own country. Wider reading is often restricted to journals taken on personal subscription, and further reading in libraries then tends to concentrate on publications cited in those papers or located in such secondary sources as abstract journals and bibliographies.

Despite the obvious importance of the process by which published material is read by an audience and so leads to further progress in the subject, the details have been little studied. A series of investigations by the *American Psychological Association* [1] have not been widely read by geographers, while the analysis of citations as a bibliographic source has been neglected, exceptions are papers by Manten [2] and Stoddart [3]. In general, the impression gained from discussion with colleagues is that most geographers feel that the process works, although they are usually well-disposed towards any attempts to improve communication among research workers by such devices as conferences, symposia, and secondary journals. The extent to which articles in primary journals reach the intended readership is not generally available in any case, but it is commonly held that most „good” articles reach a wide audience quite quickly. The satisfaction felt by most authors on seeing their publication in print seems to stifle any enquiry about how far it has been read. Few authors are aware of the very limited circulation figures of most geographical journals, and my own enquiries suggest that too few authors choose the journal in which a particular article is published with sufficient care.

This myth that all is well with regard to the present pattern of communication through the written word in geography seems to me to be full of dangers. The problem that runs right through the present situation is that we are functioning at a dangerously inefficient level. Anyone who reads really widely in his own field must be aware of the parallel and repetitive pattern of much of the research, and of the repetition of erroneous views because earlier work published in another country has not been seen. Besides being inefficient, this system may preclude the type of rapid advance that has been described for research on mast cells by Goffman [4], who showed that the surge of interest in this topic recorded by the number of publications could be described mathematically as an epidemic process. In his case, a threshold of some 100 publications a year represented the point at which rapid growth occurred — below this level the number of papers on this subject remained stable from one year to the next. A similar threshold may exist in geography, for analysis of the bibliography of Central Place Studies by Berry and Pred [5] shows a sudden upsurge, from 86 papers in the period 1945–1949 (about the same as the 76 recorded for 1935–1939), to 182 papers between 1950 and 1954. A peak of 242 was reached between 1955 and 1959, and the 1960–65 total seems to have been no higher. An epidemic can be produced by an increase in the number of papers published on any one topic, but we might also remember that it can be induced by a rise in the population at risk — by an increase in the readership of those publications, so that a smaller total may have the same effect if our pattern of communication is efficient.

As we have already mentioned, information on the actual readership of articles is hard to come by. The American Psychological Association enquiry provides some data bearing on this point. The bulk of the literature in that field (perhaps 85%) is American, and a typical article seems to reach between 100 and 200 readers [6]: few reach more than this. Foreign articles have a very much lower readership. A further difficulty is the very loose definition of „read” adopted by their enquiry — if no more than the title of the article was read (other than on the Contents page) a „reading act” was recorded. Skimming the abstract and reading the last paragraph fall well within this broad definition of „read”. A more reasonable measure may be provided by the average number of requests for separates of an article received by authors — falling in the range 11–15 [7]. We might also note the finding that 45% of the recipients of a journal failed even to open it to read the list of contents!

Most of the guesses at the readership of articles rely on the extent to which publications are cited in the list of references at the end of a paper. No doubt these citations are fewer in number than the total number of readers, although they must form a rough measure of the number of readers who have made effective use of the article in their own work. There are many problems and reservations attached to the use of citations, in analysis of this type, but the ready availability of the data makes them an attractive source of information. The science literature

is analysed in this way in *Science Citation Index* [8], a bibliographical publication that has no parallel in the field of geography. From this Index it can be established that any one article is most likely *not* to be cited in any one year, and that of those that are referred to in other publications, most are cited only once. The most popular articles in this field reach a citation frequency in any one year of about 500.

The citations of over 250 geographical articles published in the United States, Great Britain and Poland in 1965 were analysed to see what pattern emerged. This sample was drawn from the major geographical and geological periodicals of these countries in order to select articles that were aimed at a geographical and geomorphological audience. In the case of the geological periodicals, only articles on geomorphology were included in the survey. Table I shows the pattern of citation averages separately for each of the three countries. Differences in the total number of citations probably reflect in part the inclusion of several monographs in the Polish sample. Thus the average number of citations in the *Przegląd Geograficzny* and the *Czasopismo Geograficzne* is only 21.8, or about the average figure for the United States. The general pattern is quite straightforward, with a strong preponderance in all cases of citation of the literature of the same country as the publication being examined. The dominance of the home literature is less marked for Poland than the USA, reflecting the combined effects of the smaller size of the Polish literature (only 15–20% of the number of articles published annually in the United States or Great Britain) and the greater ability of Poles to read in foreign languages. Similarly the reversed positions of the French and German literature for the United States and Great Britain reflects the difference in language teaching between the two countries.

TABLE 1. The average number of citations, by country in which the articles cited were published, for papers from the main journals of the UK, the USA and Poland

	UK	USA	Poland	Germ.	France	USSR	Other	Total	No of papers
UK	9.80	2.74	0.17	0.39	0.94	0.01	2.13	16.18	100
USA	1.89	14.68	0.01	0.89	0.31	0.10	3.55	21.45	73
Poland	0.67	1.50	16.56	3.05	0.57	2.75	2.96	28.10	86

The proportions for citations of Polish publications represent a numerical total of 17 citations from the sample of 100 British articles, and only one from the sample of 73 United States papers. This one United States citation of a Polish paper was in *Economic Geography*, in a paper by Kirk H. Stone, „Focus for the Geography of Settlement”, which included (among 77 citations from many different countries) the paper by Maria Kielczewska-Zaleska, „New trends

in historical geography of rural settlement" written in Polish, and published in the *Przegląd* for 1963. With a total of some 200 Polish papers each year, the chances of being cited in the United States literature (and so perhaps of being read by those following up the bibliography of an article) are obviously extremely small.

The 17 citations found in the British sample came from four papers. All of these were geomorphological. This is not just chance, nor does it simply reflect the considerable Polish output in this field. Experience with both subscribers and abstractors in connection with *Geographical Abstracts* suggests that an interest in the literature of the subject is well developed in geomorphologists. They seem more actively concerned than are economic geographers, while social geographers are perhaps least interested in publications in their own field. It is hoped to investigate this more fully by some type of questionnaire investigation, but this is the impression we have. Of the 17 British citations of Polish articles, 10 were in the paper by Eric Brown (*Geography*, 1965) that arose out of his visit to Poland. Two articles in the *Proceedings of the Geologists' Association* carried references to Polish work published in connection with the 1961 INQUA Conference, while the last article we are concerned with here, from the same journal, was on a periglacial topic and carried four references to the *Biuletyn Peryglacjalny*, and one to a paper by Professor Dylik in the Łódź series. The significance of the INQUA Conference publications (most of the citations by Professor Brown came from this source) show how important such publications can be in reaching an audience beyond the frontiers of the host country.

Most of the citations of Polish publications abroad are of research articles, usually either written in English, French or German, or carrying an English summary. Books, which rarely carry such summaries, are not often cited. Conversely, citation by Poles of publications from Great Britain or the United States are quite commonly of textbooks, i.e. secondary sources which would normally rarely be quoted in their own countries. Another tendency that can be seen is the quotation of foreign works of a classical nature, although recent foreign publications on the same subject are ignored. Inevitably, foreign publications are older than home publications, although at times the gap is very large indeed. To give examples of these points, the citations of British literature in A. Wróbel's study, *The concept of economic region and the theory of geography*, *Prace Geograficzne* 48, 1965, are all of the older literature. The latest is 1953, and in sum they represent a dated view of the regional concept and not one that is very fashionable in Britain today. Similarly, the long list of references to Niewiarowski's article on kames in the Toruń series [9] shows considerable contrasts in the date of publication of references when these are subdivided by country. The citations from the United States, Soviet and Polish literature are plotted by date of publication in Fig. 1. (There were no citations of British literature). The two oldest papers quoted are American publications, while the most recent citation from the United States literature is the 1st. edition of R.F. Flint's *Glacial Geology and the Pleistocene*

Epoch. The second edition of this was published (with a revised title) in 1957. In contrast the time lag of perhaps one or two years between the Polish and Soviet citations seems an acceptable result of the pattern of dissemination of published material.

If we list the obstacles to communication in geography, we find that most of them are common to other subjects. We have already mentioned the size of the literature, the small circulation of papers, their small readership, and the problem of language. Further difficulties include the limited coverage of the

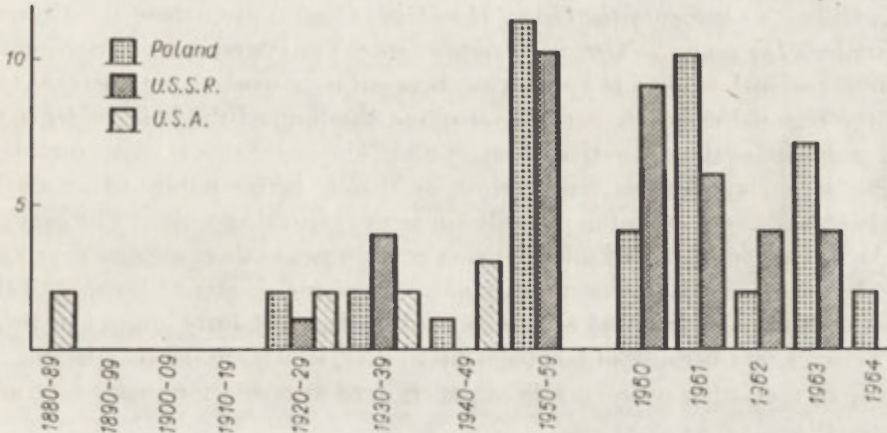


Fig. 1. Date of publication of citations in article on kames by Niewiarowski [9] subdivided by country. Data for U.S.A., U.S.S.R. and Poland only. There were no citations from the UK literature. Note change of time scale after 1960

secondary abstracting journals, the limited provision of abstracts or summaries in a language other than that of the paper, excessive length, poor style, misleading titles, and so on. In addition there seems to be a group of related problems that are peculiar to geography, or perhaps shared with such subjects as geology. A good deal of what we write is necessarily concerned with local regions. Some of this is indeed „mere” description. Most of it applies concepts already used and described elsewhere, although the topic to which these are applied may be wholly original. Some of this local literature further develops these concepts or techniques while a part develops and applies wholly new methods of analysis. Naturally most of this material is written in the local language of the area concerned and is not intended for a wide audience. Alongside this local or regional literature are articles of a systematic nature that are concerned with techniques or with generalisation and may not even quote regional examples. Such articles are commonly thought to be of interest to a wider audience, are more likely to be written in a major world language, such as English or German, and are frequently read at international conferences. It is commonly held that

such articles deserve wide circulation, in contrast to the regional literature, which is of local or at best national appeal.

This view is very dangerous, and if it becomes accepted as a working rule strikes at the whole structure of geography. If we look back at those publications that have been of great significance and have been repeatedly cited over the years, we do not find that these are exclusively of a general nature, indeed regional studies predominate. The concept of regional superimposition was most convincingly described by Douglas Johnson in *Stream Sculpture on the Atlantic Slope*, while much of what was best in early twentieth century geomorphology came from G. K. Gilbert's *Geology of the Henry Mountains*. Central place theorists still quote Christaller's *Die zentralen Orte in Süddeutschland*. The systematic theory contained in these regional studies is convincing because it is worked out in connection with the regional example. It is the satisfying handling of the field evidence that gives authority to the theoretical background. The paper that is most suitable for reading at an international conference, or that is better published in English than Polish or Czech, is not necessarily one with a general approach. The examples may be taken from two Yorkshire parishes or from two or three villages near Zakopane, but the paper is not necessarily only of provincial interest because of that. Soviet work in my own field of geomorphology has very little impact in Britain and America, not because of language difficulties, which are not insuperable, but because of the rather overgeneralised papers read at recent international conferences [10].

At an international symposium where the Polish participants have graciously (if practically) consented to use the language of their visitors, it would perhaps be pertinent to say something about the problem of language in communication. In countries like Poland or Sweden, it is necessary to read and write in foreign languages if there is to be any appreciable communication with the audience outside the home country, and this wider audience is clearly desirable. In the USSR, the UK and USA, the size of the home audience (or, in the case of UK, the world English-speaking audience) makes international communication seem less vital. In the UK and the USA day to day use of foreign languages is not particularly common and is usually limited to either German or French. The citation figures confirm this. Since about half of the world literature in geography is written in English, it is also a common language for summaries or abstracts of papers written in other languages. Yet these English summaries are not always effective means of communication, and paradoxically the increasing use of English in foreign publications is increasing the problem. Englishmen and Americans are well aware of the differences in usage between their countries, but now we have many other brands of English. It is not only language *per se* that divides us, but also differences in terminology and conceptual outlook. *Exaration* is understood by Poles, Germans and Russians, but not by most Englishmen. *Suffusion* is not used in England in the sense in which it is used by Polish geomorphologists.

A recent series of articles in Hungary describe *derasional* valleys, a term found only in the English abstracts accompanying those papers [11]. Terms such as *kame* and *drumlin* come to have national, rather than international connotations. Obviously similar examples could be drawn from other fields of geography.

Perhaps I should conclude on a practical note. My purpose in trying to establish the present limitations to communication between geographers is to establish an awareness that some improvement is needed. I think it must be more widely understood that it is one thing to publish in geography [12], and another to communicate. My own contribution to the problem, the publication of one of the several journals of abstracts of articles in geography, is no doubt a help in the wider spread of published information. However, it seems necessary to envisage some reorganisation in the primary journals, with the aim of producing some few journals with really extensive circulations. It is perhaps too much to hope that any of the smaller journals will disappear, and it is probable that many of them play a very useful role in areal description. Journals with a really extensive circulation would need to be international in character and perhaps rely to some extent on the provision of cover-to-cover translations rather than simply providing summaries in other languages. Once established, their wider readership would ensure that they attracted a large number of papers. Cooperation between some of the largest societies would seem the best way of producing such a journal; a commercial undertaking would probably find costs of distribution and the foreign exchange problems of its potential subscribers a limiting factor. At the same time journals should do all they can to improve their readership. Many fail to provide abstracts with articles, or summaries in other languages. Those that do publish in foreign languages should perhaps do more to secure competent (and this probably means foreign) editing of their material. International collaboration between geographical societies, still the main publishers in our field, can surely be increased.

In conclusion, it would seem that the number of people who ought to read any one article is not usually very high, although for the whole world it is surely well over one hundred. We would be reasonably efficient writers and publishers in geography if we reached some 75% of that potential audience. All the evidence I have seen suggests that at present we are often reaching no more than 10–20% of that potential audience, and that most of these live in the same country as the author.

University of East Anglia

BIBLIOGRAPHY

- [1] Garvey, W. D., Griffith, B. C., *Reports of the American Psychological Association's Project on Scientific Information Exchange in Psychology*, American Psychological Association, Washington D. C., Vol. 1, 1963, Vol. 2, 1965,

- [2] Manten, A. A., Van Dilst, F. J. H., The problems of language in earth sciences, *Atlas*, 1 (2) (1965), 22—23.
- [3] Stoddart, D. R., Growth and structure in geography, *Trans. Institute of British Geographers*, 41 (1967) 1—19.
- [4] Goffman, W., Mathematical approach to the spread of scientific ideas — the history of mast cell research, *Nature*, 212 (5061), (1966), 449—452.
- [5] Berry, B. J. L., Pred, A., *Central Place Studies. A bibliography of theory and applications*. Regional Science Research Institute, Bibliography Series, No. 1, Philadelphia 1961; Barnum, H. G., Kasperson, R., Kiuchi, S., *Supplement through 1964* to the above. Regional Science Association, Philadelphia 1965.
- [6] Garvey, W. D., Griffith, B. C., *American Psychological Association's Project on Scientific Information Exchange in Psychology*, Report No. 9, 1963. The use of scientific journals by psychologists and the readership of current journal articles, p. 245 *et seq.*
- [7] Garvey, W. D., Belver, C., Griffith, B. C., Scientific information exchange in psychology, *Science*, 146 (3652), (1964), 1655—1659.
- [8] *Science Citation Index*, published by Institute for Scientific Information, Philadelphia.
- [9] Niewiarowski, W., Kemy i formy pokrewne w Danii oraz rozmieszczenie obszarów kemo-wych na terenie Peribalticum w obrębie ostatniego zlodowacenia, *Zeszyty Naukowe Uniwersytetu M. Kopernika w Toruniu, Geografia*, 4 (1965).
- [10] For a good example see: *Proceedings Permafrost International Conference, 11—15 November 1963, Lafayette, Indiana*. National Academy of Sciences, Washington D. C. Publication No. 1287, n.d. (?1966), 115—185.
- [11] See, for example: Marosi, S., A deráziós völgyekről, *Földrajzi Ertesítő*, 14 (1965), No. 2; Szilárd, J., A magyarországi periglaciális deráziós völgyképződés egyes kérdései, *Földrajzi Közlemények*, 13 (1965), No. 3.
- [12] *Geographical Abstracts*, Volumes A—D and Annual Indexes, published by Geo Abstracts, University of East Anglia, Norwich Nor 88C, England.

ON GEOGRAPHICAL SCIENTIFIC INFORMATION IN POLAND AND ITS DIFFICULTIES

KAZIMIERZ DZIEWOŃSKI

Modern times create special difficulties in the spread of scientific information. This may seem paradoxical, for we live in an age of mass-media for its distribution. But one important consideration in the present situation is simply the sheer amount of information one has to digest. Another lies in the wide range of concepts and terminology developed by continuously increasing specializations. Somebody has said that during a mathematical congress no more than one tenth of the participants are able to understand what any of the speakers are talking about. I am afraid that the situation is equally bad among geographers. Finally, in geography there is another problem (which does not exist to the same extent among mathematicians), the problem of the different languages in which information is published. The situation worsens considerably with the spread of geographical research throughout the world. One hundred and fifty years ago perhaps nothing of great importance in geography was published outside Germany, France and England. Today, in order to be *au courant* with the full extent of geographical thought, a knowledge of at least five or six major languages is necessary with, in addition, a score of minor ones.

In my paper I shall try to present a report on how we in Poland are trying to cope with this overwhelming flood of geographical studies, and where we see some further possibilities of improvement. I shall start with the problem of languages and then pass on to the more fundamental and more formidable task of dealing with the multiplication of concepts, terms, methods and data.

The problem of languages is better understood by Polish geographers (representing a relatively small nation and a language which is effectively limited in use to its national territory in Central Europe) than by geographers of countries whose culture and language has spread throughout the world. For a scientific worker in Poland it is a *conditio sine qua non* to know at least two foreign languages. Usually these will be English and Russian or French and Russian, but sometimes English and French or French and German. But this is clearly not enough. A limited knowledge of foreign languages deeply influences the whole outlook of a worker and his scientific concepts and methods used. To overcome this and to create a common basis of knowledge the Institute of Geography of the Polish Academy of Sciences has produced, ever since its foundation, a duplicated series, called: *Przegląd Zagranicznej Literatury Geograficznej* (*Review of Foreign Geographical*

Literature). There are usually 3—4 volumes per year, each volume containing a series of articles translated from foreign geographical publications (in various languages), centred around a specific topic, theory or method, dealing with the basic problems of one branch or specialization of the geographical sciences.

At the same time, to introduce Polish geographical achievements to the foreign reader, our main serial publications and periodicals, i.e. *Prace Geograficzne* (*Geographical Studies*), *Przegląd Geograficzny* (*Polish Geographical Review*) and *Czasopismo Geograficzne* (*Polish Geographical Journal*) supplement all their articles by résumés in foreign languages, usually English and Russian. We have also published some occasional special numbers, usually connected with international congresses or other scientific gatherings and published only in foreign languages, mainly English. Since 1963 we have introduced a new series, called *Geographia Polonica*, which contains articles written in one of two foreign languages: English or French.

Parallel to the publication of translations from foreign languages, and of Polish studies in other languages, there is the continuous work of submitting Polish geographical literature to various current bibliographical series — international and others. The Institute of Geography is responsible for Polish publications as presented in *La Bibliographie Géographique Internationale*, *Bibliographie Cartographique Internationale* (both published in Paris by Armand Colin), partly in *Referativny Zhurnal* (published by the U.S.S.R. Academy of Sciences in Moscow), *ZIID Referatekartei Geographie* (published in Leipzig) and *Geographical Abstracts* (published by Prof. Clayton in England).

We are also trying to establish a full bibliography for all Polish geographical publications. The volumes for the years: 1936—44, 1945—1951, 1952—53, 1954, 1955—60, 1961, 1962 and 1963—64 are published. Those for recent years are with the publishers or in preparation. The volume for the years 1918—1927 is practically finished. Material for the years 1928—1935 was collected by W. Ormicki and published as a supplement in the Cracow geographical periodical *Wiadomości Geograficzne* in three parts: 1928—1933 in vol. 13 (1935), and 1934 and 1935 in vol. 14 (1936). They are now being checked and will be republished in standard form as a special volume.

To show the growth of Polish geographical publications I should mention that in 1938 (the last year before the war) there were 1123 separate bibliographical entries; for 1950—527; for 1955—1065; for 1960 — with more selective classification — 2079, and for 1965 — over 2500. The number of editorial sheets used for publication by the Institute of Geography alone, has grown from 54.8 in 1953 (the first year of its existence) to 275.0 in 1965.

In dealing with the flood of published material abroad the first answer may naturally be found in bibliographies, but even these have to be selective because of the wealth of material. Critical bibliographies are easier to use but are more

difficult to prepare. My own experience within the Commission on Methods of Economic Regionalization of the International Geographical Union shows that it is practically impossible to obtain data and material of equal value from sources varying — of necessity — in organization and qualification of their staff. In preparing such bibliographies, moreover, the growing differences in concepts, terms and methods become very evident. Possible solutions are several. The first, and seemingly the simplest, would be to establish a glossary of common, strictly-defined terms with their equivalents in all languages — to create some kind of „Basic English” for geographers. The amount of work involved in such an enterprise would be enormous and, in addition, a number of preparatory studies would have to precede it. Two methods of approach seem to be possible: inductive and deductive.

The Inductive approach would be to adopt and to follow the example of modern philology and lexicography, i.e., to study the meaning of words and terms as actually used by geographers in the past and at present. This would demand a new kind of geographer, specialized in the history of geography, languages and philological methods. In Poland we have one study of this kind, a monograph on one term, by Dr. J. Szewczyk, now being published by Institute of Geography of the Academy. Although it is an historical study (in fact concerned with the late mediaeval period) it shows very well the kind of research and methods which have to be applied. For modern studies, some kind of mass analysis with specific methods would have to be developed.

The Deductive approach implies development of a general theory of the geographical sciences. Even when limited to a single branch of specialization, the construction of a generally acceptable theory presents such difficulties that it may be considered as practically impossible. In fact, the existence of alternative theories is of great importance for the development of science. When one theory breaks down in face of reality, another is usually able to take over and continue research without denying the existence of the given phenomena. The concepts and term should be, therefore, always defined within, or for, specific theory. Yet here one condition should be made — terms should never be formed and used in opposition to their — perhaps rather loose — meaning in everyday language. All efforts to connect a specific word to a concept widely different from, or narrower than, its commonly used meaning are lost from the beginning and only lead to additional confusion. In practice, therefore, the deductive approach has always to be complemented by the inductive. In this way one point is clearly established: we geographers need many studies on concepts and terms commonly used and of long historical standing. These are still not properly developed.

In my opinion studies of the development of basic concepts or methods (and also of their applications) are very important indeed for dealing with the present flood of information and research. Perhaps you will allow me to quote an example

from my own work. This is a monograph on the development of concepts, methods and application of the urban economic base and functional structure; this has shown me the value of such historical, comparative studies. Another case is the monograph on the use of potential and gravitation models in geography prepared by Z. Chojnicki. It seems to me that co-ordinated effort in this direction on the international level would bring comparatively quick results, in form of a series of such monographs dealing with various basic geographical concepts and their terminology. This would be a real „corpus of geographic knowledge and theories” of our time, guaranteeing better understanding and improved exchange of information.

Another solution, perhaps an easier one, and, at the same time, an intermediary and preparatory step to the implementation of the first one, lies in making selective critical bibliographies on important specific subjects. Such a bibliography, relating to economic regionalization, was prepared for the leading nations in geographical research (France, Germany and the German speaking countries, Great Britain and the Commonwealth Countries, the Soviet Union, the United States), within the framework of the Commission on Methods of Economic Regionalization of the international Geographical Union. Each bibliography has an introduction on the historical development of the regional concept in the given country.

Yet another way would be through a co-ordinated effort for the interpretation on a comparative basis of the geographical terminology and vocabulary, both among the various disciplines and specializations, and on the international level. This task is of such dimensions that I had serious doubts whether to discuss it at all. Any general solution of this kind seems to be quite impossible. The only practical steps in this direction seem to me as follows. First, to pay more careful attention in our work to the concepts and terms used in the neighbouring sciences, disciplines and specializations. Such an approach would be truly geographical because — by any opinion — geographers are, or should be, characterized by their sensitivity to the existence of environmental relations. The second step may consist in developing more international and interdisciplinary working groups, which, in dealing with specific phenomena, would develop common concepts and terms, and at the same time would review together the part achievement in research.

Quite another problem in the present state of geographical information is easy access to the publications. The amount of material is such that no library possesses all publications. Everywhere the storerooms are overflowing with books and periodicals. To overcome the critical saturation points libraries have to apply stronger criteria in selection and in the end they have to limit the inflow by specialization in their collections. In Poland this problem of access is aggravated by the war losses of libraries and by the scarcity of the so-called “hard currencies”, leading to a limitation in the import of foreign books and periodicals.

The technical answer to such difficulties is in inter-library exchange and the

wide use of microfilms and photocopies, but for that prior knowledge of where the given publication may be found is necessary. For this reason we have published a general catalogue of the periodicals relating to the earth sciences published up to year 1953 and to be found in Polish scientific libraries. The preparation of the second volume, for the years 1954—1970, is already beginning. We have also published a Central Catalogue of our cartographical collections. Three volumes, listing atlases published before 1918, have been published. The fourth one, on XXth century atlases, is already printed, and we are beginning to prepare the volume on sets of historical topographical maps in our collections.

There were, and are, some efforts undertaken by the National Library in Warsaw to create a general catalogue of publications in Polish Libraries, but they have not been very successful. The Committee for Geographical Sciences, a body connected with the Polish Academy of Sciences, has recently undertaken some efforts for the co-ordination of the work of the libraries of geographical scientific institutions. In fact, in 1966 we had a conference in Warsaw of the heads of the libraries concerned. Various questions were discussed and specific proposals for the exchange of information and for co-ordination in gathering publications were adopted. However, their extent was, and still is, seriously limited by the scarcity of adequately qualified library staff.

Institute of Geography PAN, Warsaw

BEHAVIOURAL POSTULATES AND THE CONSTRUCTION OF THEORY IN HUMAN GEOGRAPHY

DAVID W. HARVEY

Human geographers have always made assumptions of some kind about human behaviour. In some cases these assumptions have been the subject of explicit investigation, but for the most part they have been implicit in geographic analysis rather than explicitly formulated. Similarly, human geographers have resorted to theory in their search for adequate modes of description and explanation. But for the most part the structure and form of these theories have not been explicitly discussed. Most theoretical formulations function as "hopeful explanations" rather than as the controlled hypothetico-deductive systems characteristic of scientific explanation.

The aim of this paper is to discuss the structure and form of scientific theory, to examine the function of such theories in geographic explanation, and to discuss the implications of this mode of analysis for the future development of research in human geography. In the process it will become evident that at least some of the postulates from which appropriate geographic theory may be derived must necessarily refer to human behaviour. An explicit development of scientific theory in geography automatically entails, among other things, an explicit analysis of the behavioural postulates on which that theory can be based.

THEORY

Geographers have frequently appeared to subscribe to the Baconian view of scientific explanation which assumes that "the facts" can be formulated independently of some theory. More recently opinion has tended to suggest that such independence is illusory. "I doubt", writes Kemeny (1959, p. 89), "that we can state a fact entirely divorced from theoretical interpretation." Burton (1963, p. 156) similarly states: —

"The moment that a geographer begins to describe an area ... he becomes selective (for it is not possible to describe everything), and in the very act of selection demonstrates a conscious or unconscious theory or hypothesis concerning what is significant."

The current view seems to be, therefore, that "facts, measurements, and theories are methodologically the same." (Churchman, 1961, p. 71).

Geographers have traditionally regarded themselves as being concerned with the description, interpretation, and explanation of sets of events distributed over the earth's surface. They have automatically been concerned, therefore, with theory of some kind. Theory serves to facilitate and expedite description and explanation. Its role in explanation can most easily be demonstrated by reference to Hempel's covering law model of explanation. Hempel (1965) maintains that all explanation that presumes to be rational can be reduced to the following fundamental form: —

$$\left[\begin{array}{l} \{C_1, C_2, \dots, C_n\} \\ \{L_1, L_2, \dots, L_t\} \end{array} \right] \rightarrow \text{therefore } E$$

(a set of initial conditions)
(a set of laws)
(the event to be explained)

A set of initial conditions, taken in conjunction with a set of law statements which can be referred to those conditions, provide a deductive explanation for a particular event. Not all explanations are rendered in this manner but Hempel maintains that in so far as a set of statements constitutes an explanation, it can potentially be rendered in this form or in its probabilistic equivalent. The function of scientific theory is to provide a consistent, reasonable, and powerful set of law statements which we can use with comparative confidence in such explanations. Even in the absence of explicit scientific theory, certain statements made in explanation will function as-if they are laws. If explanations are to be consistent it is necessary to ensure that the law-like statements used are consistent and reasonable. And this amounts to constructing a corpus of scientific theory.

Not all, of course, accept the Hempelian covering law model. Indeed, historians have been notoriously reluctant to accept that historical explanation necessarily involves establishing historical laws (Dray, 1957; 1964, gives a good summary of this argument). Probably many geographers would not accept the model either, although there has been no methodological debate on this specific issue (the only reference remotely associated with geography is that by Sprout and Sprout, 1966). It is not my purpose to discuss this issue here, or to discuss the wider question of laws in geography. I shall therefore assume that the covering law model provides a powerful explanatory form of relevance to geography, and assume too that the "quest for an explanation is the quest for a theory" (Zetterberg, 1965 edn., p. 11).

To understand the implications of this view we must examine the nature of scientific theories. Theories are essentially intuitive creations. Any speculative assertion might thus be regarded as a theory of some sort. But scientific theory is a particularly controlled form of speculation. The controls ensure the consistency of any theory relative to other theories, the reasonableness of the theory with reference to a particular type of phenomena, and the power of the theory with respect to the range of the phenomena it may be applied to. Scientists try to satisfy these criteria by observing certain rules — rules which prescribe how a theory should

be stated, how a theory should be related to empirical phenomena, and how a theory may be verified. All of these rules are vital for the scientific enterprise, but for the moment we shall consider only one set of rules — those which govern the way in which theory is structured.

(1) THE STRUCTURE OF SCIENTIFIC THEORIES

A scientific theory „may be considered as a set of sentences expressed in terms of a specific vocabulary.” (Hempel, 1965, p. 182). A theory may thus be regarded as “a language for discussing the facts the theory is said to explain.” (Ramsey, 1960 edn., p. 212). The nature and function of such language systems have been discussed by philosophers and logicians. Carnap (1942; 1958) suggests that such languages may be decomposed into three major features. Pragmatics describes the conditions under which a language is formulated; semantics describes the relationship between the abstract signs and symbols contained in the language and empirical phenomena; syntax describes the logical system of relationships contained within the language. With the help of these concepts a paradigm of the structural form of scientific theories may be devised. In particular, it becomes possible to differentiate between the logical structure of scientific theories and their empirical interpretation (see Rudner, 1966, for a full discussion on this point).

By following prescribed syntactical rules the logical coherence and consistency of any theory may be ensured. Briefly summarising, any theory contains a number of undefined primitive terms and a number of defined terms which may be derived from the primitive terms according to certain rules. The sentences contained in the theory may similarly be divided into primitive sentences — axioms — and derived sentences — theorems. In addition the theory will contain a set of rules which govern the formation of theorems. In general these rules are those of deductive inference. An abstract syntactical system — usually termed a calculus — has logical coherence but no empirical content. For empirical usage the calculus requires an interpretation. This is provided by devising a semantical system the same in structure but which possesses in addition a set of designation rules which relate abstract symbols to empirical concepts, and a set of rules which determine the truth condition of the sentences contained in the calculus.

This particular view of a scientific theory raises a number of interesting methodological problems. One of the most interesting is that of providing appropriate empirical interpretations for abstract calculi such as those developed in mathematics. If mathematics be regarded as abstract analytic knowledge, then defining the rules whereby geographic problems can be mapped into such abstract calculi is a profoundly interesting methodological problem. This type of problem is discussed in a general way by writers such as Nagel (1961), Hempel (1965),

Brown (1963) and Rudner (1966) and is examined in a geographic context by Chojnicki (1969) and Harvey (1969). The present paper is concerned, however, with just one aspect of the general problem of theory development — namely the nature of the axiomatic statements — or the basic postulates — from which theorems — or laws — may be derived to facilitate and expedite geographic explanation and description.

(2) THEORY IN GEOGRAPHY

Scientific theory in geography has only weakly been developed. The basic postulates of geographic theory could be exposed by examining the geographical literature and prising out of substantive studies the basic assumptions implicit in their analysis. Such a task would be Herculean and, in addition, one suspects that no clear view of geographic theory would emerge simply because in many cases it is difficult to specify exactly what is being postulated and what is being derived. To get round this problem it seems reasonable to take a normative view and discuss what the basic postulates of geographic theory ought to be, given some generally accepted view of the nature of geography. The justification for this procedure is simply that there is a very strong connection between metaphysical speculation and theory formation. Such speculations have, to use Körner's (1955) terms, acted as "directives" or "regulative principles" in the search for scientific theory. In geography we possess a multitude of such directives in our traditional literature and these have been effectively summarised in general views of the nature of geography.

To provide a generally acceptable definition of the nature of geography is no easy task. For the purposes at hand, however, it suffices to adopt a very general and firmly traditional view. This view has been expressed by Hettner:

"Geography can be an independent science only as chorology; that is, as knowledge of the varying expression of the different parts of the earth's surface. It is, in the first place, the study of lands; general geography is not general earth science; rather, it presupposes the general properties and processes of the earth, or accepts them from other sciences; for its own part it is oriented about their varying areal expression." (quotation from Sauer, 1963, p. 317).

I do not propose to debate the validity of this view here although it is worth noting that it commands wide support among many geographers (Hartshorne, 1959; Sauer, 1963). I shall confine attention to the implications of this view for the construction of geographic theory.

It follows from this view of the nature of geography that theory in geography may be of two distinctive types. On the one hand there will be theories which are special cases of theory developed by other empirical sciences; these theories, it may be suggested, relate entirely to temporal processes. On the other hand there will be theories indigenous to geography; these relate entirely to spatial form.

In terms of basic postulates, therefore, we are presented with an interesting dichotomy between derivative process postulates and indigenous morphometric postulates. The notion of synthesis in geography may then be interpreted as an attempt to derive spatial form laws from process laws. Complete synthesis here will amount to the specification of general theory. This suggested dichotomy is therefore worth examining in greater detail.

(A) Indigenous Postulates

There are plenty of "concepts" and "principles" in geography, which could function as postulates for theory. When examined in this light, however, most of these principles turn out to be either actually or potentially derivable as a special case of some more all-inclusive theory of another discipline. In cases where no such derivation can be given, it may either be because theory in some other discipline is not yet sufficiently advanced for such a derivation to be provided, or it may be that the concept is itself inappropriate.

But there is one distinctive group of concepts which resist reduction. These concepts are essentially to do with location, nearness, distance, pattern, morphology, and so on. They make reference therefore to the rather special relationships which exist between geography and geometry. This relationship is special in the sense that geometry may be regarded as abstract analytic knowledge and geography may be regarded as no more derivative of it than marginal economics is derivative of calculus. Geometry thus provides a syntactical system ready made for the construction of geographic theory. This key relationship between geography and the different forms of geometry has perhaps been lost to view until recently (see Bunge, 1966 edn., chapters 8 and 9; Haggett, 1965, pp. 9—17). Indigenous geographic theory may be regarded therefore as an attempt to state the laws of spatial form in the specialised languages of geometry or topology, or in the more general form of spatial statistics. Such laws describe spatial form, areal associations, areal co-variances, and so on. They are explanatory in the sense that given one side of an Euclidean triangle and two angles all the other features of the triangle are determined. Such laws rest firmly on the interpretation of abstract mathematical concepts — such as point, line, plane figure, lattice, and so on, in a geographical context. These interpretations form the basic postulates for indigenous geographic theory.

(B) Derivative Postulates

The laws of spatial pattern do not necessarily tell us anything as regards process. Relatively complete explanations of chorological form demand an explicit consideration of process and at this point geographic theory becomes potentially

derivative of theory in some other discipline. The truth of this could be demonstrated in all areas of geographic research. King (1966, p. 328) has thus stated that the morphometric approach in geomorphology provides "an objective method of establishing fundamental empirical relationships between various attributes of the landscape", but goes on to point out that "this type of analysis is one stage in the development of a theory of landscape development; by itself it cannot explain the significance of the relationships". As King later points out, the theory must rest firmly on an understanding of process and in this the postulates of physics play an all-important part. Barry (1967) has similarly pointed out that work in meteorology and climatology rests on the process postulates of physics.

In human geography the derivations are less clear partly because the social sciences as a whole do not possess well-ordered sets of postulates. But there are many instances where the geographer's notion of process is clearly derivative of one or another of the social sciences. Christaller (1966 edn) in his initial formulation of central place theory, for example, rested his analysis entirely on the postulates of economics. But in general the interrelationships are not very easy to identify. The rest of this essay, therefore, is specifically concerned with the behavioural postulates for geographic theory.

THE BEHAVIOURAL POSTULATES

Human geographers have always made *some* assumptions about human behaviour. The determinist viewpoint postulates a mechanistic response of individuals and groups to physical environment, the possibilist view stresses the inner psychology that governs such response, and the ecological view emphasises the complex interaction between environmental stimuli, human perception and decision making to maximise largely culturally determined goals. Such themes have never been ignored in human geography (Kirk, 1951; Brookfield, 1964). In other cases geographers have accepted the notion of "economic man" to specify location theory, only to become disillusioned with the rather too rigorous bounds placed on behaviour by such a normative concept. The history of such assumptions about behaviour is an intriguing one — unfortunately it yet remains to be written. In some cases such assumptions have been developed indigenously only to be subsequently sharpened and clarified by reference to some other discipline. In other cases completely new assumptions have been introduced. A good example of the former case is the history of the idea of "satisficing behaviour". Wolpert (1964) recently re-introduced the concept into geography, deriving the notion mainly from the psychologist Simon (1957; 1966). Wolpert (1964, p. 558) thus wrote of the farmers of Central Sweden that:

"The concept of the spatial satisficer appears more descriptively accurate of the behavioural pattern of the sample population than the normative concept

of Economic Man. The individual is adaptatively or intendedly rational rather than omnisciently rational."

By referring to the psychological literature Wolpert is able to clarify and partly operationalise the concept of satisfaction. But the concept is not new to geography. In 1940 Sauer (1963, p. 360) commented that the mode of living developed in any culture area was "its way of maximising the satisfactions it seeks and of minimising the efforts it expends. That is perhaps what adaptation to environment means." In 1912 Brunhes (1920 edn., p. 605) wrote that it was the task of the human geographer to "distinguish the natural and general psychological effect that (geographical) facts produce upon men, upon men obeying certain instinctive or traditional suggestions, seeking the satisfaction of certain needs. . . ." Doubtless the same idea could be further traced in traditional geographic literature. For the most part, then, we seek to clarify and sharpen concepts by relating them to theoretical formulations in one or other of the social sciences.

The behavioural concepts which may be so related to the postulates of social science theory may be crudely classified under a number of headings. We may thus refer to concepts regarding: —

1. Cultures.
2. Economies — national, regional, local.
3. Complex organisations — e. g. governments.
4. Simple organisations — e. g. firms.
5. Groups — religious, political, ethnic, social.
6. Individuals.

Each social science tends to be concerned with a specific group of concepts — anthropology with culture concepts; economics with certain attributes of economics, organisations, groups and individuals; sociology with attributes of organisations and groups; psychology with attributes of individuals. But there is considerable overlap and it may well be better to think of a unified social science rather than different disciplines each pursuing separate aims. This suggests the ultimate possibility of defining a small number of basic postulates upon which all the theoretical formulations of social science may be based.

Given the present state of knowledge it is unlikely that such reduction will be accomplished in the near future. The issue is, therefore, largely a metaphysical one, but it does have some methodological significance. In the natural sciences, where physics reigns supreme, a considerable proportion of the concepts may be defined by reference to a few basic postulates of physics. Such theoretical definition of the idealisations of natural science give natural science theory an important practical advantage over the idealisations of social science. In the social sciences it seems that if such reduction is possible it will relate all social science concepts to concepts regarding individual behaviour. Some, such as Carnap (1956), have suggested that further reduction to purely physical postulates is both feasible and desirable. Such a chain of reduction, although not operationally

possible at the present time, holds out the tantalising prospect, according to Freeman (1966), of a value-free and truly objective social science. It may also allow the theoretical definition of the idealisations of social science. This will facilitate the verification of social theory and the important practical differences which currently exist between idealisations of, say, economics and physics, will disappear (see Hempel, 1965, chapter 7, for a discussion on this point). It would naturally follow from this that geographic theory based on economic postulates, such as central place theory, would be susceptible to precise empirical test, instead of relying on intuitive appeal and rough empirical test for its justification as at present.

On a more restricted scale, such reduction does suggest an increasing interlocking of social science concepts. Such a prospect does not please all social scientists. Devons and Gluckman (1964, p. 241) object on the grounds that "the different human and social sciences may be different realms, in whose borderlands trespass is dangerous save for the genius". The Gestalt psychologist and the field theorist would object that there are many concepts which cannot be reduced. Similarly many anthropologists, such as Kroeber (1944), regard culture and its associated concepts as providing a self-sufficient analytic framework — a view which has been turned into methodological credo by White (1947) when he states that "the most realistic and scientifically adequate interpretation of culture is one that proceeds as if human beings did not exist." But in the search for a deeper and more penetrating understanding of cultural dynamics more and more anthropologists are turning to psychological concepts (G. and L. Spindler, 1963). Economists in their search for a more adequate understanding of concepts of value and utility have increasingly turned to psychology for enlightenment (Simon, 1966). Psychologists have also become aware that anthropological, sociological, and economic theories have implications for psychological theory (see Koch, 1963). These growing interrelationships are of tremendous importance to analytic human geography, since they indicate a rapid shift in the nature of the process postulates on which geographic theory may be based.

A DECISION-MAKING FRAMEWORK

Decision making may be regarded as the most convenient focus for the discussion of behavioural postulates in a geographic context. We may regard spatial patterns in human geography as being the immediate consequence of human decisions. The complexity of interacting behavioural forces may be summed up in the act of decision. In a sense, organisations, such as firms, decide, and even cultures and economies can be examined with the assistance of decision-making models. Models of conflict behaviour, for example, can be applied to individuals, firms, nations, complex organisations, groups, and so on. The

application of such models requires care but considerable insight can also be attained into the human processes governing spatial form.

We may usefully divide our understanding of decision making processes into three approaches: —

1. Normative — how people ought to behave given certain goals.
2. Analytic — how people ought to act given our knowledge of human behaviour.
3. Descriptive — how people in fact behave in response to certain stimuli.

The differences between these viewpoints may in some instances be quite small. As Savage (1954, p. 20) has remarked, even logic itself “can be interpreted as a crude but sometimes handy empirical psychological theory.” The specification of a normative model may involve deep analytic understanding and the resultant difference between the normative model and a descriptive one may appear almost nominal. Nevertheless the distinction between the approaches is a convenient one for discussing behavioural postulates.

(1) NORMATIVE MODELS

Microeconomic theory provides us with a highly sophisticated normative model of entrepreneurial behaviour under conditions of certainty and under the assumption that the entrepreneur is concerned to maximise returns or, on occasion, to minimise costs. From such a concept of marginal behaviour location theory has been derived. The von Thünen model is a classic case of the development of a set of spatial laws from such a model of decision-making behaviour. Similarly, the Löschian, interregional equilibrium, and central place models, may be developed from the same set of postulates as regards behaviour, although the initial conditions are specified differently. The Weberian model is rather different since it is concerned with cost minimisation rather than profit maximisation and it has, thus, probably a less general applicability to spatial systems.

The nature of the spatial systems which have been deduced from microeconomic theory are too well documented to discuss further here (see Isard, 1956). Nevertheless these deductions form the starting point for much of our discussion. Geographers have sometimes objected to such derived models on the grounds that they are far too normative to be descriptively useful. Economists, on the other hand, have sought to modify economic theory to make it more realistic. In particular they have sought to introduce the notions of (1) risk (2) uncertainty and in connection with (2) have re-examined (3) their assumptions about the state of information available to a decision maker. Economists such as Shackle (1961) have thus placed the emphasis on expectation in decision making, and expectation, it turns out, is difficult to measure except in terms of some value system. The importance of psychology and sociology in such a determination is self-evident.

Most normative decision models now attempt to specify how an entrepreneur (or a government leader or an individual) ought to proceed in the face of incomplete information about future states of the environment (natural or social) (Isard and Dacey, 1962). Game theory has been specifically developed to deal with this kind of situation. Two types of game have special significance.

(a) Games against nature are particularly important in the context of resource-use decision-making in an uncertain environment. The importance of such games in an agricultural context has been discussed elsewhere (Dillon and Heady, 1960; Gould, 1963; Harvey, 1966). Games of this type have the following structure; a number of strategies, f_i , and a number of states of nature, s_j , are defined and each combination is associated with an outcome, o_{ij} , and the matrix of outcomes forms the pay-off matrix. The decision maker chooses among the strategies (or a mixture if conditions allow) and the normative model assumes that the decision maker chooses so as to maximise some quantity. Precisely how he does this depends upon the way in which the decision model is set up and the criteria adopted for making a choice.

In attempting to provide a rigid yet reasonably realistic decision model a number of difficult but interesting problems have emerged; particularly problems of defining utility, value, and so on. Before considering these a brief example will be developed. Suppose a farmer is faced with determining whether to grow hops or fruit on a given acre of land and the quantity he is concerned to maximise is profit. Suppose that a pay-off matrix of the following form can be identified:

		relevant states of nature			
		s_1	s_2	s_3	s_4
Strategies	Hops	0	0	200	25
	Fruit	20	40	20	40

Associated with each state of nature is a probability of occurrence. We may thus define a row vector of probabilities $p(sj) = (2/7, 1/7, 1/7, 3/7)$ for example. One simple way to solve this is to evaluate Bayesian expectation by the matrix manipulation $[o_{ij}] \cdot p(sj)^T$, (where $p(sj)^T$ is the transpose of the probability vector) which yields an expected income of 49.3 units for hops and 31.4 units for fruit (see Jeffrey, 1965, pp. 1—13, for a discussion of the Bayesian decision rule). Numerous other decision rules could be used to solve such a problem. Dillon and Heady (1960) probably give the best review in an agricultural context but Luce and Raiffa (1957) discuss the problem in general.

(b) Competitive games are important in the field of conflict of interest between individuals or organisations. The conflict of interest may be among nations (Rapoport, 1960), regions (Isard and Smith, 1966 and 1966B), firms and individuals (see Luce and Raiffa, 1957; Shubik, 1964; and Becker and McClintock, 1957; for surveys). Again there are a number of ways in which such games may be specified and given a normative solution and each has kind of implication for

spatial form. An interpretation of location theory in terms of game theory leads to interesting results which are frequently at variance with the theory of Losch or von Thünen. Hotelling (1929) pointed out that three firms locating under inelastic demand would locate at the centre of the market, while with more than three firms the solution to the location problem becomes undeterminate. The generality of this view has been demonstrated by Stevens (1961) using game theoretic concepts. Specifying games and evaluating their spatial implications is instructive therefore, since it allows an experimental treatment of processes which for far too long have been assumed as monolithic. Consider, for example, a special class of game, usually termed Prisoner's Dilemma, which has been intensively studied by Rapoport and Chammah (1965). It is a non-zero-sum game in which pure competition between players is not the only strategy choice. A player may co-operate or compete, but if one player co-operates and the other competes, then the pay-offs are so structured that the player who competes maximises his pay-off and the co-operating player receives minimum reward. But if both players compete they receive less than if they both co-operate. Again, we may treat this in a spatial context with some interesting results. Consider two players who think of two possible locations for a retail store — centre and suburb. The structure of pay-off might typically be as follows: —

		Player A	
		centre	suburbs
Player B	centre	10	20
	suburbs	2	12

Assume each player makes his choice independently. This amounts to specifying a probability vector of the following form: $p(sj) = (p, 1 - p)$. Using the Bayesian decision rule both players find the expected outcome for a central location to be 20 units and for a suburb location 20 units irrespective of the value given to p . An examination of the four possible outcomes indicates that both players in a central location is an inefficient solution and that the best solution in net returns is a suburb location. This highly simplified prisoner's dilemma game demonstrates how a rational decision process under conditions of pure competition may lead to disadvantageous central city agglomeration rather than to a more worthwhile dispersal of economic activity. A socially optimal location process can only be determined under conditions of co-operation among entrepreneurs. The location pattern in this case is dependent upon the degree of co-operation between players.

The point of demonstrating these two types of game is to show how various decision models may be used to predict spatial form. A large number of decision models can be developed and each may have distinctive spatial implications. A normative analysis of spatial pattern thus requires an evaluation of alternative

decision models and at this point we need some deeper understanding of the way such models may be specified and solved. In particular it involves an inspection of how strategies are identified (in many cases a large number of possible strategies may exist), how the relevant states of nature may be identified (again there may be many), how probabilities are attached to such states, and how some value system governing choice behaviour may be specified. In some cases game theory has itself been used to identify some aspects of the decision process. The modern definition of utility, for example, depends upon game theory for its specification (von Neumann and Morgenstern, 1964 edn; Fishburn, 1964). Given that people in fact make choices between a finite number of strategies and attach subjective probabilities to certain outcomes, we can, under certain assumptions, construct a utility scale for individuals. This kind of situation has been studied empirically (e. g. Davidson and Suppes, 1957) and the results suggest that many of the assumptions — regarding transitivity and comparability of choice for example — are sometimes reasonable (see Becker and McClintock, 1967, for full discussion). But as Simon (1966, pp. 6—8) has pointed out, complicated choices do not appear to be described so well and even some simple situations show marked departures (e.g. one armed bandit experiments indicate that people often develop event-matching behaviour rather than profit-maximising behaviour). This has led many workers to abandon the theory of utility maximisation on the grounds that it has little relevance to real choices.

In spite of attempts to specify normative search procedures to identify alternative strategies (Charnes and Cooper, 1958), attempts to identify relevant states of nature normatively, attempts to show how one should learn from experience normatively (using the Bayesian rules for identifying *a posteriori* probabilities), and so on, this whole problem of specifying the decision model and specifying its mode of solution appears to require a non-normative approach. We shall, therefore, turn to a brief discussion of analytic and descriptive models of decision.

(2) ANALYTIC MODELS

The attempt to understand how people in fact decide has two aspects. Psychologists have thus developed analytic concepts regarding the drives, motives and stimuli which lie behind decision making as well as descriptive measures of actual decision-making behaviour under experimental conditions. Both aspects have significance for geographic theory.

Analytic studies, such as those of Hull (1952) and Spence (1956), have largely been geared to the construction of theory and it is possible to deduce certain features of spatial behaviour from such a system. Hull (1952, p. 268) develops two specific concepts to discuss spatial behaviour — approach or *adient* behaviour and avoidance or *abient* behaviour towards objects located in space. Both forms of behaviour “have gradients of reaction potential which are high near the objects in question and decline with distance from the objects, probably roughly accord-

ding to a negative growth function." This observation fits Zipf's (1965 edn.) notion of least effort spatial behaviour and suggests an underlying rationale for numerous models of spatial interaction (e. g. gravity and diffusion models). Unfortunately, many of Hull's statements are not proven either deductively or empirically.

Similar comments may be made regarding perception. People react to perceived stimuli and the structure of their perceptions materially affects the decisions they take. Visual perception of space, for example, is non-Euclidean and exhibits negative curvature (Luneburg, 1947; Roberts and Suppes, 1967). The mental maps, as Gould (1966 A) calls them, of the individual decision maker are a mixture of elements determined by the negative curvature of spatial vision, the intuitive perception of space and elements of non-spatial imagery attached to place names. Certainly, such maps exhibit warping and discontinuity. According to Lee (1963) people's vision of space may be conceptualised as schemata — a concept which may be extended from the neighbourhood unit which Lee was specifically concerned with to the structure of shopping areas and so on. That people have highly structured images regarding the space that surrounds them cannot be doubted (Lynch, 1960) but the impact of this structure has been almost ignored in the analysis of behaviour in space. Do people simply shop indiscriminately within some imaginary boundary? Problems such as this play an important role in determining spatial patterns as Peterson (1967) has recently suggested in a study on the perception of residential desirability.

But we are not simply concerned with the perception of space since the perception of what that space contains is also significant. In part, as Hallowell (1951) has pointed out, these perceptions are governed by cultural factors. Thus Firey (1960) has attempted to build a theory of resource use based on the three-fold relationship of what is ecologically feasible, what is conceptually feasible, and what is economically gainful. Certainly the perception of environment, as workers such as Kates (1962) and Saarinen (1966) have pointed out, has tremendous implications for how that environment is used.

Gestalt psychologists have been particularly concerned with perception and the particular branch of it termed field theory (Lewin, 1951; Yinger, 1965) has special relevance to a discipline concerned, as is geography, with the world-wide ecosystem. The general parallel between field theory and the ecological approach in geography has not been exploited, although Kirk (1951) has pointed to the significance of gestalt concepts to geography and Wolpert (1965) has more recently indicated how field theory may be relevant to migration studies. Thus Lewin's concept of a life-space — which is the surface over which an organism can move and which is "dependent upon the needs, drives, or goals of the organism and its perceptual apparatus" (Wolpert, 1965, p. 163) — has much in common with the ecological formulations given to geography by such writers as Brookfield (1964) and Stoddart (1967).

The other elements in field theory — the drives, needs, and goals, provide another sphere of analysis of relevance to geography. That people act to satisfy (or reduce) needs may appear self-evident, but it is not always easy to state what those needs are or how the action relates to such needs. Simon (1957) has thus suggested a concept of satisfaction—a concept which is difficult to make operational but which does serve to indicate that individuals act in response to different needs. Individual entrepreneurs may be regarded as satisficers instead of economic optimisers (Wolpert, 1964), but perhaps even more significant for location theory is an analytic understanding of the behaviour of firms and other complex organisations. According to Simon (1945), March and Simon (1958) and Cyert and March (1963) a firm may be regarded as a coalition between investors, managers, workers, and customers, but each group within this coalition has a rather different set of objectives. Managers, for example, may be concerned to maximise their salary and minimise their effort and only indirectly via pressures from other elements within the coalition will they be concerned to maximise the profits of the firm. Normative location theory has usually assumed an entrepreneur concerned to maximise profits, and this may have been a reasonable assumption in nineteenth century capitalist societies. But at the present time such an assumption is not so reasonable since the majority of enterprises are corporate and the objectives of the different groups within a corporation may be radically different. Behavioural theories of the firm are thus likely to lead to the formulation of a location theory which is more in accordance with reality. Similarly, an analysis of complex organisations (Etzioni, 1961) will yield much greater insight into important political and institutional decisions which again, in the mid-twentieth century context, are extremely important in governing the development of spatial patterns. Theories developed to explain such behaviour should clearly include analytic concepts to do with motivation, goal identification, stress behaviour, and the like, if they are to be realistic.

(3) DESCRIPTIVE MODELS

Some psychologists, behaviourists in the strict sense, suggest that the analytic concepts so far developed are not powerful enough to be treated theoretically. Skinner (1953) thus insists on a rigorously descriptive and a-theoretic approach to human behaviour. Some of the most interesting work to stem from this conception is learning theory. Guthrie initially developed a simple contiguity of stimulus reponse learning and, as Golledge (1967) has pointed out, such models have direct applications to geographic problems. Golledge thus uses the stimulus-response-reward type of model developed by Bush and Mosteller (1955) to examine the emergence of a system of market areas. The behaviour of the consumer is here regarded as a stochastic learning process operating over space — a conceptual view which has also been explored by Gould (1965, 1966 B). Certainly,

the enormous amount of work on statistical learning theory has conceptual implications, if not technical applications, in our attempt to define the processes which create spatial form (Atkinson, Bower and Crothers, 1965, provide a good introduction to such learning models).

The difficulty of such formulations, however, is to adapt models developed to describe very simple experimental situations to apply to aggregate phenomena. This problem may be partly overcome by developing more generalised descriptive measures of human behaviour. Probability distributions appear particularly suitable for such a purpose and, once identified, a particular probability distribution may itself be used as a postulate about behaviour from which further theory can be developed. Undoubtedly the most important of these probability distributions is the Poisson (together with its generalised or compound versions). The Poisson distribution describes a random process occurring at a given intensity (or density) and it has been fitted to all kinds of data such as the frequency of telephone calls arriving at an exchange, cars arriving at traffic control points, customers in a shop, and so on. The distribution, as Coleman (1964, 291) points out, has peculiarly appropriate attributes for the description of aggregate human behaviour. It has the added advantage of also being applicable to spatial series. Probability theory thus provides at least one way of accomplishing a time-space transformation simply because the same calculus is directly applicable to both process and spatial form. For example, suppose the arrival of cars at a traffic control point on a road is described by a Poisson distribution with given density and we know the average speed of cars along that stretch of road, then, if the process remains constant, we can infer the probability distribution of cars along sections of the road.

Poisson processes have been widely discussed in geographic research. Dacey (1962, 1966) has used the Poisson distribution in various situations to describe spatial pattern and has also indicated the relationship between such descriptions and simple process postulates. Curry (1962, 1964, 1967) has also indicated how the Poisson distribution might be used as a basic behavioural postulate for the construction of geographic theory. He seeks for a more adequate foundation of central place theory than that provided by normative economic postulates and provides this foundation by assuming that customer shopping behaviour can be described as a Poisson process. Such an approach appears much freer than that of traditional central place theory, which is too closely bound to the idealisations of economics to be empirically justifiable. In his latest work Curry (1967) attempts a more sophisticated time-space transformation and indicates one possible channel along which the synthesis from process postulates to spatial form might proceed. Curry assumes purchasing behaviour to be a stationary stochastic process over time. By the assumption of ergodicity, a stationary stochastic series over time is assumed to exhibit the same statistical properties as an ensemble of events distributed over space. This amounts, there-

fore, to a direct statement of the statistical properties of a spatial series of settlements from an understanding of the statistical properties of a temporal purchasing process. The assumption of ergodicity is a brave one, but it provides the simplest possible transformation. It is characteristic of such general treatments to discard much of our analytic understanding of decision processes in order to discuss the impact of aggregate decisions on spatial form with greater clarity and ease. But at this stage in the search for some general theoretical framework, it seems doubtful if the problem can even begin to be discussed without making some very general assumptions. The assumption of ergodicity will certainly be challenged, but it would be wrong to dismiss this approach at the outset on the grounds that the assumptions are unrealistic. Probabilistic postulates about aggregate human behaviour will undoubtedly form an important basis for a growing volume of geographical theory and this theory will have the distinctive advantage of being aggregative and general rather than individual and specific.

CONCLUSION

At the present time the level of understanding and control which behavioural science as a whole gives us over decision-making processes is not sufficient for the construction of geographical theories of any great power. As behavioural science itself progresses so we may expect more powerful behavioural postulates to be generated. Perhaps a first conclusion to be reached is that an awareness of these changes is a *sine qua non* for the geographer seeking to understand spatial form, since the concepts of behavioural science can sharpen and fashion geographic thinking in a most productive way. In terms of research effort this survey suggests two modes of approach to the general problem of deriving an understanding of spatial form from a knowledge of temporal process. We may thus seek to apply specific normative or analytic concepts to certain situations, or we may attempt to explore the dimensions and form of general theory using the mathematics of probability theory, which seem peculiarly adapted to such a task. Whichever path we follow, it seems certain that geographic theory can no longer rest content with implicit assumptions about human behaviour. The sooner these assumptions are made explicit the better.

University of Bristol

BIBLIOGRAPHY

- [1] Atkinson, R. C., Bower, G. H. and Crothers, E. J., *An introduction to mathematical learning theory*, Wiley, New York 1965.
- [2] Barry, R. G., Models in meteorology and climatology, in Chorley, R. J. and Haggett, P. (eds.), *Models in geography*, Methuen, London 1967.

- [3] Becker, G. M. and McClintock, C. G., Value: behavioral decision theory, *Annual Review of Psychology*, 18 (1967), 239—286.
- [4] Brookfield, H., Questions on the human frontiers of geography, *Economic Geography*, 40 (1964), 283—303.
- [5] Brown, R., *Explanation in social science*, Routledge and Kegan Paul, London 1963.
- [6] Brunhes, J., *Human geography*, Harrap, London 1920 (Engl. edition).
- [7] Bunge, W., *Theoretical geography*, Lund Studies in Geography, (2nd edition) Series C, No. 1, 1966.
- [8] Burton, I., The quantitative revolution and theoretical geography, *Canadian Geographer*, 7 (1963), 151—162.
- [9] Bush, R. R. and Mosteller, F., *Stochastic models for learning*, Wiley, New York 1955.
- [10] Carnap, R., *Introduction to semantics*, Harvard Univ. Press, Cambridge 1942.
- [11] Carnap, R., The methodological character of theoretical concepts, *Minnesota Studies in the Philosophy of Science*, 1 (1956), 38—76.
- [12] Carnap, R., *Introduction to symbolic logic and its applications*, Dover Paperbacks, New York 1958.
- [13] Charnes, A. and Cooper, W. W., The theory of search: optimum distribution of search effort, *Management Science*, 5 (1958), 450—458.
- [14] Chojnicki, Z., Some theoretical aspects of the application of mathematical models in economic geography, *Geogr. Polon.*, 18, 1969.
- [15] Christaller, W., *Central places in southern Germany*, Prentice Hall, Englewood Cliffs, New Jersey, 1966 (Engl. edition).
- [16] Churchman, C. W., *Prediction and optimal decision*, Prentice Hall, Englewood Cliffs, New Jersey 1961.
- [17] Coleman, J. S., *Introduction to mathematical sociology*, Free Press of Glencoe, New York 1964.
- [18] Curry, L., The geography of service centres within towns: the elements of an operational approach, in Norborg, K., (ed), *I. G. U. symposium in urban geography*, C.W.K. Gleerup, Lund 1962.
- [19] Curry, L., The random spatial economy — an exploration in settlement theory, *Annals Ass. Am. Geogr.*, 54 (1964), 138—146.
- [20] Curry, L., Central places in the random spatial economy, *Journ. of Regional Science*, 7(1967), 217-238.
- [21] Cyert, R. M. and March, J. G., *A behavioral theory of the firm*, Prentice Hall, Englewood Cliffs, New Jersey 1963.
- [22] Dacey, M. F., Analysis of central place and point patterns by a nearest neighbor method, in Norborg, K. (ed.), *I.G.U. Symposium in urban geography*, C.W.K. Gleerup, Lund 1962.
- [23] Dacey, M. F., A probability model for central place locations, *Annals Ass. Am. Geogr.*, 56 (1966), 550—568.
- [24] Davidson, D. and Suppes, P., *Decision making: an experimental approach*, Stanford Univ. Press, Stanford 1957.
- [25] Devons, E. and Gluckman, M., Models and consequences of limiting a field of study, in Gluckman, M., (ed.), *Closed systems and open minds*, Oliver and Boyd, Edinburgh 1964.
- [26] Dillon, J. L. and Heady, E. O., *Theories of choice in relation to farmer decisions*, Iowa Agricultural and Home Economics Experimental Research Station, Bulletin No. 485, Ames, Iowa.
- [27] Dray, V., *Laws and explanation in history*, Oxford Univ. Press, London 1957.
- [28] Dray, W., *Philosophy of history*, Prentice Hall, Englewood Cliffs, New Jersey 1964.
- [29] Etzioni, A., *A comparative analysis of complex organisations*, Free Pres of Glencoe, New York 1961.

- [30] Fishburn, P. C., *Decision and value theory*, Wiley, New York 1964.
- [31] Firey, W., *Man, mind and land : a theory of resource use*, Free Press of Glencoe, Ill. 1960.
- [32] Freeman, D., Social anthropology and the scientific study of human behavior, *Man*, New Series, 1 (1966), 330—342.
- [33] Golledge, R., A conceptual model of the market decision process, *Journ. of Regional Science*, 7(1967), 239-258.
- [34] Gould, P., Man against his environment: a game theoretic framework, *Annals Ass. Am. Geogr.*, 53 (1963), 290—297.
- [35] Gould, P., *A bibliography of space searching procedures*, Research Note, Dept. of Geography, The Pennsylvania State University, 1965.
- [36] Gould, P., *On mental maps*, Discussion Paper No. 9, Michigan Inter-University Community of Mathematical Geographers, Ann Arbor 1966-A.
- [37] Gould, P., *Space searching procedures in geography and the social sciences*, Working Paper No. 1, Social Science Research Inst. Univ. of Hawaii, 1966-B.
- [38] Haggett, P., *Locational analysis in human geography*, Edward Arnold, London 1965.
- [39] Hallowell, A. I., Cultural factors in the structuralization of perception, in Rohner, J. H. and Sherif (eds.), *Social psychology at the crossroads*, Harper, New York.
- [40] Hartshorne, R., *Perspective on the nature of geography*, Rand McNally, Chicago 1959.
- [41] Harvey, D. W., Theoretical concepts and the analysis of land use patterns, *Annals Ass. Am. Geogr.*, 56 (1966), 361—374.
- [42] Harvey, D. W., *Explanation in geography*, Edward Arnold, London 1969.
- [43] Hempel, C. G., *Aspects of scientific explanation*, Free Press of Glencoe, New York 1965.
- [44] Hotelling, H., Stability in competition, *Economic Journ.*, 39 (1929), 41—57.
- [45] Hull, C. L., *A behavior system*, Yale Univ. Press, New Haven 1952.
- [46] Isard, W., *Location and space economy*, Wiley, New York 1956.
- [47] Isard, W. and Dacey, M. F., On the projection of individual behavior in regional analysis, — Parts I and II, *Journ. of Regional Science*, 4 (1962) (1), 1—34 and (2), 51—83.
- [48] Isard, W. and Smith, T. E., A practical application of game theoretical approaches to arms reduction, *Peace Research Society Papers*, 4 (1966-A), 85—98.
- [49] Isard, W. and Smith, T. E., On the resolution of conflicts among regions of a system, *Papers*, Regional Science Ass., 17 (1966-B), 19—46.
- [50] Jeffrey, R. C., *The logic of decision*, McGraw-Hill, New York 1965.
- [51] Kates, R. W., *Hazard and choice perception in flood plain management*, Research Paper No. 78, Dept. of Geography, Univ. of Chicago, 1962.
- [52] Kemeny, J. G., *A philosopher looks at science*, Van Nostrand, Princeton 1959.
- [53] King, C. A. M., *Techniques in geomorphology*, Edward Arnold, London 1966.
- [54] Kirk, W., Historical geography and the concept of the behavioural environment, *Indian Geogr. Journ.*, Silver Jubilee edition, 1951.
- [55] Koch, E., *Psychology : a study of a science*, Vol. 6., McGraw-Hill, New York 1963.
- [56] Körner, S., *Conceptual thinking*, University Press, Cambridge, Mass. 1965.
- [57] Kroeber, A. L., *Configurations of cultural growth*, Univ. of California Press, Berkeley 1944.
- [58] Lee, T. R., Psychology and living space, *Transactions*, Bartlett Society, 2 (1963), 9—36.
- [59] Lewin, K., *Field theory in social science*, Harper, New York 1951.
- [60] Luce, R. D. and Raiffa, H., *Games and decisions*, Wiley, New York 1957.
- [61] Luneburg, R. K., *The mathematical analysis of binocular vision*, University Press, Princeton, New Jersey 1947.
- [62] Lynch, K., *The image of the city*, M.I.T. Press, Cambridge, Mass., 1960.
- [63] March, J. J. and Simon, H. A., *Organizations*, Wiley, New York 1958.
- [64] Nagel, E., *The structure of science*, Harcourt, New York 1961.

- [65] Peterson, G. L., A model of preference: quantitative analysis of the perception of the visual appearance of residential neighborhoods, *Journ. of Regional Science*, 7 (1967), 19—32.
- [66] Ramsey, F. P., *The foundations of mathematics and other logical essays*, Littlefield Adams, Paterson, New Jersey 1960 edn.
- [67] Rapoport, A., *Fights, games and debates*, Univ. of Michigan Press, Ann Arbor 1960.
- [68] Rapoport, A. and Chammah, A. M., *Prisoner's dilemma: a study in conflict and co-operation*, Univ. of Michigan Press, Ann Arbor 1965.
- [69] Roberts, F. S. and Suppes, O., Some problems in the geometry of visual perception, *Synthese*, 17 (1967), 173—201.
- [70] Rudner, S., *Philosophy of social science*, Prentice Hall, Englewood Cliffs, New Jersey, 1966.
- [71] Saarinen, T. F., *Perception of the drought hazard on the Great Plains*, Research Paper No. 106, Dept. of Geography, Univ. of Chicago, 1966.
- [72] Sauer, C. O., *Land and life*, Univ. of California Press, Berkeley 1963.
- [73] Savage, L. J., *The foundation of statistics*, Wiley, New York 1954.
- [74] Shackle, G. L. S., *Decision, order and time*, University Press, Cambridge 1961.
- [75] Shubik, M., (ed.), *Game theory and related approaches to human behavior*, Wiley, New York 1964.
- [76] Simon, H. A., *Administrative behavior*, Free Press of Glencoe, New York 1945.
- [77] Simon, H. A., *Models of man*, Wiley, New York 1957.
- [78] Simon, H. A., *Theories of decision making in economics and behavioral science*, *Surveys of Economic Theory*, Vol. 3, MacMillan, London 1966.
- [79] Skinner, B. F., *Science and human behavior*, Free Press of Glencoe, New York 1953.
- [80] Spence, K. W., *Behavior theory and conditioning*, Yale Univ. Press, New Haven 1956.
- [81] Splinder, G. and L., Psychology in anthropology; applications to culture change, in Koch, S. (ed.), 1963, *op. cit.*
- [82] Sprout, H. and M., *The ecological perspective on human affairs*, Univ. Press, Princeton 1965.
- [83] Stevens, B., An application of game theory to a problem of location strategy, *Papers*, Regional Science Ass., 7, 1961, 143—158.
- [84] Stoddart, D. R., Organism and ecosystem as geographical models, in Chorley, R. J. and Haggett, P. (eds.), *Models in geography*, Edward Arnold, London 1967.
- [85] von Neumann, J. and Morgenstern, O., *Theory of games and economic behavior*, Wiley, New York 1964 edn.
- [86] White, L. A., Culturological vs. psychological interpretations of human behavior, *Am. Soc. Review*, 12 (1967), 686—698.
- [87] Wolpert, J., The decision process in spatial context, *Annals Ass. Am. Geogr.*, 54 (1964), 537—558.
- [88] Wolpert, J., Behavioral aspects of the decision to migrate, *Papers*, Regional Science Ass., 15 (1965), 159—169.
- [89] Yinger, J. M., *Toward a field theory of behavior, personality and social structure*, McGraw-Hill, New York 1965.
- [90] Zetterberg, H., *On theory and verification in sociology*, Bedminster Press, Totawa, New Jersey, 1965.
- [91] Zipf, G. K., *Human behavior and the principle of least effort*, Hafner, New York 1965.

SOME THEORETICAL ASPECTS OF THE APPLICATION OF MATHEMATICAL MODELS IN ECONOMIC GEOGRAPHY

ZBYSZKO CHOJNICKI

It is the purpose of this paper to discuss some theoretical aspects of the application of mathematical models in economic geography. I do not intend to describe the possible range of mathematical models that can be used in the processes of scientific discovery in economic geography, but rather to consider some presuppositions and implications in the use of mathematical models. I think that a full understanding of the theoretical basis of mathematical models helps us to avoid some misunderstandings concerning their use.

As examples of mathematical models I shall use gravity and potential models. Whilst these models are in no way particularly sophisticated, they certainly constitute the simplest and at the same time the most widely used interaction models in economic geography. I hope that they will illustrate some general theoretical aspects of the use of mathematical models. There are many examples in the geographical literature of the use of mathematical models. The best analytical review is by P. Haggett (1965).

At this point it is necessary to consider the general meaning of a model, although this is not an easy task. The results of studies presented by methodologists show us that we cannot hope to give one unique structural definition for the concept of a model in different disciplines. Given that the model concept is variously interpreted and performs a variety of functions, the only way of reaching a definition of it is by way of formal pragmatics: system A is a model of system B if the study of A is useful for understanding B and if there is no direct or indirect interaction between A and B (L. Apostel, 1960, p. 180). The systems must therefore resemble one another and the resemblance is in terms of the pattern or order exhibited in each system. More specifically, the models have an isomorphic relationship, and the term analogue can be used as a generic term for both conceptual and physical isomorphs (A. Kaplan, 1964, p. 263).

R. L. Ackoff et al. (1962, p. 108) have suggested a simple three stage classification of models into iconic, analogue and symbolic models, in which each stage represents a higher degree of abstraction than the last. Iconic models represent properties at a different scale; analogue models represent one property by another; symbolic models represent properties by symbols. R. Chorley (1964) carried this classification process further and created a "model of models", illustrating it with examples from both physical and human geography.

Models can also be classified on the basis of their language. Physically a model is a non-linguistic system analogous to some other system being studied. Conceptually, if the model and its prototype are both general systems or structures one has the most abstract case of an algebraic model. If one of the terms of the relationship is represented by a language one has a semantic concept of a model (that is, relating a language to an arbitrary domain). If both terms are languages one has a syntactic concept (that is, relating two languages to each other). Very broadly speaking a mathematical model falls into the category of a semantic model.

The original field (e. g. the geographic reality we are concerned with) is thought of as being “projected” upon an abstract mathematical domain (made up of sets and functions defined on sets), such that the causal or structural relationships which hold in the original field can be said to be “modelled” by the relationships that hold in the mathematical domain. It is sometimes held that the mathematical model functions as a kind of ethereal analogue model, in which the mathematical equations are regarded as if they express some inherent property — an invisible mechanism — of the original system under investigation. This last suggestion has been rejected as an illusion by M. Black (1962, p. 224). A mathematical model has only one thing in common with analogues — the property of “as-if-ness”. A mathematical model also rests on appeals to look at events “as if” the underlying causes or interactions had a certain structural analogy to the model proposed.

K. J. Arrow (1951) defines a mathematical model as a set of quantitative relationships expressed in the language of mathematics and describing the interaction of phenomena. Starting with the proposition that such a model consists of a system of equations, each of which is meant to measure in quantitative terms some distinct relationship, I must emphasize that this concept departs in many ways from the classical concept of the model as an isomorph. However, the two concepts have some similarity.

A mathematical model is really a simplification in mathematical language of a certain cognitive problem, so that this problem may be more readily understood and a solution attempted. Any mathematical model is an approximation of a given real-world situation. It is generally simpler than the situation it represents, for these situations are usually so complex that an exact representation would lead to futile mathematical complexity. Moreover, the simplifying assumptions should be made explicit, so that one can determine how they falsify the cognitive problem. R. L. Ackoff et al. (1962, p. 117) proposed that the complete justification for such an approximation required a comparison between the “cost” arising from mathematical complexity and the “cost” of lost performance in following a course of action selected by use of the model as compared with a course of action selected by use of a less approximate model.

The conversion of the simplifying assumptions into mathematical forms is conditioned by the choice of variables and the functional form of the model.

This process can be carried out on the basis of common sense, or by reason of more sophisticated theoretical considerations, or by analogy.

The choice of variables leads to the omission of some of them, because their impact on the performance of the model is quite small, and their contribution to the mathematical complexity large. Also, the mathematical characteristics of the variable may be changed to simplify its handling; for example a continuous variable may be treated as discrete, or vice versa. However, the chief problem in the construction of a model is the choice of mathematical function. As a general rule the simplest function is desired. Linear functions are usually the simplest to deal with and nonlinear functions are frequently approximated by linear functions. Furthermore, if the function is not specified by a hypothesis or theory, it has to be determined by empirical search.

An interpreted mathematical model has two fundamental cognitive functions: the hypothetico-deductive and hypothetico-empirical. The hypothetico-deductive function of a model facilitates the easy and accurate deduction of conclusions. The hypothetico-empirical function leads to the concretization of the model, i.e. giving numerical values to the parameters of the model. The process of concretization of the model is based on the theory of statistical estimation. Here the researcher faces a difficult choice between the most efficient method of estimating the parameters of the model and other methods which allow reasonable estimates to be made with a minimum amount of computation. Thus regression procedures are frequently used to estimate relationships, but the appropriateness of such procedures, particularly with respect to socio-economic phenomena, has been called into question by some writers. A discussion of this issue, however, involves a discussion of the whole rationale of statistical methods. The estimation of parameters requires a sound knowledge of the mathematical properties of the different methods of estimation and cannot rely upon intuition and good sense alone.

Models may also be classified as stochastic or deterministic. However, there is a controversy about this classification, a controversy which can be found in almost any field, except where the model can be formulated only in one of the two ways. The main argument for the probabilistic or stochastic approach is that social or economic processes are in fact probabilistic, so that a deterministic process is only a poor substitute. But, as J. Coleman (1964, p. 427) suggests, a deterministic model has some advantages. It can reflect in a more simple manner the same basic process as does the stochastic model. Thus it is possible to treat processes of a far greater degree of complexity than the cumbersome stochastic equation would allow. In addition one of the tendencies in the use of a probabilistic model is to develop a kind of "know-nothing" approach towards the behaviour of the system that the model is intended to reflect. Such a model has few or no definite statements about cause and effect or the relations between variables.

In the realm of social science mathematical models can be classified also into (1) a descriptive or class phenomena model, and (2) a normative or decision model. A descriptive or class phenomena model simply expresses significant causal or structural relationships stripped of the irrelevancies and complexities of the real world. A normative model specifies what ought to happen (or be made to happen) if an optimum (or satisfaction) is desired.

Because economic geography is a discipline predominantly concerned with cognition and not with the strategies of decisions it should concentrate upon the construction of descriptive models as one of the methods allowing the formation of quantitative generalizations, and, in my opinion, a descriptive model can play a more important role in economic geography than can a normative one. Such descriptive models provide a basic framework for a predictive and theoretical science in which the emphasis is placed on empirical statements. The procedure of statistical estimation of the parameters of the model will then allow one to state specific empirical relations.

The main stream of geography is concerned with the collection, classification and ordering of data. This descriptive, old-fashioned view of geography played, in the early stages of the discipline, an important role, but now it is not sufficient. In essence, descriptive economic geography does not focus on the most necessary goals of modern science, namely explanation and prediction. The application of a mathematical model is strictly associated with the role of economic geography as a predictive or theoretical discipline, with emphasis placed on making general statements. In this context an essential point is that a model provides the framework on which we can build a theory, and mathematical models and statistical methods are means to help achieve this end.

In the present part of the paper I wish to discuss gravity and potential models. They consist of simplified assumptions for studying interaction in space, and they are the simplest and most widely used interaction models. Summaries of the literature can be found in a review article by G.A.P. Carrothers (1956) and in the works of W. Isard (1960), G. Olsson (1965) and Z. Chojnicki (1966).

The gravity and potential models should be considered in the first place from the point of view of their expediency for space economics. Such expediency, however, is dependent on the cognitive value of the models, which consists of a proper description of reality. Gravity and potential models can be applied in many fields, for they are important as generalizing descriptions and as elements of various more complex models. According to W. Isard (1960), the models in question are of substantial significance for regional planning, since they complement other analytical and forecasting methods, such as input-output and linear programming, which have gained general application in research.

The models in question do not have uniform constructions, so they will be dealt with separately.

THE GRAVITY MODEL

The initial statement of the gravity model in a form analogous to the basic notions of Newtonian physics should be treated as an attempt to describe stable, structural space relations in the behaviour of populations. This model has purely heuristic meaning, since in the present state of knowledge it seems to be impossible to establish a transformation which would "project" the properties of a physical field into economic space. Such an analogy, however, may be treated as an initial step in formulating a simplified assumption about actual spatial behaviour. The proper gravity model, or better, the gravity model of interaction in space, represents such an assumption. The model focuses on the structural interdependence concerning interaction of masses (populations) and omits the motivation of individual behaviour. The model is also empirical, because the mechanism connecting distance with the frequency of interaction cannot be explained directly by a centrally symmetric and strictly monotonic function of distance, and because mass must be discriminated in relation to the type of mutual interaction.

These assumptions are formulated as an equation: —

$$I_{ij} = G \frac{(w_i M_i)^{\alpha_i} (w_j M_j)^{\beta_j}}{D_{ij}^b}$$

where

- I_{ij} = interaction between places i and j ;
- M_j = mass of inhabitants in place j ;
- M_i = mass of inhabitants in place i ;
- D_{ij} = distance between places i and j ;
- w_i = weight attached to M_i ;
- w_j = weight attached to M_j ;
- G = empirically derived constant;
- α_i, β_j, b = empirically derived exponents.

The gravity model formulated as above poses, however, conceptual and technical problems, namely: —

- (1) The choice of adequate measures of mass and distance, and
- (2) The choice of exponents attached to such variables.

When this equation is taken as the general assumption for investigations concerning interaction with respect to a determined range of empirical phenomena, the problem is placed in the domain of a mathematical model. This enables a statistical estimation of the parameters of the model, once the form of the appropriate function has been established.

It should be noticed that most mathematical models describing quantitative relations between various economic and non-economic magnitudes are the expression of general economic laws. But the lack of an adequate theory makes it impossible to specify general regularities in space economic research. In the case of

the gravity model the normal course of establishing a mathematical model is reversed. Starting with a function similar to the notion of gravitation energy one endeavours to determine analytically the nature of the assumptions necessary for investigating economic interrelations in space.

Empirical research on the part played by distance in interaction demonstrates that the function may be approximated by a linear function. Available results of research also demonstrate that the gravity model can be presented as a simple regression of $\log I_{ij}/M_i M_j$ as the dependent variable on $\log D_{ij}$ as the independent variable. That is, we have: —

$$\log \frac{I_{ij}}{(w_i M_i)^{\alpha_i} (w_j M_j)^{\beta_j}} = \log G - b \log D_{ij}^b$$

Such a formulation, which may be called the regression form of the gravity model, enables one to make a statistical estimation of the parameters of the model. Such a statistical analysis is relatively simple, since: —

1) the identification problem does not occur here, i.e. the parameters of the model can be always estimated statistically,

2) estimation may be done by means of the least squares method, although other methods of estimation should not be excluded. It should be stressed that the estimation of the model parameters, formulated above, enables one to determine the exponent of the distance variable, and to verify the model by means of variance analysis. The model thus defined, fulfilling the conditions of verification, may be treated as an empirical regularity, having very definite restrictions in time and space. Those restrictions are determined by the framework of statistical data, within which the parameters of the model have been estimated and verified. Simplicity and easy estimation are basic virtues of the model thus determined.

It must be stated, however, that correct estimation of the parameters of the model is limited, in the first instance, by its assumptions. That is why the cognitive value of the model is dependent on the variables of the model and their conceptual identification, that is to say, on the determination of the types of interaction of mass and distance. This means that the basic gravity model is superimposed upon reality. The appropriateness of this procedure has not been closely evaluated from a methodological point of view, although it is clearly of crucial significance in evaluating the applicability of the model. The difficulty of identifying *a priori* a problem situation to which the gravity model may be applied, is the weakest point in the methodological analysis of the applicability of the model to social and economic phenomena. Thus far, research has largely amounted to the simplified evaluation of the influence of distance and mass, those two last elements being taken as basic factors which determine the quantity interpreted as spatial interaction.

Adequate statistical data, in terms of their amount and precision, are the second essential element in correct estimation. Previous research, especially the investi-

gations of M. Helvig (1964), demonstrate that if the above assumptions are met, essential interrelations describing interaction are detected. One should, however, point to the necessity of keeping a certain balance in determining the degree of detail within the notions of interaction and mass variable. Excessive detail may lead to distortions in the structural unity of the dependence being considered, and it may also cancel the assumption concerning the mass character of the investigated phenomena.

It should also be stressed, that the gravity model cannot be applied correctly to such spatial phenomena or processes which do not correspond to its assumptions, or which are of a more complex character. However, the review of recent research by Z. Chojnicki (1966) has also shown that the gravity model may be used to determine other variables, such as effective distance, social residue etc., which extends the range of its application. Empirical relations or hypotheses obtained on the basis of the model are a step towards setting up an empirical theory of spatial structure.

THE POTENTIAL MODEL

The concept of the potential model is derived from the gravity model. It is represented thus: —

$$V_i = G \sum_{j=1}^n \frac{(w_j M_j)^{\beta_j}}{d_{ij}^b}$$

where G , W_j , M_j , β_j , d_{ij} and b are defined as before, and where there are 1, 2, ... n places j , and V_i is the measure of the potential exerted at place i by the n places j . This procedure is an example of the deductive role of a mathematical model. The potential model is concerned with the influence of all masses on one point, while the gravity model concerns interaction of pairs of masses. In the potential model, contrary to the gravity model, the empirical correspondent of the value of interaction, expressed by the potential at a given point, cannot be strictly determined. That is why evaluation of the parameters of this model, and especially of the parameter of distance, is impossible. Thus, the potential model cannot be empirically verified in a direct way. It may, however, be verified indirectly, when determining the degree of correlation between the magnitude of the interpreted potential model and other phenomena, e.g. as the potential of population or of income with other social or economic phenomena.

Criticism of the potential model is based on the fact that the notion of potential cannot have a strict interpretation in relation to the spatial structure of social and economic phenomena, and that it cannot be statistically approximated. This criticism is based on the assertion that the model in question assumes an unlimited spatial continuum, and only describes the tendencies toward spatial equilibrium.

However, such criticism does not seem justified. From the point of view of its logical construction the potential model should be interpreted as a model of measurement, serving as a measurement standard and not as an empirical hypothesis. In relation to spatial research on social and economic facts, in which the actual systems are isolated to a certain degree (countries, continents), the application of the model may be limited to an integrated description of such system. The role of the potential model is based on the fact that each of the elements of the system (i. e. areal unit, region) is characterized in relation to all other elements (and to itself), considered from the point of view of their spatial economic interrelations, measured by distance within the system. The range of empirical interpretations of the model is expressed by the specification of the notions of mass and distance. Thus, the potential model enables one to find a measure of an empirically interpreted mass (of population or income), in a given place, from the point of view of spatial differentiation of other masses. When the potential is determined for all masses under consideration, the effect of the situation of each of those masses within the system in question can be taken into account. It follows that the potential model as a research assumption fulfils the postulate of treating the social and economic reality in this spatial aspect as a structural whole.

Adam Mickiewicz University, Poznań

BIBLIOGRAPHY

- [1] Ackoff, R. L., *Scientific method*, New York, 1962.
- [2] Apostel, L., Towards a formal study of models in the non-formal sciences, *Synthèse*, 12 (1960), 125—161.
- [3] Arrow, K. J., Mathematical models in the social sciences, in: D. Lerner, H. D. Lasswell (eds.), *The policy sciences*, Chapter VIII, New York 1951, 129—130.
- [4] Black, M., *Models and metaphors*, Ithaca, 1962.
- [5] Brodbeck, M., Models, meaning and theories, in: L. Gross (ed.), *Symposium on sociological theory*, Evanston, III., 1959.
- [6] Carrothers, G. A. P., A historical review of the gravity and potential concepts of human interaction, *Journal Am. Inst. of Planners*, 22 (1956), 94—102.
- [7] Coleman, J., *Introduction to mathematical sociology*, New York, 1964.
- [8] Chojnicki, Z., *Zastosowanie modeli grawitacji i potencjału w badaniach przestrzenno-ekonomicznych* (Sum.: The application of gravity and potential models in spatial economic research), Studia, KPZK PAN, Vol. 14, Warszawa 1966.
- [9] Chorley, R., Geography and analogue theory, *Annals of the Ass. Am. Geogr.* 54 (1964), 127—137.
- [10] Czerwiński, Z., O interpretacji równań ekonometrycznych (Sum.: On the interpretation of economic relations), *Przegl. Statyst.*, 13 (1966), 229—243.
- [11] Dąbbska, I., Le concept de modèle et son rôle dans les sciences, *Revue de Synthèse*, Paris 1959.
- [12] Haggett, P., *Locational analysis in human geography*, London 1965.

- [13] Helvig, M., *Chicago's external truck movements : spatial interactions between the Chicago area and its hinterland*, Univ. of Chicago, Dept. of Geography, Research Paper 90, 1964.
- [14] Isard, W., *Methods of regional analysis*, New York 1960.
- [15] Kaplan, A., *The conduct of inquiry*, San Francisco 1964.
- [16] Lukermann, F. and Porter, P.W., Gravity and potential models in economic geography, *Annals Ass. Am. Geogr.*, 50 (1960), 493—505.
- [17] Olsson, G., *Distance and human interaction*, Philadelphia 1965.
- [18] Pawłowski, Z., *Modele ekonomiczne równań opisowych* (Sum.: Econometric models), Warszawa 1963.
- [19] Warntz, W., *Toward a geography of price*, Philadelphia 1959.

RECENT BRITISH CONTRIBUTIONS TO THE STUDY OF POPULATION GEOGRAPHY

JAMES H. JOHNSON

This paper presents a progress report on British¹ research into population geography during the last decade. This work is notable for its quantity and diversity, reflected both in the variety of periods and areas studied and in the purposes for which research has been undertaken. Yet, looking over recent work, three themes are particularly noticeable. One is a concern with the geographical study of population data from the underdeveloped countries, particularly those in Africa. Another is a recurring interest in population found among British historical geographers. The third is the analysis of contemporary population problems within the British Isles.

POPULATION STUDIES IN OVERSEAS LANDS

The historical accident that until recently Britain was a major colonial power, with the responsibility for initiating higher education, largely explains the interest of British geographers in many overseas countries. The staffing of university institutions in former colonial territories has led to the export of many British geographers and this contribution has not yet been extinguished. Many of these overseas universities were established after the end of the Second World War and geographers on their staffs have naturally turned their attention to their local environments, where the equation between population and resources is often a more direct one than in Europe or North America. In many of these countries, too, censuses of passable accuracy were being taken for the first time in the post-war decades. Hence it is not surprising that many workers in these lands were attracted by problems connected with the geography of population.

¹ In the preparation of this paper 'British' research has been defined rather liberally. The work of British-born geographers who are now permanently resident in the United States, Australia and New Zealand has been excluded, except where their research refers to geographical problems in Britain or was published before their departure. On the other hand, the work of Commonwealth geographers with permanent academic appointments in Britain has been included. Research by British geographers who are at present serving in universities in the underdeveloped countries has also been included.

In some fortunate cases geographers were able to influence the actual planning of censuses and in particular were able to ensure that enumeration areas were clearly delimited in map form before the censuses were taken. Often this influence was at an informal level and has not been documented, but J. M. Hunter has described the manner in which enumeration districts within urban areas in Ghana were planned as part of a preparatory exercise for the 1960 census [1].

Sometimes, either through default or because of the sheer difficulties of the task, census results in the underdeveloped countries are not too clearly tied to well-defined areas and the geographer's contribution has been more of a rescue operation. An interesting example of the problems involved and the manner in which they have been solved is given in K. M. Barbour's, account of the preparation of his „Location of Omodias Map" for Sudan [2]. The preparation of this map was necessary before the construction of a detailed population map, since neither the boundaries of the smallest census areas nor their relative position were known for the country as a whole. What did exist were lists of villages in each *Omodia*, and from this information it was possible to reconstruct the approximate boundaries for each of these units, which were then checked in the field. In practice, changes in place names caused a considerable number of difficulties, but at least a working draft could be produced from such topographical maps as existed, and this draft was then elaborated and corrected by local enquiries.

Preparatory work of this kind and the general interest of geographers in the demographic problems of the former colonial territories led at first to the preparation of a series of population distribution maps, which represented an important contribution to the inventory of human resources in these lands and presented the information on population distribution in the census returns with a precision never before achieved. A few examples of the great variety of maps produced include those of K. M. Barbour for the Sudan, R. M. Prothero and J.H. Jennings for Nigeria, W. T. W. Morgan for Kenya, T. E. Hilton for Ghana, J. R. V. Prescott for Rhodesia, R. G. Ward for Fiji, and D. L. Niddrie for Puerto Rico [3]. Some of these maps were also converted into density maps, using the precise distributional patterns revealed on the dot maps as a basis for this work [4].

Curiously enough, in the light of this considerable cartographic activity and the social and economic context provided by the underdeveloped countries, relatively little fundamental work has been undertaken by British geographers on the concept of population density. Certainly they have easily avoided the trap into which some economists and demographers have stumbled, that if one area of a country supports a certain density of population, then another area cannot be "overpopulated" if it supports fewer people per unit area. But while geographers have naturally been keenly aware of the inherent variations in the physical endowments of different areas, they have been less efficient in dealing with the problem of objectively equating people as producers and consumers of resources. Such

ideas as the use of Standard Nutritional Units have probably a useful application in making allowances for such factors as the varying food requirements of different populations [5], but they provide a much less satisfactory key to the cultural differences which complicate this kind of calculation. Similarly, no satisfactory solution has been advanced for the problem of equating different economies, technologies and energy resources, which is required in the cross-cultural comparison of population densities.

But if most British population geographers have not grappled frontally with these intractable problems, at least they have been aware of them; and, as they have usually restricted their attention to the relationship between population and resources within individual countries or regions, they have avoided some of the hazards which are involved in the comparison of different cultural contexts, although often subjective methods of assessing the degree of population congestion have had to be used. J. R. V. Prescott, for example, has studied overpopulation and overstocking in Matabeleland, where he has established an interesting contrast between the congested African areas and the "underdeveloped" farmlands held by Europeans [6]. Elsewhere A. T. Grove has also examined the links between agriculture and changing population densities, his work being undertaken in Northern Nigeria [7].

Recently, too, some attempts have been made to apply objective formulae in assessing population density in relation to resources. For example, J. M. Hunter has experimented with the problem of ascertaining population carrying capacity under traditional systems of agriculture [8]. This work was conceived with special reference to conditions in Ghana and it remains to be seen whether it can be applied elsewhere. Hunter has also approached this basic problem from a somewhat different direction by studying the incidence of seasonal hunger in northern Ghana, a condition which he sees as an indication that population pressure exceeds the capacity of the agricultural system [9].

More elaborate work of this kind shows that British geographers are moving beyond the simple description of population distribution, although it can be argued that this cartographic activity was an essential step following the burst of census-taking in the late 1950s and early 1960s. The elaboration of the earlier studies is shown, not only by studies of population density, but also by analyses of recent population changes, which are adding depth to our understanding of contemporary distributions. In many of the emergent countries this is a topic of the greatest importance, but it is one that is particularly difficult to study because of the problems involved in comparing censuses of shaky reliability.

Nevertheless the attempt has been made. K. M. Barbour has described population changes in Sudan during the twentieth century, although handicapped by the fact that there has been only one full census enumeration in that country [10]. In Ghana, aided by a more satisfactory history of census-taking, T. E. Hilton has examined population changes between 1921 and 1948; and, working in

more detail, he has also studied the connection between soil exhaustion and depopulation in the Upper Region of Ghana [11]. In Ghana, too, J. M. Hunter has examined the regional patterns of population growth between 1948 and 1960 [12]. The general importance of population growth in the world today has also attracted some British geographers to study a number of areas outside Africa, where population growth is a particularly significant factor. For example, D. L. Niddrie has studied the population problems of Puerto Rico and Tobago [13], and H. C. Brookfield has analysed the severe demographic problems of Mauritius from the geographer's viewpoint [14].

Population movements of a variety of types and distances have also been examined as a complement to studies of distribution. In particular, there have been a number of studies by British geographers of temporary and permanent migration in different parts of Africa. R. M. Prothero has summarized his own extensive work on this topic and has collected together many of the other findings concerning African population mobility [15]. Prothero has also discussed the implication of these movements in the spread of malaria, a field in which he has acted as a consultant for the World Health Organization [16].

Other studies have examined the movement of population from rural to urban areas. In particular Hong Kong has been subjected to intensive study by S. G. Davis, D. J. Dwyer and T. D. Vaughan [17]. Here immigration from south China has been an important influence on the growth of population, which has in turn encouraged the expansion of manufacturing. However, the prosperity created by population growth is now threatened by its continuance, which poses problems for housing and water supply. The housing problems caused by immigration to Asian cities has also been examined by Dwyer, in a comparison of the situation in Manila and Victoria-Kowloon [18]. Similar urban problems in the West Indies were explored by C. G. Clarke in a study of Kingston, Jamaica, which examined the demographic and other social distributions found within this tropical city [19].

RECENT POPULATION STUDIES BY BRITISH HISTORICAL GEOGRAPHERS

A second group of population studies has been associated with research in historical geography. In describing this geographical work rough justice must be done to the general body of research on the history of British population, since geographers have made an important, but nevertheless only a partial contribution to that work. Moreover, it must be said that those who have undertaken this research often would not look upon themselves as demographic geographers in the strict sense of the term, but have become involved with the study of population because of its direct connection with changing regional landscapes and economies.

The population studies of historical geographers have largely been focussed

on the British Isles during the nineteenth century, largely because of the availability of census results which form a firm base for studies of regional variations at a variety of scales. Some geographers, however, have undertaken population studies of pre-census times. In particular the monumental geographical analysis of the Domesday Book of 1086, under the direction of H. C. Darby, is producing information on the medieval English population, although this is just one of the many topics which the geographical exploration of this famous document is revealing [20].

Information which is more susceptible to more precise demographic analysis becomes available after the middle of the sixteenth century, when parish registers were first kept in England. Adopting the methods of some French researchers, A. E. Wrigley has been pioneering in England the reconstitution of the demographic history of individual families from the information in these records [21]. The work is laborious, but this technique allows a precise demographic explanation of population changes in a way which other methods do not permit, particularly where migration may have taken place. So far Wrigley has published an analysis for one parish in which there is a particularly satisfactory collection of records [22]. Here he has studied such features as fertility changes, age of marriage, size of completed families and intervals between births; and this sample study has revealed that, even in pre-industrial England, there was a very flexible demographic response to economic and social conditions. Although this research requires expansion to other areas and times, it is likely to be one of the most profitable lines of approach to the study of the population of pre-census Britain.

A number of other workers have attempted to bridge the gap between the sixteenth century and the nineteenth. Again, of necessity, they have had to undertake local studies at a variety of scales and with a variety of purposes, often reflecting the nature of the information available for study. R. A. Butlin, for example, has assessed the population of Dublin in the late 17th century, when this city, with 70,000 inhabitants, was second only to London among the towns of the British Isles [23]. In 17th century Dublin the distribution of social classes showed affinities with surviving traditional cities rather than western cities². An example of a different type of study of the same period is provided by D. R. Mills' examination of the effects of the Poor Laws on the distribution of population in parts of Lincolnshire in the seventeenth and eighteenth centuries [24]. Yet another type of problem was explored by Nora A. McIntosh in her study of the changing population of part of Renfrewshire in relation to the decline in agriculture and the development of manufacturing in the eighteenth and early nineteenth centuries, using the superior information on eighteenth century population which is available in Scotland [25]. In Scotland, also, R. A.

² The residences of the wealthier section of the population lay inside the walls, with the poorer classes being located in a peripheral location.

Gailey has studied the connection between settlement and population changes in Kintyre in the second half of the eighteenth century [26].

Studies concerned with the nineteenth century possess more of a family likeness because of the similarity of their source material, although there were detailed variations in the censuses of England and Wales, Scotland and Ireland. Although no hard and fast line can be drawn, these studies can be divided into those which examine the population geography of the nineteenth century for its own inherent interest and those which are concerned with the evolution of modern conditions from the situation which existed in the nineteenth century.

Much of the work of R. Lawton falls into the first category, with its concern for the redistribution of population in Victorian England. For example, Lawton has used the published census reports to discuss population movement in Midland England during the middle of the nineteenth century and to examine the destination of Irish immigrants to England and Wales at the same time [27]. Such work has led him into a consideration of the general problems which are presented by the analysis of population mobility in England and Wales in the later nineteenth century, concerning himself, in particular, with the role of migration in causing population change [28]. A parallel theme has been studied by C. M. Law, who has examined the growth of urban population from 1801 to 1911 [29].

Similar topics have been investigated for nineteenth century Ireland. Here S. H. Cousens has examined the distribution of population change and certain other demographic features in a series of articles which cover the period between 1821 and 1891 [30]. Other workers have preferred to work on a narrower scale. J. H. Johnson, for example, has examined the population geography of one Irish county from a number of points of view, studying such features as the changing age of marriage, the effects of the Great Famine of the 1840s and pre-Famine emigration [31]. Other studies have sought to link population change with other geographical phenomena. V. B. Proudfoot and T. D. Vaughan have explored the connection between the changing rural settlement pattern of northeastern Ireland and population change, and J. H. Johnson has investigated the same topic in Londonderry [32]. Similarly Johnson's analysis of harvest migration to Britain from nineteenth century Ireland has explored how this movement was closely integrated with the annual time-table of rural life, both in Ireland and in Britain, and was closely linked with the socio-economic status of the migrants [33].

The connections between population and other geographical features has also been investigated by E. A. Wrigley, who has worked on a much broader canvas in his examination of population changes in western Europe in the second half of the nineteenth century [34]. This study analyses the relationship between coal resources, economic growth and population change in the coalfield areas of Belgium, Northern France and Western Germany between 1850 and 1914, and compares the demographic experiences of a number of economically similar regions

in France and Germany. These comparisons suggest that differences in national character were of relatively little importance in explaining the differences in population growth in these two countries and that similarity of economic background was much more important than similarity of nationality in determining demographic trends.

At another extreme, perhaps, are the very detailed analyses of British census returns which are being made possible by the gradual release of the manuscript documents on which the printed census reports were based. This information is proving particularly useful for solving detailed problems in local sample areas. A good example of the use of these enumerators' returns is June Sheppard's careful study of the agricultural labour force in part of eastern Yorkshire in the mid-nineteenth century [35]. Using this source she has been able to demonstrate the varying importance of family workers, farm servants and agricultural labourers in different parts of the countryside, many of whom were shown to travel considerable distance to work. The returns are also most useful in urban areas. In particular, R. Lawton has studied Liverpool using this information, in order to shed light on the origins of the urban population [36].

A second group of historical studies have been primarily concerned with the evolution of contemporary conditions. Again these investigations have been undertaken at a number of scales in various parts of the British Isles.

In Scotland R. H. Osborne has examined population movements between 1851 and 1951, while K. Walton has focussed his attention simply on the north-east [37]. At an even more detailed level J. R. Coull has compared the contrasting demographic development of the Faroe and Shetland Islands; and the same author has also made a similar study of the Orkneys [38]. Work on the development of the population of modern Ireland shows similar variations in the level of investigation. J. H. Johnson has traced the population changes in the country as a whole from 1841 to 1961; in much greater detail R. A. Gailey has examined population change on the Aran Islands; and R. H. Buchanan has concentrated on an Irish townland of only 807 acres [39]. Even an area as small as this townland contains internal variations in the success of different families in maintaining their family farms, owing to the sheer inability of some of the inhabitants to reproduce themselves.

In England the same variation in the scale of work can be seen. R. H. Osborne has undertaken the laborious task of analysing migration during the twentieth century; C. D. Morley has restricted his attention to the population of Northamptonshire (a county which has been dominated by the development of one town and by a group of smaller settlements); and R. Lawton has made an analysis of Lancashire and Cheshire at about the same scale [40]. In still greater detail (and again using the manuscript census enumerators' returns, as well as personal investigation of the contemporary situation), June Sheppard has examined the population changes in three small sample areas of Yorkshire since 1851 and has

shown the intricate nature of the changes associated with rural depopulation [41]. A similar theme has been investigated by W. M. Williams in the different social and economic context provided by a Devon parish in south-west England, where the occupation and ownership of land has been characterized by a large number of complex changes. Here the process of rural depopulation was only one element in a number of complex demographic changes, with short-distance migration within the local area being an important factor in the alterations which took place [42].

This variety of scale of approach is probably a strength rather than a weakness in these studies. It is clearly necessary to uncover the evolution of the national and regional situation; and local studies are also required to provide an important corrective to over-facile generalization and explanation. Yet it seems a valid comment that up to the present British geographers have not evolved completely convincing methods of choosing the appropriate scale at which to examine differing problems in population geography.

STUDIES OF CONTEMPORARY POPULATION PROBLEMS WITHIN THE BRITISH ISLES

In the study of the contemporary geography of the British Isles, much of the effort of geographers has been aimed at the processing of the most recent census statistics of Great Britain and Ireland. In particular, population distribution maps have been prepared for Scotland [43], Ireland [44], and England and Wales [45]. Only in this last case was the map the result of the efforts of a government agency, but one in which the expertise of geographers has been influential. The Population Studies Group of the Institute of British Geographers has provided an important catalyst for work of this kind, particularly with its scheme for a cooperative analysis of information from the 1951 and 1961 censuses. Some of the national maps which were produced by this work have now been published by the Institute; more detailed regional maps may be used in personal research projects and some have already been published [46].

Studies of population distributions have been supplemented by various analyses of population change in the decade between the last two official censuses. M. F. Tanner has studied this topic in England and Wales [47], H. A. Moisleley in Scotland [48], J. H. Johnson in Ireland [49], and interesting cartographic experiments in showing information of this kind have been undertaken by geographers associated with the Ministry of Housing and Local Government [50].

These studies of recent changes have been backed by analyses which push back earlier into the twentieth century and which attempt to assess the relative importance of permanent migration and natural changes in altering the population map. The sequence of population changes which occurred in Britain before and

after the Second World War is most succinctly described by E. C. Willatts and M. G. C. Newson, and the relative importance of natural changes for part of this period has been ingeniously analysed by J. W. Webb [51]. The analysis of this topic has been taken through to the most recent census by R. H. Osborne, although measured in terms of the major regions of the country, rather than the smallest areas possible [52]. Yet the detail of this study is quite sufficient to reveal that the interregional movement of population in Britain shows remarkably consistent trends, which are unlikely to be greatly altered until official planning reshapes Britain's economic pattern.

A number of these studies of population distribution and change have been applied to specific planning problems. The most noteworthy example of this work, undertaken by academic geographers rather than government officials, is the series of studies (sponsored by the Ministry of Labour) which are being undertaken on migration and mobility in the north-east of England by J. W. House and E. M. Knight. Three reports have been published so far: one is concerned with population changes in rural areas between 1951 and 1961 in relation to general level and nature of employment in these areas; a second considers the age, sex and other characteristics of migrants in the north-east of England during the same period; and a third examines the links between population and employment in one sample employment exchange area, chosen because of the grave problem it exemplifies [53]. G. D. Symes has undertaken more restricted work of a similar kind with reference to western Ireland, where the problem of rural depopulation is complicated by the desire of the Irish government to preserve the Gaelic language [54].

The population problems of conurbations as well as of more rural area have received study by geographers. One population movement of considerable importance is the flow of immigrants to Britain from outside the British Isles. This topic has been explored by G. C. K. Peach, who has analysed the factors affecting the distribution of West Indians in Great Britain [55]. The immigrants, he points out, show a tendency to congregate in urban areas with a declining population; but, in fact, these are the inner areas of conurbations, which are experiencing population decline because of the redevelopment of decaying inner areas and the outward spread of low density dwellings beyond the statistical limits of the inner cities. An example of this outward spread of population has been studied by D. J. M. Hooson in his examination of population changes in Hertfordshire in the twentieth century [56].

Not all important population movements recorded in the census are permanent. British geographers have become particularly interested in the analysis of the journey to work, which is becoming an ever-increasing element of everyday life in Britain. The extent of that increase has been shown by R. Lawton's analysis of the journey to work information in the 1921 and 1951 censuses, dates which were forced on him by the need to study comparable statistics [57]. Lawton has

shown that in the 1920s many town-dwellers in England and Wales had considerable journeys to work, but that, measured in terms of movements over local authority boundaries, rural areas were often relatively self-contained. In 1921 there was a considerable number of local authority areas where the journey to work was very important, but these areas were concentrated in and around the major urban centres. By 1951 areas with the same intensity of movements covered a broad axial belt stretching from Lancashire to the south coast of England and also covering large sections of south Wales and north-east England.

This broad description has been elaborated by more local studies. For example, G. Humphrys has examined the journey to work in industrial south Wales, where there has been an increasing divorce of residential communities from the employment which supports them and a related expansion in the areas from which individual employment centres draw their labour; and J. C. Dewdney has undertaken similar work in county Durham [58]. R. E. Pahl has explored part of the fringe of metropolitan London, which is only partly connected with the central city, has a complex pattern of journeys to work and perhaps might be considered "a city in its own right" [59]. In even more detail R. T. Dalton and W.S.G. Thomas have examined the journey to work in Lincoln, and have explored the tendency for particular means of transport to dominate in particular residential areas; and H. B. Rodgers has studied the special case of an overspill community [60]. The movement has also been studied in a rural context by G. J. Lewis, who has examined its influence in Cardiganshire, a sparsely populated area suffering from severe rural depopulation [61].

The published census returns give only a limited and coarse-grained view of the population of contemporary Britain and recently attempts have been made to resolve this difficulty. One line of approach is the use of unpublished returns for enumeration districts, which are the areas surveyed by one census enumerator. In recent years this information has been made more easily available (a trend which it is hoped will be continued with the 1971 census); and this detailed information has been applied with particular success in urban research.

A pioneer analysis by a British geographer using this information is E. Jones' study of Belfast, where he has explored not merely the facts of population distribution, but also such features as the detailed demographic characteristics and religious affiliations of the population, a topic of vital importance in that city [62]. Elsewhere Jones has stressed the general importance of considering detailed social features as well as total numbers in studying the geography of population [63]. The analysis of enumeration districts plays an important part in the recent *Atlas of London and the London Region*, which has been edited by Professor Jones and D. J. Sinclair. The first of the two volumes which are to form this atlas has now been published and shows the social complexity exhibited by the population of much of London; and this greater detail "not only confirms the

generally known distributions but draws attention to a wealth of starting points for new generalizations" [64].

Another source which has been exploited in an attempt to find greater detail is the electoral registers, which are compiled annually. Unfortunately these list only people over 21 years old, so that they provide a source that must be used with care; but their detail and frequent publication adds to their value. This source has been used in the study of both rural and urban areas. H. R. Jones has used the registers in a study of rural migration from central Wales (which also gathered information from school admission registers and personal interviews) [65]. R. J. Johnston has used them in a study of population change in Nidderdale in Yorkshire, as has also G. C. Dickinson [66]. The value of this kind of analysis, which is particularly laborious, is the wealth of data it reveals about local as well as long-distance movements of population. The electoral registers have also been used in urban areas, particularly for the task of preparing urban population distribution maps. A. Hunt and H. A. Moisley have prepared maps of Sheffield and Glasgow using this source [67], although now that enumeration district information is becoming available this method is less necessary than it was before.

The general movements taking place within Britain can largely be explained in economic terms. For example, the analysis by E.M. Rawstron and B. E. Coates of regional levels of affluence shows the manner in which income and employment opportunities vary considerably in different parts of the United Kingdom [68]. Maximum incomes were found in London and south-eastern England, with the minimum values being found in Northern Ireland and parts of Wales; and these income levels are related to the general direction of inter-regional population movements. But not all population movements are guided by purely economic motives. W.K.D. Davies, for example, has used British Health Service records to assess the regional preferences of doctors who are applying for new practices [69]. The status of different regions of the country in the eyes of doctors strongly influences the competition for available practices, which is scarcely related to economic gains. In this context the work of P. Gould and R. White is particularly interesting, with its attempt to analyse the attitudes to different parts of Great Britain held by children who are about to leave school [70].

THE INFLUENCE OF MODERN METHODS OF COMPUTING

The revolution in data handling techniques which is sweeping the world has been relatively late in affecting the work of British geographers, but it is now having an increasing influence on the work being undertaken, not least in the field of population geography. This new line of research cannot be described as a major

theme of work during the past decade, but it will be surprising if it does not form one in the years ahead.

Some of the applications of computers are simply a logical extension of work which has previously been undertaken by conventional methods, but with more elaborate statistical techniques being made possible by labour-saving methods of computation. For example, B. T. Robson has studied the socio-economic groupings found in the population of Sunderland [71]. For this he used the technique of component analysis, which is designed to select objectively groups of variables which tend to vary independently of one another. A similar method is used by Janet B. Henshall in her analysis of the role of demographic factors in shaping the agriculture of Barbados [72]. Here the technique of factor analysis was used on 32 variables selected for study after the collection of data from 116 random farms.

Other research seems to be moving towards new directions as a result of the more efficient methods of handling data; and although it is too early yet to know whether these new techniques will bring revolutionary findings, a growing number of British geographers are experimenting with these methods. In the field of population geography, for example, D. Thompson has used multivariate analysis in his study of spatial variation in population growth in the United States, while P. A. Compton has applied a stochastic model to the analysis of migration in Hungary [73].

A further piece of work, which depends on the ability of the computer to handle vast quantities of data rather than to undertake sophisticated calculations, is the analysis of migration in prehistoric Polynesia being undertaken by R. G. Ward and J. W. Webb [74]. The basic problem here is whether these voyages could have been accidental; and this problem is being tackled by simulating a large number of random voyages on a computer. These voyages are started from a number of critical locations and are then allowed to proceed by chance, but influenced by such factors as the distribution of Pacific islands and the known incidence of winds and currents at different times of the year. It is hoped to compare these simulated "accidental" voyages with journeys which are known to have taken place. The method will not provide a categorical answer to this interesting problem in population migration, but it will establish the statistical likelihood of whether or not the voyages took place by accident or according to some pre-arranged plan.

CONCLUSION

The diversity of the work which has been described in this paper makes it difficult to claim that British research in population geography forms a compact, well-integrated body of investigation united by a single methodological aim. This can be said in spite of the fact that at least two British geographers have

echoed Trewartha's plea that the study of population should hold a central position in human geography [75]. D. J. M. Hooson, for example, has claimed that "the distribution of population acts as a master thread, capable of weaving into a coherent pattern the otherwise disparate threads of the subject and expressing its philosophical unity, particularly in the context of regional geography" [76]. E. A. Wrigley has taken what appears to be a less extreme position, but he too sees population as a convenient nexus for geographical research "into which all strands can be seen to lead" [77].

Despite these two eloquent statements of intent, it requires the eye of faith to see British research in population geography as the point of reference from which all other geographical elements should be observed. Rightly or wrongly, British population geographers have simply preferred to identify specific problems connected with population in regions which they know and then to study them from the various points of view which interest them individually. Yet, even with this empirical approach (and with a concern, more often than not, for practical ends rather than the advance of the methodology of academic geography) most of these population studies at least seek to fit population distributions and changes into the social, economic and physical environments of various periods and places. Hence, in this sense at least, these studies have a geographical flavour which serves to distinguish them from the parallel work of economists and demographers.

University College, London

BIBLIOGRAPHY

- [1] Hunter, J. M., An exercise in applied geography: geographical planning in urban areas for the 1960 census of Ghana, *Geography*, 46 (1961), 1—8.
- [2] Barbour, K. M., Maps of Africa before Surveys: the Location of Omodias Map of Sudan, *Geogr. Review*, 51 (1961), 71—86.
- [3] Barbour, K. M., maps 1 and 2, Appendix I, in K. M. Krotki, *21 Facts about the Sudanese*, Khartoum 1958; Prothero, R. M., *Northern Region, Nigeria, 1 : 1 m, Distribution of Population (1952 Census)*, Directorate of Overseas Surveys, Miscellaneous maps no. 237, London 1960; Jennings, J. H., *Nigeria Eastern Region Population (1953 Census)*, 1 : 500,000, Federal Survey Department, Lagos 1957; Morgan, W. T. W., *Population distribution map of Kenya, 1962*, Department of Geography, Royal College Nairobi, 1963; Hilton, T. E., *Ghana Population Atlas*, Edinburgh 1960; Prescott, J. R. V., Population distribution in Southern Rhodesia, *Geogr. Review*, 52 (1962), 559—565, map on p. 561; Ward, R. G., The population of Fiji, *Geogr. Review*, 49 (1959), 322—341, maps on pp. 326—327; Niddrie, D. L., The problems of population growth in Puerto Rico, *J. Trop. Geogr.*, 20 (1965), 26—33, map on p. 26.
- [4] See, for example, Prothero, R. M., Population maps and mapping in Africa south of the Sahara, and Barbour, K. M., Population mapping in Sudan, in Barbour and Prothero (eds.), *Essays on African population*, London 1961, 63—82 and 99—114.

- [5] The idea of Standard Nutritional Units was advanced in Stamp, L. D., The measurement of land resources, *Geogr. Review*, 48 (1958), 1—15.
- [6] Prescott, J. R. V., Overpopulation and overstocking in the native areas of Matabeleland, *Geogr. Journ.* 127 (1961), 212—225.
- [7] Grove, A. T., Population densities and agriculture in Northern Nigeria in Barbour, K. M. and Prothero, R. M. (eds.), *Essays on African population*, London 1961, 15—136.
- [8] Hunter, J. M., Ascertaining population carrying capacity under traditional systems of agriculture in developing countries, *Prof. Geogr.*, 18 (1966), 151—154. See also Hunter, Population pressure in a part of the West African savanna, *Ann. Ass. Am. Geogr.*, 57 (1967), 101—114.
- [9] Hunter, J. M., Seasonal hunger in part of the West African savanna: a survey of the body-weights in Nangoni, north-east Ghana, *Trans. Inst. Br. Geogr.*, 41, 1967, 167—185.
- [10] Barbour, K. M., Population shifts and changes in Sudan since 1898, *Middle Eastern Studies*, 2 (1966), 98—122. See also Barbour's analysis of ethnographic distribution in: North and south in Sudan, a study of human contrasts, *Ann. Ass. Am. Geogr.* 54 (1964), 209—226.
- [11] Hilton, T. E., *Ghana Population Atlas*, Edinburgh 1960, figs. 9, 10, p. 14; Hilton, Depopulation and population movements in the Upper Region of Ghana, *Bull. Ghana Geogr. Ass.*, 11 (1966), 27—47.
- [12] Hunter, J. M., Regional patterns of population growth in Ghana, 1948—1960, in Whittow, J. B. and Wood, P. D. (eds.), *Essays in geography for Austin Miller*, Reading 1966, 172—190.
- [13] Niddrie, N. L., (1965), op. cit.; Niddrie, *Land use and population in Tobago, an environmental study*, World Land Use Survey, Regional Monograph 3, Bude 1961.
- [14] Brookfield, H. C., Mauritius: demographic upsurge and prospect, *Popul. Stud.*, 11 (1957), 102—122; Brookfield, H. C., Pluralism and geography in Mauritius, *Geogr. Stud.* 5 (1958), 3—19; Brookfield, H. C., Population distribution in Mauritius: an enquiry into pre-determinants of distribution in a tropical sugar land, *Journ. Trop. Geogr.*, 13 (1959), 1—22.
- [15] Prothero, R. M., Continuity and change in African population mobility, in: Steel, R. W. and Prothero, R. M., (eds.), *Geographers and the Tropics : Liverpool essays*. London 1964, 189—234. A recent additional contribution is Harrison, R. S., *Migration as a factor in the geography of Western Libya*, Unpublished Ph. D. thesis, University of Cambridge, 1966; see also Harrison, Migrants in the City of Tripoli, Libya, *Geogr. Review*, 57 (1967), 397—421.
- [16] Prothero, R. M., *Migrants and malaria*, London 1965; see also: Prothero, R. M., A geographer with the World Health Organization, *Geogr. Journ.*, 128 (1962), 477—493.
- [17] Davis, S. G., The rural-urban migration in Hong Kong and its new territories, *Geogr. Journ.*, 128 (1962), 328—333; Vaughan, T. D. and Dwyer, D. J., Some aspects of post-war population growth in Hong Kong, *Econ. Geogr.*, 42 (1966), 37—51.
- [18] Dwyer, D. J., The problem of in-migration and squatter settlement in Asian cities: the two case studies, Manila and Victoria-Kowloon, *Asian Studies*, 2 (1964), 145—169.
- [19] Clarke, C. G., Population pressure in Kingston, Jamaica, a study of unemployment and overcrowding, *Trans. Inst. Br. Geogr.*, 38, 1966, 165—182.
- [20] Darby, H. C., (ed.), *The Domesday geography of England*, 6 vols., Cambridge 1952. Five volumes have now appeared in a number of which Professor Darby has been joined by co-editors. A gazetteer is also being produced to accompany the series.
- [21] Wrigley, E. A., Family reconstitution, in Wrigley, E. A., (Ed.), *An introduction to English historical demography from the sixteenth to the nineteenth century*, London 1966, 96—169.
- [22] Wrigley, E. A., Family limitation in pre-industrial England, *Econ. Hist. Rev.*, (second series), 19 (1966), 82—109.

- [23] Butlin, R. A., The population of Dublin in the late 17th century, *Irish Geogr.*, 5, (2) (1965), 51—66.
- [24] Mills, D. R., The Poor Laws and the distribution of population c. 1600—1800, with special reference to Lincolnshire, *Trans. Inst. Br. Geogr.*, 27, 1959, 185—195.
- [25] McIntosh, N. A., Changing population distribution in the Cart basin in the eighteenth and early nineteenth centuries, *Trans. Inst. Br. Geogr.*, 22, 1956, 140—157.
- [26] Gailey, R. A., Settlement and population in Kintyre, 1750—1800, *Scot. Geogr. Mag.*, 76 (1960), 99—107.
- [27] Lawton, R., Population movements in the West Midlands, 1841—1861, *Geography*, 43 (1958), 164—177; Lawton, R., Irish immigration to England and Wales in the mid-nineteenth century, *Irish Geogr.*, 4 (1959), (1), 35—54.
- [28] Lawton, R., Population changes in England and Wales in the later nineteenth century, *Trans. Inst. Br. Geogr.*, 44, 1968, 55—74.
- [29] Law, C. M., The growth of urban population in England and Wales, 1801—1911, *Trans. Inst. Br. Geogr.*, 41, 1967, 125—143.
- [30] Cousens, S. H., The regional pattern of emigration during the Great Irish Famine, 1846—1851, *Trans. Inst. Br. Geogr.*, 28, 1960, 126—129; The regional variations in emigration from Ireland between 1821 and 1841, *Ibid.*, 37, 1965, 15—30; The regional variation in mortality during the Great Irish Famine, *Proc. R. Ir. Acad.*, 63 (1963), section C, 127—149; The restriction of population growth in pre-Famine Ireland, *Ibid.* 64 (1966), section C, 85—99; Emigration and demographic change in Ireland, 1851—1861, *Econ. Hist. Rev.*, (second series), 14 (1961—62), 275—288; The regional variations in population changes in Ireland, 1861—81, *Ibid.*, 17 (1964—65), 301—321.
- [31] Johnson, J. H., The population of Londonderry during the Great Irish Famine, *Econ. Hist. Rev.* (second series), 10, (1957—58), 273—285; Marriage and fertility in nineteenth century Londonderry, *Journ. Stat. and Social Enquiry Society of Ireland*, 20 (1957—58), 99—117; Population movements in pre-Famine Derry, *Proc. R. Ir. Acad.*, 60 (1959), section C, 141—162.
- [32] Proudfoot, V. B. and Vaughan, T. D., Changes in settlement and population in Northern Ireland, 1835—1860, *Ulster Folklife*, 5(1959), 20—26; Johnson, J. H., The disappearance of clachans from County Derry in the nineteenth century, *Irish Geogr.*, 4 (1963), (6), 404—414.
- [33] Johnson, J. H., Harvest migration from nineteenth-century Ireland, *Trans. Inst. Br. Geogr.*, 41, 1967, 97—112.
- [34] Wrigley, E. A., *Industrial growth and population change*, Cambridge 1961.
- [35] Sheppard, J. A., East Yorkshire's labour force in the mid-nineteenth century, *Agric. Hist. Rev.*, 9 (1961), 43—54.
- [36] Lawton, R., The population of Liverpool in the mid-nineteenth century, *Trans. Hist. Soc. of Lancs. and Cheshire*, 107 (1955), 89—120.
- [37] Osborne, R. H., The movements of people in Scotland, 1851—1951, *Scottish Studies*, 2 (1958), 1—46; Walton, K., Population changes in the north-east of Scotland, 1696—1951, *Scottish Studies*, 5 (1961), 149—180.
- [38] Coull, J. R., A comparison of demographic trends in the Faroe and Shetland Islands, *Trans. Inst. Br. Geogr.*, 41, 1967, 159—166; Coull, J. R., Population trends and structures on the island of Westray, Orkney, *Scottish Studies*, 10 (1966), 69—77.
- [39] Johnson, J. H., Der Bevölkerungswandel Irlands im 19 und 20 Jahrhundert, *Geogr. Rundsch.*, 16 (1964), 221—230; Gailey, R. A., Settlement and population in the Aran islands, *Irish Geogr.*, 4 (1959), (1), 65—78; Buchanan, R. H., Rural change in an Irish townland, 1890—1955, *Adm. Sci., Lond.*, 14, no. 56 (1958), 291—300.
- [40] Osborne, R. H., Migration trends in England and Wales, 1901—1951, Problems of Applied

- Geography II, *Geogr. Polon.*, 3, 1964, 137—162; Morley, C. D., Population of Northampton and the Ise Valley, 1801—1951, *E. Midld. Geogr.*, 2 (1959), (11), 20—29; Lawton, R., Population trends in Lancashire and Cheshire from 1801, *Trans. Hist. Soc. of Lancs. and Cheshire*, 114 (1962), 189—213.
- [41] Sheppard, J. A., Rural population changes since 1851: three sample studies, *Sociological Rev.* (New Series), 10 (1962), 81-95.
 - [42] Williams, W. M., The social study of family farming, *Geogr. Journ.*, 129 (1963), 62—74; and for full discussion see Williams, *A West-Country village : Ashworthy*, London 1963.
 - [43] Caird, J. B. and Diamond, D. R., (eds.), *Scotland, Distribution of population, 1961*, Glasgow 1965, map of the scale of 1 : 500,000.
 - [44] Map of the distribution of population in Ireland, 1961, scale 1 : 1,250,000, issued with *Geogr. Rundsch.* 16 (1964). See also more detailed regional maps of Irish population in Freeman, T. W., *Ireland: a general and regional geography*, London 1965, (3rd edition).
 - [45] Ministry of Housing and Local Government, scale 1 : 625,000, 2 sheets, forthcoming.
 - [46] Osborne, R. H., *Atlas of population change in the East Midland counties*, University of Nottingham, Dept. of Geography, 1966; Institute of British Geographers, Special Number on Population Maps of the British Isles, *Trans. Inst. Br. Geogr.*, 43 (1968).
 - [47] Tanner, M. F., Population changes in England and Wales, 1965—61, *Geography*, 46 (1961), 357—360; see also; Osborne, R. H., Population changes in England and Wales, 1951—61, *E. Midl. Geogr.*, 2 (1961), (15), 41—50. There have been other detailed studies in post-graduate theses, for example, Anita Sparks, *Changes in the distribution of population in the Manchester conurbation, 1951—1961*, unpublished M. A. thesis, Univ. of Manchester, 1965.
 - [48] Moisley, H. A., Population changes and the Highland problem, 1951—61, *Scottish Studies*, 6 (1962), 194—200; see also: Moisley, H. A., The deserted Hebrides, *Scottish Studies*, 10 (1966), 44—68.
 - [49] Johnson, J. H., Population changes in Ireland, 1951—1961, *Geogr. Journ.*, 129 (1963), 167—174.
 - [50] Ministry of Housing and Local Government, *Great Britain, Population Changes, 1951—61*, 2 sheets, scale 1 : 625,000.
 - [51] Willatts, E. C. and Newson, M. G. C., The geographical pattern of population changes in England and Wales, 1921—1951, *Geogr. Journ.*, 119 (1953), 431—50; Webb, J. W., The natural and migrational components of population changes in England and Wales, 1921—31, *Econ. Geogr.*, 39 (1963), 130—48.
 - [52] Osborne, R. H., Changes in the regional distribution of population in Great Britain, *Geography*, 49 (1964), 265—273.
 - [53] House, J. W., *Rural north-east England, 1951—1961*, Papers on migration and mobility in North-East England, No. 1, Dept. of Geography, University of Newcastle-upon-Tyne, 1965; House, J. W. and Knight, E. M., *Migrants of north-east England*, Papers on migration and mobility in North-East England, no. 2, 1965; House, J. W. and Knight, E. M., *People on the move*, Papers on migration and mobility in North-East England, no. 3, 1966.
 - [54] Symes, D. G., *The population resources of the Connemara Gaeltacht*, Dept. of Geography, University of Hull, Miscellaneous Series, 1, 1965; *The population resources of the Gaeltacht, co. Donegal*, Miscellaneous Series, 2, 1965.
 - [55] Peach, G. C. K., Factors affecting the distribution of West Indians in Great Britain, *Trans. Inst. Br. Geogr.*, 38, 1966, 151—163.
 - [56] Hooson, D. J. M., The recent growth of population and industry in Hertfordshire, *Trans. Inst. Br. Geogr.*, 25, 1958, 197—208.
 - [57] Lawton, R., The journey to work in England and Wales: forty years of change, *Tijdschr. Econ. Soc. Geogr.*, 54 (1963), 61—69. See also: Lawton, R., The daily journey to work

- in England and Wales, *Town Planning Review*, 29 (1959), 241—257. Lawton, R., The journey to work in Britain: some trends and problems, *Regional Studies*, 2 (1968), 27—40.
- [58] Humphrys, G., The journey to work in industrial south Wales, *Trans. Inst. Br. Geogr.*, 36 (1965), 85—96; Dewdney, J. C., The journey to work in county Durham, *Town Planning Review*, 31 (1960), 107—124.
- [59] Pahl, R. E., *Urbs in Rure: the metropolitan fringe in Hertfordshire*, London School of Economics, Geogr. Papers, 2, 1965.
- [60] Dalton, R. T. and Thomas, W. S. G., The journey to work in Lincoln, *E. Midld. Geogr.*, 4 (1966), 46—51;
- [61] Rogers, H. B., Employment and journey to work in an overspill community, *Sociological Review*, 7 (1959), 213—229.
- [62] Jones, E., *A social geography of Belfast*, London 1961.
- [63] Jones, E., Sociological aspects of population mapping in urban areas, *Geography*, 46 (1961), 9—17.
- [64] Jones, E., The London Atlas, *Geogr. Journ.*, 131 (1965), 330—343; quotation on p. 340; Jones, E., and D. J. Sinclair (eds.) *Atlas of London and the London Region* (Oxford, 1968).
- [65] Jones, H. R., A study of rural migration in central Wales, *Trans. Inst. Br. Geogr.*, 37, 1965, 31—45.
- [66] Johnston, R. J., A reconnaissance study of population change in Nidderdale, 1951—61, *Trans. Inst. Br. Geogr.*, 41, 1967, 113—124; Dickinson, G. C., The nature of rural population movement — an analysis of seven Yorkshire parishes based on electoral returns from 1931—1954, *Yorkshire Bull. of Econ. and Social Research*, 10 (1958), 95—108.
- [67] Hunt, A. and Moisley, H. A., Population mapping in urban areas, *Geography*, 45 (1960), 78—89. See also Hunt, Urban population maps, *Town Plann. Rev.*, 23 (1952), 238—248.
- [68] Rawstron, E. M. and Coates, B. E., Opportunity and affluence, *Geography*, 51 (1966), 1—15.
- [69] Davies, W. K. D., Latent migration potential and space preferences, *Prof. Geogr.*, 18 (1966), 300—304.
- [70] Gould P. R. and White, R. R., The mental maps of British school leavers, *Regional Studies*, 2, 1968, 161—82. For a discussion on the complex method used in assessing preferences see Gould, *On mental maps*, Discussion Paper 9, Michigan Inter-University Community of Mathematical Geographers, Univ. of Michigan, Sept. 1966.
- [71] Robson, B. T., *The human ecology of Sunderland: a study of social structure and attitudes towards education*, (Unpublished Ph. D. thesis, University of Cambridge, 1966). See also Robson, An ecological analysis of the evolution of residential areas in Sunderland, *Urban Studies*, 3 (1966), 120—142.
- [72] Henshall, J. B., The demographic factor in the structure of agriculture in Barbados, *Trans. Inst. Br. Geogr.*, 38, 1966, 183—195.
- [73] Thompson, D., *A multivariate approach to population growth in the U.S.A., 1950—60*, and Compton, P. A., *The interpretation of some migration streams in Hungary by the use of a stochastic model*, Mimeogr. papers read to the Institute of British Geographers, Sheffield 1967. See also Compton, A régiók közötti vándorlás vizsgálatára matrix-módszerrel, *Demográfia*, 4 (1966), 475—98.
- [74] Ward, R. G., *The simulation of pre-historic voyages in the Pacific*, Unpublished paper read to the Institute of British Geographers, Sheffield 1967.
- [75] Trewartha, G. T., The case for population geography, *Ann. Ass. Am. Geogr.*, 43 (1953), 71—97.
- [76] Hooson, D. J. M., The distribution of population as the essential geographical expression, *Can. Geogr.*, 17 (1960), 10—20.
- [77] Wrigley, E. A., Geography and population, in: Chorley, R. J. and Haggett, P., (eds.), *Frontiers in geographical teaching*, London 1965, 62—80, ref. on p. 69.

THE INTERNAL MIGRATION OF POPULATION IN POLAND, 1961—1965 *

LESZEK KOSIŃSKI

INTRODUCTION

Previous study of the components of population growth in Poland in 1951—1960 showed that there were in this country four types of population change¹. In the map of types an interesting spatial pattern emerged. Many towns and industrial poviats were characterised by an excess of migration gain over natural increase. They were surrounded by areas where natural increase exceeded immigration, and these were encircled by areas where natural increase exceeded migration loss. In only a very few poviats was population decline the result of an excess of migration loss over natural increase.

The above-mentioned study became the point of departure for the present work, in which population changes in the succeeding five years are analysed. Basic reference units are again 396 poviats. Using the method suggested by J. W. Webb², the relationship between the natural and migration components of growth was determined and the types of changes defined.

Population data furnished by the census of December 6, 1960 were published for individual poviats, but part of the population (370,000) was centrally registered. However, in the official estimate of the Central Statistical Office of December 31, 1965 all the population was divided into poviats [3]. In order to make the data comparable it was necessary to apportion the centrally registered population of 1960 among the poviats. This was done on the assumption that the allocation should be proportional to the census population of the poviats. Another assumption was that since the census was taken so near to the end of the year the data for December, 1960 could be considered as representing the situation at the beginning of the 5-year period 1961—5. By comparing the data for 1960 and 1965 it was possible to establish the total population change in the poviats over

* Statistical work was done by Mgr A. Żurek and maps and diagrams were drawn by Mgr A. Gawryszewski. Their contributions are gratefully acknowledged.

¹ The results were summarised by the author (L. Kosiński, Studies on the structure of changes, potential and concentration of population of Poland in 1950—1960, *Geogr. Polonica* 7, 1965, 81—94).

² J. W. Webb, The natural and migrational components of population changes in England and Wales, 1921—1931, *Econ. Geogr.*, 39 (1963), 2, 130—148.

five years. By comparing total change with natural increase [5, 6] the balance of migration was established. The average annual natural increase and migration balance were compared with the average population during the five-year period and ratios per 1000 inhabitants were computed.

The data were used for constructing a classification diagram in which all the poviats are represented by points on a Cartesian co-ordinate graph (Fig. 1). Their location depends on the relationship of natural increase (+y) or decrease (-y) to migration gain (+x) or loss (-x). The diagram is a basis for the classification of poviats into types of population change. In this type of approach migration change and natural change are considered as independent variables, although their interdependence in the long run is beyond any doubt.

The aim of the study, however, was not only to analyse population change but also to attempt an explanation of the reasons for migration. In particular, two factors for which the statistical data are available were taken into account — capital expenditure and industrial employment. Data on capital expenditure in 1961—1965 have recently been published [2], thus making it possible for the first time to correlate migration and new investment. However, there is one basic question which has to be answered in such a study. To what extent does migration reflect present or past investment? It is obvious that the full impact of a new project can be felt only after several years. Unfortunately, lack of statistical data prevented us from interpreting migration in 1961—1965 with reference to former investment outlays. On the other hand, a period of five years was long enough for some projects, even if they began during these years, to have some impact on migration, unless they were begun towards the end of the period. When the study was almost finished, a new publication of the Central Statistical Office appeared. This contained additional data by poviats which had previously been unavailable [4]. Only the data on industrial employment were taken into account in this study.

The possibility of employing other information, such as the nature of the infrastructure (e. g. level of education, housing conditions) and per capita income, was also considered. However, the former has probably little value in Polish conditions and information on income is not available.

POPULATION GROWTH IN 1961—1965

In the years 1961—1965 the population of Poland increased by 1,658,000, from 29,893,000 to 31,551,000. Average annual growth was 1.2% but this rate of growth was less than in previous periods. Average annual growth in the years 1951—1955 had amounted to 1.9% and from 1956 to 1960 to 1.6%. Nevertheless, the rate of population growth in Poland was one of the highest in Europe, much higher than in adjacent Czechoslovakia (0.7% in 1961—1965) and the German Democratic Republic (-0.2%). The decline of the rate of population growth

in Poland was mainly due to a fall in natural increase, which in turn was influenced by a decline in the birth rate from ca. 31 per thousand in 1950—1951 to 22.6 in 1960 and 17.4 in 1965. Mortality rates were all the time rather low. They decreased from 11.6 per thousand in 1950 to 7.4 in 1965. A decrease of fertility was caused partly by structural changes (the smaller cohorts born during the Second World War were now entering the reproductive age-group) and partly by the introduction of methods of birth control, reflected in the decline of the age-specific fertility rates. In 1950 there were 10.9 births per 100 women aged 15—49, while in 1964 the rate was only 7.5. The decline of this rate in the towns, from 9.9 to 6.0, was even more spectacular.

In contrast to experience in the previous decade, international migration was not very significant. The number of immigrants in the years 1961—1965 amounted to 1,489 and the number of emigrants to 120,751. The net migration loss was thus equal to 119,262 persons, or 7% of total population growth during that period.

On the other hand internal migration played an important role. According to the registration data giving changes of address, the number of in-migrants into towns and cities amounted to 2,467,000 in the years 1961—1965. At the same time the number of out-migrants was 1,890,000, thus leaving a gain from migration of 577,000. Since the urban population increased by 1,280,000, from 14,401,000 to 15,681,000, in-migration represented 43.5% of the total growth. The rest of the growth consisted of natural increase, since administrative changes and external migration cancelled each other. It should be emphasised that migration flows were not limited to rural-urban migration, since there were also substantial movements between towns as well as between villages. The total number of persons who changed the locality of their residence reached 5,070,000 during the five years, the annual figures showing a slight decline from 1,180,000 in 1961 to 920,000 in 1965.

TYPES OF POPULATION CHANGE IN THE YEARS 1961—1965

The analysis of the relationship between natural change and migration indicates that all poviats in Poland can be classified into 5 types.

In the majority of poviats (98.7%), total population increased and only in five poviats (1.3%) was there depopulation. Most of the growing poviats (71%) were those in which growth was due to an excess of natural increase over migration loss (Type A). In this group, covering 87% of the area of the country and representing 60% of the national population, rural poviats predominated.

The second largest group consisted of poviats in which natural increase exceeded migration gain (Type B). These poviats, which were mostly urban, represented 9% of the area and 20% of the population in Poland.

TABLE 1. Types of population change in the years 1961—1965

	Total	Population growth by type area				Popu- lation decline
		A Natural increase exceeds out-mi- gration	B Natural increase exceeds in-mi- gration	C In-mi- gration exceeds natural increase	D In-mi- gration exceeds natural decrease	H Out-mi- gration exceeds natural increase
Administrative units						
City voivodships						
Number	5	—	1	4	—	—
%	100.0	—	20.0	80.0	—	—
Urban poviats						
Number	74	9	43	21	1	—
%	100.0	12.2	58.1	28.3	1.4	—
Rural poviats						
Number	317	272	33	7	—	5
%	100.0	85.8	10.4	2.2	—	1.6
Total number	396	281	77	32	1	5
%	100.0	70.9	19.4	8.1	0.3	1.3
Absolute population and surface area						
Population (thous)	31,023.5	18,657.7	6204.4	5708.4	23.7	429.3
%	100.0	60.1	20.0	18.4	0.1	1.4
Area (thous.km²)	311,730	271,323	28,983	6413	13	4998
%	100.0	86.9	9.3	2.1	0.1	1.6

The third group, consisting almost entirely of urban units, included those poviats in which migration gain exceeded natural increase (Type C). This group represented 2% of the area and 18% of the population of the country. Migration gain exceeded natural increase only in one case, and in five cases migration loss more than off-set natural increase.

On the analytical graph those points representing rural poviats are concentrated to the left and those representing urban ones to the right of the ordinate. Poviats situated within the industrial districts³ are also concentrated on the right-hand side of the graph. Hence, it is the economic character of an area which has a decisive influence upon migration. The range of migration rates is considerable. They vary from -15.5% (Kolno) to +40% (Lublin).

³ Industrial districts were those defined by M. Najgrakowski, whose classification was based on 1960 data (*Geografia Powszechna*, vol. III, Warszawa 1965, p. 592).

The amount of natural increase is shown by the distance of a point from the abscissa. The difference between the northern and western areas, settled after the war, and the remaining parts of the country persisted. Maximum values were found in the north (Gołdap 24.8%, Pisz 24.7%, Szczytno 23.2%) while the lowest values approximated to 5% in many Upper Silesian poviats and were as low as 3.1% in Łódź and 2.7% in Żyrardów. Only in one case (Cieszyn —0.7%) was there a natural decrease. Hence the range of variation in natural increase was more restricted than that found in migration.

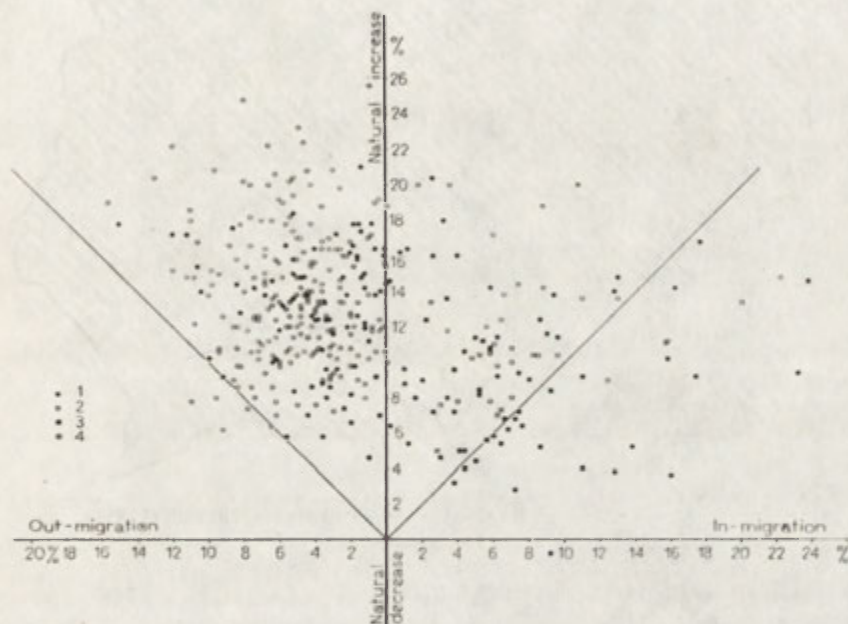
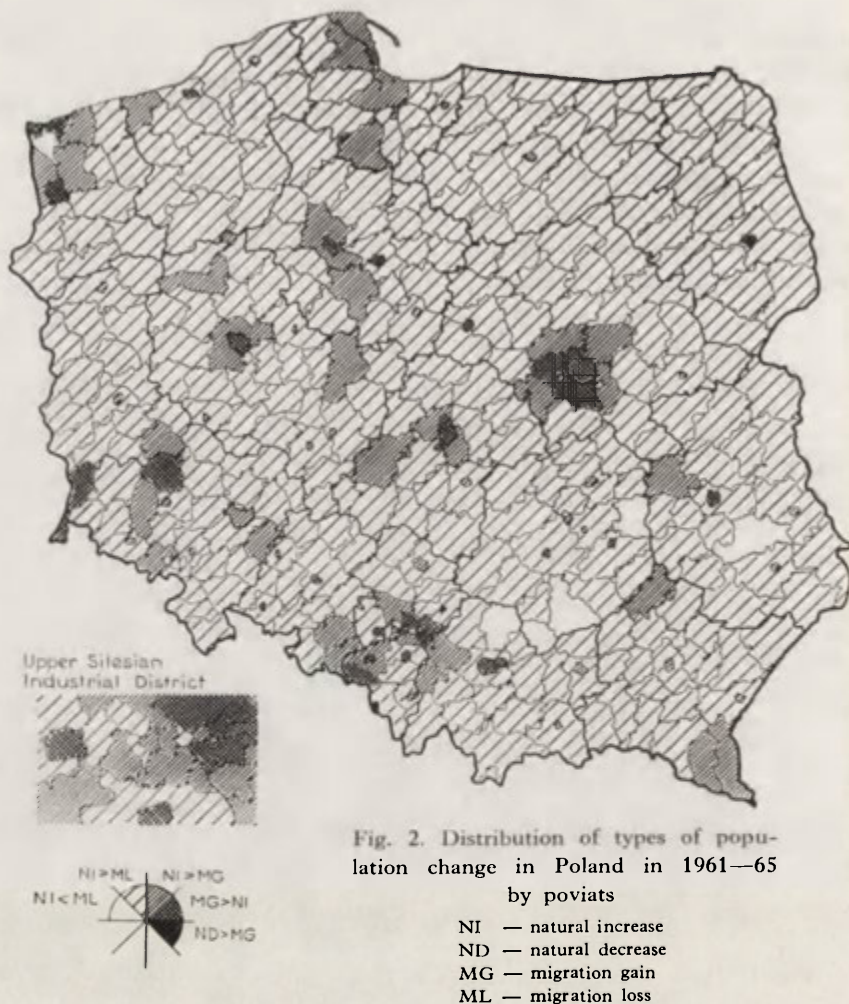


Fig. 1. Types of population change in Poland in the years 1961—65

1 — rural poviats situated in industrial districts, 2 — remaining rural poviats, 3 — urban poviats situated in industrial districts, 4 — remaining urban poviats

The map of types of population change shows the existence of a dynamic situation around the largest cities, which were surrounded by groups of poviats typified by in-migration (B and C). The biggest groups are to be found around Warsaw, Szczecin, Gdańsk, in Upper Silesia, and around Bydgoszcz and Poznań. Usually migration gains also dominate in the principal city encircled by the group, while in the outer poviats migration gains give way to natural increase as the dominant element. In the largest cities natural increase played a more important role than migration only in two recently re-settled cities — Gdańsk and Wrocław — and also in some of the settlements in the Upper Silesian Industrial District. In many cases, e.g. in Kraków, Lublin, Białystok, Opole, migration gains dominate only in the city and the out-migration zone (A) immediately adjoins. Migra-



tion gains were important, although only rarely dominant, in some other towns and industrial poviats. The very few poviats with declining populations were situated in south-eastern Poland.

If the results of the present classification of poviats are compared with those for the period 1951—1960 certain general conclusions can be drawn. There was an increase in the number of out-emigration (A) poviats (from 259 to 281) and of those where migration gain was dominant (C) (from 23 to 32). On the other hand, the number of poviats in which natural increase exceeded migration gain (B) decreased (from 107 to 77). To a certain extent, therefore, the more extreme types of change are replacing the intermediate ones. Previously the more important factor was natural increase; now the decisive role is taken by migration.

CAPITAL EXPENDITURE AND INDUSTRIAL EMPLOYMENT AS FACTORS EXPLAINING POPULATION GROWTH

As already mentioned, the aim of this study is to analyse the factors causing population change and, in particular, the contribution by migration. It was intended to carry out a correlation analysis, in which net migration and capital expenditure would be compared. Points representing poviats were distributed within Cartesian co-ordinates, whereby net migration (ordinate) was plotted against per capita capital expenditure (abscissa). The results of an initial rough analysis were discouraging. Inspection of the scatter of points made it clear that, although higher expenditure was associated with high migration gains, thus indicating positive correlation, lower expenditure was also associated with high migration gains or losses, which therefore made pointless any correlation analysis for the

TABLE 2. Capital expenditure and industrial employment in different types of poviat

	Total	Types of poviat				Popu- lation decline
		Population growth				
		A Natural increase exceeds out-mi- gration	B Natural increase exceeds in-mi- gration	C In-mi- gration exceeds natural increase	D In-mi- gration exceeds natural decrease	H Out-mi- gration exceeds natural increase
Average population in 1961—1965 Thousands %	31,023 100.0	18,658 60.1	6204 20.0	5708 18.4	24 0.1	429 1.4
Capital expenditure in 1961—1965 total Millions zł ^a %	611,326 ^b 100.0 ^b	261,658 42.8	146,202 23.9	183,276 30.0	409 0.1	3956 0.6
Capital expenditure in 1961—1965 Per capita thousands zł ^a	19.7	14.0	23.6	32.1	17.3	9.2
Industrial employment in 1965 Thousands %	3758 100.0	1316 35.1	1194 31.7	1230 32.7	11 0.3	7 0.2

^a Capital expenditure at current prices.

^b Including 15,825,000 zł (2.6%) of expenditure not distributed by areas (e. g. rolling stock in railways etc.).

whole group. Consequently only average values for particular types of poviats were computed.

Those areas with migration losses (A) have a smaller share of capital expenditure and industrial employment than they have of the national population (Table 2), which indicates that they are less industrialized and have a lower investment activity than other parts of the country. These features characterise poviats with a declining population (II) to an even greater extent. It should be emphasised, however, that for both types their percentage of capital expenditure is higher than their share of industrial employment.

In the poviats where natural increase exceeds migration gains (B) and, in particular, in those where migration gains dominate (C), the shares of both capital expenditure and industrial employment were higher than the share of population.

An analysis based on per capita capital expenditure emphasizes the differences among the types of poviats even more strongly. The highest average occurred in types C, followed by types B and D. Thus the top positions were occupied by in-migration areas, whereas the lowest values were registered in out-migration poviats and in particular in those where population declined (type H).

Rate of migration and investment outlay per head are shown in Figure 3.

Out-migration areas occur throughout the country. Those where investment outlay is extremely low are concentrated in the east and south-east. Areas with high investment rates occur on the fringe of the Upper Silesian Industrial District and around places with heavy investment (e.g. poviats Bolesławiec, Płock, Gorzów). It seems that the last-named group occurs as the result of the location of certain accompanying investments in the vicinity of centres attracting population. In some cases capital expenditure has been undertaken only in recent years and therefore its impact has not yet been felt. On the other hand, out-migration may have been terminated by new projects but the overall result for five years was still negative.

In-migration areas are less numerous. The poviats with high investment rates are concentrated around Warsaw, the Upper Silesian Industrial District and (in particular in its eastern part) in the valley of lower Vistula and at the mouth of the Odra. High investment figures are also to be found in large cities and, of course, in those poviats where new big industrial complexes are being developed (e.g. poviats Puławy, Tarnobrzeg, Konin, Lubin, Zgorzelec). On the other hand, those poviats in which migration gains are associated with a very low per capita capital expenditure are located in the vicinity of large cities (e.g. E of Warsaw or around Łódź) or in the Lower Silesian industrial areas (poviats Głogów, Złotoryja). Alternatively these conditions occur in the towns themselves (e.g. Legnica, Świdnica, Grudziądz, Siedlce). In some cases the concentration of population in the vicinity of large cities may have reflected the difficulties encountered by those people who wanted to settle in these cities but who were refused



Fig. 3. Capital expenditures (in thousands zł) and net migration in the years 1961—65

permission. Consequently they moved into the suburban zones, as long as there were no restrictions there.

Cartographic analysis thus helps to explain certain relationships, such as association of migration gains with low investment rates or migration losses with high investment rates. The present analysis can be considered as only a first step in research into the causes of migration and of different types of population change. There is a need for more detailed study of capital expenditure, in which not only

totals but also the structure of expenditure should be analysed. Clearly, different kinds of investments can have varying impacts upon migration.

In further studies the importance of other factors must also be taken into account. In particular, it is necessary to explore the impact of non-economic factors upon the pattern of migration. The pull of cities undoubtedly is based partly on their attraction as centres of social and cultural life. Nevertheless, in Poland a non-economic reason is never sufficient to prompt a decision to move. For example, there is practically no movement in Poland of retired people into pleasant, quiet, or inexpensive places. Employment is undoubtedly an important factor, but this includes not only employment in manufacturing but also in other sectors. In fact, manufacturing industry accounts for only about 40% of the recent employment increase and its share has remained stable. It is beyond the scope of the present discussion to attempt an analysis of the distribution of different migration factors. Why the new investments are located where they are or what accounts for the growth of labour requirements in different places both need a separate study.

Institute of Geography PAN, Warsaw

STATISTICAL SOURCES

- [1] *Ludność Polski w latach 1945—1965* (Population of Poland in the years 1945—1965), GUS, Studia i prace statystyczne, 1, Warszawa 1966.
- [2] *Nakłady inwestycyjne według powiatów 1961—1965* (Capital expenditures by poviats 1961—1965), GUS, Statystyka regionalna, 4, Warszawa 1966.
- [3] *Rocznik statystyczny 1966* (Statistical yearbook), GUS, Warszawa 1966.
- [4] *Rozwój gospodarczy powiatów w latach 1950—1965* (Economic development of poviats in the years 1950—1965), GUS, Statystyka regionalna, 5, Warszawa 1967.
- [5] *Statystyka ludności — małżeństwa, urodzenia, zgony 1961—1964* (Demographic statistics — marriages, births, deaths 1961—1964), GUS, Statystyka Polski, 113, Warszawa 1965.
- [6] *Statystyka ruchu naturalnego ludności 1963—1965* (Vital statistics 1963—1965), GUS, Statystyka Polski — Materiały statystyczne, 11 (133), Warszawa 1966.

INVESTIGATIONS OF URBAN LAND USE IN POLISH GEOGRAPHICAL STUDIES

JULITTA RAKOWICZ-GROCHOLSKA

Early post-war investigations of land use in towns and cities were a continuation of the landscape studies which had been started at the beginning of this century. The first attempt after the war to present a new approach to the study of land use in general was made by A. Jahn [1]. It included an outline of a project, based on certain British studies, which contained suggestions as to the development of a map of land use for the whole country. The land use of the whole territory of Poland was also the subject-matter of studies carried out by F. Uhorczak. The settlement pattern was abstracted from maps scaled at 1 : 100,000 and afterwards was reduced photographically. The results, in the form of a 1 : 1,000,000 survey map of land use for the whole of Poland, were published in 1957 [2].

Since that time studies of land use have followed two different directions — those concerning land use in towns, and those dealing with land use in rural areas. The number of detailed studies has also been growing.

A survey of these detailed studies ought to begin by mentioning K. Bromek's investigations of land use in Kraków and its vicinity, which he began in 1948. His initial methodological model was A. Jahn's study quoted above. A report on the first phase of the work, covering the mapped area of Kraków within its administrative boundaries of the period, was published in *Przegląd Geograficzny* [3]. On the basis of a cadastral map and of the land use map that had been constructed, K. Bromek made a comparative analysis of the changes in land use in this large city, as well as in its vicinity, in the course of the last hundred years. This enabled him to draw conclusions as to the trends in, and the determinants of, the changes observed. He noticed both the impact of the geographical environment on the effectiveness of a given mode of land use and the impact of a given mode of utilizing the land on particular elements in the geographical environment. As far as field-work was concerned, he emphasized the necessity to take into account the local economic factors that affect the manner of utilizing the land.

An interesting attempt to provide a specialised study of a single town was made by H. Leonhard (Fig. 1). This monograph, on the town of Trzebnica [4] differs from the "classical" studies of its type both in the approach adopted and in its handling of the subject. On the basis of detailed analyses carried out previously, H. Leonhard made an attempt to synthesize the landscape of the town. In effect, she made a division of the town into functional units of identical historical origin,

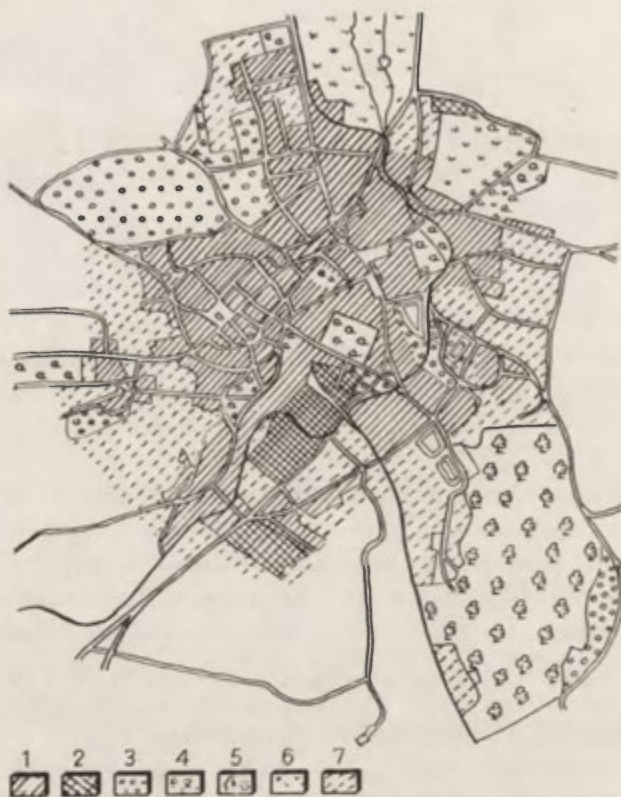


Fig. 1. Trzebnica: land use in 1948

1 — built-up areas, 2 — industrial areas, 3 — orchards, 4 — municipal parks and gardens, 5 — woods and forests, 6 — meadows and pastures, 7 — arable land

Source: H. Leonhard [4]

and with common characteristics in use of land. In conclusion the study contains an appraisal of the role of the town in the economy of the region.

In the mid-fifties, on the initiative of Professors K. Dziewoński and J. Kostrowicki, investigations into the use of land in urban and rural areas were started in the Institute of Geography of the Polish Academy of Sciences. The first collective investigations were conducted in the powiat of Mrągowo and included a detailed analysis of types of land use within the town of Mrągowo, based on the characteristics of the geographical environment and the historical development of the town (Fig. 2).

From the point of view of the method adopted, a similar analysis was made by the author of the present paper in an investigation of the land use in the town of Trzcińsko-Zdrój, which was a part of a larger study of the town [5]. In the case of Trzcińsko closer attention was paid to an analysis of the mode of utili-

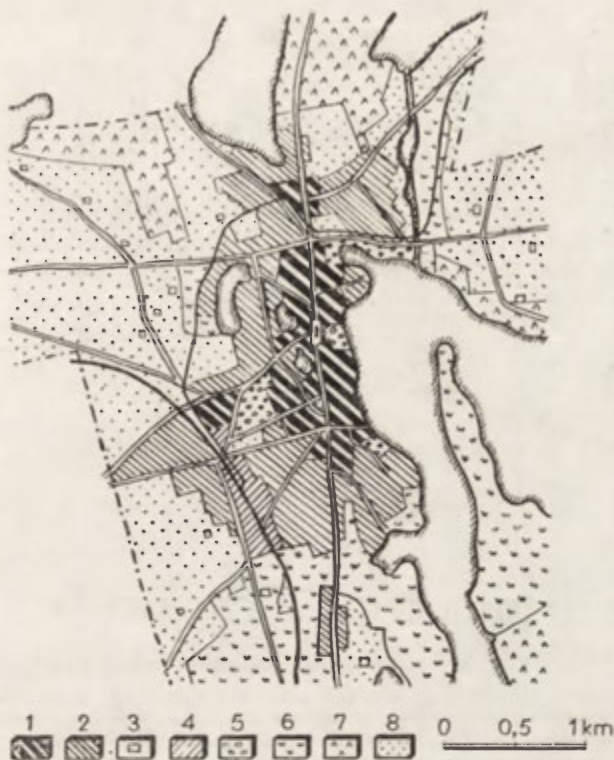


Fig. 2. Mragowo: land use in 1955

1 — urban mixed residential-commercial uses, 2 — low density residential areas with gardens and allotments, 3 — rural houses, 4 — industrial-storage areas, 5 — municipal parks and recreational facilities, 6 — meadows and pastures, 7 — woods and forests, 8 — arable land

zing the rear of the building-plots. This is an example of a very detailed study of land use. Next, by direct field-observation, data on the use of land in 8 towns were collected and, subsequently, elaborated. Because the method adopted proved to be too laborious, especially in the cases of the larger towns and cities, A. Werwicki suggested a simplified method when he started his investigations of Lower Silesian towns. His method consisted of a direct mapping of the elements selected for recording. Because the age of the buildings had been left out, the element of time was lost and the results obtained thus failed to give a dynamic presentation of the phenomenon. Being aware of this deficiency, A. Werwicki introduced the element of time into his plotting method in his more recent studies.

Short descriptions of land use are also contained in some monographs analysing the economic development potential of small towns [6]. For instance, in her study of the town of Lipno, E. Romahn-Kwiatkowska outlines briefly the mode of land use and divides the town into several zones in the course of describing the spatial structure and the development of investment in the town. The

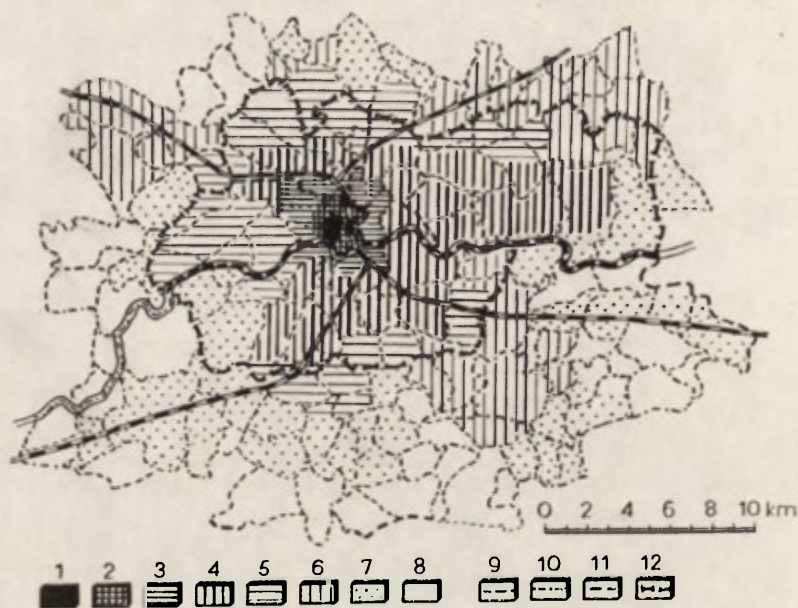


Fig. 3. Cracow: land utilization zones 1960

1 — zone I city centre, 2 — zone II city interior, 3 — zone III city exterior, 4 — zone IV ante-urban, 5 — zone V interior suburban, 6 — zone VI intermediate suburban, 7 — zone VII exterior suburban, 8 — zone VIII the periphery, 9 — villages and cadastral quarters boundaries, 10 — communities boundaries, 11 — counties and administrative quarters boundaries, 12 — boundaries of Cracow

Source: K. Bromeek [4]

descriptive part is supplemented by a summary of land uses as well as a map of the built-up zones against the background of the use of land.

Further methodical progress was made by K. Bromeek's study presenting the use of land in Cracow and in the suburban belt in about the year 1960 (Fig 3) [7]. This study concludes one phase of a broad-scale project planned by its author (cf. note 3).

The objective set up by K. Bromeek was to investigate how, in its economic activity, a community utilizes the land-surface in current socio-economic conditions and within the technical, social and economic setting of a large city and its suburban zone. To answer this the author attempts to define the economic characteristics of the community in the area under examination. Thus he analyses the development of the population in the particular cadastral units, and, afterwards, the size of individual farms, the forms of land ownership etc. He draws attention to the concentricity of the patterns of the phenomena observed. The part devoted to the analysis of the geographical environment is concluded by an evaluation of its natural conditions and resources from the standpoint of the needs of the different branches of the national economy. Next, the fundamental types of land use

and the changes observed in the course of the last hundred years are discussed. From the standpoint of method, the concluding chapter, which contains a division into zones of land use, is the most interesting. On the basis of his investigations, he recognizes that the distance from the core of the city is the decisive factor. By analysing the trends of the regression-lines (which illustrate how the distance from the centre of the city affects the share of different types of land use in the total area) K. Bromek obtains a geometrical picture of the extent of the particular zones; subsequently, he compares the latter with geographical zones. To establish the effects of the influence of the city-core, which are expressed by the degree of "urbanity" and "suburbanity" of the land use, the author selected several criteria. According to them eight zones of land utilization in the city of Kraków have been distinguished. This is a further contribution to the investigation of land use in that it provides a new and more precise approach as well as interesting comparative data that may be of use in further typological studies.

On the basis of this review of studies and from my own experience I propose a systematization of studies in land use. The first problem to be settled is the selection of the method by which the measurements will be carried out. So far, the method of direct observation on the spot has been the most frequent; it consists of the simultaneous filling-in of special forms and the mapping of the data collected. In effect, a graphical presentation of a careful inventory of land use is obtained, while the whole study is of a descriptive character. This type of procedure is very tedious and it is applicable in practice only in small towns, or else with the help of a whole team of collaborators. No doubt, the time devoted to survey can be shortened substantially by adopting the sample method. For this purpose, a number of samples ought first to be selected, and afterwards thoroughly analysed. The examination of a definite number of samples (this number can be determined by statistical methods) will furnish sufficient material to verify the interrelationships discovered in other studies; it will enable the student to identify, order, and find out the relative importance of the particular factors that affect land use, as well as to draw conclusions referring to the whole settlement pattern under analysis.

Another method for accelerating the measurements can be attained by designing special forms which would simplify and shorten as much as possible the field-work by tabulating or plotting on the spot only those elements which had been previously selected for this purpose. In this case, the final results of the investigation will, to a great extent, depend on what particular elements are selected.

A substantial shortening of the initial phase of the work can also be attained by employing the rich empirical material giving a fairly accurate illustration of the state of land use in towns, which can be found in town and country planning offices or in geodetic survey offices. In this case, the researcher's work could be reduced to verifying the source material. The analyses of this material which have

been carried out so far has shown that the most frequent discrepancies arise from faults of classification, which may reflect the arbitrary judgements of the authors who may fail to adopt exact categories. This problem must be given close attention in the analysis of such source material. On the whole, however, the latter constitute a valuable "data bank" for the student of those problems.

A further method to simplify and accelerate measurement is a change in the scale adopted. So far, the basic scale employed in drawing up detailed land use maps has been 1 : 5,000. Though it is still practicable to use this scale in studies of small towns, any study of medium-sized and large towns or cities will require a smaller scale.

A change in scale necessitates generalization. This is the next logical step in investigations of land use. Generalization is possible if it is preceded by a number of detailed studies. In the course of investigations the fundamental generalized classes of land use are established. In spatial analysis this means the aggregation — or generalization — of the units (working-units or others) established previously. The single plot or building then ceases to be the unit of reference and is replaced by the so-called "block", i.e. a compact area delimited by the streets or other boundaries. The employment of the concept of a "block" necessitates the introduction of the category of "mixed land use", which is one of the characteristic traits of land use in the central area of the town.

The next phase of the work is classification. This must start with an attempt to establish the size and the degree of complexity, as well as the general functional characteristics, of the settlement system under examination. Suggestions in this respect are contained in K. Dziewoński's study of the morphological typology of towns, in which methods of morphological analysis are presented and classification procedure is discussed [8]. The classification of land use necessarily demands also a general and comprehensive view of these problems, because it is in the occupational structure of the population, in the urban fabric, and in the forms of land use that the functional character of the town or settlement finds its reflection (*op. cit.*, p. 443).

After these preliminary operations the proper classification of land use can be started. The closest attention of the student will be focused on the establishment of adequate methods of comparison, that is on the selection and the definition of features for classification. One such method consists of the selection of features by the correlation method, which, in this case, could be employed by examining the correlation between the characteristic features that are specific for an area, and comparing the areas one with another afterwards. This is a multi-factor correlation, and its results may permit one to construct a classification from the units adopted. Similar investigations based on the taxonomic method are being conducted in Poland by S. Lewiński.

The final phase of the research is typology. It is a synthesis of a number of detailed studies. The characteristic features and relationships discovered and

subsequently examined by mathematical analysis furnish sufficient comparative material, which — after verification in a large number of analogous cases — will make possible the establishment of the typological features of land use that are specific for the given systems. At present it is difficult to set the number of the various types of land use. In current investigations from several to more than ten types are provisionally assumed.

To conclude, it must be emphasized that before starting studies on land use in a large city the researcher ought first to acquaint himself with the materials gathered by urban and topographical survey. If there is a sufficient stock of detailed material it is advisable not to undertake new field-work but to have recourse to the existing material instead. Of course, the remarks concerning the reliability and the classification of these materials made above ought to be borne in mind. Moreover, it seems to be pointless to analyse the whole city, because the sample method, verified by different examples, renders fairly satisfying results. The selection of samples could be carried out by laying a square geometrical grid upon the plan of the city so that the node of the grid is placed in the middle of the urban system. Then, for each point or area the geometrical co-ordinates could be established (X , Y and an additional co-ordinate Z for the housing areas, which should denote the number of the storeys of the buildings) and the particular cells for analysis could be selected either by a random method or by observing a definite principle of succession. Such a formalized frame of reference enables the student to disregard the administrative divisions, which are variable both in time and in space. Further investigations should aim at providing detailed problem-studies of selected samples, the results of which will reveal any tendency to subjective error. In analysis, the relative importance of the particular factors influencing land use ought to be established. This analysis must be based upon the knowledge of the socio-economic conditions of the area on which the town under investigation is situated. The general thesis of the study will be contained in the proposition that the land use in town A at place i depends primarily on such factors as

- distance from the city-core,
 - distance from local centres,
 - distance from municipal transport nodes and principal transport lines,
- and that this system is modified by the factors of
- the character of the contiguous area, and
 - natural conditions,

and the role of these factors will vary in dependence on local conditions.

Next, on the basis of a knowledge of these factors the particular urban areas must be evaluated from the standpoints of different classes of land use. From our classification of uses, we shall obtain the generalized classes of land use with a measure of intensiveness of area-utilization and precisely determined functions performed for the benefit of the whole of an urban organism.

The next task consists in comparing and explaining the differences between

a theoretically derived system and the actual picture of land use. For this purpose statistical analysis ought to be employed. The ultimate purpose of such studies will be a proposal for a typology of land use in a large city.

Institute of Geography PAN, Warsaw

BIBLIOGRAPHY

- [1] Jahn, A., *Studia nad użytkowaniem ziemi w Polsce, stan dotychczasowych badań i projekt instrukcji* (Studies on land use in Poland: current state of investigation and an instruction project), mimeographed, 1947.
- [2] Uhorczał, F., *Polska przeglądowa mapa użytkowania ziemi w skali 1 : 1.000.000* (A general map of land use in Poland), Warszawa 1957.
- [3] Bromek, K., Opracowanie szczegółowej mapy użytkowania ziemi dla Krakowa (Sum.: Elaboration of a detailed map of land use for Cracow), *Przegl. Geogr.*, 27 (1955), 589—604.
- [4] Leonhard, H., Trzebnica; studium geograficzno-planistyczne, monografia użytkowania ziemi (Trzebnica — a geographical planning study), *Czasopismo Geogr.*, 21/22 (1950/51), 301—344.
- [5] Rakowicz, J., Trzcińsko-Zdrój, użytkowanie ziemi w mieście (Land use in town), *Dokum. Geogr.*, 2, 1958.
- [6] Studia geograficzne nad aktywizacją małych miast (Geographical studies on activization of small towns), Polish Ac. Sc., Inst. of Geogr., *Prace Geograficzne*, 9, 1957.
- [7] Bromek, K., Użytkowanie ziemi w Krakowie i przyległych częściach powiatu krakowskiego około 1960 roku (Land use in Cracow and in the contiguous areas of the neighbouring poviat about 1960), *Zeszyty Naukowe Uniw. Jagiell.*, 128, *Prace Geogr.*, 14, Kraków 1966.
- [8] Dziewoński, K., Zagadnienie typologii morfologicznej miast w Polsce (Sum.: The problem of a morphological typology of towns in Poland), *Czasopismo Geogr.*, 33 (1962), 441—457.

SOME METHODS OF DETERMINING LAND USE AND AGRICULTURAL "ORIENTATIONS" AS USED IN POLISH LAND UTILIZATION AND TYPOLOGICAL STUDIES *

JERZY KOSTROWICKI

1. INTRODUCTION

As the history, methods and techniques of the Polish Land Utilization Survey have been already presented to the 1st Anglo-Polish Seminar [38] and also published several times elsewhere [9, 17, 18, 31, 32, 35, 37, 41, 45, 64, 69, 70], I feel free to omit all these questions from the present report except for some remarks on the more recent developments of the survey.

Despite all our efforts the Polish survey has not yet attained the stage reached by the first British survey, or even the second, as regards covering the whole country. One of the reasons, inherent in the Polish method itself, might be that by being much more detailed in the mapping of various aspects of land use and thus more labour absorbing, a detailed land use survey made on the scale of 1 : 10 000 or 1 : 25 000 required considerable effort and large funds, not so much for field mapping but rather for the drawing and printing of maps. At the research institute we have neither sufficient staff, nor have we succeeded in obtaining sufficient funds. For these reasons, despite the fact that quite large areas have been mapped already, only a few sample printed maps are available. To overcome this difficulty the method and technique of the simplified land-use map on the working scale of 1 : 100 000 and published at 1 : 200 000 has been recently devised and some thousand square kilometres mapped. It is now intended to cover the whole of Poland in this way, leaving the detailed study for some regions that are of particular interest, either for scientific or practical reasons. In collaboration with geographers of the other East-Central European countries we intend to start common mapping of all these countries. The problem has already been discussed within the special regional subcommission of the IGU Commission on World Land Use Survey [66] and the methods were discussed in 1969 at the 3rd sub-commission meeting in Yugoslavia.

* Partly based on, and examples taken from: W. Biegajło, K. Bielecka, W. Gadomski, *Methods of determination of specific typological characteristics of agriculture in the research work of the Agricultural Geography Department*. 17 pp. (mimeographed).

Both the detailed and more general land-use studies carried out in Poland and in the other East-Central European countries, with whom we constantly exchange teams of field workers, have collected rich and interesting material. In a number of studies based on this material (1—3, 6—9, 12—14, 19—24, 27—30, 55—63, 65, 67, 71, 73—74, 76—78, 81—86, 93) numerous methods and techniques have been introduced, some of which have found their application in the works of the other IGU Commission on Agricultural Typology.

To this Seminar I would like to present some of the methods and techniques which are used in the Polish Land Utilization Survey and which are associated with the determination of the so-called “orientation” in land utilization and agriculture.

2. THE CONCEPT OF ORIENTATION

Some explanation of the terms used should precede the discussion of the methods and techniques concerned and particularly of the term “orientation” itself.

The term “orientation” is widely used in some languages other than English to define the tendency or inclination of a particular holding or area to use the land for particular purposes, or to grow particular crops or to produce particular agricultural commodities either for home consumption or for sale. The idea of orientation does not, however, concern all elements but only the main elements, such as leading land uses, leading crops or leading products. The problem that immediately arises is how to distinguish the main from the secondary elements.

3. ARABLE-LAND ORIENTATION

The idea of orientation, borrowed from agricultural economics, was first applied to land use surveys [35, 71, 73]. The introduction of this idea, which had been used only in Polish land-use surveys, resulted from the following.

1. Both in Poland and in the other countries where small-scale village agriculture predominates, plots of land under particular crops are too small to be presented on a map with a scale that can also accommodate more extensive areas.

2. On most of the arable areas particular crops are grown each year on different plots of land, following more or less regular crop rotation. So, in contrast to other land uses, such as perennial crops, permanent grasslands, forests, etc., the picture of crop distribution changes every years. What is changing but slowly are the proportions of land occupied by the leading crops. Such information can be obtained by less labour-absorbing techniques than the field mapping of each individual plot of land or even by the interpretation of air photographs.

The method that was accepted is based on the analysis of agricultural statistics for each village or large-scale holding to measure the proportions of the main crops. In this case the notion of orientation does not differ much from that of crop combination as introduced by J. C. Weaver [87—90] and followed by many others [25, 26, 68], including J. T. Coppock for Great Britain [15], except that the more sophisticated method used by them is more labour absorbing, while not necessarily being more accurate than the one used in the Polish studies (94).

On the other hand, following the same line of thinking that had led Weaver to his rather pessimistic conclusions [89], an inquiry had to be made during field studies as to whether the crop combinations established for whole villages reflected the real situation in particular holdings or operational units, i.e. whether they represented the true orientation or showed only the village averages for a number of diverse orientations. In the latter case, particularly when the area under study is highly differentiated, either from the point of view of natural conditions or size of farms, the result may not reflect the real situation. The results may be even more suspect when crop combinations are defined for larger units. In such situations crop combinations obtained from an analysis of statistical data should be critically studied and, if necessary, corrected until they are recognized as arable orientations. In particular, some orientations defined for the units situated between larger areas with different orientations may show crop combinations which would disappear if the study were to be based on the smaller units. This is so, not because of their transitional character, which can sometimes be the reason, but because individual parts of the area under study have different orientations similar to those practiced on the neighbouring areas.

For this reason, particularly when larger units are involved, a special procedure based on the study of natural and other conditions is necessary to cope with units showing uncommon crop combinations. Only sometimes do they reflect the real differentiation in arable land utilization and as such should be recognized. In most cases, however, the units have to be split into two or more parts and incorporated in neighbouring orientation areas. As a result of such a procedure the number of orientations can usually be greatly reduced and the map thus becomes clearer and simpler.

The arable-land orientation has recently been accepted [49, 58] as one of the important indices denoting type of agriculture. The number of meso- and macro-scale studies using this idea is growing [10, 11, 14, 22, 44, 46, 50, 52, 72, 74].

Some explanations now follow of how the technique is used in Poland to determine orientations of arable land utilization. Because in the small-scale village agriculture of a semi-subsistence character with little specialization, numerous crops are usually grown within the area of a particular holding or village, and because some crops are of a similar or complementary character, it was decided not to study the role of each of them separately but first to group them and then to study the dominant elements in each group. As a primary concern of land use

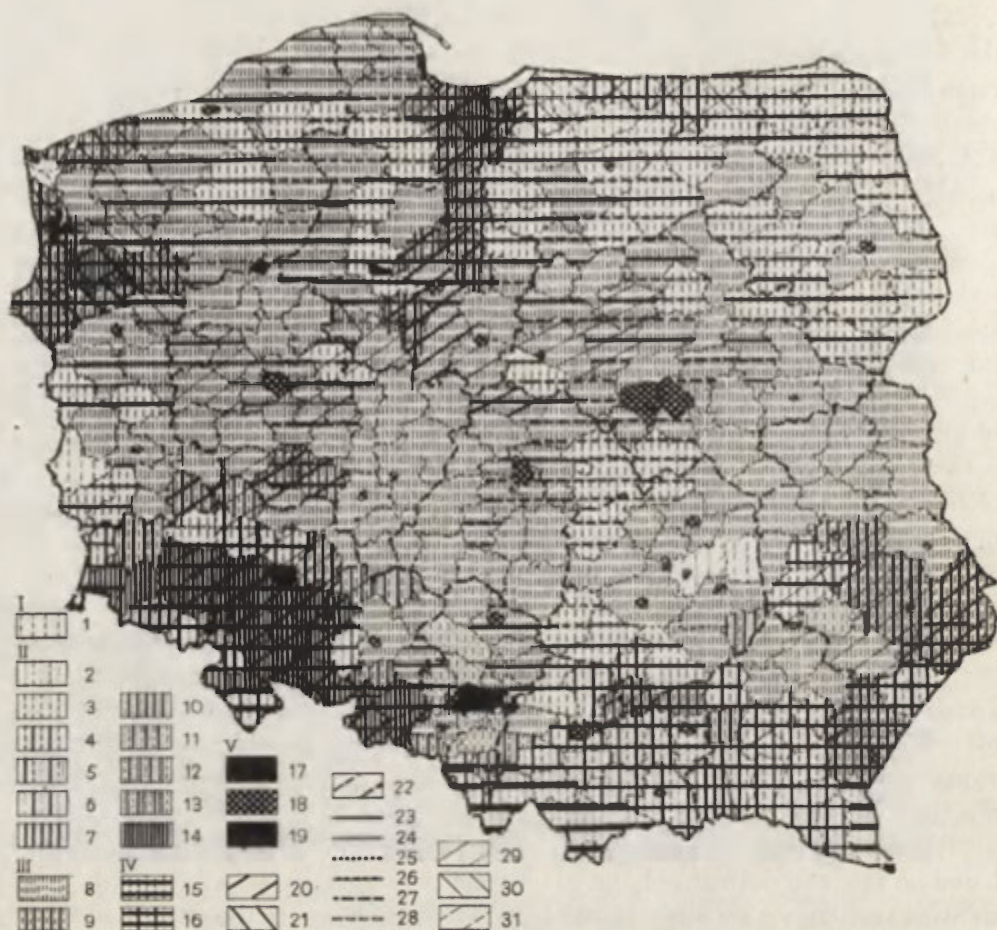


Fig. 1. Poland: Orientations of arable land utilization (crop combinations), by R. Kulikowski

I. Mainly cereal: 1 — mainly rye with potatoes

II. Cereal: 2 — rye-oats with potatoes, 3 — rye with potatoes, 4 — rye — wheat and wheat — rye with potatoes, 5 — wheat — rye — oats and rye — wheat — oats with potatoes, 6 — wheat — oats with potatoes, 7 — wheat with potatoes

III. Cereal — ridged-up crops: 8 — rye — potatoes, 9 — rye — wheat — potatoes and wheat — rye — potatoes, 10 — wheat — oats — potatoes and oats — wheat — potatoes, 11 — wheat — rye — oats — potatoes and rye — oats — wheat — potatoes, 12 — rye — wheat — barley — oats — potatoes, 13 — rye — wheat — barley — potatoes, 14 — wheat — potatoes

IV. Cereal — leguminous crops: 15 — oats — clover with potatoes, 16 — wheat — oats — clover with potatoes and oats — wheat — clover with potatoes

V. Ridged-up crops: 17 — potatoes, 18 — potatoes — vegetables and vegetables — potatoes, 19 — vegetables

Orientations with industrial crops: 20 — with sugar beet, 21 — with rape-seed, 22 — with tobacco

Orientations with leguminous crops: 23 — with clover, 24 — with lucerne, 25 — with mashlum, 26 — with vetch or field pea, 27 — with serradella, 28 — with lupine

Industrial crops exceeding 30 percent of the area under ridged-up crops: 29 — with sugar beet prevailing, 30 — with rape-seed prevailing, 31 — with tobacco prevailing

studies is the more rational use of the natural properties of land by better, more efficient methods and techniques, the grouping of crops was based on their agroeconomic similarities, particularly with regard to their requirements in terms of natural conditions, labour input, fertilization, place in crop rotation etc. Based on these criteria the following three groups of crops were distinguished [38, 39, 43, 45, 47].

1. *Intensive crops*, requiring much labour, care in cultivation, and fertilization. As the result of these requirements their cultivation enriches the soil in humus and minerals, thus increasing the soil fertility, so that they are considered as good forecrops. Ridged-up or row crops, vegetables, and most of the industrial crops are classified here.

2. *Structure-forming crops*, not requiring much labour or fertilization but nevertheless enriching the soil with nitrogen and helping to create a good crumb structure. They are also good forecrops. All leguminous crops, both for human consumption and for stockfeed, as well as those used for green manure, are included here.

3. *Extractive (or exhausting) crops* are those that exhaust the soil most. To restore its natural fertility after extractive crops the soil requires more intensive fertilization and/or proper succession in crop rotation, to prevent its becoming depleted or degraded. Mostly cereals and other crops with similar requirements have been included here.

The classification of world crops according to their agronomic properties and economic aims has been published elsewhere [39].

After grouping all crops grown on a particular holding or area into these three groups, the next step is to study the proportion between these three groups. For many years the following technique has been applied. The group with 20% or less of the total area under a crop is disregarded. The role of the others is determined as follows:

Percentage of the area under a particular crop	Role	Rank	Orientation (for example)
over 80	dominant	4	(mainly wheat)
60—80	preponderant	3	(wheat with ...)
40—60	equal	2	(wheat and ...)
20—40	accompanying	1	(... with wheat)

As in Poland the share of intensive crops (among which sugar beets and potatoes dominate) seldom exceeds 40% or falls below 20%, in order to emphasize their important role an additional threshold of 30% was introduced in some studies, thus increasing the number of ranks to 5. In the countries south of the Carpathians, where maize is a dominant element in this group, and differentiation in the role of intensive crops is much greater, such an additional threshold is not necessary.

TABLE 1. Arable land utilization (examples)

Specification	Sym- bol	<i>Gromadas</i> (communes)									
		Próchna		Zebrzydowice		Jaworzynka		Brenna		Koniaków	
		ha	%	ha	%	ha	%	ha	%	ha	%
Extractive											
crops	E	495	51.4	438	52.6	264	27.5	346	42.6	208	30.3
wheat	tv	136	14.1	186	20.2	—	—	12	1.5	—	—
rye	sc	117	12.1	133	16.0	—	—	46	5.7	—	—
barley	hs	42	4.4	6	0.7	11	1.2	25	3.1	6	0.9
oats	av	200	20.8	131	15.7	253	26.3	263	32.3	202	29.4
Intensive											
crops	I	236	24.5	253	30.4	202	21.0	195	24.0	129	18.9
oleaginous	h	21	2.2	6	0.7	—	—	—	—	—	—
fibre plants	t	3	0.3	9	1.1	—	—	—	—	—	—
maize	zm	20	2.2	17	2.0	1	0.1	1	0.1	—	—
sugar beet	bs	15	1.5	3	0.3	—	—	—	—	—	—
fodder root											
crops	pr	40	4.1	38	4.6	7	0.7	7	0.9	3	0.4
potatoes	st	131	13.6	162	19.5	185	19.3	184	22.6	123	18.1
vegetables	lg	6	0.6	18	2.2	9	0.9	3	0.4	3	0.4
Structure											
forming crops	S	232	24.1	141	17.0	495	51.5	272	33.4	349	50.8
vetch	vs	2	0.2	6	0.7	1	0.1	8	1.0	1	0.1
clover	ti	135	14.1	119	14.3	373	38.8	257	31.6	131	19.1
lucerne	ms	15	1.5	—	—	—	—	1	0.1	—	—
serradella	os	55	5.7	—	—	120	12.5	1	0.1	215	31.3
mashlum	px	25	2.6	16	2.0	1	0.1	5	0.6	2	0.3
Sown area		963	100.0	832	100.0	961	100.0	813	100.0	686	100.0

Crop combinations:

Próchna	$E_3 av + I_1 st + S_1 ti$	oats with potatoes and clover
Zebrzydowice	$E_3 av + I_1 st$	oats with potatoes
Jaworzynka	$E_1 av + I_1 st + S_3 ti$	clover with oats and potatoes
Brenna	$E_3 av + I_1 st + S_1 ti$	oats with potatoes and clover
Koniaków	$E_1 av + S_3 ti$	clover with oats

After the proportions between groups are established, the dominant element in each group is determined. A crop occupying more land than any other within a given group is considered as dominant; those, if any, occupying over 80% of the area under the dominant crop, in the detailed land use studies, and 60% in the macroscale studies, are considered as co-dominant and marked on the map.

The resulting orientation is expressed in a notation or formula in which the groups are shown in capital letters, dominant or co-dominant crops in small letters, and the rank of particular groups of crops by figures. The symbols for

particular crops are based on their Latin (botanical) names. The orientation is called after the dominant crop (s).

The following example is taken from the powiat (county) of Cieszyn in the Western Carpathians and shows the orientations common for various mountain zones. The crop combinations were established for *gromadas* (communes), the lowest administrative units, usually containing 2—4 villages.

As can be seen from Table 1, in Próchna extractive crops together occupying 51.4% of the total area under crops, are ranked third; the intensive crops, with 24.5%, rank as 1; and the structure-forming crops, occupying 24.1%, are also ranked as 1. Oats is a dominant crop in the group of extractive crops, the next one, wheat, occupying less than 80% of the first. Potatoes dominate clearly among intensive crops, and among the structure-forming crops clover is most important. The crop combination can then be expressed by the following formula — $E_{3av} + I_1 st + S_1 ti$ — and is defined as “oats with potatoes and clover”. Brenna has a similar orientation, despite the share of structure-forming crops (with dominant clover) being higher. In Jaworzynka, however, which is situated higher in the mountains, the group of structure-forming crops comes to the forefront, with rank 3, and the share of other groups does not exceed rank 1. So the crop combination is defined as “clover with oats and potatoes”. In Zebrzydowice, lying lowest, the share of structure-forming crops falls below 20%, so the resulting crop combination is “oats with potatoes”. The orientation “clover with oats” also applies in the case of Koniaków, the highest area, where intensive crops do not exceed 20% of the crop land. Of course a revision of the crop combinations has to be made to eliminate the averages for two or more orientations.

Once determined, orientations are then marked on the land use maps, using stripes whose width is proportional to the rank of group, with the colour showing the dominant crop [16, 42].

As one can easily see from the above example, the sum of the rank figures is sometimes 5 and sometimes 4, depending on the role of the group that is omitted, and if the additional threshold of 30% is accepted it varies even more. This unnecessarily increases the number of crop combinations and is troublesome for drawing the map, as the total width of stripes is variable.

To avoid these difficulties and to make the procedure even less laborious, a new technique has been recently applied which may be called the technique of “successive quotients”. It is largely modelled on the *d'Hondt* system used in some countries in elections. The area under particular groups of crops as established above is divided successively by 1, 2, 3, 4 etc., until the final number of classified places (quotients) is obtained. To get a proper degree of differentiation the number of quotients should not be too small, but also not too large, as the resulting picture should not be too complicated. As the orientation reflects to some degree the crop rotation cycle and should illustrate most of the situations the number of quotients should be as much as possible divisible. For these reasons

6 successive quotients have been usually accepted for more detailed studies or less differentiated areas and 4 successive quotients for less detailed studies or more differentiated areas.

Using the same examples from Table 1, the procedure for the determination of crop combinations using the successive quotients technique would be as follows (Table 2).

TABLE 2. Gromada Próchna

Divisors	Extractive crops (E)		Intensive crops (I)		Structure-forming crops (S)	
	hectares	quotients	hectares	quotients	hectares	quotients
1	495	1/	236	2/	232	3/
2	247	4/	118		116	
3	165	5/	78		77	
4	123	6/	59		58	
5	99		47		46	

Taking 4 successive quotients only and using the same technique of determining dominant crops as in the percentage method the following crop combination results:

$E_2 av + I_1 st + I_1 ti$ — i.e. oats with potatoes and clover.

Taking 6 successive quotients we obtain a formula — $E_4 av + I_1 st + I_1 ti$ — that does not change the name of orientation but emphasizes rather the dominant role of extractive crops (cereals).

TABLE 3. Gromada Zebrzydowice

Divisors	Extractive crops (E)		Intensive crops (I)		Structure-forming crops (S)	
	hectares	quotients	hectares	quotients	hectares	quotients
1	438	1/	253	2/	141	5/
2	219	3/	126	6/	70	
3	146	4/	84		47	
4	109		63		35	
5	87		51		28	

Using 4 quotients the result is as follows: $E_3 av I_1 st$ — an “oats with potatoes” combination, the same as obtained by the percentage method. Using 6 quotients one obtains:

$E_3 av I_2 st S_1 ti$ — a much more complicated combination showing the preponderance of oats, the secondary role of potatoes and accompanying role of clover.

TABLE 4. Gromada Koniaków

Divisors	Extractive crops (E)		Intensive crops (I)		Structure-forming crops (S)	
	hectares	quotients	hectares	quotients	hectares	quotients
1	208	2/	129	4/	349	1/
2	104	6/	64		174	3/
3	69		43		116	
4	52		32		87	
5	41		25		69	

Using 4 quotients the result is:
 $E_1 av + I_1 st + S_3 ti$ — i.e. the three-fold “clover with oats and potatoes” in contrast to the two-fold result obtained by the percentage method.

Using 6 quotients one comes to $E_2 av + I_1 st + S_3 ti$ — in which clover is preponderant, oats a secondary, and potatoes accompanying element.

The sum of ranks — quotients — is always the same and may always be marked on the map by the same total width of the stripes; the calculation is easier as no percentage counting is necessary and a number of crop combinations that may differ by the number of ranks only is usually less. The technique thus seems to be more practical.

4. OTHER LAND-USE ORIENTATIONS

The same technique could be applied for the determination of orientation in agricultural land uses other than arable, for instance for permanent grassland and perennial crops. The first could be based either on a division by composition of grassland vegetation or on their mode of use (whether cut for hay or grazed) and then on particular systems of grassland utilization. The method has not been tried as yet. On the other hand some attempts have been made to determine orientation for perennial crops, either from the proportion of area they occupy or from the number of trees, shrubs or vines. Of course, the productivity of a tree, shrub or vine of various species differs greatly, and the density of trees or shrubs also differs. Some positive results showing the dominance of particular perennial crops within highly complicated “coltura promiscua” on the Dalmatian coast of Yugoslavia have been achieved [55].

The method has recently been applied with more success to determine general land-use orientation.

It is particularly useful either for macroscale studies where the actual pattern of land use could hardly be marked on the map, or where the pattern of land use is very complicated. Also the dynamics of land use in the areas with complicated and changing use of land can be studied using this method as, for instance, the

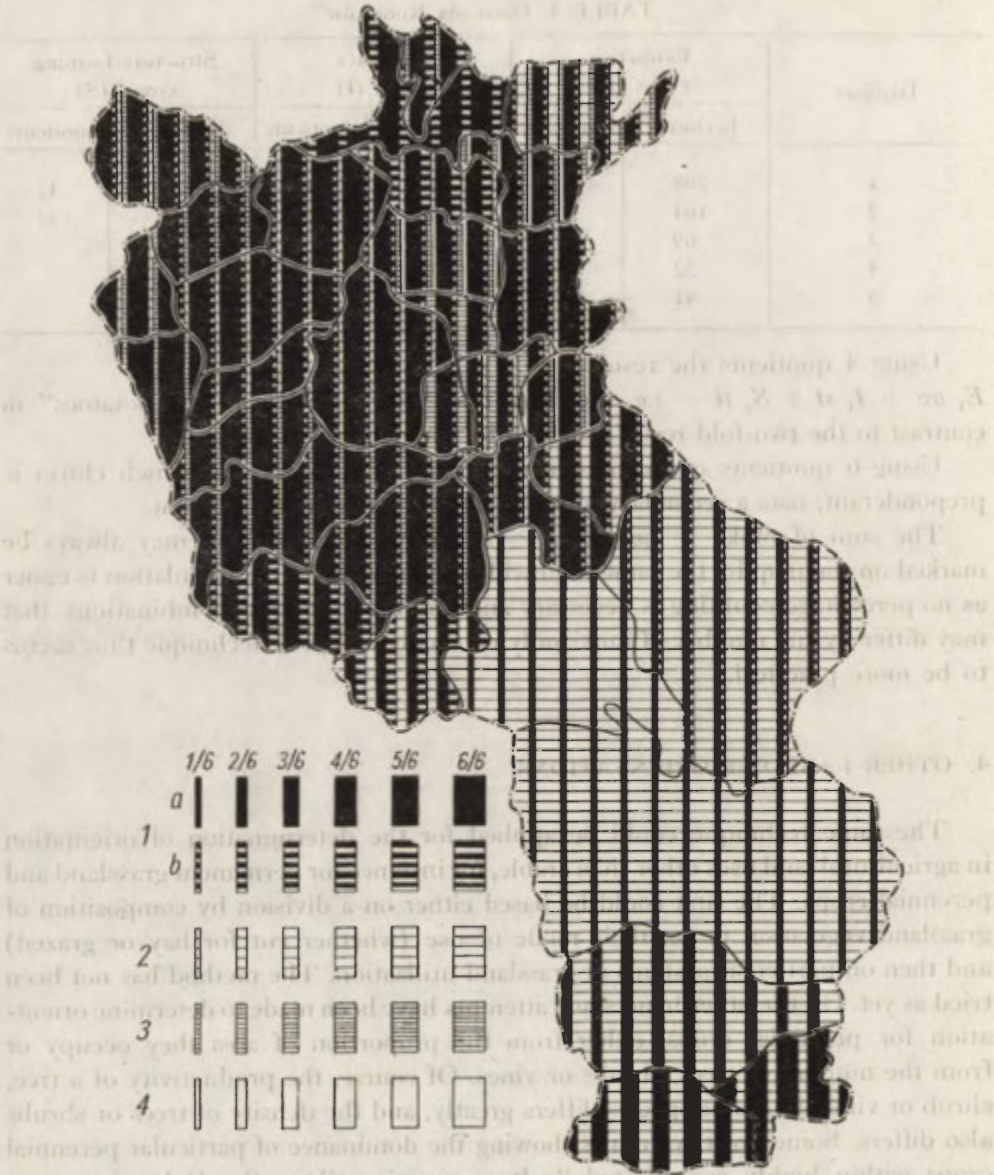


Fig. 2. Poviát of Cieszyn, Poland. Orientations of land utilization 1961
(by W. Gadomski)

1 — Agricultural land: a — arable land, b — grassland; 2 — Woodland; 3 — Settlements; 4 — Waters

changing boundary between agricultural and residential areas which is very important nowadays, or in some areas the changing boundary between agricultural land and forests. The idea might also be useful for typological studies of agriculture to show the place it occupies within the total land used or available.

As the basis for determining land use orientation, the total amount of land as divided between the main land uses (agricultural land, forests, waters, settlement and unproductive land) has been adopted and then their subdivisions, as for instance that of agricultural land according to arable, permanent grassland and perennial crops, or sub-divisions of urban land by reference to residential, industrial, transport, recreational uses etc. The question is not yet resolved as regards how such uses as forests or waters may be subdivided in a similar way. In the studies made up to the present no such subdivision has been undertaken.

The method is here presented using the examples from the same gromadas as in Table 1. Both techniques (the percentage method and that of successive quotients) might be used to determine land use orientations. Using the percentage technique and presenting the results in a formalized way, the following land-use orientations could be obtained.

Próchna: $R_3(o_2 + p_1ps) + F_1$ — i.e. Agricultural, field (arable) with pastoral and forest orientation.

Zebrzydowice: $R_2(o_2) + F_1 + D_1$ — Agricultural, field with forests and settlement orientation.

Jaworzynka: $R_2(o_2) + F_2$ — Agricultural, field — forest orientation.

Brenna: $R_1(p_1ps) + F_3$ — Forest with agricultural, pastoral orientation.

TABLE 5. Land Utilization (examples)

Gromada (commune)	Sym- bols	Total area	Agricultural Land						For- ests	Settle- ment areas
			total	arable land	peren- nial crops	permanent grassland				
						total	mea- dows	pastu- res		
			R	o	v	p	pt	ps	F	D
Próchna	ha	2174	1476	999	21	457	172	285	439	268
	%	100.0	67.6	45.7	1.0	20.9	7.8	13.1	20.1	12.3
Zebrzydowice	ha	2065	1132	833	10	289	120	169	439	494
	%	100.0	54.9	40.4	0.5	14.0	5.8	8.2	21.2	23.5
Jaworzynka	ha	2223	1035	949	17	69	49	20	1066	122
	%	100.0	46.6	42.7	0.8	3.1	2.2	0.9	47.9	5.5
Brenna	ha	7816	2294	815	8	1471	514	957	5152	370
	%	100.0	29.3	10.3	0.1	18.8	6.6	12.2	65.9	4.8
Wisła	ha	9757	1780	834	21	925	598	327	7760	217
	%	100.0	18.3	8.6	0.2	9.5	6.1	3.4	79.5	2.2

Wisła: F_3 — Predominantly forest orientation.
The sum of elements (ranks) varies between 3 and 4.
Using the successive quotients technique for both the main land uses and their subdivisions, the land use orientations for the same communes could be determined as below.

TABLE 6. Gromada Próchna

Divisors	Agricultural land (R)		Forests (F)		Settlement (D)	
	hectares	quotients	hectares	quotients	hectares	quotients
1	1476	1/	439	4/	268	
2	738	2/	219		134	
3	492	3/	146		89	
4	369	5/				
5	295	6/				

With 4 quotients the orientation would be $R_3(o_2 + p_1pe) + F_1$ — Agricultural, field with pastoral and forest orientation.
With 6 quotients — $R_5(o_4 + p_1ps) + F_1$ — Highly agricultural, field with pastoral and forest orientation.
Zebrzydowice — in the same way — $R_2(o_2) + F_1 + D_1$ or $R_4(o_3 + p_1ps) + F_1 + D_1$
Jaworzynka — $R_2(o) + F_2$ or $R_3(o) + F_3$
Brenna — $R_1(o_1) + F_3$ or $R_2(o_1 + p_1ps) + F_3$
Of course the problem of averages and real orientations as discussed in the section on crop combinations should be kept in mind here as well.

5. ORIENTATION IN LIVESTOCK FARMING

The method presented above could also be used to determine the orientation of livestock farming. The problem is to find the common denominator for various kinds of animals that would correspond to area in land use. As the common measure the so-called large stock units widely applied in many countries to show the density of livestock could be adopted. They are based on the average weight of particular animals, but according to the purpose, some other units, such as fodder or manurial units, could be tried.
Both techniques (percentage counting and successive quotients) to find the share of particular kinds of livestock (cattle, pigs, sheep, poultry, etc.) within the sum of large stock units could be of use here. The problem is only whether non-productive livestock, such as horses, asses, mules etc., as well as oxen, are to be included in the study of orientation or not. Of course, if the problem of the

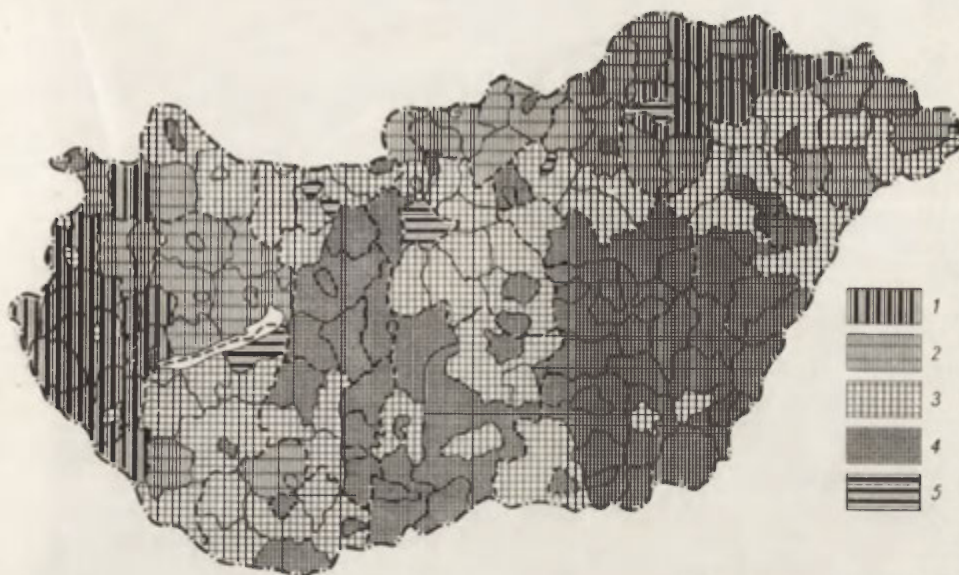


Fig. 3. Hungary. Orientations of livestock farming (by R. Szczęsny)

1 — mainly cattle, 2 — cattle with pigs, 3 — cattle — pigs or pigs — cattle, 4 — pigs with cattle, 5 — mainly pigs

fodder base is considered, they participate greatly in its consumption, to the detriment of the productive livestock. If, however, the aim is to study the tendency of individual holdings or areas to keep particular kinds of livestock for productive purposes, then the draught and pack animals should be eliminated, as they really form part of the motive power and transport equipment, together with tractors, trucks etc. They should not be included with the productive animals, therefore, except when they are kept for breeding purposes.

6. ORIENTATION OF AGRICULTURAL PRODUCTION

The orientation of agricultural production reflects the tendency or disposition of a particular holding or area to produce specific agricultural commodities. It may first be characterized by the proportions of crop and livestock production and then the proportions of particular principal crops or livestock products.

The determination of the orientation of agricultural production is more difficult and involves several problems. First, there is that of the common measure of agricultural production. The natural measures, such as weight, are meaningless and cannot be applied. The other two solutions are to use monetary units or conventional units. Both of these measures have their advantages and disadvantages (see discussion [39]). The use of monetary units is the easiest and the most comprehensive measure, but has to be based on prices, which are rather difficult to establish for non-commercial production and are subject to change over both time



Fig. 4. Poland. Orientations of agricultural production (by R. Szczesny)

1 — rye — oats with potatoes, hay and cattle, 2a — rye with potatoes, hay and cattle, 2b — rye with sugar beet and cattle, 3 — rye — wheat with potatoes, hay and cattle, 4 — wheat with sugar beets, hay and cattle, 5 — rye — potatoes with hay and cattle, 6 — rye — hay with potatoes and cattle, 7 — wheat — oats — hay with cattle, 8 — rye — potatoes — hay with cattle, 9 — potatoes with rye, hay and cattle, 10 — potatoes — vegetables with rye, hay and cattle, 11 — vegetables with rye, hay, cattle and pigs, 12 — clover or pasture with rye, hay and cattle, 13 — oats — cattle — hay with potatoes, 14 — rye — oats — cattle with potatoes, 15 — rye — cattle — potatoes with hay, 16 — rye — wheat — cattle with potatoes and hay, 17 — clover — oats — wheat — cattle, 18 — hay — cattle with cereals and potatoes, 19 — pig breeding as associated element

and space. Time changes can be eliminated by the introduction of so-called fixed prices, which are more or less artificial. Prices, however, differ also in space, both according to the distance from the market and, more important, between various countries, as a result of governmental policies (tariffs, subsidies, etc.). As a result agricultural prices, seldom based on the true interplay of demand and supply, do not reflect the real value of agricultural products.

Still another problem is whether gross agricultural production or final production is to be used to define the orientation of agricultural production [39]. The use of gross production involves the double-counting of some agricultural products, such as fodder and livestock production or green manure and crops grown after it. On the other hand, final production, while eliminating these disadvantages, could not be used to calculate agricultural characteristics connected with labour or capital inputs, as they are both used separately for fodder and livestock production. In most cases, as there are usually no data on the use of non-commercial products, final production is more difficult to calculate. For these reasons, despite all deficiencies, gross production expressed in grain equivalent units is used in most of the Polish and many other studies.

Table 7 shows the estimated gross production and its components calculated for the villages of Borysówka in the Białystok Voivodship (North-Eastern Poland), as well as for Belski Vrh and Goričak in Eastern Slovenia, Yugoslavia.

Unlike arable land-use orientation, crops are here grouped according to the reasons for which they are produced (food, fodder and industrial crops) [39]. The grouping is also clear-cut; there are, however, some difficulties with crops which are used in various ways, as for instance potatoes or barley, which may be used partly for food, partly as fodder or partly for industrial purposes. When basing these calculations on statistical data it is often difficult to establish what part of these crops is used for each of those purposes.

As far as livestock production is concerned the grouping may be either by products or by kinds of livestock, or both. In the present example animal production is first sub-divided by livestock and then by products but the reverse would also be acceptable.

The result is presented by a formula in which V means “vegetal”, i.e. crop and other plant production, and A “livestock (animal)” production. The determination of orientation could proceed either by the percentage or the successive quotients technique.

First let us try the percentage technique, similar to the determination of orientation of arable-land utilization. The proportion between two main groups of agricultural production is first established in the following way:

Percentage share of crop production in total gross production	Name	Formula
over 80 per cent	predominantly crop	V_4
60 — 80 per cent	crop	V_3A_1
40 — 60 per cent	mixed	V_2A_2
20 — 40 per cent	livestock	V_1A_3
under 20 per cent	predominantly livestock	A_4

TABLE 7. Gross agricultural production (examples)

Specification	Sym- bol	Borysówka		Belski Vrh		Goričak	
		grain units	%	grain units	%	grain units	%
I Crop production	V	3092	53.4	2047	75.6	1723	65.5
A. Food crops	a	1289	22.2	1323	48.9	1013	38.5
wheat	tv	120	2.1	58	2.2	122	4.6
rye	sc	756	13.1	18	0.6	22	0.8
barley	hs	—	—	—	—	19	0.7
maize	zm	—	—	57	2.1	137	5.2
other cereals	g	28	0.5	—	—	—	—
leguminous (pod crops)	px	27	0.5	60	2.2	—	—
potatoes	st	180	3.0	65	2.4	104	4.0
vegetables	lg	68	1.2	50	1.9	24	0.9
fruit	vr	110	1.8	390	14.4	447	17.0
grapes	vv	—	—	625	23.1	138	5.3
B. Fodder crops	p	1735	30.0	724	26.7	710	27.0
grains	g	526	9.1	62	2.3	126	4.8
succulent forage (silo, root)	s	272	4.7	9	0.3	—	—
meadow hay	pt	937	16.2	653	24.1	584	22.2
C. Industrial crops	i	68	1.2	—	—	—	—
sugar beet	bs	19	0.3	—	—	—	—
fibre plants	t	49	0.9	—	—	—	—
II Animal production	A	2694	46.6	661	24.4	908	34.5
A. Cattle	bt	1030	17.8	404	14.9	680	25.8
milk	l	693	12.0	267	9.9	405	15.4
meat (beef and veal live weight)	m	337	5.8	137	5.0	275	10.4
B. Pigs (live weight)	ss, m	880	15.2	198	7.4	165	6.3
C. Sheep	ov	468	8.1	—	—	—	—
meat (live weight)	m	274	4.7	—	—	—	—
wool	t	194	3.4	—	—	—	—
D. Poultry	k	316	5.5	59	2.1	63	2.4
meat	m	114	2.0	10	0.3	11	0.4
eggs	oo	202	3.5	49	1.8	52	2.0
Gross production		5786	100.0	2708	100.0	2631	100.0

The orientation of agricultural production is then established as follows:

For example, in Goričak crop production accounts for 65.6% of the gross agricultural production and receives 3 (V_3). Within crop production food crops occupy 38.4% and fodder crops 27% of the total gross production and to both of them is ascribed rank 1.

Orchard fruits prevail among food crops and meadow hay in fodder products. Livestock production does not exceed 34.4 per cent of gross production and is



Fig. 5. Poland. Orientations of commercial production (by R. Szczęsny)

I. Mainly livestock: 1 — mainly cattle with pigs and poultry, 2 — mainly cattle and pigs

II. Livestock: 3 — cattle with rye and potatoes, 4 — cattle with wheat and sugar beet, 5 — cattle — pigs with rye and potatoes, 6 — cattle — pigs with wheat and sugar beet, 7 — pigs with sugar beet and rye or wheat, 8 — pigs with potatoes, 9 — pigs with vegetables and fruit

III. Crop — livestock: 10 cattle — sugar beet — wheat, 11 cattle — vegetables (or fruit), 12 cattle — pigs — rye — potatoes, 13 cattle — pigs — sugar beet — wheat, 14 cattle — pigs — sugar beet — rye, 15 cattle — pigs — tobacco, 16 cattle — pigs — vegetables (or fruit)

IV. 17 — Crop: Fruit — vegetable with cattle and pigs

ranked as 1. Cattle are responsible for predominant livestock production and their share in gross agricultural production is 25.8 per cent, with milk production clearly prevailing. Finally the production orientation for the village of Goričak could be defined as crop (food-fodder), fruit—meadow hay, with dairy cattle and expressed in the following notation: $V_3(a_1 vr + p_1 pt) + A_1(bt_1 l)$.

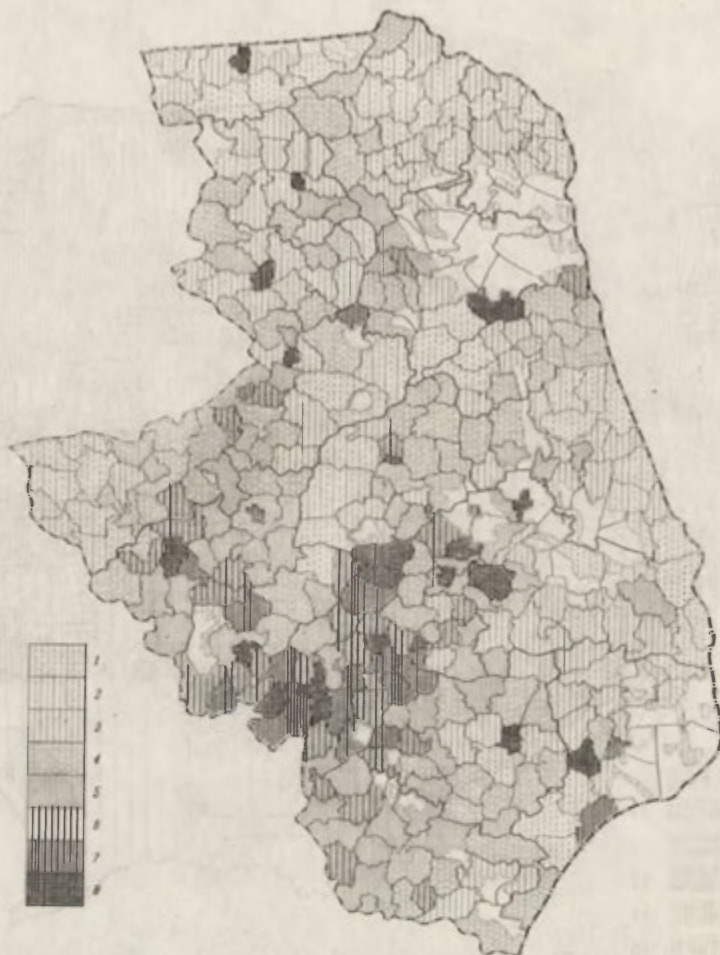


Fig. 6. Białystok Voivodship, North-Eastern Poland. Land productivity (by W. Biegajło)

Gross agricultural production in grain units per hectare of agricultural land 1 — less than 18, 2 — 18—21, 3 — 21—24, 4 — 24—27, 5 — 27—30, 6 — 30—33, 7 — 33—36, 8 — over 36

The orientation of agricultural production for the other villages was determined in the same way, as follows (12).

Belski Vrh — $V_3(a_1 vv + p_1 pt) + A_1(bt, l)$ crop, (food-fodder), grapes-hay with dairy cattle.

Borysówka — $V_2(a_1 sc + p_1 pt) + A_2(bt, l + ss, m)$ mixed, (food-fodder) rye — meadow hay — dairy cattle (meat) — pork.

The procedure of using the successive quotients technique to determine agricultural production orientation would be as follows.

First, the proportion between crop (plant) and livestock production is established as V_3A_1 , using 4 quotients, or V_4A_2 using 6 quotients, then in the same

TABLE 8. Example: Goricak Village

Divi- sor	Total crop V production		Total animal A production		Groups				Cattle bt		Pigs ss	
					food crops a		fodder crops p					
	grain units	quo- tients	grain units	quo- tients	grain units	quo- tients	grain units	quo- tients	grain units	quo- tients	grain units	quo- tients
1	1723	1/	908	2/	1013	1/	710	2/	680	1/	165	
2	861	3/	454	5/	506	3/	355	4/	340	2/	82	
3	574	4/	303		337		238		227		.	
4	430	6/	227		253		177		170		.	
5	344		181		202		142		136		.	

way the proportion between the various groups of crops is established as a_3p_1 (together V_3) in the first case and a_2p_2 (together V_4) in the second. Among live-stock products cattle provide most by either calculation (bt_1 or bt_2). Finally, the dominance or co-dominance of particular elements within each group is established by the usual method.

Among food crops orchard fruits (Vr) clearly dominate and among fodder crops, meadow hay (pt). In the production provided by cattle, milk (l) dominates.

Using 4 quotients the orientation of agricultural production in Goricak will be: $V_3(a_2vr + p_1pt) + A_1(bt_1l)$ i.e. crop (food with fodder), fruit with meadow hay and dairy cattle.

Using 6 quotients the notation is as follows: $V_4(a_2vr + p_2pt) + A_2(bt_2l)$, meaning almost the same — crop, (food-fodder), fruit — meadow hay — dairy cattle orientation.

Using the same technique the following orientation for the other villages were established.

Belski Vrh — $V_3(a_2vv + p_1pt) + A_1(bt_1l)$ — Crop (food with fodder), grapes with hay and dairy cattle — or $V_5(a_3vv + p_2pt) + A_1(bt_1l)$ mainly crop (food with fodder), grapes with hay and dairy cattle.

Borysówka — $V_2(a_1sc + p_1pt) + A_2(bt_1l + ss_1m)$ — Mixed (food-fodder) rye — hay — dairy cattle (meat), pork, or $V_3(a_1sc + p_2pt) + A_3(bt_2l + ss_1m)$ — Mixed (fodder with food), hay with rye — dairy cattle with pork.

7. ORIENTATION OF COMMERCIAL AGRICULTURAL PRODUCTION

Using the same techniques the orientation of commercial production i. e. specialization of agriculture understood as a tendency towards the selling of particular agricultural produce. The question which arises here first is whether particular cash products are to be grouped or not and if so in which way when commercial production is under consideration. In view of the fact that usually

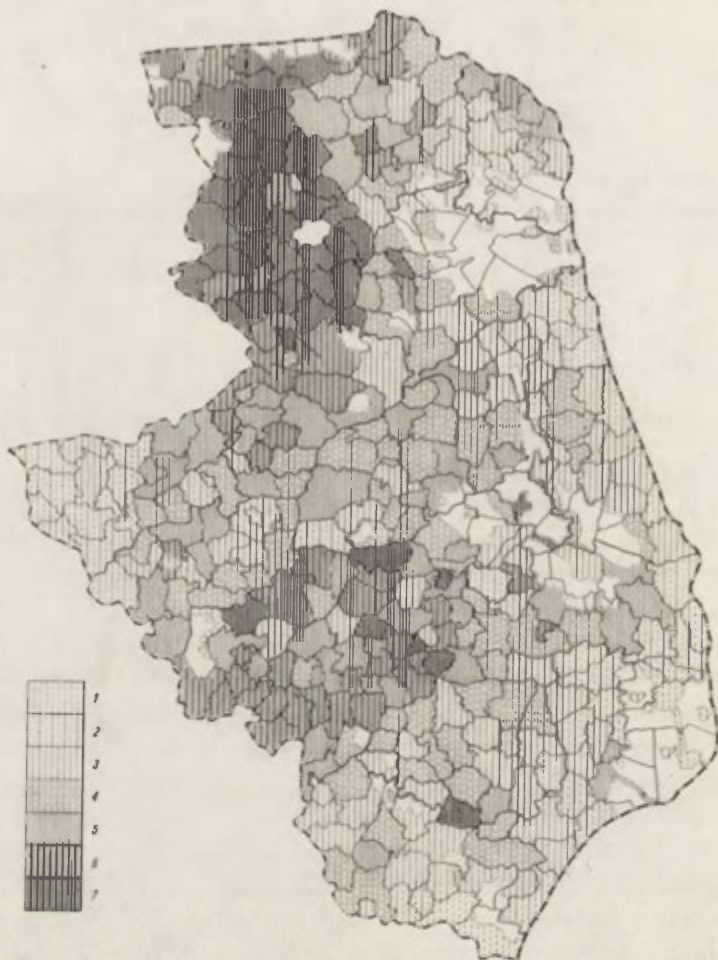


Fig. 7. Białystok Voivodship. North-Eastern Poland. Labour productivity (by W. Biegajło)
Gross agricultural production in grain units per person employed in agriculture 1 — less than 60, 2 — 60—70, 3 — 70—80, 4 — 80—90, 5 — 90—100, 6 — 100—120, 7 — over 120

they are not very numerous the grouping of commercial products may not be necessary. The next question is whether agricultural cash products should be expressed in conventional or monetary units. It seems that in this case monetary units may have more advantages. If so, however, what prices ought to be used in the calculations?

Having estimated gross production it is easy to determine land productivity as given by agricultural production per unit of agricultural land, and labour productivity, i.e. agricultural production per number of persons employed in agriculture [39, 49, 58].

Having estimated commercial production one can establish the degree of

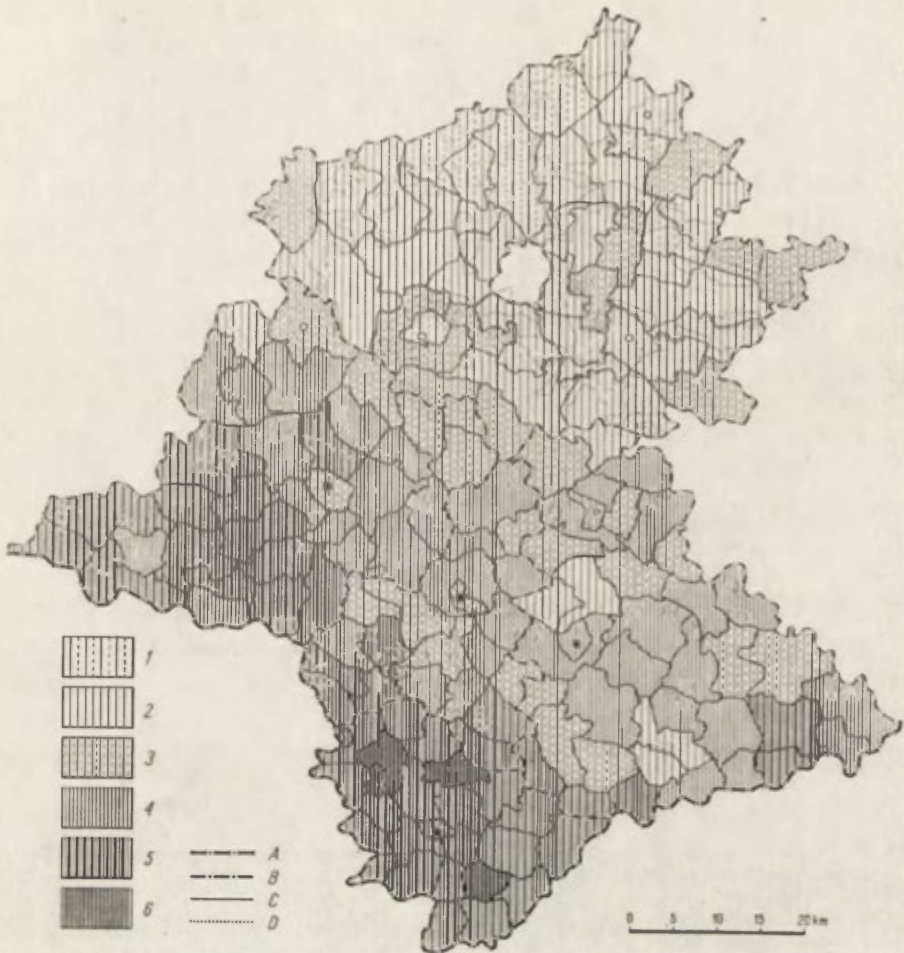


Fig. 8. Ponidzie, Southern Poland. Degree of Commercialization (by W. Stola)

Commercial agricultural production in per cent of gross production: 1 — less than 15 per cent, 2 — 15—20 per cent, 3 — 20—25 per cent, 4 — 25—30 per cent, 5 — 30—40 per cent, 6 — over 40 per cent

A — voivodship boundaries, B — poviat boundaries, C — gromada boundaries, D — limits of major forests

commercialization, i.e. the share of commercial production in gross agricultural production, or the level of land commercialization, i.e. amount of commercial agricultural production per unit area of agricultural land, and labour commercialization, i.e. commercial agricultural production per person employed in agriculture [39, 40, 58]. In fact many studies of this type have been already made in Poland [10, 75, 79, 89, 91, 92]. Some examples of the maps are included in this paper. What still has to be considered is how the degree of specialization i.e. the degree to which commercial production is concentrated on the smallest number of products, can be expressed in a quantitative way.



Fig. 9. Hungary. Level of commercialization (by R. Szczęsny)

Commercial agricultural production in grain units per hectare of agricultural land



Fig. 10. Hungary. Commercialization of labour (by R. Szczęsny)

Commercial agricultural production per person employed in agriculture

SUMMARY

The specific character of the methods and techniques in use in Polish agricultural geography can be presented as follows (46, 50, 52):

1. There is a strong tendency towards a more synthetic study of land use and agriculture, wherever possible using quantitative, and thus measurable and comparable, methods and techniques, without, however, excessive generalization and resultant over-simplification.

2. Numerous synthetizing ideas and indices have been introduced and tried in studies at various scales for measuring such aspects of agriculture as land and labour productivity and degree and level of commercialization.

3. For some non-measurable aspects of land use and agriculture an attempt has been made to present them in a formalized and thus comparable way.

Among them are methods and techniques to define land use or agricultural orientations, as already described, and here set down in summary form:

a. elements are grouped according to their properties rather than analysed separately,

b. only leading elements contribute to orientation (or combination),

c. the results are expressed in notations, easy to compare with each other in order to arrive at a typology or regionalization,

d. not only crop combinations within the total of field crops are analysed, but the method has been extended to determine

— land-use orientation

— orientation of agricultural production

— orientation of commercial agricultural production

e. most of the methods and techniques have already been checked, using not only data from land-use surveys but also material based on statistical data for Poland and other countries.

4. As the several methods and techniques discussed above are important for agricultural typology, together with the others proposed by the Commission [49, 58], they are now being checked as to their universal applicability by a number of case studies distributed throughout the world. Together with other methods and techniques they will be eventually used to build up world and regional typologies of agriculture.

5. As agricultural planning or programming consists largely in establishing desirable types of agriculture for a given time and space, characterized by their synthetic properties, and then in working out ways of transition from present to future types, the methods and techniques developed by land use and typological investigations of agriculture and the results of these could be of practical importance for planning agricultural development [46, 50, 51]. This is, however, a problem which is treated more extensively elsewhere [54].

BIBLIOGRAPHY

- [1] Biegajło, W., Gospodarka rolna na Żuławach; wieś Radunica (Sum.: Farming in Gdańsk Żuławy; the village of Radunica), *Przegl. Geogr.*, 31 (1959), 345—360.
- [2] Biegajło, W., Gospodarka rolna w powiecie gdańskim (Farming in the Gdańsk County), *Dokumentacja Geograficzna*, 1, 1960, 1—59.
- [3] Biegajło, W., Borysówka, Grodzisko and Hruskie villages in the North-Eastern undeveloped corner of Poland, in: Land utilization in East-Central Europe; case studies, *Geogr. Polon.*, 5, 1965, 29—60.
- [4] Biegajło, W., Vyzkum využití pudy v Polsku (Land utilization investigations in Poland), *Sborník Českosl. Spol. Zem.*, 70 (1965), (3), 251—265.
- [5] Biegajło, W., Z badań nad użytkowaniem ziemi w Prowansji, gmina Banon (Sum.: From the study on land utilization in Provence, France; the case of the commune of Banon), *Przegl. Geogr.*, 37 (1965), 109—141.
- [6] Biegajło, W., Z badań nad użytkowaniem ziemi w Dalmacji: Jesenice, gmina (općina) Omiš, (Sum.: Land utilization in Central Dalmatia; the cadaster unit Jesenice, South of Split), *Dokum. Geogr.*, 2/3, 1966, 130—153.
- [7] Biegajło, W., Z badań nad użytkowaniem ziemi we wschodniej Słowenii; Wieś Belski Vrh. (Sum.: Land utilization in Eastern Slovenia; the village of Belski Vrh), *Dokum. Geogr.*, 2/3, 1966, 50—79.
- [8] Biegajło, W., Recherches sur l'utilisation du sol dans la region de Forcalquier, commune de Banon (Basses-Alpes), *Cah. du Centre d'Etudes des Soc. Médit.*, 1 (1966), 131—171.
- [9] Biegajło, W., Polish land utilization survey in the years 1960—1964, in: Land utilization in Eastern Europe, *Studies in geography in Hungary*, 4, 1967, 28—34 (with sample maps of land utilization of the communes: Łomianki, Warsaw suburban zone, Poland; Pacanów, Oleśnica and Kwaczała, South of Poland; Kačice, North-West of Prague, Czechoslovakia; Badacsony, Western Hungary; Żeleznik, South-West of Belgrade, Serbia, Yugoslavia; the key of symbols to the detailed land utilization maps is also included).
- [10] Biegajło, W., Rolnictwo (Agriculture), in: *Województwo Białostockie, monografia geograficzno-gospodarcza*, Lublin 1967, 76—302.
- [11] Biegajło, W., Types of agriculture in North-Eastern Poland, Białystok voivodship, *Geogr. Polon.*, 14, 1968, 275—282.
- [12] Biegajło, W., Kowalczyk, D., and Piskorz, H., Land utilization in Nieborów, in: *Problems of Applied Geography*, Polish Ac. Sc., Institute of Geography, Geogr. Studies 25, Warszawa 1961, 56—68.
- [13] Biegajło, W., and Paulov, J., Studium využitia zeme na príklade obce Velke Blahovo na Žitnom Ostrove (Sum.: Study of land utilization exemplified by the village Velke Blahovo in Žitni Ostrov), in: Aspects of the study of regional geographical structure, *Acta Geol. et Geogr. Univ. Comenianae, Geographia*, 6, Bratislava 1966, 225—248.
- [14] Bromek, K., Użytkowanie ziemi w Krakowie i przyległych częściach powiatu Krakowskiego około 1960 roku (Sum.: The utilization of land in Cracow and in the adjoining parts of the Cracow County around 1960), *Zeszyty Naukowe Univ. Jagiell., Prace Geograficzne*, 14, 1966.
- [15] Coppock, J., Crop, livestock and enterprise combinations in England and Wales, *Economic Geography*, 40 (1964), 65—81.
- [16] *Detailed Map of land utilization*, Sheet N 34—54 C-a, Chroberz, Instytut Geografii PAN, 1964.
- [17] Dziewoński, K., Detailed survey of land utilization in Poland, *Przegl. Geogr.*, 28 (1956), Suppl., 26—31.
- [18] Dziewoński, K. and Kostrowicki, J., Detailed land use survey in Poland, U.G.I., *Comptes*

- Rendus du XVIII^e Congrès International de Géographie*, Rio de Janeiro 1956, vol. 4, Rio de Janeiro 1966, 108—113.
- [19] Enyedi, I., The „Kossuth” Collective Farm of Bekescaba in the Southern Part of the Great Hungarian Plain, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 407—420.
- [20] Gadomski, W., Użytkowanie ziemi i rolnictwo w Kotlinie Koszyckiej na przykładzie Spółdzielni Drienovec (Sum.: Land utilization and agriculture in the Kosice Basin, East Slovakia; the case of the Drienovec collective farm), *Dokum. Geogr.*, 2/3, 1966, 27—47.
- [21] Gadomski, W., Z badań nad użytkowaniem ziemi w regionie górnej Nitry; Spółdzielnia Nitranske Sučany (Sum.: Land utilization in the Upper Nitra Valley, Slovakia; the case of the collective farm of Nitranske Sučany), *Dokum. Geogr.*, 1, 1967, 5—26.
- [22] Gluziński, J. and Rzymowski, S., Użytkowanie ziemi i gospodarka rolna powiatu Elbląskiego (Sum.: Land utilization and agriculture of the County of Elbląg), *Zeszyty Geogr. WSP w Gdańsku*, 7, 1966, 73—151.
- [23] Hauzer, S., Użytkowanie ziemi i gospodarka rolna w gromadzie Łomianki, (Land utilization and agriculture in the commune of Łomianki), *Dokum. Geogr.*, 5, 1965, 39—63.
- [24] Hauzer, S., Użytkowanie ziemi i gospodarka rolna we wsiach Žvab i Runec (Słowenia Wschodnia), Jugosławia (Sum.: land utilization in Eastern Slovenia: the case of Žvab and Runec villages), *Dokum. Geogr.*, 2/3, 1966, 105—127.
- [25] Islam, M. A., *Crop combination regions in East Pakistan*, Ph.D. thesis, Clark University, Worcester, Mass., 1965 (unpublished).
- [26] Islam, M. A., Crop combination regions in Pakistan, *The Oriental Geographer*, 9 (1965), (1), 1—16.
- [27] Jelic, D., Jeršič, M., Lojk, J., and Vojvoda, M., The cadastrian commune of Trebijovi in the Karstland of Hercegovina, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 267—284.
- [28] Jelonek, A., Z badań nad użytkowaniem ziemi w powiecie limanowskim, (From the studies on land utilization in the Limanowa County), *Zeszyty Naukowe Uniwersytetu Jagiellońskiego, Prace Geogr.*, 4, 1961, 71—96.
- [29] Jeršič, M., Lojk, J., Olas, L., and Vojvoda, M., The village of Sebeborci on the Slovenian fringes of the Great Pannonian Plain, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 215—234.
- [30] Klemenčič, V., The village of Podgorje in Slovenian Sub-Alpine region, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 195—214.
- [31] Kostrowicki, J., Polskiye issledovaniya ispolzovaniya zemel, *Izvestiya Akademii Nauk SSSR, Ser. Geogr.*, 1958, 4, 131—134.
- [32] Kostrowicki, J., Badania nad użytkowaniem ziemi w Polsce (Sum.: Research studies on land utilization in Poland), *Przegl. Geogr.*, 31 (1959), 517—533.
- [33] Kostrowicki, J., Brief summary of the report on land use studies in East-Central Europe and the USSR, *IGU Newsletter*, 11, (1960), (1/2), 41—42.
- [34] Kostrowicki, J., Land utilization survey as a basis for geographical typology of agriculture, *Przegl. Geogr.*, 32 (1960), Suppl., 169—183.
- [35] Kostrowicki, J., The aims, concept, and method of Polish land utilization survey, *Dokum. Geogr.*, 3, 1960.
- [36] Kostrowicki, J., (ed.), *Land utilization, methods and problems of research, Proceedings of the International Conference, Poland 1960*, Polish Ac. Sc., Inst. of Geogr., Geogr. Studies 31, Warszawa 1961.
- [37] Kostrowicki, J., Le survey polonais de l'utilisation du sol, in: *Land utilization, methods and problems of research*, Polish Ac. Sc., Inst. of Geography, Geogr. Studies, 31, Warszawa 1961, 31—58.

- [38] Kostrowicki, J., Polish land utilization survey, in: *Problems of applied geography*, Polish Ac. Sc., Inst. of Geography, Geogr. Studies, 25, Warszawa 1961, 45—56.
- [39] Kostrowicki, J., Geographical typology of agriculture in Poland, methods and problems, *Geogr. Polon.*, 1, 1964, 11—146.
- [40] Kostrowicki, J., Geographical typology of agriculture, principles and methods; an invitation to discussion, *Geogr. Polon.*, 2, 1964, 159—168 and *Revista Geografica*, 33 (1964), 15—24.
- [41] Kostrowicki, J., The Polish detailed survey of land utilization, methods and techniques of research, *Dokum. Geogr.*, 2, 1964.
- [42] Kostrowicki, J., (ed.), Land utilization in East Central Europe; case studies, *Geogr. Polon.*, 5, 1965.
- [43] Kostrowicki, J., Land utilization, Case studies: origins, aims, methods, techniques, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 7—28.
- [44] Kostrowicki, J., An attempt to determine the geographical types of agriculture in East-Central Europe on the basis of the case studies on land utilization, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 453—498.
- [45] Kostrowicki, J., *Polskie zdjęcie użytkowania ziemi, metoda i technika badawcza* (Polish land utilization survey, methods and techniques), Warszawa 1965, (mimeographed).
- [46] Kostrowicki, J., Savremenite problemi i metodi na Polската selskostopanska gheografija (Contemporary problems and methods of the Polish agricultural geography), *Izvestia na Bulgarskoto Gheografsko Drujhestvo*, 5 (15), 1965, 115—125.
- [47] Kostrowicki, J., Metody opracowania materiałów zdjęcia użytkowania ziemi, (Sum.: On the methods applied to the elaboration of the materials collected by the land utilization survey), *Dokum. Geogr.*, 2/3, 1966, 1—23.
- [48] Kostrowicki, J., On the study of agricultural typology, *Annals of the Assoc. of Econ. Geogr.*, 1966, 70—78, (in Japanese).
- [49] Kostrowicki, J., *Principles, basic notions and criteria of agricultural typology*, IGU Com-mision for Agricultural Typology, Warszawa 1966, (mimeographed).
- [50] Kostrowicki, J., Problemy i metody polskiej geografii rolnictwa, (Sum.: Contemporary problems and methods of the Polish agricultural geography), *Czasopismo Geogr.*, 37 (1966), 289—304.
- [51] Kostrowicki, J., Sur l'application de la géographie de l'agriculture aux besoins pratiques de l'aménagement du territoire, in: *La géographie appliquée dans le monde*, Commission de la Géographie Appliquée de l'UGI, Actes de la réunion à Prague du 13—18 Septembre 1965, Prague 1966, 66—70.
- [52] Kostrowicki, J., Die Agrargeographie in Polen; Entwicklung, Forschungsgebiete und -ergebnisse, *Mitteilungen für Agrargeographie, Landwirtschaftliche Regionalplanung und Aus-ländische Landwirtschaft*, 12 (1967), 165—176.
- [53] Kostrowicki, J., Methods applied in elaborating the material of land utilization survey. in: Land utilization in Eastern Europe, *Studies in Geography in Hungary*, 4, 1967, 9—18.
- [54] Kostrowicki, J., Agricultural typology. Agricultural regionalization. Agricultural develop-ment, *Geogr. Polon.*, 14, 1968, 265—274.
- [55] Kostrowicki, J., Kowalczyk, D., Barsko Polje on the Adriatic Sea Coast of Montenegro, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 285—344.
- [56] Kostrowicki, J., Hauzer, S., Velchev, I., and Borisov, Z., The collective farm of Petarch in the suburban zone of Sofia, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 345—372.
- [57] Kostrowicki, J., Stola, W., Velchev, I., and Borisov, Z., The collectivized village of Der-mantsi on the Northern Foothills of the Balkan Range, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 373—406.

- [58] Kostrowicki, J., and Helburn, N., *Agricultural typology, principles and methods : preliminary conclusions*, Boulder, Col., 1967, (mimeographed).
- [59] Kowalczyk, D., Użytkowanie ziemi w powiecie suwalskim (Land utilization in the Suwałki County), *Dokum. Geogr.*, 6, 1962, 41—77.
- [60] Lutovac, M., The village of Ritopek on the Danube in the suburban zone of Belgrade, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 235—266.
- [61] Matusik, M., Niektóre problemy użytkowania ziemi a rzeźba terenu na Pojezierzu Kaszubskim (Sum.: Some problems of land utilization and the relief of the Cassubian Lakeland), *Zeszyty Geogr. WSP w Gdańsku*, 7, 1966, 153—178.
- [62] Piskorz-Skocka, H., Użytkowanie ziemi w powiecie koszalińskim (Land utilization in the Koszalin County), *Dokum. Geogr.*, 6, 1962, 7—39.
- [63] Piskorz-Skocka, H., The commune of Miłogoszcz on the Pomeranian Baltic Sea Coast, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 157—194.
- [64] Remy, G., Les travaux polonais sur l'utilisation du sol, *Etudes Rurales*, 12 (1964), 134—136.
- [65] Sarfalvi, B., The village of Csepreg in Western Hungary, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 437—452.
- [66] Sarfalvi, B., (ed.), Land utilization in Eastern Europe, *Studies in Geography in Hungary*, 4, 1967.
- [67] Simon, L., Nyiradony — the village in North-East of Great Hungarian Plain, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 421—436.
- [68] Singh, H., Crop combination regions in the Malwa Tract of Punjab, *Deccan Geographer*, 3 (1965), 21—30.
- [69] Stamp, L. D., A report on land use studies in East-Central and Eastern Europe 1960—1964, *IGU Newsletter*, 15 (1964) (1/2), 37—40.
- [70] Stamp, L. D., Land use in Eastern Europe, *Geogr. Review*, 54 (1964), 115—116.
- [71] Stola, W., Gospodarka rolna w strefie podmiejskiej Warszawy; wieś Bielawa (Sum.: Rural economy in a suburban area on the example of Bielawa village), *Przegl. Geogr.*, 34 (1962), 173—184.
- [72] Stola, W., Kierunki użytkowania gruntów ornych w województwie białostockim (Sum.: The orientations in arable land utilization in the voivodship of Białystok), *Przegl. Geogr.*, 36 (1964), 143—155.
- [73] Stola, W., The commune of Czersk in the Warsaw suburban zone, in: Land utilization in East Central Europe, *Geogr. Polon.*, 5, 1965, 87—124.
- [74] Stola, W., Użytkowanie ziemi i stosunki gospodarcze w powiecie pińczowskim (Land utilization and economic problems in the Pińczów County), *Dokum. Geogr.*, 5, 1965, 3—38.
- [75] Stola, W., Agricultural typology of a mezo-region based on the example of Ponidzie (Central Poland), *Geogr. Polon.*, 14, 1968, 283—290.
- [76] Stola, W. and Jankowski, W., Użytkowanie ziemi i gospodarka rolna w Badacsonyomaj (Sum.: Land utilization and farming in Badacsony vineyard district; the case of Badacsonyomaj), *Dokum. Geogr.*, 1, 1967, 51—70.
- [77] Szczęsny, R., Gospodarka rolna w Beskidzie Niskim; gromada Cergowa (Sum.: Farming in Lower Beskid; village Cergowa), *Przegl. Geogr.*, 29 (1959), 630—644.
- [78] Szczęsny, R., Użytki rolne w powiecie mrągowym (Agricultural land uses in the Mrągowo County), *Dokum. Geogr.*, 1, 1950, 1—24.
- [79] Szczęsny, R., Próba określenia kierunków produkcji rolniczej w Polsce (Sum.: An attempt to determine the orientations of agricultural production in Poland), *Przegl. Geogr.*, 36 (1964), 143—155.
- [80] Szczęsny, R., The orientations in agricultural production of Poland, *Geogr. Polon.*, 2, 1964, 169—177.

- [81] Szczęsny, R., The Carpathian commune of Cergowa, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 61—86.
- [82] Szczęsny, R., and Stola, W., Rolnicze użytkowanie ziemi w międzyrzeczu Dunaju i Cisy na przykładzie wsi Harta (Sum.: Agricultural land utilization in the interfluvium of the Danube and Tisza; the case study of the village of Harta), *Dokum. Geogr.*, 1, 1967, 27—50.
- [83] Tyszkiewicz, W., The commune of Kruszwica on the Cuiavian Plain in Central Poland, in: Land utilization in East-Central Europe, *Geogr. Polon.*, 5, 1965, 125—156.
- [84] Tyszkiewicz, W., Z badań nad użytkowaniem ziemi we wschodniej Słowenii; Wieś Goričak (Sum.: Land utilization in Eastern Slovenia; the village of Goričak), *Dokum. Geogr.*, 2/3, 1966, 81—103.
- [85] Tyszkiewicz, W. and Szczęsny, R., Z badań nad użytkowaniem ziemi w gminie Svetina (Słowenia), (Sum.: Land utilization in Central Slovenia; the case of cadaster unit Svetina), *Dokum. Geogr.*, 1, 1967, 71—92.
- [86] Tyszkiewicz, W. and Szczęsny, R., Użytkowanie ziemi i rolnictwo na przykładzie gminy Izola (Istria), (Sum.: Land utilization in Istria; the case of cadaster unit Izola), *Dokum. Geogr.*, 1, 1967, 93—125.
- [87] Weaver, J. C. Changing pattern of cropland use in the Middle West, *Econ. Geogr.*, 30 (1954), 1—47.
- [88] Weaver, J. C., Crop combination regions of the Middle West, *Geogr. Review*, 44 (1954), 175—200.
- [89] Weaver, J. C., The county as a spatial average in agricultural geography, *Geogr. Review*, 46 (1956), 536—565.
- [90] Weaver, J. C., Hoag, L. P. and Fenton, B. L., Livestock units and cropland regions in the Middle West, *Econ. Geogr.*, 32 (1956), 277—289.
- [91] Kostrowicki, J., Problemy geografii rolnictwa (Problems of agricultural geography), in: S. Leszczycki, L. Kosiński (eds.), *Zarys geografii ekonomicznej Polski* (Outline of Economic Geography of Poland), Warszawa 1967, 147—182.
- [92] Kostrowicki, J. and Szczęsny, R., Rolnictwo (Agriculture), in: *Struktura przestrzenna gospodarki narodowej* (Areal structure of national economy), Warszawa 1969, 17—124.
- [93] Paulov, J., The agricultural belt-line of Košice, in: K. Ivanička (ed.), *The geography of the region of East-Slovakian ironworks*, Bratislava 1964, 231—256, and Land utilization map of the whole region, accompanying this volume.
- [94] Siddigi, M. F., Combinational analysis. A review of methodology, *The Geographer*, 14 (1967), 81—99.

THE QUANTITATIVE ANALYSIS OF LAND USE PATTERNS
WITH SPECIAL REFERENCE TO LAND USE MAPS:
SHAPE ANALYSIS, WITH AN APPLICATION *

CHRISTOPHER BOARD

In Britain and many parts of Europe work on land use has been dominated by a desire to complete an inventory of the use to which land is put. The greater part of this task has been concerned with rural, mainly agricultural, areas. This is understandable in small countries where the loss of agricultural land to other competitive uses may be important, or in other countries where the need to increase agricultural production is crucial. Considerations such as these have encouraged workers to complete the coverage of land use maps, with the result that land use studies have remained mainly descriptive. In some cases the very effort of accumulating information on land use and recording it on maps is so time consuming that any effective analysis of land use patterns depicted on the maps is at best perfunctory, or at worst is abandoned altogether.

By contrast, in the United States and in some large, developing countries, probably because of the difficulty of achieving a complete coverage of land use information on a uniform basis and in map form, work has concentrated instead on specimen areas and sample studies (Board, 1965). The U.S. National Inventory of Soil and Water Conservation Needs includes a sample survey of land use (Anderson, 1961). Several recent agricultural censuses in Africa have obtained land use information by means of sampling techniques. However, in the latter instances (for example in Lesotho) a map of land use was not required. Some workers have felt free to experiment with rigorous methods for analysing land use patterns. These include Haggett (1963, 1964), Latham (1963), Proudfoot (1942), Reid (1963) and Wood (1955). However, it is Berry's study (1962) on the land use of flood-plains that presents a commentary on sampling designs and simple statistical analysis in the context of the choice of methods of field work appropriate to rapid analysis and the storage of information in electronic computer

* The author thanks Dr. E. C. Willatts and Professor P. Haggett for reading and commenting on an early draft of this paper. The author has also benefited from suggestions made during discussion which followed the presentation of an earlier version of the paper at the Madingley Lectures, 1967. The author gratefully acknowledges the help of Miss S. Brown for advice on statistical matters, Miss J. Tagg and Miss D. Dubury for typing the manuscript and the staff of the Drawing Office for drawing some of the illustrations. Calculations were done on the computer at the London School of Economics.

systems. Significantly, while a point-sampling design will yield a substantial quantity of information capable of immediate analysis, he recognizes that it is no substitute for a land use map. The development of widely available aerial photography has made possible many more land-use studies, incorporating land-use maps of considerable parts of the earth's surface. Indeed the International Geographical Union already has a Commission on the Interpretation of Aerial Photographs which includes a working party which has collected data on rural land use (see *Photogrammetria*, 20, 1965).

There is now an increasing tendency for geographers to prefer a problem-oriented approach to land-use studies, rather than a comprehensive approach along the lines of a regional survey, which was prevalent hitherto. This change in methodology has emancipated geographers interested in land use from the servitude of completing laborious mapping tasks before attempting explanations of land-use patterns. It is now possible for them to concentrate instead on smaller areas and attempt a deeper analysis of processes resulting in particular land-use patterns in the framework of specific research problems. There is now a great range of rapid methods, many with appropriate measures of their reliability, which can be employed to collect land-use data. Furthermore, quantitative methods of analysis can be applied equally well to such data (Haggett, 1964; Reid, 1963; Vellin, 1965).

That studies of land-use pattern have so frequently treated area as the only variable may lie in the very complexity of the pattern itself. Because it is continuous and continually changing, with elements which interlock, land use pattern has been compared with *facies*, the changing lithology of the geologist (Wool-dridge, 1945, p. 395). Only where a category of land use is dominated by a "back-cloth" or "matrix" of another category is it generally possible to recognize isolatable, discrete fragments of that category. For the most part it is possible only to recognize trends in areas covered by particular categories. Recent studies of shape in geography (Boyce and Clark, 1964; Stoddart, 1965) have been concerned with the description of the outlines of distinct features like Central Business Districts or coral atolls. The description of land use pattern has therefore taken refuge in analogies and metaphors. Veatch in a study of patterns of swamp land in Michigan (1938) calls patterns dendritic and likens them to mazes or chains. Latham (1963, p. 195) refers to "fernlike shapes", "lakes on land" and "fiorded coast" to describe the pattern of cropland in Pennsylvania.

The land-use pattern depicted on maps at scales from 1 : 25,000 to at least 1 : 125,000 is at the first stage of aggregation from, or generalization of, the basic units of which it is composed. On such maps it is not possible to distinguish the working units whose use is determined by the farmer or the community's decision. The farms, estates or holdings, which represent the basic economic units, are rarely mapped together with land use. The large number of land-use maps now available and the wealth of information from vertical aerial photographs

invite the geographer to make as much use of them as possible. This temptation is more easy to yield to because of the geographer's traditional concern with macro-studies rather than process oriented, micro-studies.

Building upon studies in the fields of agriculture and forestry geographers have investigated alternative methods of accumulating land-use information. These have most often concentrated on the primary variable *area* to the exclusion of other attributes of land-use pattern. Quantitative "explanations" and regionalization of pattern are nearly always based on the *area* of land covered by a particular land-use category. In contrast other attributes of land use pattern (texture and orientation) have been ignored, possibly because of difficulties involved in describing them. It is unlikely that such work has been neglected for the reason that it is not important, because many students of land-use patterns have commented on the apparent differences in their texture and orientation. In his classic report on land use in the North Riding of Yorkshire, Wooldridge (1945, pp. 388—396) makes use of a measure of texture, the arable "site", qualified by adjectives large or small, to demonstrate a connection between land-use and soil conditions. "The recognition of characteristic and significant sites, nevertheless remains a prime objective of analysis even in mixed farming regions and serves to bring to light those areas where the physical characteristics of the lands enforce or strongly encourage a particular mode of utilisation" (p. 389). In another instance Board (1962), describing the pattern of land use in part of the Eastern Cape Province (South Africa), uses the term "block" for a continuous stretch of arable land, often composed of several "fields"¹, and furthermore attempts to distinguish between the patterns of land use in adjacent regions on a basis both of area of arable land and of the character of the outlines of blocks of arable land. The rectangular outline of arable land on farms occupied by German settlers is a result of 19th century property surveys, whereas the irregular outline of neighbouring African peasant-worked arable lands results from the aggregation of numerous small unsurveyed "lands" or fields, now usually surrounded by a stock-proof fence.

A glance at the terminology used by authors of the county reports of the first Land Utilisation Survey of Britain reveals the difficulty they faced when describing land-use patterns in words. Most authors make use of indefinite and subjective terms such as "dominant", "almost unbroken expanse", "by no means continuous stretch", "tiny isolated fields", "the ground-mass of the land use pattern" and "generally speaking arable cultivation tends towards concentration..." in this area or that. One author (Thomas, 1938, p. 38) in a moment of despair writes "the distribution of Permanent Grassland as mapped by the Survey can best be seen from the accompanying map . . . more clearly that it can be descri-

¹ "Fields" or "lands" (amasimi in Xhosa) are the working units of cultivation similar to strips in unenclosed arable fields in Europe.

bed". Although the map is the best descriptive vehicle for spatial patterns different map readers usually see the patterns in different ways.

This paper is one of several which have attempted to resolve this contradiction, in the belief that mapped land-use patterns represent a considerable storehouse of information which can never be collected again. If ways can be found to release some of the information imprisoned on the face of the map, land use studies will have acquired a valuable tool for research.

Although it is never possible entirely to eliminate the subjective element from the description of land use, it is possible to reduce it considerably. The employment of a set of techniques rigorously applied by different operators, the use of measurements of length, area and quantity rather than verbal descriptions should increase the objectivity of land-use analysis. As will be discussed later geographers have also failed to develop satisfactory, objective methods for defining land-use regions.

The object of the techniques put forward in this paper is to provide a simple set of indices which may be obtained by any map user to describe four attributes of land-use pattern. These are (1) area, (2) fragmentation, (3) intricacy and (4) directional orientation. If acceptable indices can be derived by some fairly rapid and simple procedure so that they predict realistic attributes of the pattern, this would be better than presenting *only* some new, reproducible indices which are more objective than verbal descriptions.

During the search for suitable objective methods, point samples were discarded because they yielded little information besides area. The line traverse sampling design was preferred because it seemed more versatile and promised to furnish not only data on area but indications of *other* attributes of land use pattern. In earlier work on land-use sampling Haggett (1963) observed that line sampling yielded successively more accurate results for a given expenditure of time, whereas point and area sampling were both inferior. A study by Latham (1963) made use of line traverse sampling to analyse the cropland pattern of the state of Pennsylvania, U.S.A., from a map on the scale of 1 : 1/2 million. In that case, however, the traverses were rotated to six different directions, greatly increasing the information on pattern and taking into account the effect of orientation on estimates of area. The results were given in terms of the frequency of *intercepts* (uninterrupted distances) of different length classes. Latham used two graphic methods to demonstrate the characteristics of the pattern of cropland: graphs of the accumulated percentage of intercepts of progressively greater length; and orientation roses of median and quartile intercept lengths for each of six directions. Latham, working at a relatively small scale and with cropland areas "one quarter mile or more in minimum dimension" (at least 0.4 km.) inevitably studied coarser-grained aspects of land-use patterns. For those areas where detailed, large scale, land-use maps are available the investigator can approach more closely the realities of the agricultural landscape and view it as a complex of

fields each used in a particular way. Continuous stretches of land under one use are readily interpreted as blocks of fields having the same use. The total number and size of fragments under one category of use in a unit area bear a direct relation to the "sites" referred to by Wooldridge (1945). Again where the alignment of field boundaries is fairly obvious, the orientation of a land use pattern can be discerned more precisely than by employing six arbitrary directions for traverse sampling. Furthermore the intricacy of the pattern is also closely related to the sizes of fields and the density of settlement.

Other indices of land-use pattern characteristics which are more appropriate to agricultural landscapes as detailed as those in Britain are also worth presenting. The "One inch" (1 : 63,360) maps of the Land Utilisation Survey provide an unrivalled source of information for an analysis of this kind, but experiments already carried out by the author on his land-use map (1 : 125,000) of part of the Eastern Cape Province suggest that there the scale and degree of detail also permit an analysis at a similar level.

Because it is the intention of this paper to explore attributes of land-use pattern other than that of area, the method of estimating area described by Haggett and Board (1964) was accepted in virtually unchanged form. That method was published as a comment on Latham's rotational traverse sampling method (1963) for obtaining data on area and orientation of land use patterns. It was pointed out by Haggett and Board that the progressive rotation of traverse sampling lines does not by itself accumulate a better estimate of area than does intensifying the sampling fraction by multiplying the number of parallel traverse sampling lines. In fact great gains in accuracy could be achieved by intensifying sampling to 5 miles per square mile (3.1 km./km.²). Latham (1964) in a reply comments that his method is not properly tested because his system of traverses along six directions involves a sampling intensity of 12 miles per square mile (7.5 km./km.²) and in any case is intended to extract the data which measure and describe shape, orientation and dispersion of areal patterns. Certainly, the rotational traverse method using an intensity of up to 6 miles per square mile (3.8 km./km.²) does not show anything like the same improvement in mean percentage error as that provided merely by intensifying parallel traverses without rotating them.

Although Latham's intensive technique probably yields more accurate estimates of land use area and orientation, it is too laborious to be useful as an exploratory, descriptive tool. The investigator of land-use patterns needs a versatile technique yielding estimates of several attributes derived if possible from a single set of measurements. If these are reasonably accurate and as precise as possible they may then be used to delimit regions of homogeneous land use (land-use regions) more objectively. Sampling by parallel traverses at an intensity of 4 miles per square mile (2.5 km./km.²) enables estimates of area with a mean error of less than 2% to be obtained. In the investigation that followed estimates of area were

not checked against more accurate methods of measurement. To this extent, when the attribute *area* is mentioned in the technique set out here, it has been estimated by parallel-traverse sampling.

FRAGMENTATION

Experiments were carried out on sheet 40 and 48 (Grimsby, Louth and Skegness) of the *One Inch Land Utilisation Survey of Britain*. A pilot investigation and graphic plot of results suggested that the technique was worth trying on

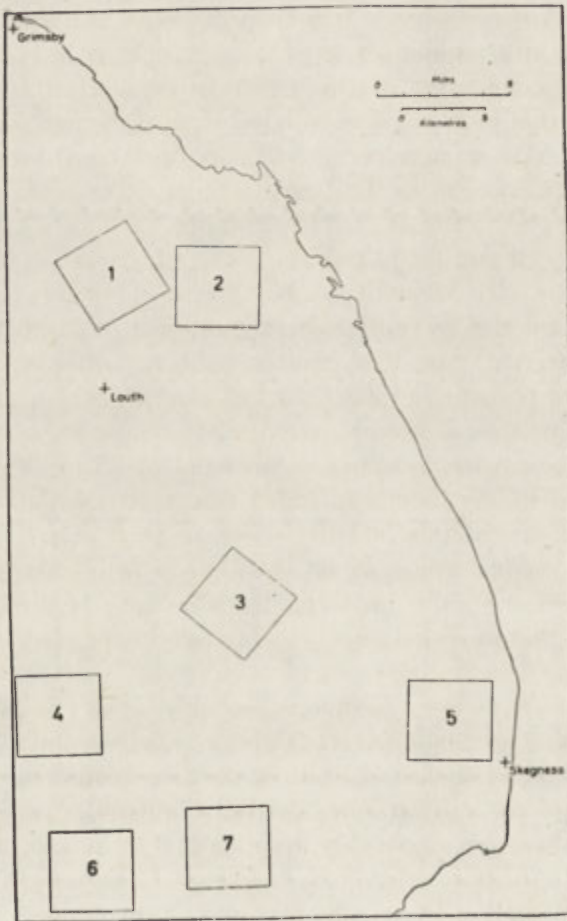


Fig. 1. Location of sample quadrats which provided land use data for the pilot survey

a larger scale ². The sheet selected included a good range of patterns in two major land-use categories: arable and permanent pasture. In order to test the applicability of the method in regions of rather different proportions of arable and grassland, areas characterised by different proportions and distinctive spatial arrangement were first identified visually. Next a quadrat approximately 10 square miles (26 km.²) in extent was placed within each area so that the land-use pattern within it was representative of the characteristics of the region as a whole (Fig. 1). Each square, a sort of quota sample, was sampled by parallel traverses at intervals of 1/4 inch (6.3 mm.) with an intensity of about 4.1 miles per square mile (2.6 km./km.²). The arable pattern and the intercepts where traverses crossed the arable pattern are shown in Figs. 2 and 3 respectively. Within each square the total length and number of all intercepts were ascertained. In addition for each square the number of separate fragments of arable (or pasture) was counted to provide a real measure of texture. The intricacy of the pattern was obtained by counting in the outline of arable land all angles of less than 160°. Furthermore the orientation of the pattern was found by placing a tracing of pattern of traverse lines over the pattern in such a way that the lines would cross arable land as much as possible in uninterrupted stretches.

For the more intensive study operators were shown several contrasting land use patterns on sheet 40 and 48 of the Land Utilisation Survey of Britain and asked to select the position and orientation of each of the four sample squares they studied. By deliberately increasing variations in the patterns studied, it can be assumed that the results are more representative of different conditions. For ease of working, the size of sample squares was reduced to 3 × 3 miles (4.8 km. × 4.8 km.) and parallel traverses a quarter of an inch apart (6.3 mm.) were used, to give a sampling intensity of 4 miles per square mile (2.5 km./km.²). For each sample square chosen by the individual operator both the arable and permanent pasture pattern were studied. This had the effect of widening the scope of the investigation to include different land-use categories with markedly different pattern types: the ground-mass of pasture with scattered small arable "sites" (Fig. 2 : 5); the dominantly arable peat-fen with very small patches of pasture (Fig. 2 : 7); and the largely interconnected, but dissected block of pasture broken up by a complex bundle of elongated strips and blocks of arable (Fig. 2 : 2).

A record was made on transparent paper overprinted with a tenth inch (2.5 mm.) grid of the total length and number of intercepts, where traverse lines crossed a fragment of the appropriate land use type. This was done for each sample square and both arable and pasture uses. The total number of distinct fragments of each category was also recorded. Blocks of the same category sepa-

² The arduous task of sampling 80 squares for both arable and pasture patterns was done by the author's practical class at the London School of Economics in the summer of 1966. Their help is gratefully acknowledged.

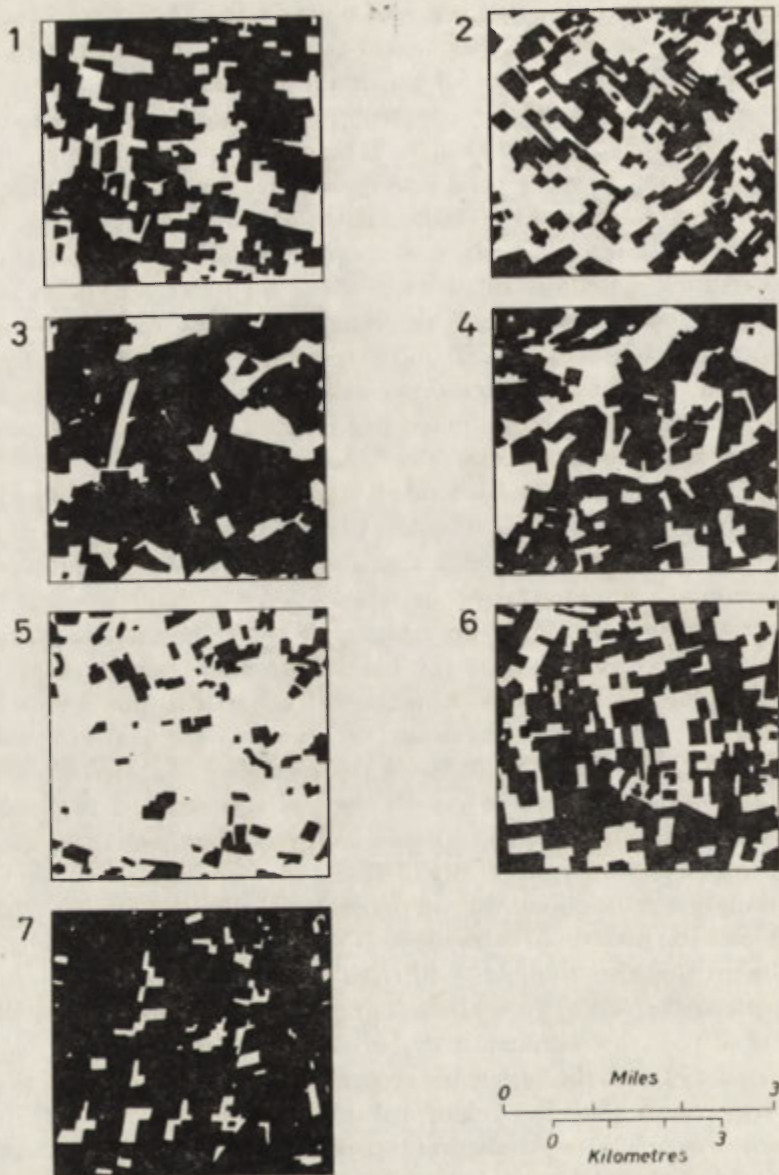


Fig. 2. The pattern of arable land in the sample quadrats studied in the pilot survey of land use pattern shape

rated only by a road or stream were regarded as parts of the same fragment. Blocks which touched only at their respective corners were treated as separate fragments. Estimates of intercept length were made to the nearest tenth of an inch (2.5 mm.).

The area covered by each land use category was computed from total length of intercepts as a percentage of total length of traverse lines related to the nine square miles (23.3 km.²) of the sample square. At the scale of one inch to one mile (1 : 63,360) the formula is:

$$\text{Area of land use category in square miles} = \left(\frac{l}{36} \times 100 \right) \times \frac{9}{100} \text{ where } l \text{ is the total length of intercepts in inches. Or expressed in metric units } \left(\frac{l \text{ mm.}}{914.4 \text{ mm.}} \times 100 \right) \times \frac{23.31}{100}.$$

In order to predict the number of fragments into which the particular land use category was broken it was necessary to appreciate the relationship between the number of intercepts and the number of fragments. The former represents changes in land use observed when scanning the sample area. The latter are the blocks of uniform use which Wooldridge called "sites". It is quite possible, for instance, (Fig. 2.3 and 2.5) for a pattern of a few large blocks and one of many small fragments to possess the same number of intercepts. In the former case intercepts are long and the gaps are small, in the latter intercepts are short and the gaps long. In general, large areas of one type of land use occur in few large fragments and small proportions in many small fragments. Other things being equal numerous tiny fragments, by being surrounded by zones of land use of another kind, must take up less room on the land surface. Since it is the fragments (or the units (fields) which comprise them) that are basic to the variation in the agricultural landscape, the area covered by a particular kind of land use should be used only as a weighting factor in predicting fragment number. *Area* alone cannot "explain" fragment number, but *area per intercept* or *average length of intercept*, to which it is related by a constant, has been found to predict fragment number reasonably well. When plotted on arithmetic graph paper, fragment number plotted against acres per intercept produces a curvilinear relationship. With a double logarithmic transformation the curve was straightened out and the product moment correlation coefficient became $r = .932$ for arable, $.866$ for pasture and $.902$ for both arable and pasture. All subsequent calculations and data refer to observations made on both the arable and pasture pattern.

Since the aim was the prediction of fragment number, regression analysis was carried out on the transformed data. The best fit regression line took the form:

$$\log Y = 4.3804 - 2.0166 \log X$$

where Y is *fragment number* and X is *acres per intercept*. Some 87 of the total variation is "explained" by the correlation and, ignoring the standard error of X , which is very small, the standard error of $\log Y = .25$. At the 95% confidence level one can predict fragment number to within ± 4 or thereabouts. One

TABLE 1. Measurement and indices obtained from a pilot study of arable pattern on Sheet 40 and 48 of the *Land Utilisation Survey of Britain*

(a) 13 traverses totalling 40.7 in. (1,033.8 mm.) over 9.8 square miles (25.38 km²)

Quad-rat No.	Num-ber of inter-cepts	Per-cent of total line length	Area in Sq. Miles	Area in km ²	Area in Acres	Number of Frag-ments	Number of Corners	Sq. Miles per intercept	km ² per intercept
1	71	61.4	6.11	15.82	3910	6	395	.0861	.2230
2*	112	42.2	4.13	10.70	2643	46	580	.0369	.0956
3	49	76.4	7.48	19.37	4787	3	224	.153	.3962
4	83	57.6	5.65	14.63	3616	6	332	.0681	.1764
5	49	15.7	1.54	3.99	986	57	317	.0314	.0813
6	87	57.5	5.64	14.61	3610	12	439	.0648	.1678
7	53	88.9	8.71	22.56	5574	2	301	.1645	.4260

(b) 18 traverses totalling 80.46 in. (2,043.7) over 20 square miles (51.8 km²)

1	125	54.2	10.84	27.97	6937	19	not	.0867	.2245
2	177	39.5	7.9	20.46	5056	82	counted	.0446	.1155
3	124	62.8	12.56	32.50	8038	20102	.2642
4	126	55.6	11.12	28.80	7116	120883	.2287
5	92	16.8	3.36	8.70	2150	1000365	.0945
6	152	54.2	10.84	27.97	6937	380714	.1749
7	100	81.6	16.32	42.22	10444	3163	.4222

* For this region, where visual inspection suggested that the pattern might well be most strongly oriented in one of two directions, roughly at right angles, 93 intercepts were recorded at a bearing of 49½° and 84 intercepts at a bearing of 139½°. The latter was tentatively accepted as being the direction of strongest orientation.

of the features of the scatter graph (Fig. 4) is the relatively large number of observations where the fragment number is one and which occur in a fairly wide range of area values. Because it is quite easy, by visual inspection, to see whether the land use pattern under consideration forms one connected fragment in the sample area, it was thought possible that the omission of observations where there was a single fragment would not matter and might indeed improve the predictive power of the regression. When this was done the best fit regression line took the form:

log Y = 4.0911 — 1.8099 log X

Although only about 76 percent of the total variation was now “explained” the standard error of log Y was reduced to .205, or slightly more than ± 3 at the 95% confidence level. It is very easy to make mistakes when counting the number

of fragments on a map of land use, particularly when they are numerous or when the pattern has a maze-like complexity. Hence a technique which will allow fragmentation to be estimated within such limits is useful, particularly when it is but one by-product of a method of assessing area.

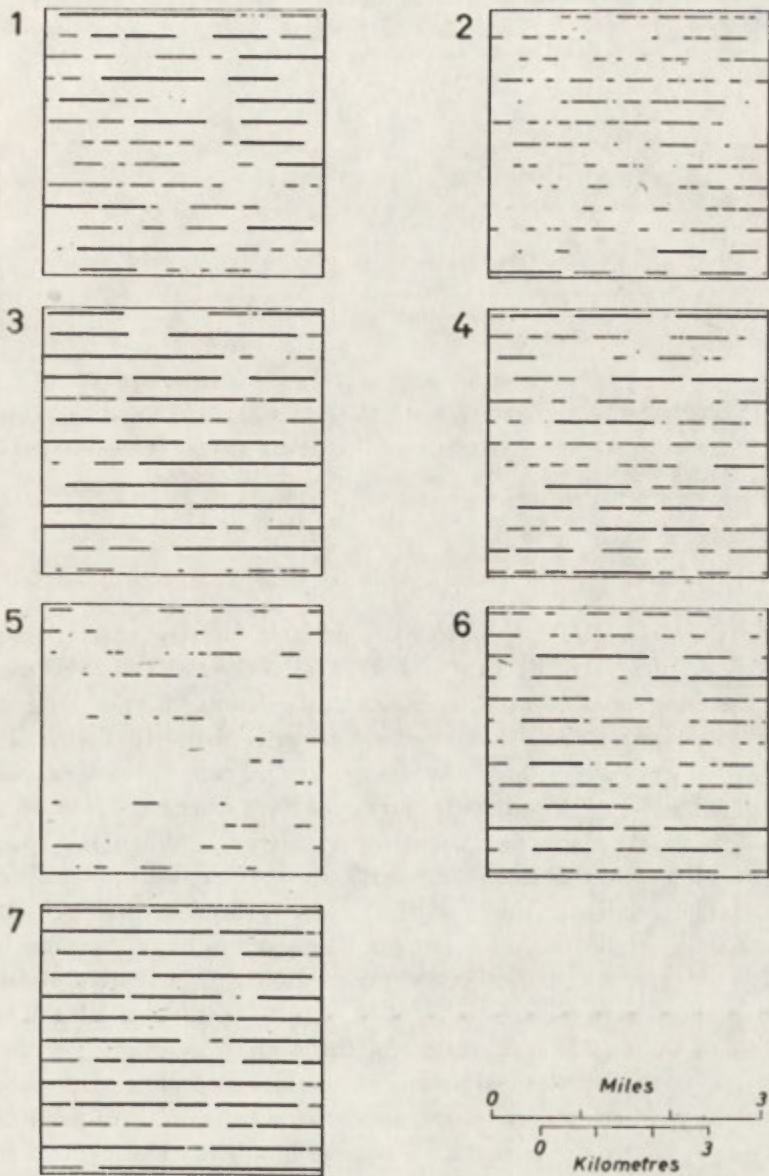


Fig. 3. Arable intercepts resulting from line sampling of the arable pattern in the pilot survey

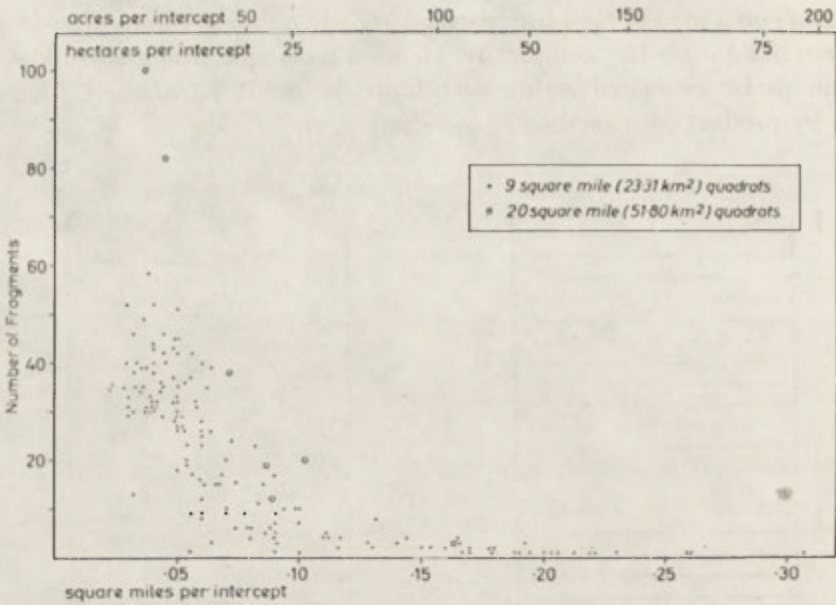


Fig. 4. Graph of the relationship between area per intercept and number of fragments of the same land use category, for two sizes of sample quadrat

INTRICACY

Following the description of the texture of land use patterns by estimating fragmentation, the degree of detail or intricacy of the pattern may be measured in terms of the number of corners or angles in the land-use type being studied. In a pilot investigation of the seven sample areas depicted in Figure 2 angles greater than 160° were disregarded. An arithmetically scaled graph of *number of intercepts* against *number of corners*, with 13 traverse lines spaced at $1/4$ in. (6.3mm.), suggests that there are about five corners per intercept. When the spacing of traverse lines for the same area was at $3/10$ in. (7.6 mm.) there were about six corners per intercept. The minimum number is, of course, three corners per intercept when traverse lines are far enough apart for fragments to be intercepted by only one line. This index, unlike the others, depends more on the system of sampling used. By enlarging the sample area one would expect to increase both the number of intercepts and the number of corners in the same proportion. On the other hand the effect on the estimation of fragment number of doubling the sample area is to yield a lower fragment number for the same area per intercept. It also, of course, yields much greater fragment numbers. The general form of the relationship, however, remains the same as that established for a 3 mile \times 3 mile (4.8 km. \times 4.8 km.) sample square.

Lastly an index of the orientation or "graining" of land-use pattern was developed. For areas of woodland which were mainly discrete Haggett and Board (1964) assessed orientation by comparing highest and lowest estimates of area from rotational traverses. Indeed it is difficult to imagine any way of assessing orientation without employing rotation of sampling lines or grid. Latham's method automatically yields information on orientation because it depends upon rotating traverse lines. It is here contended that the main lineaments of a land use pattern can be discerned by visual inspection, but that the precise direction of an orientation may not be so easily obtained. It is proposed that a sample square of exactly the same dimensions and with the same number of traverse lines as used for the other attributes of pattern be centred on the original mid-point of the sample area. It should then be rotated until the observer is satisfied that the traverse lines intercept as much arable (or other land-use type) as possible. The intercepts are then traced and counted. The smallest possible number of intercepts will occur when the traverse lines lie in the direction which coincides with the strongest orientation of the arable (or other) pattern. To check that the minimum number of intercepts has, in fact, been recorded, the parallel traverse lines can be rotated 10° each way from the original orientation, or can be checked by further trials. There quickly comes a point where the gain in accuracy is far outweighed by the extra labour of tracing and counting intercepts. In this respect a rotational traverse sampling system is superior because it always yields a minimum number of intercepts. However, it is wasteful by not making maximum use of preliminary visual study of the pattern and has the disadvantage that orientation is limited to the directions arbitrarily decided upon before the investigation begins. Neither of the other two methods gives results more accurately than to the six directions 30° apart. The merit of the technique suggested here is that it is more rapid and less laborious and furthermore bears a simple relationship to measures used for other attributes of pattern.

It is not difficult to imagine applications for these descriptive indices. At a fairly coarse-grained scale and over a wide area it would be possible to employ such indices as fragmentation and intricacy in a trend-surface analysis. In that way major trends in the pattern may be sifted from the local irregularities. Thus it might be possible to take Stamp and Willatts' generalisations (1935) about the transitions observable in the land-use pattern of part of North-east England, describe them more precisely and seek out local areas where the trends are not maintained. The measure proposed for orientation would probably not be suitable for such a study of trends and anomalies, except that the strongest orientation would, of course, be indicated. Second and third order trends in orientation might better be obtained from Latham's method.

Another application of these relatively objective measures lies in the field of regional delimitation. It has long been accepted that land-use regions are an important objective of study, providing a framework for explanation and for

studies of agricultural processes on specimen farms. Delimitation, however, has usually been achieved by subjective criteria, generally the visual appearance of land use patterns, backed occasionally by assessments of the area occupied by particularly significant categories of land use. In many cases it has been acknowledged that delimitation purely on a basis of land use itself has been difficult. Marshall's study of Oxfordshire (1943, p. 225) makes just this point and Davies in her study of Pembrokeshire (1939, p. 148) writes "in view of the commingling of types of land use it may at first appear that there is no adequate basis for the definition of distinctive land use regions" and goes on to invoke criteria from physical geography to make the division. The present author, faced with similar problems in the Eastern Cape Province of South Africa, preferred to distinguish between regions with superficially similar patterns, when they were known to have fundamentally distinct tenurial, ethnic and economic characteristics.

This same region in South Africa provides a simple example to show how at last two descriptive measures may be used to simulate regional delimitation based on one crop pattern. One of the most significant and productive regions in the mid-1950's was that around East London, which concentrated on pineapple cultivation. In the original report on land use in the Border Region (1962) the author based the delimitation of the "Pineapple Belt" upon the high proportion of land under pineapples. Visual inspection of the map of pineapple cultivation (Fig. 5) suggests that the region dominated by pineapples lies south-west of East London in an arc some 9 or 10 miles (15—16 km.) from Kidd's Beach. To confine the pineapple belt to that area, however, leaves out many other areas where that fruit was cultivated. Even in the main area it is by no means easy to decide upon the exact position of the boundary. Should it include only the large fragments between the Gulu and Chalumna rivers and those immediately to the east of the Gulu? Should the small detached concentration half way between the Gulu and Buffalo rivers be included and the less prominent group of lands south-east of the Laing Dam? What should be done with the extensive areas and concentrations east of the Nahoon river? The boundary maker is often confronted by questions such as this. The fact that different individuals tend to approach the task of delimitation in different ways suggests that chance plays a part in drawing boundaries on maps. If this process can be regarded as stochastic in character, it may perhaps be simulated with the help of a Monte Carlo model similar to that used in studies of diffusion (Hagerstrand, 1967). By deriving the probability of an area being included in the pineapple region from attributes of the land-use pattern suggested in this paper regional delimitation may at least be approached from a common base.

It is commonly accepted that it is relatively easy to recognise the core of a region, where its most typical characteristics are to be found. In the simulation of regional delimitation presented below, two core areas are employed. The main one lies inland from Kidd's Beach and a smaller one north of East London.

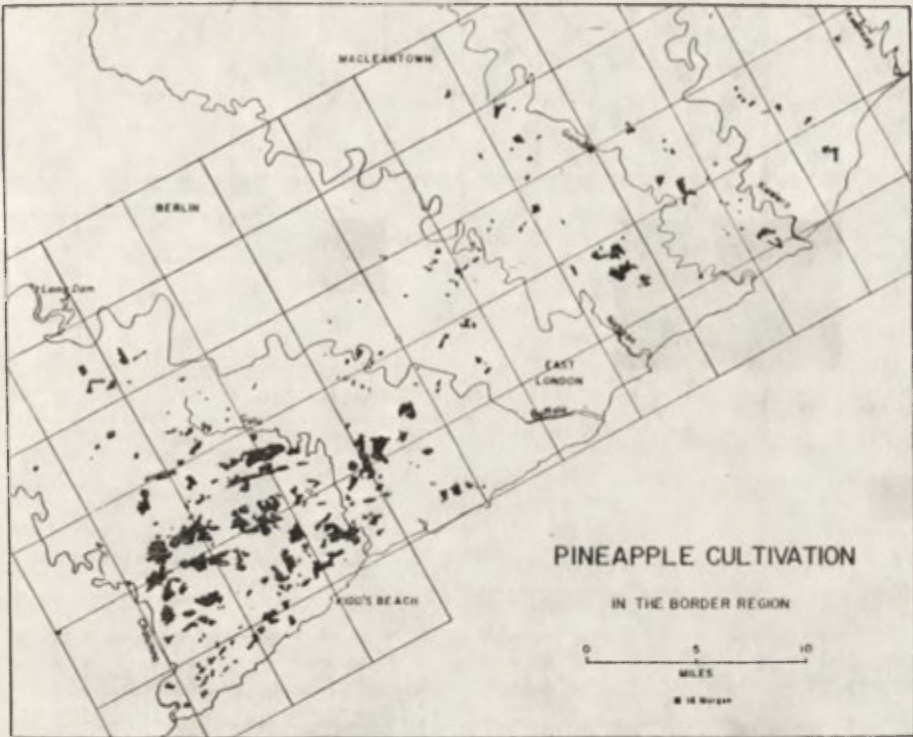


Fig. 5. Relationship between the grid used for land-use pattern measurements and the pattern of pineapple cultivation in part of the Eastern Cape Province, South Africa

A grid of 4 miles \times 4 miles (6.4 km. \times 6.4 km.) squares was laid over the area which had pineapple cultivation, in such a way that the core regions were concentrated in as few squares as possible. Area of pineapples and number of fragments of pineapple land provided the weightings for squares outside the core areas. This model differs from a diffusion model in the sense that distance from the centre of diffusion is built into it only by demanding that the core areas may grow only by adding *contiguous* squares³. The probability Q_i that any square i be added to the cores is given by:

$$Q_i = \frac{A_i \cdot F_i}{\sum_i^n A_i \cdot F_i}$$

where A is the area of pineapple land in a square and F the number of fragments into which it is broken, and n the number of squares under consideration (only

³ Contiguous includes those squares touching only at corners.

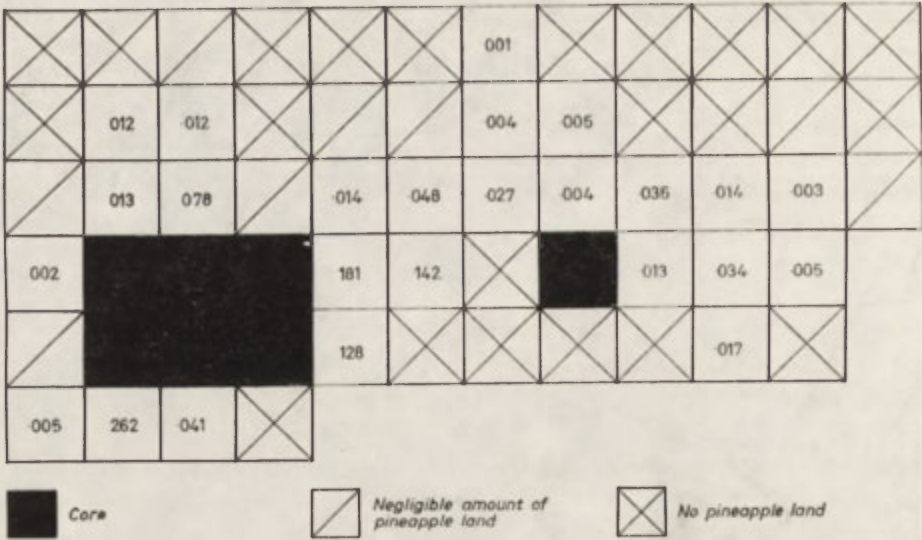


Fig. 6. Probability field based on the pattern of pineapple land in Figure 5; used for Monte Carlo simulation of regional delimitation

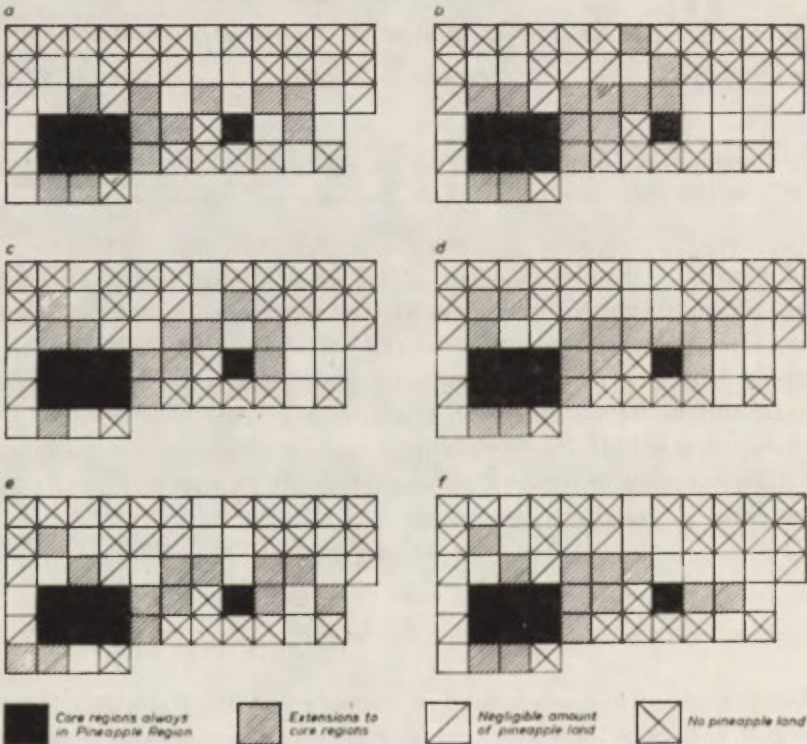


Fig. 7. Six different simulations of regional delimitation using data on the pattern of pineapple land in Figure 5

The diagonally shaded squares indicate cells in the grid which form part of the Pineapple Region after simulating fifty inspections of the pattern, to see what parts of the area around the core regions should be included

those with some pineapple land). In allocating the ranges of random numbers to cells, eight squares had probabilities lower than 0.0005 and so were disregarded. The probability field outside the core areas is presented in Fig. 6. In each of six sample simulations of regional delimitation, 50 random numbers, representing fifty decisions taken after visual inspection of the pattern, were used. Once a square was included within the region, random numbers selecting it again are regarded as wasted. The perimeter of each new region appears different, but there are common features in the squares whose probability of being selected is high. The effect of employing a Monte Carlo model to simulate regional delimitation is clearly seen in Fig. 7, where differently shaped regions emerge after each simulation run. It is contended that this may well represent the approaches of several different individuals to boundary making in spite of the fact that land use pattern is the basis of delimitation.

Once one region is delimited attention may be turned to an adjacent one. If the same probability approach is adopted, the second region can be delimited in the same way. It is one matter to delimit one region separately, but quite another to divide two regions. However, by making use of the way each region "grows" as a result of a set number of "inspections" (ten random numbers represent ten inspections), each region can be made larger in stages. The peripheral zones around two neighbouring regional cores will at some stage begin to meet or intersect. The inter-regional boundary may then be drawn along an indifference zone or at some other median position where the characteristics of the respective cores are equally weakly represented.

London School of Economics

BIBLIOGRAPHY

- [1] Anderson, J. R., Toward more effective methods of obtaining land use data in geographic research, *Professional Geographer*, 13 (6), (1961), 15—18.
- [2] Berry, B. J. L., *Sampling, coding storing flood plain data*, U. S. Department of Agriculture Handbook No. 237, Washington D. C., 1962.
- [3] Board, C., *The Border Region : natural environment and land use in the Eastern Cape*, Cape Town 1962.
- [4] Board, C., Field work in geography, in *Frontiers in geographical teaching*, ed. R. J. Chorley and P. Haggett, London 1965.
- [5] Boyce, R. R. and Clark, W. A. V., The concept of shape in geography, *Geogr. Review*, 54 (1964), 561—572.
- [6] Davies, M., Pembrokeshire, Part 32 of the *Land of Britain, the report of the Land Utilisation Survey of Britain*, ed. L. D. Stamp, 1939.
- [7] Hägerstrand, T., On Monte Carlo simulation of diffusion, in *Quantitative Geography*, part 1, *Northwestern Studies in Geography*, No. 13, 1967, ed. W. L. Garrison and D. F. Marble.

- [6] Haggett, P., Regional and local components in land-use sampling: a case study from the Brazilian Triangulo, *Erkunde*, 17, (1/2), (1963), 108—114.
- [9] Haggett, P., Regional and local components in the distribution of forested areas in South-East Brazil: a multivariate approach, *Geogr. Journal*, 130 (1964), 356—380.
- [10] Haggett, P. and Board, C., Rotational and parallel traverses in the rapid integration of geographic areas, *Annals Ass. Am. Geogr.*, 54 (1964), 406—410.
- [11] Latham, J. P., Methodology for an instrumented geographic analysis, *Annals Ass. Am. Geogr.*, 53 (1963), 194—209.
- [12] Latham, J. P., Comment on a note by Haggett and Board, *Annals Ass. Am. Geogr.*, 54 (1964), 410—411.
- [13] Marshall, M., Oxfordshire, Part 56 of the *Land of Britain*, the report of the Land Utilisation Survey of Britain, 1943.
- [14] Photogrammetria, 20 (1965).
- [15] Proudfoot, M. J., Sampling with transverse lines, *Journ. Amer. Sta. Ass.*, 26 (1942), 265—270.
- [16] Reid, I. D., *The application of statistical sampling in geographical studies*. Liverpool University, M. A. thesis, 1963.
- [17] Stamp, L. D., and Willatts, E. C., *The Land Utilisation Survey of Britain: an outline description of the first twelve one-inch maps*, 1935.
- [18] Stoddart, D. R., The shape of atolls, *Marine Geology*, 3 (1965), 369—383.
- [19] Thomas Alwyn N., Glamorgan, Part 31 of the *Land of Britain, the report of the Land Utilisation Survey of Britain*, 1938.
- [20] Veatch, J. O., Patterns based upon the distribution of swamp land in Michigan, *Papers of the Michigan Academy of Science, Arts and Letters*, 24 (Geography), 1938.
- [21] Vellin, J. L. C., *Land use in Mauritius*, London University, M. Sc. thesis, 1965.
- [22] Wood, W. F., Use of stratified random samples in a land-use survey. *Annals. Ass. Am. Geogr.*, 45 (1955), 350—367.
- [23] Wooldridge, S. W., The North Riding of Yorkshire, Part 51 of the *Land Utilisation Survey of Britain*, 1945.

RANDOM SAMPLING TECHNIQUES IN SOCIAL GEOGRAPHY

H. BRIAN RODGERS

We are all aware that geography, like sociology, psychology and the other behavioural sciences, is passing through a quantitative revolution. This is a revolution in technique to which there are two sides, firstly the collection and secondly the processing of data. To take the latter first, geographers have borrowed and developed the techniques of applied statistics, in order to derive a more precise meaning from the bodies of distributional data that they handle. This has led geographers through simple tests of association, to more complex multi-variate analyses and ultimately to the intricacies of model-building. But no amount of elaborate processing will get a clear meaning from unclear data, and some geographers have applied themselves to the prior problem of exact measurement, the collection of precise, clear and carefully defined statistical information. This almost inevitably involves some form of "sampling" operation.

THE NEED FOR SAMPLING IN SOCIAL GEOGRAPHY

It may be thought that in any country with an efficient national census the geographer need hardly concern himself with social measurement. This is — or should be — done for him. But almost any census, by its very nature, has many weaknesses. It is the sad fate of censuses to be obsolete before they are published. The last volumes of the British census of 1961 were not published until 1966: these were some of the most interesting and significant of the tabulations — internal migration and the journey-to-work, especially — but by the time they were printed they were evidence for the social historian. Our new 10% sample census, held mid-way through the ten year interval between full censuses, ought to provide quicker results: indeed the 1966 tabulations are already partly published. But unfortunately the census becomes obsolete most quickly in precisely those areas for which its results are needed most urgently as evidence for the shaping and re-casting of planning policies: in the New Towns, Expanded Towns and other "overspill" sites and growth points¹. The social situation in these areas can

¹ "Overspill" is used here in its most general sense, that is the planned transfer of population from congested cities and conurbations, either to "short-range" sites which take the form of peripheral estates built a little beyond the city's legal boundaries or to "long-range" sites in the form of New Towns or "Expanding Towns" growing under the provisions of the 1952 Town Development Act.

change radically in two years and be transformed beyond recognition in ten. If for no other reason, social sampling is needed to provide current data, weeks rather than years old.

The British census (like almost all others) is not only out-of-date on publication: it is also geographically insensitive in its application to the social patterning of towns. Until 1961 the census provided most of its more useful information only for entire local authority areas: simple population and housing data were given for wards (electoral divisions of towns, with populations ranging from several hundreds to many thousands) but virtually all the data that had a bearing on the functioning of the town as an economic and social entity were unavailable at this local level. There came a great improvement in 1961 with the availability (given prior notice and at substantial cost) of the full range of census data for Enumeration Districts. These are census units, each with a population in the range 750 to 1,000. Here, it may be thought, is social material related to a close mesh of areal units and thus likely to reveal the structure of the urban community in fine detail. But even the Enumeration District data have limitations. The districts themselves are not intended to be socio-geographical units: often they cut right across the pattern of social contrast in towns. And unfortunately the availability of Enumeration District data coincided with a change in census technique. In 1961 much of the most significant social and economic data (for example occupational, educational and journey-to-work material) was gathered not from the entire population but from a 10% sample of households. This means that this material in a typical Enumeration District is based upon responses from perhaps only thirty households, far too small a sample to be valid (especially in sub-divided form) for some of the purposes to which one would like to put this material. Moreover the 10% data are known to contain some bias. In short, for such tasks as a study of journey-to-work patterns within a town the census is almost useless; and only a social survey, using a geographically rational pattern of sampling sub-areas, can provide the basic data.

Lastly, there is much that the census cannot or does not attempt to record. For example the British census has only once included data on religion (1851). This is by no means merely an academic weakness. In many New Town and Town Development projects, especially in the strongly Catholic north-west of England, a large minority — or even a majority — of the incoming population is of Roman Catholic faith. Their birth-rate and their average family size are self-evidently higher than in the national or regional population as a whole. But how much higher? Since we ask no religious question in the census, we do not know, with any precision, what the “excess” birth-rate is within the Catholic community: still less do we know how this excess varies with age, income or occupation, or other social variables. This means that it is possible to attempt only an approximate population projection for a strongly Catholic New Town. Nor can the demand for houses of various sizes be evaluated (since the family size character-

istics among Catholics are unknown) and calculations of future demand for school places must be inaccurate in the absence of an exact measure of fertility among Catholic families of different ages. In short, all the social mathematics on which the development plan for a New Town depends are reduced to broad estimates in the absence of a crucial item of demographic information. Again, the gap can be filled only by a social survey. Many other types of information necessary to the planner fall into the same category. The census provides no material on shopping patterns, recreational habits or, prior to 1966, on car ownership and use. Each of these is involved in an understanding of the way in which the urban society operates, within the framework of the urban environment. Without the data one is more than half blind in trying to see those patterns of movement within towns which are the clearest indicators of the functioning of the urban machine.

None of this should be regarded as criticism directed peculiarly at the British census. Though it shares most of the weaknesses which are inseparable from census-taking, it is probably better than most: it gives both a very full range of social and economic data, and it provides all of this, on request, for smaller unit areas than most censuses in comparable countries. Moreover, there is reasonable hope that future censuses of Great Britain will take fuller advantage of the immense power of the modern computer to sort and index data. There is a distinct possibility that data will be tabulated for 100 metre grid-squares, as well as for the arbitrary local authority framework. Perhaps, too, the fact that individual responses to the census questionnaire become completely anonymous once coded on computer cards or tape may lead to the release of data at the level of the individual or the household. This would open up quite new horizons in the social analysis of census data. But however the census may be improved it must always remain a very general multipurpose economic and social stock-taking. More specialised enquiries must therefore continue to depend, in whole or part, on social surveys.

WHAT SORT OF SAMPLE? SIMPLE RANDOM, STRATIFIED, CLUSTERED AND QUOTA SAMPLES

To be of the slightest use a sample must be entirely representative of the total population from which it is drawn: it must not be biased towards any components within, or characteristics of, that population. Only a perfectly random sample, of adequate size, drawn from a complete and up-to-date sampling frame will meet these requirements. It is axiomatic that in any random sampling operation each member of the total population has an equal probability of selection to become part of the sample. It therefore follows that the sample must be chosen randomly from a complete and accurate list of the individuals (or households)

who form the population. Merely by applying a table of random numbers to such a list (the sample frame) a perfectly random (and therefore representative) sample is achieved ².

This is commonly known as a simple random sample. It has little or no application in social geography, unless one has a "captive" population neatly listed (for example the labour force of a factory for which one wishes to construct a journey-to-work pattern). In Great Britain (and in most other countries) there is no simple, up-to-date list of the individuals in the population. The nearest approach is the register of electors. This is never quite accurate, and it covers only voters, i.e. adults of 21 years and over ³. It is a poor sampling frame for the total population, but perhaps an acceptable one for a sample of households. But, increasingly, there are households in which neither husband nor wife has yet reached 21. These would be missed by a random sample based on the electoral registers, and the families left out are presently or potentially the most fertile in the population, with obvious effects on the demographic accuracy of the survey results. For want of a better alternative, the electoral registers have become a widely used sampling frame in Britain, one suspects because of their easy availability. In some senses the local authority rating lists (in which all residential dwellings are listed by address) are a better frame for the sampling of households, though in this case there are problems when two or more families share a single dwelling.

There is not space to elaborate on all the practical and theoretical difficulties which follow from any attempt to construct a simple random sample from an inadequate frame. Indeed there is no need to consider these problems in detail, since the fundamental purpose of any social survey by a geographer (the investigation of areal contrast) provides both a more acceptable sampling frame and an easier technique. A stratified sample, using as the basis for stratification a framework of geographically rational sub-areas, is likely to yield both a technically better sample and one which is geographically more meaningful.

The stratified random sample has always been an attractive alternative to the simple random sample. In stratifying one divides the population to be sampled into "layers", for example age-groups or income-groups, and each "layer" is sampled randomly. Each stratum — let us say each age-band — is thus given its correct representation in the total sample: if 15% of the population fall in the age-band 0—10 years, then this stratum is "allocated" 15% of the total sample. In short, the sample is designed to be representative in one important respect,

² An excellent account of the principles of random sampling is given in Moser, C. A., *Survey methods in social investigation*, London 1958. Many geographical applications are reviewed in Gregory, S., *Statistical methods and the geographer*, London 1963. Both these authors develop the mathematical basis of sampling, which is largely ignored in the present paper.

³ Recent proposals to reduce the voting age to 18 in the United Kingdom will improve the reliability of the electoral registers as a sampling frame for households.

and so one possible source of bias is eliminated. In general, its standard error will be lower than for a simple random sample.

The principle of stratification can be adopted in most types of geographical application. In urban survey work, the structure of the city itself provides a ready-made framework for stratified sampling. Most British cities are fairly sharply divided into sub-communities, which differ from each other in income level, occupational type, educational standard and many demographic qualities like fertility and family size. Type of housing provides a crude guide to these community contrasts: the small terraced houses of the nineteenth century, the modern municipal housing estates of the interwar and post-war periods, the privately owned houses of intermediate size, the large, detached houses of the highest income groups — each of these is occupied by a distinctive sub-community. Thus a type of housing map may provide an acceptable (if rather crude) framework for a stratified sample. By taking each type of housing area and randomly sampling, say every tenth household within it, one has assembled a sample which is certainly representative of house type and so — since type of house is socially diagnostic — a sample which ought to be acceptably representative of the full population in respect of many other socio-economic characteristics. An application of this sort of sample design is considered in detail in a later section.

It may be argued that type of house is too uncertain an indicator of community contrast to be used validly as a sample frame. It is possible (though again with certain reservations) to use the value of housing as the basis of the social framework. Each house (or apartment) in England and Wales is assessed for "rates", a local tax based upon the annual rental value of the premises⁴. With some labour, maps of the variation in the rateable value of houses may be constructed for towns and the resultant pattern may then be used as a sample frame. If house-value is a better clue to social contrast than house-type, this should give more representative results. A much more complex approach to the problem of constructing an urban sampling frame is to adopt the Shevky-Bell system of social area analysis developed in the U.S.A. and recently applied to a British case study⁵. This uses several socio-economic indicators in combination (not a single index like house type or value) to distinguish a pattern of community contrast. It should, therefore, at least in theory, provide an almost ideal frame for a stratified sample. But the time and labour involved (and the doubts that attach to some of the social assumptions made) prohibit the use of the Shevky-Bell approach merely in order to arrive at a sampling framework.

Given an adequate sampling frame, whatever its basis, the process of drawing a random sample (of, let us say, 10% of the total) is easy enough. Each unit-area

⁴ For an account of the use of rateable values in Urban Geography, see: Herbert, D. T. and Williams, W. M., Some new techniques for studying urban sub-divisions, *Problems of Applied Geography* II, *Geogr. Polon.*, 3, 1964, 93—117.

⁵ Herbert, D. T., Social area analysis: a British study, *Urban Studies*, 4 (1967), 41—60.

within the frame becomes a sampling sub-area, in which a randomly chosen 10% of the population is selected. This selection could be made by applying a table of random numbers to an address list: it is much simpler merely to adopt a standard sampling interval (every tenth household) and to allow the interviewer to construct his own sample, using this interval and accepting no substitutes, as he follows a route round the sub-area which includes every address within it. There are theoretical objections to such an "interval sample": no neighbours are ever selected, and once the starting point has been chosen, arbitrarily, it is no longer true that every household has an equal chance of selection. But the British census uses this method and it works well enough in practice.

To reduce the costs of the operation, or to give a better "depth" to local samples, the "cluster" principle may be adopted as a modification to this type of stratification by social area. This involves a two stage sampling process: firstly the selection of a representative sample of sub-areas and secondly the selection of a random sample of households (or persons) within each chosen sub-area. For example, if a map of house-values were being used as the sample frame, it might be thought unnecessary to sample each individual sub-area in a value category, but instead to make a random choice of a number of sub-areas, within which, alone, random sampling would take place. Each stratum in the sample, therefore, is composed of a number of clusters not spread over all the sub-areas in a given valuecategory but confined to a few. Obviously there is a problem: each stratum (that is, each class of houses falling in a particular value range) must be sampled in such a way that it makes a correct pro-rata contribution to the total sample, and this may need rather arbitrary adjustment to achieve. Usually, the cluster principle is applied to small samples of large populations. If it were proposed to take a sample of 1,000 people in a city of 1 million inhabitants (a sampling fraction of only 0.1%) it would be very laborious and expensive to attempt interval sampling, selecting an individual from every 1,000th household. There would be much to be said for making a 10% selection of house-value sub-areas, and then a 1% sample within those areas.

In every form of random sampling — simple, stratified or clustered — it is crucially important that an interview be secured from every sample address, or at least that the "no-response" proportion is kept to the minimum. If interviewing is confined to the ordinary working day, and if a high "no-response" fraction is accepted (more than 10% of the sample) very serious bias is introduced. The non-responding households will tend to be those without very young children and where the wife works. So the age structure, the female "activity rate" and even family size, car-ownership and shopping habits, will be seriously inaccurate in the sample data. There are statistical methods, involving some type of weighting, for dealing with a high "no-response" factor; but the problem is best avoided by persistent calls until a response is achieved.

The last type of sample with an application in social geography, "quota samp-

ling", avoids the no-response problem completely, but is open to many other objections. An interviewer assembling a quota sample is sent out to question specified numbers of individuals falling in pre-determined categories. He must "collect" given numbers of men and women, with a specified distribution over the age-ranges and perhaps income-groups or occupational categories. The quota for each age, income and occupational group is so calculated that the sample is — at least in these respects — perfectly representative of the total population. These are the quota controls, a sort of sieve designed to select a correct sample. But in other respects quota samples are often wildly unrepresentative. The actual selection of individuals to be approached (but rejected if they do not fit a quota) is left to the interviewer. Thus relatively accessible and approachable individuals tend to be over-represented, while unapproachable or inaccessible individuals are under-represented. For various reasons attractive girls, idlers and people walking or travelling by bus tend to be over-sampled; while mothers with children-in-arms, drunkards, taxi-cab users, miners or merchant seamen tend not to be sampled.

Depending upon its purpose, a quota sample may be biased or entirely adequate, but one never knows. Despite its cheapness and speed, it is therefore somewhat suspect and little used. Yet in some cases it is the best frame that is available. For example, the only practicable way of plotting, with real precision, the urban hinterland of a town is to make a sample survey of people shopping in its central area, or using other central facilities. By asking about their home address, the mode of transport they used to come shopping, how often they come, where else they go, what they buy on their visits to the town a complete appraisal may be made of the strength of a town in attracting trade to itself from its hinterland, in competition with its rivals. But what sort of frame can be used for keeping unconscious bias out of such a sample? In fact such samples are often taken entirely without controls, and they can be anything but random. In this sort of case a quota system is better than nothing, though it may not be very good.

WHAT SIZE OF SAMPLE?

The short (but misleading) answer to this question is "as large as possible", for the higher the sampling fraction the greater the chance of perfect representativeness. The standard error of a random sample (in effect a probability measure of its margin of error) varies inversely with the square root of sample size. Thus, to halve the likely errors in a sample one needs, theoretically, to increase the sample size fourfold. In practical terms, therefore, one can spend a great deal of money in enlarging the sample size, but one "buys" only a small (and perhaps only theoretical) increase in accuracy. This leads to a simple, practical conclusion: it is far better to spend money on perfect randomness than on modest increases in

sample size. A 10% sample in which no-response is kept to the minimum by persistent re-calls at "difficult" addresses is likely to be much more valid than a 20% sample which suffers from a 25% no-response rate (though the actual number in the sample is much higher in the latter case).

It is easy in theory (but often impossible in practice) to estimate the sample size needed for any survey. If one can decide what margin of likely error is tolerable (i.e. arrive at a maximum acceptable standard error) then merely by inverting the standard error formula one has a means of calculating the required sample size ⁶. But it must be remembered that, in the sort of application that geographers are likely to be concerned with, there is really no overall standard error for a sample, but rather individual standard errors for each quantity that the sample yields. These quantities will vary very greatly: in a sample of 1,000 households in a New Town population, 750 may be the families of factory workers, but only 50 the families of professional men. The acid test of the sample is whether it measures the smaller of these two quantities with reasonable accuracy: thus any estimate of the sample size needed must be based upon the probable size of minorities in the population. If attributes of these minorities are to be investigated (for example if journey-to-work patterns and shopping habits of professional families are to be compared with those of factory operatives) then even the minority elements in a population must have substantial samples — and this, in fact, means a wasteful over-sampling of majority elements. If the population to be sampled contains no or few minorities, a much smaller sample would be perfectly effective. The point can be put theoretically: where variability in a population is slight a small sample is acceptable, since standard error rises with standard deviation.

The need to have a sample large enough to give minorities adequate representation is quickly limited by questions of cost. A large national survey of outdoor recreation in Britain, intended to provide a background for recreational planning, has recently been made. For this a national sample of some 3,200 was adopted, largely on the grounds of cost. This can be shown to have covered most majority attributes of the population quite adequately ⁷. From it one could measure the present and potential demand for many relatively popular recreations and so

⁶ The standard error of a proportion (p) in a population is given by:

$$S.E._p = \frac{p(1-p)}{n} \quad (\text{where } n = \text{size of sample}).$$

$$\text{Thus } n \text{ (sample size)} = \frac{p(1-p)}{S.E._p^2}$$

But this involves, apart from specifying a standard error, making an estimate of p , which may be quite impossible.

⁷ See Rodgers, H.B., *Pilot National Recreation Survey, First Report*, London 1967, p. 93 (Appendix) for a discussion of these problems.

derive conclusions about the allocation of land and investment to recreational purposes. It did not, and could not, adequately cover recreations followed by only one or two per cent of the sample population. To have investigated these adequately would need not a doubling or trebling, but at least a tenfold increase in sample size, which would have raised costs towards the £ 100,000 level, a quite unthinkable sum. In short, practical considerations, rather than theoretical estimates, must often govern sample size.

In surveys of smaller populations (overspill communities in expanding towns, for example) sample size is controlled, in practice, by the fineness of sample sub-division made necessary by the analytical work intended. It may be intended to compare fertility or family-size a) by age-groups, b) by occupational classes within age-groups, and c) by religion within occupational classes. If so, very small numbers will exist within some of the cells of such a complex tabulation. Only a very high-ratio sample will keep these small-cell values large enough to be usable. In short, local sampling of this type must always be macroscopic in terms of sample ratio; sampling fractions in the range 10% to 30% are often necessary. National surveys are bound to be microscopic in sample ratio: even a sample of 0.01% of the British population is very costly. And so calculations of sample size needed for any case, using the inverted standard error formula, tend to be rather academic, though they may provide a useful guide to the limits imposed on subsequent analysis. In fact, deciding on a sample ratio becomes more a matter of judgement governed by experience, than calculation.

So far, this paper has attempted to outline the principles of random sampling techniques, treated practically rather than theoretically, and to review their general application to social geography. In what follows two particular problems involving the use of random samples are summarised, by way of example.

COMMUNITY CONTRAST IN AN "EXPANDING" TOWN

The small town of Winsford in Cheshire (1961 pop. 12,600) was the creation of a now almost extinct salt industry: by 1960 it had become economically moribund, but by agreeing to accept "overspill" population from both Manchester and Liverpool it had begun to grow again by the mid-1960s and to attract new industry. With excellent communications (close to the main M.6 motorway and served by the electrified London-Liverpool railway) it is an attractive growth point, and in 1966 a decision was taken, in principle, to expand Winsford to a population in the range 60—70,000, thus making it one of the chief foci of planned population and industrial growth in Northwest England. Consultants were appointed to produce a master-plan to guide this growth. But inadequate social and demographic data were available to provide a statistical backcloth for the plan: Winsford had changed too much, by absorbing its new "overspill" community,

for the 1961 census data to be anything but obsolete. The Department of Geography at Keele University was therefore asked to carry out a social survey, with an emphasis on community contrast within the town.

It was decided to use a stratified form of random sample keyed to the contrasted types of community which the town contained. Seven "Community-type" areas were identified, and these served as the sampling frame. They were as follows:

- I A well-established "overspill" community, which had been brought to Winsford, chiefly from Manchester, some three to four years prior to our survey, and which could be expected to have "settled down" in the town. (A total of c. 500 households).
- II A more recently arrived "overspill" community, of mixed Manchester and Liverpool origin, which had been in the town for less than two years and so was still passing through a phase of adjustment to its new environment (c. 500 households).
- III Inter-war municipal housing built for local families, known to be occupied by a largely ageing and now childless community (214 households).
- IV Post-war municipal housing, occupied by local families but with a more normal age-balance (c. 900 households).
- V Small nineteenth century houses, chiefly terraced, many obsolete and ready for clearance. These contain a chiefly elderly community, the social survival of the "old" Winsford (c. 1,200 houses).
- VI Recently built estates of owner-occupied houses, the privately-built equivalent of "overspill" (c. 300 houses).
- VII Areas and ribbons of mixed house-type, chiefly medium-sized houses of various styles, but containing most of the town's houses falling in the middle and upper parts of the rateable-value range. These, essentially, are the middle-income areas of the "old" Winsford (c. 1,600 households).

Each of these community-type areas is easily identifiable "on the ground" as several separate components. Each was divided into survey sub-areas of about 250 households. A 10% sample ratio was adopted, so that each sub-area yielded 25 households selected randomly by using a standard interval along a pre-set route. Thus the structure of the sample is governed by the share of each community type in the total population of the town. Not only did this procedure give a reasonable hope of acceptable validity in the sample, but also it meant that the results of the survey were pre-sorted to reveal the pattern of community contrast which was our focus of interest.

Though the survey was a fairly broad one, and yielded a mass of data on the general pattern of community contrast in Winsford, there were four crucial questions to which it was intended to supply answers. These were as follows:

a) What differences in family size and age, and what contrasts in fertility, exist between the various parts of the town.

TABLE 1. Winsford: age distribution of total population
(Percentages of sub-area population)

Community Type Areas	0 to 4	5 to 9	10 to 14	15 to 19	20 to 29	30 to 39	40 to 49	50 to 59	60 and over
I "Old" Overspill (Total sample pop. 211)	19	19	8	3	17	20	9	3	3
II "New" Overspill (Total sample pop. 240)	23	17	7	4	22	17	5	4	3
III "Old" Local Authority (Total sample pop. 72)	13	7	13	6	10	13	19	14	9
IV "New" Local Authority (Total sample pop. 311)	12	13	10	10	10	14	16	7	6
V "Terraced" Areas (Total sample pop. 152)	8	4	3	6	11	10	14	18	25
VI "New" Private Estates (Total sample pop. 115)	22	19	4	3	19	23	4	3	2
VII Areas of „Mixed" Type (Total sample pop. 347)	9	7	5	8	10	14	11	11	26

In some community-type areas small proportions of the sample were "lost" because of reluctance to divulge ages.

b) It was known from an earlier survey that the "new" population, of chiefly young families, had a birth-rate about twice the national average. Would this fall, once the families were established, to lower levels, and if so how quickly; and would there be a significant difference between Catholic and Non-Catholic families?

c) How does the journey-to-work pattern of the "new" population (I, II and III in the list) differ from that of older established communities in length, destination and type of transport used?

d) How do shopping patterns differ between component communities, and do these differences have a relevance for the design of the retail-model for the New Town?

It is impossible to summarise the results of the survey in a paragraph or two, but some of the key tables are given here (Tables 1—3)⁸. It was found that the

⁸ Full results have been published in: Rodgers, H. B. and Old, F. M. T., *The Winsford Master Plan Survey* (1967). Fuller details of sampling and validity testing are given.

TABLE 2. Winsford: mode of travel to work by (a) men (b) women (c) children. (Percentages of employed population)

The average for all workers is given in the right-hand columns. Children are defined as sons and daughters of working age living with parents

Community Type Areas	Bus		Train		Car		Scooter or Motor Bike		Bicycle		Walk		No Reply		Activity Rate among		
															Men	Wo- men	Chil- dren
I	(a)	9		0		74		9		2		2		0	99		
	(b)	24		0		35		0		0		29		12		32	
	(c)	25	14	0	0	50	62	0	6	0	1	25	13	0	3		8
II	(a)	27		2		58		5		0		7		2	98		
	(b)	72		0		12		6		0		6		6		29	
	(c)	50	38	0	1	17	45	0	5	0	0	33	9	0	3		5
III	(a)	19		6		31		0		25		19		0	85		
	(b)	43		0		0		0		21		37		0		61	
	(c)	44	33	0	3	0	13	0	0	11	21	44	31	0	0		30
IV	(a)	7		3		56		9		17		8		0	90		
	(b)	11		0		5		11		11		63		0		22	
	(c)	28	12	0	2	14	39	11	10	25	18	18	19	4	1		18
V	(a)	32		3		39		11		13		3		3	71		
	(b)	6		0		6		0		19		62		6		26	
	(c)	2	23	0	2	45	31	0	6	11	14	22	20	0	3		20
VI	(a)	6		0		66		6		9		13		0	100		
	(b)	20		0		60		0		0		20		0		15	
	(c)	0	8	0	0	50	64	0	5	0	8	50	15	0	0		3
VII	(a)	6		1		55		1		19		11		7	82		
	(b)	33		0		20		0		10		36		0		23	
	(c)	33	17	0	1	33	44	0	1	20	17	14	17	0	4		27

TABLE 3. Winsford: occasional shopping for durable goods
Centres used (% of mentions)

a — Centres regularly used; b — Centres visited in previous month outside Winsford

Community Type Areas		Winsford High St.	Winsford New Town Centre	North- wich	Sand- bach	Mid- dle- wich	Crewe	Che- ster	Liver- pool	Man- chester	War- ring- ton	Others	% of sample who had shopped outside Winsford in the month prior to survey
I "Old" Overspill	a	17	3	12	4	0	27	1	3	34	0	0	34
	b			7	10	0	28	3	17	41	0	0	
II "New" Overspill	a	14	5	10	0	0	13	2	29	23	0	2	49
	b			10	0	0	17	0	48	25	0	0	
III "Old" Local Authority	a	29	3	20	3	0	20	6	6	9	0	0	38
	b			50	0	0	25	25	0	0	0	0	
IV "New" Local Authority	a	22	7	11	1	0	32	7	5	7	3	0	43
	b			24	0	0	52	5	4	4	7	0	
V "Terraced" Areas	a	33	5	12	0	1	28	7	4	5	1	0	20
	b			27	0	0	36	27	9	0	0	0	
VI "New" Private Estates	a	12	2	16	0	0	23	0	14	19	4	7	45
	b			28	0	0	41	0	7	7	3	15	
VII Areas of "Mixed" Type	a	31	3	12	1	1	24	9	7	6	0	4	23
	b			20	0	0	54	9	4	7	0	6	

demographic contrasts, area by area, within the town were extreme. The three communities which make up the "new" Winsford, with "immigrant" populations (I, II, and III), all have distorted age-structures, with most adults in the 25—35 age band and huge "excesses" of children below 10. Birth-rates here averaged between 16 and 19 surviving children per 100 households per annum. In contrast, some of the communities of the "old" Winsford (especially III and V) contain almost "relic" populations, of dominantly old, small and childless families. Here birth-rate is only about 5/100 households per annum. But the new families of "overspill" origin seem quickly to revert to a normal — or even sub-normal — birth-rate once established in the town: families re-visited after a two-year interval as part of a follow-up enquiry to an earlier survey showed a drop in fertility to about 5/100 households per annum⁹. Surprisingly, Catholic fertility and family size (but in a dangerously small sub-sample) were not markedly above average. For the town as a whole there were 182 children per 100 Catholic families compared with 114/100 for others, but this contrast disappears if the data are adjusted for age-differences.

The "old" and the "new" populations of Winsford are as different in the economic as in the demographic sense. In community areas I and II (the "overspill" population) men work chiefly in the town's new industries, have a short daily journey, and use cars in getting to work much more than does the established population. In the "older" communities of the town (especially the low-income areas III, IV, and V) employment is chiefly in the town's older and now declining industries: here substantial proportions of men travel out of the town each day to the chemical industries of Northwich. Among the men of these "older" communities the journey-to-work is most commonly made by bus, by bicycle or on foot, rather than by car. The journey-to-shop varies from community to community just as much as travel to work. In the older parts of the town most shopping is local and short-range: the old (and now quite inadequate) town centre serves most needs, and relatively small proportions of families leave the town to shop in higher-rank rivals 8 to 15 kilometres away. But shopping among the "new" communities is much less self-contained within the town: here families make far more use of a new and still developing town centre (which is almost ignored by the "old" population), but they also make more frequent shopping trips elsewhere.

This is the merest summary of a few conclusions from a lengthy report. Short though it is, it perhaps shows how material of practical relevance in shaping the design of what is virtually a New Town may be gathered, quickly and cheaply, by a social survey, using the pattern of community contrast within the town as the basis for stratification in taking a random sample.

⁹ This earlier survey was a 1 in 3 random sample of the overspill community only. See Rodgers, H. B. and Herbert, D. T., *Overspill in Winsford*, 1964.

COMMUNITY BUILDING IN A NEIGHBOURHOOD UNIT

Some two years ago it was decided to conduct a study of a large re-housing project, developed over the previous decade on the fringes of the City of Stoke-on-Trent, at Ubbertley-Bentilee. This large estate was designed to function as a single neighbourhood unit, and to act, in some senses, as a self-contained community. It was given a quite unusually full range of shopping and social facilities at its centre, but it was also made, perhaps, too large both in population and extent to function effectively as a single, inwardly-oriented social unit. In short, there were factors at work here tending both to promote and to retard the development of a sense of community consciousness. It seemed therefore to provide an excellent opportunity for a test of the validity of the neighbourhood unit concept, on which the social aspect of British town-planning has been based since the war. The survey and the sample frame were designed with this primary object.

Almost 3,600 households (14,700 people) lived in the area at the date of survey, too large a population to sample effectively with the resources available. For this reason — but also because it fitted well with the object of the survey — an elementary type of cluster sample was chosen. The estate is long and rather linear: its western end is close to the excellent shopping and social facilities of Hanley, the dominant centre among the Pottery towns that constitute the City of Stoke-on-Trent. The eastern end of the estate is remote from any good shopping centre — either inside or outside the neighbourhood unit: while the central zone has easy access to the estate's own excellent cluster of shopping and social facilities. To test the effect of location on the use of community facilities survey areas were chosen in the west, centre and east: together they included a majority of the total population. Within the survey areas sub-areas were defined, and the sub-areas could be roughly graded by distance from the neighbourhood centre and other, external rivals to it. In each sub-area every fifth address was chosen to become part of the sample. This was a rather primitive sort of cluster sample, but the clusters, were, in a sense, "stratified" by distance from social facilities. The questionnaire was designed chiefly to measure how the use of the community facilities provided at the estate centre varied with distance, and how this, in turn, has helped to shape patterns of social life and the general growth of a sense of community consciousness. In short, this was an attempt at a critique of a piece of social planning¹⁰.

A few of the more significant tables yielded by the survey are presented here (Table 4). From these, contrasts in social integration between the east, centre and west of the estate are fairly obvious: these contrasts are somewhat blunted in the grouped data, but are very much stronger if individual survey sub-areas

¹⁰ See Rodgers, H. B. and Herbert, D. T., Space relations in a neighbourhood unit, *Town and Country Planning*, April 1967.

TABLE. 4 Measures of social integration in Ubbertley-Bentilee

(A) Use of community facilities				
	% using the specified facility			
	At least weekly	Monthly	Less often	Never
I Community Centre				
Eastern sub-areas	7	11	13	69
Central sub-areas	25	1	29	45
Western sub-areas	24	5	36	35
II Health Centre				
Eastern sub-areas	3	6	18	73
Central sub-areas	4	6	39	51
Western sub-areas	5	6	39	50
III Church				
Eastern sub-areas	23	2	12	63
Central sub-areas	28	6	26	40
Western sub-areas	15	4	53	28

(B) Use of facilities by length of residence: a) families resident here 3 years or less b) families resident here 10 years or more

	% using facilities							
	At least weekly		Monthly		Less Often		Never	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Community Centre	16	26	8	3	14	28	62	43
Church	15	23	2	7	30	38	32	62

(C) Shopping for everyday needs

	% using				
	Estate shops	Hanley	Longton	Stoke	Others
Eastern sub-areas	72	11	8	2	8
Central sub-areas	83	10	4	1	3
Western sub-areas	65	25	5	3	3

(D) Length of residence on the estate

	0—3 years		3—5 years		over 5 years
Eastern sub-areas	26		8		67
Central sub-areas	10		16		74
Western sub-areas	26		6		68

very close to the neighbourhood centre, or distant from it on the estate margins, are compared. The use made of the shopping centre varies strongly (Table 4c). This is sociologically important, since casual friendships made while shopping are one of the principal mechanisms in community building. Families living near the estate centre rely on it almost entirely for day-to-day shopping; and partly for this reason they are much more fully aware of the social activities provided by the large and well-equipped Community Centre. But families on the eastern and western margins are much less orientated towards the estate centre for shopping: in some peripheral sub-areas almost half the families do most of their daily shopping off the estate, and so feel a weaker sense of social identity with it. Not surprisingly, there was almost a gradient of satisfaction with shopping provision, from the centre to the margins of the estate.

Use of the Community Centre (and the nearby Health Centre) is perhaps the best single index of social integration within the neighbourhood unit, since the Centre was designed to act as a focus for a wide range of recreational activities and so to provide a foundation for community building. Again distance governed participation. The majority of families in the eastern survey area almost ignored the Community Centre, and in some sub-areas here few families even knew of the Centre's existence. In general, our western survey area used the Community Centre more strongly, but again families in the most marginal of sub-areas here looked elsewhere for their social satisfaction. In contrast, in the estate centre, half the population of some sub-areas use the Centre at least weekly.

Distance is not the only determinant of involvement in the life and activities of the Community Centre: length of residence on the estate has almost as powerful an effect, as Table 4 B shows. But these two factors, distance and length of residence, work together, for the average period of occupation of houses is greater in the centre than in the east or west. Only 10% of families living close to the estate centre have occupied their homes for less than three years, while one quarter of families in both the east and west are relative newcomers. This means that there is a higher rate of tenancy turnover in the west and east than in the centre — clearly one reason why these areas are so poorly integrated into the life of the neighbourhood. But this high turnover rate is itself associated with a sense of social dissatisfaction with the estate and its environment. In these peripheral areas

much smaller proportions of people are satisfied with their homes and with the estate than among families close to the centre. Moreover, in the marginal areas, only about 65% of respondents thought the estate a friendly place to live in, while over 80% thought so at the centre. In some marginal sub-areas most friendships were with people off the estate. From all of this it is obvious that the progress of community building has varied very strongly with distance from the neighbourhood centre.

These findings lead to the conclusion that the size of the estate, accentuated by the linear form of its plan, has impeded the development of a neighbourhood unit in the functional, social sense. The margins are only very feebly integrated with the centre. Our survey threw into sharp focus the planner's insoluble dilemma in the design of neighbourhood units. There is clearly a virtue in sheer size, in that this creates a local market for shopping and social facilities big enough to make a very wide variety economically viable; but there is also, paradoxically, a virtue in smallness, in that this gives the neighbourhood the best chance of developing into a genuine community. The economic arguments are for large size, the social arguments are for smallness and compactness. Perhaps the undesirable social effects of the over-large size of the Ubbertley-Bentilee unit might have been counteracted by a plan design in which there was an effective system of linkages which would render all parts of the neighbourhood easily accessible to each other and to the central services. But the chief practical lesson that could be drawn from the survey is that a neighbourhood of this size can be socially successful only if high housing densities are used to keep distances small. Otherwise what is physically a big unit becomes, in fact, operationally small through the failure of its margins to become intergrated into the whole.

The conclusions to which this paper leads are simple and brief. Random social survey techniques are a new and developing tool which the social geographer may profitably borrow from the sociologist, though he will want to use it differently and distinctively, to reveal areal contrast among urban and rural communities. Basically, this is a simple (though sometimes costly) technique, and not one which is restricted only to the fairly few geographers with a sound mathematical training. The practice of random sampling is, in fact, a good deal simpler than its theory. Since social surveys of this type gather unique data not otherwise available they have a particular relevance in that part of the field of social geography which borders on land-use planning. Indeed, this is work in which geographers and planners are being increasingly brought into fruitful contact.

University of Keele

SOME COMMENTS ON THE RELEVANCE OF MULTIVARIATE ANALYSIS TO GEOGRAPHY: A METHODOLOGICAL REVIEW

DEREK THOMPSON

Geography has experienced fairly rapid growth in the last ten years in the use of statistical and mathematical techniques. Well represented in this expansion is the group of statistical procedures known as multivariate analysis¹. This paper aims to illustrate the relevance of such techniques to geography, not by summarizing their applications as they have appeared in the literature, but by discussing their basic features and the kinds of question which they can help solve. This approach provides a more useful guide to their possible applications in geography, although it does mean that the description of the uses of particular techniques is subordinate to an account of multivariate methods in general.

THE MAIN FEATURES OF MULTIVARIATE ANALYSIS

Multivariate statistical analysis is a branch of mathematical statistics dealing essentially with a large number of variables. In such general terms multivariate includes most of statistical theory, but commonly a more restricted use is made of the term². However, rather than try to summarize a very large field by one small definition I feel it is best to convey its nature by describing its characteristic features.

Three such features seem most important.

(a) There are observations on a large number of variables, p , for a number

¹ A recent paper by Lavalle, McConnell, and Brown [21] (see bibliography for publications cited) provides some data on the use and teaching of quantitative methods in geography in the U.S.A. In the period 1954—1965 62 articles using sophisticated statistical and mathematical techniques appeared in the three major American geographical journals. In the same period, for the three chief British publications, there were only 8 such papers (information compiled by author). Various multivariate techniques account for about 85% of the techniques used in the American articles.

² Kendall [20], p. 6, has defined multivariate analysis as "the branch of statistical analysis which is concerned with the relationships of sets of dependent variates".

of individuals, n . The multivariate character ³ lies in p , not in the size of the set, n .

(b) Each individual can be located as a point in p -dimensional space, just as one can locate a point in a bivariate situation using the familiar two-dimensional Cartesian co-ordinates.

(c) The variables exhibit some degree of association amongst themselves; we cannot isolate one variable from the others because they are involved in combination as a system.

We have the means, then, in multivariate analysis of considering individuals, variables, and combinations of both. More attention will be paid to such basic features later, but first it is necessary to consider other features which allow one to distinguish amongst the great variety of techniques which exist.

A review of standard texts ⁴ reveals so many different methods of classification based on parameters, assumptions, mathematical procedures, concepts, complexity, and uses, that one achieves little by suggesting one as the "best" grouping system. Accordingly, the scheme used here is put forward as a guide to some of the basic concepts and uses only, as it is here that our main interest lies. The classification is given in Table 1, in which the most commonly encountered techniques ⁵ are described in terms of seven bipolar characteristics, of which the first six are derived from the three basic features previously mentioned..

Most of the techniques analyse the relationships amongst the variables, whilst some methods consider the differences between groups of individuals for two or more variables (criterion D in Table 1). This is a basic contrast between, for example, analysis of variance and correlation-based techniques ⁶, but neither approach deals with individuals as more than data points.

For those techniques considering the relationships amongst variables, three further criteria are identified. First, there is a distinction between the methods treating variables as separate entities, for example regression and variance analysis, and those considering groups of variables (criterion C). Amongst the latter are canonical correlation and factor analysis. Secondly, one can contrast those techniques which examine the strength of a relationship and the nature of it (criterion B). Correlation methods fall into the former group, and regression into the latter. Thirdly, one can divide the methods into two groups depending on whether they specify a particular variable as having a pre-eminent position

³ Strictly speaking if $p = 2$ we have a multivariate situation, but usually this is referred to as a bivariate situation, leaving multivariate to refer to cases with $p \geq 3$.

⁴ The chief of these are Anderson [1], Kendall [20], Rao [27], and Roy [30] for mathematical treatments, and Cooley and Lohnes [11] for computer applications.

⁵ Other techniques, such as dimensional analysis, time series, and trend surface analysis could possibly be regarded as multivariate methods, but they are usually considered as separate groups of procedures.

⁶ Burt [7] explains the basic differences of factor analysis and analysis of variance.

TABLE 1. Features of some multivariate statistical techniques

Technique	Feature: <i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>
Multiple correlation	1	1	0	1		0	1
Multiple regression	0	0	0	1		0	0
Multiple analysis of variance	0		0	0	0	0	0
Multiple analysis of covariance	0	01	0	0		0	0
Factor analysis	1	1	1	1	*	1	01
Components analysis	1	1	1	1	*	1	01
Cluster analysis	1	1	1	1		1	1
Canonical correlation	01	1	1	1		1	01
Discriminant functions analysis				0	0	0	0
<i>T</i> ² , <i>D</i> ² , and other distance measures				0	01	0	0

Note: a 0 or 1 is placed opposite a technique if it is characterised by one of the two opposing features of a particular criterion. The criteria are:
*A*₀: the technique is concerned with dependence relationships, *A*₁: the technique is concerned with interdependence;
*B*₀: concerned with the nature of relationships, *B*₁: concerned with the strength of relationships;
*C*₀: concerned with variables as separate entities, *C*₁: deals with sets of variables;
*D*₀: treats group differences rather than *D*₁: relationships within the body of variables;
*E*₀: concerned with testing existing classes, *E*₁: concerned with creating classes;
*F*₀: usually deals with a few variables (say 2—20), *F*₁: usually handles a very large number of variables (say 20—100);
*G*₀: usually used for testing hypotheses, *G*₁: usually used for suggesting hypotheses.

* Factor and component scores can be used for classifying, but the factor techniques are not designed for this purpose.

in the analysis. The two groups here are: — (i) analysis of dependence in which one examines the “effect” of several “independent” or predictor variables upon a specified “dependent” variable or criterion, and (ii) analysis of interdependence concerned with the relationships amongst all variables (criterion A). Regression and factor analysis are examples of the two types, as Figure 1⁷ illustrates.

In all the above attention is focussed on the variables rather than on the individuals. However, some techniques do consider the latter, and these may be contrasted on bipolar criterion E — those methods which group individuals into classes and those which test the placement of individuals in existing groups⁸.

The remaining two criteria concern the use of techniques in hypothesis testing

⁷ The diagram is based on Kendall [20], p. 6, and [19], pp. 60—61.
⁸ Techniques such as *D*², *T*², and analysis of variance are really extensions or generalized versions of procedures used in univariate situations. However, we are justified in regarding them as multivariate methods if only because they deal with many variables and have features in common with other techniques contained within the field.

or creating situations (G), and their capacity for handling large numbers of variables (F). Factor analysis, canonical correlation, and components analysis are designed primarily for very many indices, but analysis of variance and correlation/regression analyses work best with a smaller number. In the analysis of variance, there are great complexities of design and interaction effects with a large number of variables. For regression analysis, the linear estimation equation of the functional relationship becomes exceedingly long when p is large, because each variable is referred to separately. It is this problem which factor analysis and related techniques were designed to overcome as they treat groups of variables.

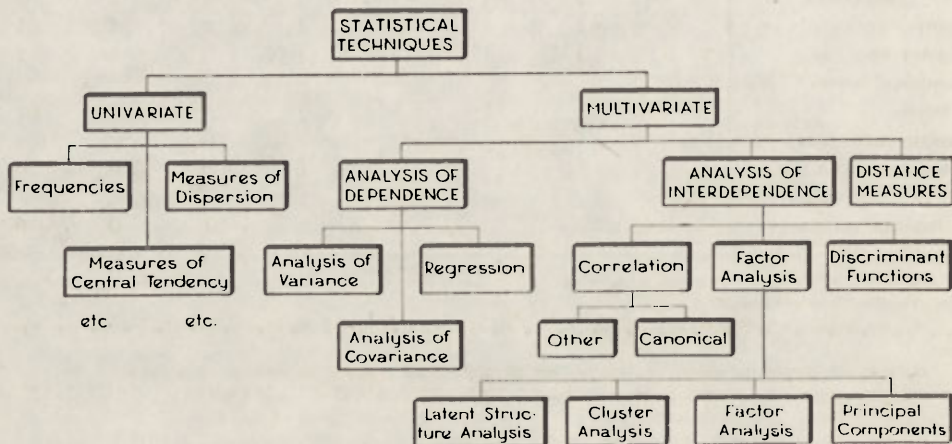


Fig. 1 Diagram of multivariate techniques

It has been seen, then, that the several techniques deal with the basic ingredients of a number of individuals scored on several variables in different ways. A brief glance at some basic procedures will help to clarify this point.

Most techniques examine the variances and covariances of the sets of observations or, more usually, the correlations between variables, as the latter are standardized covariances anyway. Factor and canonical procedures investigate the correlation matrix in different ways, whilst the variance, covariance, and discriminants methods use the dispersion matrix or consider different partitions of the variances. The basic aim, however, rests upon the relationships of numerous variables.

A second feature common to several techniques is the aim of reducing the original p -dimensions to a reduced number sufficient to describe the information contained in the original observations. This is done in factor, canonical, and discriminant analyses by some form of linear combination of variables, the nature of which is determined by the type of original information which the technique preserves. Canonical correlation calculates the maximum correlation between linear functions of two sets of variables; while factor analysis reveals the common

dimensions in the correlation matrix. One sees here one major purpose of multivariate analysis — achieving parsimonious description of complex situations.

Finally, the many multivariate techniques are most efficiently handled with the aid of electronic computers. This is due largely to the lengthy, repetitive calculations involved, and the use of matrix and vector algebra in dealing with a large number of observations. The calculation of the variance-covariance or correlation coefficient matrices and the subsequent computation of factor loadings, discriminant functions and so on are easily accomplished by appropriate manipulations of the basic vectors and matrices. In addition, some techniques require the calculation of the eigenvalues and eigenvectors of a matrix, particularly factor analysis, a procedure which can be done best by numerical iterations using a computer.

THE CHIEF QUESTIONS TACKLED BY MULTIVARIATE TECHNIQUES

Geography, like other fields, has many variables to consider, but this is hardly sufficient reason to justify the use of sophisticated multivariate techniques. We must delve deeper into the nature of the concepts of the latter to see if the procedures are appropriate to solving geographical problems. To do this, it is necessary to expand the treatment of the first section of this paper by discussing some of the important questions considered by the multivariate methods in general.

These are concerned, primarily, with analysing complex situations for two main types of problem. The first of these considers the relationships amongst variables using individuals as data points only; the second deals with the behaviour of the cases in groups or as individuals and may be referred to as the classification problem.

The study of relationships is itself varied, as previous remarks have suggested. To repeat, one may consider dependence or interdependence, strength and nature of associations, and deal with variables as groups or separate entities. These notions can be expressed somewhat differently in a series of questions which are appropriate to achieving an understanding of the usefulness of the methods in geography.

- (1) What is the nature of an environment, which is defined as the complex, or assemblage, of phenomena (i.e. variables) for a given universe?
- (2) How do the environmental elements react with each other?
- (3) Are there any features common to several variables?
- (4) Are there any common features of over-riding significance?
- (5) What processes are at work within the environment, in terms of the impact it has on specified subjects?
- (6) Are the major dimensions of the environment necessarily those of greatest impact on specified subjects?

The study of the environment as a whole is best accomplished by procedures

involving interdependence notions and capable of utilising a large number of variables. Here we find that correlational methods are the most useful — answering question (2) by calculating bivariate coefficients of correlation, and (3) and (4) using factor analytical methods. However, as we assume the existence of a complex environment, it is vital to consider relationships other than those of only two variables, and answers to question (2) provide us with relatively little useful information. Furthermore, different variables will often have features in common and will not make additive contributions to the environment. This last situation is analysed also by factor analysis in extracting factors representing the overlap of variables, which means it is possible to point to the existence of major features, whose presence might not have been suspected initially.

Investigations of environment-subject relationships (5) and (6) necessarily involve notions of dependence, which are best treated using regression analysis or even canonical correlation. We are not concerned with all elements in the environment, but with the impact of several features, perhaps some of the initial indices, or perhaps the common measures produced by factor analysis, on some subjects. Methodologically they are linked in a stimulus-response manner, involving feedback and circular connections. Therefore any question of causality or necessary connection must still be provided by interpretation, although one may deal with time sequences by a proper selection of variables.

The second type of problem encountered in multivariate analysis concerns the grouping of individuals in the p -dimensional space. Grouping in itself is a more general problem, but it is of relevance here because most classifications treat at least two variables. Some of the vital features of this grouping problem⁹ are summarized in four questions.

(7) How can one combine individuals into groups so as to achieve an optimum grouping?

(8) How can one test for the efficiency of groups?

(9) How can one locate any individual into one of several existing classes?

(10) Is it possible to combine elements of an environment into groups?

This last question is similar to (7) in creating groups, but it is separated from it because it deals with variables based on a number of cases, rather than dealing directly with individuals. Question (10) also serves to illustrate that one cannot really separate problems of relationships from those of classification, as it really is essentially the same as question (3).

In each of the above questions, however, we are dealing primarily with aggregates¹⁰ in some form, but we also maintain interest in individuals because it

⁹ Berry [2] has a few remarks on this type of question.

¹⁰ A distinction is made between aggregate and group. The former simply refers to a collection of objects, e.g. environmental variables, or cases, and may also be considered as a set. A group is a set of individuals possessing some functional relationship or sharing common features, e.g. a geographical trading area.

is possible to specify how one case is located in the p -dimensional space. If we wish to group individuals in that space such that they form homogeneous clusters we have questions of type (7). Statistically, we hope to maximize the ratio of the variations amongst groups to the variations within the groups. A slightly different problem is involved in (8), where we are measuring the homogeneity of groups already determined or the existence of group differences. This can be done simply by carrying out an analysis of variance test for a multivariate situation, or a difference of means test using a T^2 or D^2 generalized distance measure.

Question (9), however, refers to individuals in a group, where we wish to test if they are correctly assigned. Discriminant functions analysis calculates composite measures for individuals and groups and assigns an individual to that group to which it is most similar.

Whilst other questions could be added to the list of ten above, especially when considering the detailed functions of individual techniques, it is felt that they do provide an adequate summary of the purposes of multivariate analysis. This broad field covers many different methods and procedures and can fit into the various stages of a research design, from initial exploratory stages, as exemplified by questions (1) to (4), to the subsequent and perhaps more fundamental tasks of questions (5) to (10). The different techniques might have certain niches in a general classification of statistical methods, but they are united in their basic aims of describing individuals in terms of several characteristics or treating the relationships amongst those features.

EXAMPLES OF THE USE OF MULTIVARIATE TECHNIQUES

This section of the paper illustrates the two main types of problem tackled by multivariate analysis, in accordance with the stated aim of concentrating on methodologies. Although it is still not possible to cover more than two basic techniques, it is hoped that the others may be seen in the light of the examples given.

EXAMPLE A — THE STUDY OF ENVIRONMENTAL INTER-RELATIONSHIPS

This first example illustrates the use of certain techniques in an exploratory situation. We are faced with a mass of features describing an environment for a set of places and wish to be able to reduce the apparent complexity to manageable size. Factor analysis is well suited to a problem such as this and has been so used in several studies¹¹.

¹¹ For geographical applications see, for example, Berry [3], Carey [8], Henshall [16], and Wong [32]. Canonical correlation also reduces the mass of correlations amongst variables to a smaller number of dimensions, but does this for two separate sets of indices, one of which usually contains a number of different dependent variables.

The selected environment is the U.S.A. in the period 1950—1960, measured by 72 social, demographic, and economic variables. Three hundred counties (approximately 10% of all counties) selected by a random sample are the individuals being used as data points. At the moment we have no other interest in them, as we are investigating the environment in the aggregate. Nor do we postulate hypotheses about the variables, although we could do this, as we are interested solely in obtaining information about the relationships within the environment.

If we consider only bivariate relationships, such as with the product-moment or some other correlation coefficient, we have $\frac{1}{2} p(p-1)$ or 2556 associations to deal with in this particular example. We can learn something from a matrix of these correlations, but this assumes there is no influence of other variables on each pair, an assumption which is generally not tenable¹².

A factor or components analysis¹³ of the correlation matrix gives us several useful pieces of information about an environment, in particular the identification of the major dimensions in the whole mass of indices. In the example, the 72 variables were resolved into 11 dimensions, which are shown in Table 2. The magnitude of a factor is shown by the size of its eigenvalue, and its nature is shown by the initial variables highly related to it. The factors are usually obtained in descending order of size, and in practice one uses only those with eigenvalues greater than 1.0. In the example, one sees the over-riding importance of the first three dimensions, labelled level-of-living, change, and urbanization.

Each factor or component is interpreted in terms of the variables associated with it, those with the highest loadings being diagnostic. Thus factor I is labelled level-of-living because the variables with strongest associations are income and education. The loadings, which are shown for factor I in Table 3, also serve to identify those indices which are redundant (as indicated by similarly high coefficients) because they are measuring similar things. For example, one need use only one of several income variables (053, 052, or 010), or one of a number of education indices (013, 054, or 012). In such a situation factor analysis is a useful screening device in the initial stages of a piece of research.

¹² *N.B.* All the techniques mentioned in this paper are to be used only if certain fundamental assumptions are satisfied. It is not possible to consider them here, except to mention that conditions of random sampling, normal distributions, linear relationships, and homoscedasticity usually apply. The reader is referred to the appropriate books for accounts of the mechanics of the use of the various techniques.

¹³ There are many different methods included under the heading of factor analysis. Two main types are called factor analysis and principal components analysis, both of which do essentially similar things, but in somewhat different ways. The reader is referred to the vast literature on the techniques, as only a few superficial comments can be made in this paper. See, in particular Cattell [10], and Harman [15].

TABLE 2. Principal components extracted from matrix of correlation coefficients for 72 variables

Name of component	Eigenvalue ¹	Cumulative % of trace ²	Number of variables loaded ³
Level-of-living	22.30	30.97	36
Change	12.15	47.85	22
Urbanization	7.67	58.50	23
Location	3.41	63.23	6
Manufacturing	2.54	66.76	7
Race	2.25	69.89	9
Unemployment	1.83	72.43	6
Age	1.59	74.64	8
Old population difference	1.28	76.42	4
White-collar difference	1.18	78.06	2
Primary production ⁴	1.04	79.50	4

¹ An eigenvalue is a characteristic value or latent root of a matrix. In factor analysis it serves to indicate the importance of a component.

² The trace of a matrix is the sum of the diagonal elements, which, for a correlation matrix with unities in the diagonal, is equivalent to the number of variables, or, here, 72. The total variance of the matrix is 72, which is split amongst a large number of factors the total of whose eigenvalues is also 72.

³ This refers to variables with loadings of 0.30 or more only.

⁴ By convention, only those eigenvalues greater than 1.0 are used.

The overlap of variables is measured in a less sophisticated way by cluster analysis¹⁴, which provides a means of grouping indices with similar correlation coefficient profiles. (This is really the concept of question (3) above.) The technique is not described here, but profiles are illustrated as a means of showing the behaviour of one variable with respect to all others. The coefficients of correlation for a single measure are obtained from the correlation matrix and plotted on a graph containing *r* values on the y-axis and a list of variables along the x-axis. One can compare profiles for different variables in this way, and it is instructive that one can usually identify those variables that are eventually found to be connected with a mathematically — produced component.

This feature can be seen in Fig. 2, which plots the profiles of the five variables with highest loadings on the largest factor, termed level-of-living. Each profile has very little deviation from the others. Figs. 3 and 4 show other groups of profiles where the variables are less closely related to the dimension indicated. The first shows six variables with the highest loadings on factor II (change) and reveals no very great deviations, but the other, component VIII (age), has considerable

¹⁴ Cluster analysis lies midway between a visual inspection of correlation profiles and the extraction of dimensions by factor analysis. It involves the calculation of a coefficient representing the similarity of profiles. Fruchter [13] has a description of the technique, and Mayfield [23] a geographical application. Note that cluster analysis in this sense is not the same technique as the analysis of clusters of individuals in *p*-dimensional space.

TABLE 3. Variables with highest factor loadings on the largest factor, level-of-living

Short title	Index	Factor loading
Low income, 1950	053	0.90
High school, 1960	013	-0.90
Median income, 1950	052	-0.89
High school, 1950	054	-0.88
Median education, 1960	012	-0.85
Farm level-of-living, 1950	040	-0.84
Retail sales, 1948	056	-0.84
Low income, 1960	010	0.84
Farm level-of-living, 1959	039	-0.84
Houses sound, 1960	020	-0.82
Houses with phone, 1960	024	-0.81
Median income, 1960	009	-0.81
.	.	.
.	.	.
.	.	.
January temperature	067	0.48
.	.	.
.	.	.
.	.	.
Non-white, 1960	090	0.30

Note: 36 variables had loadings of at least 0.30 on this factor.
12 variables had loadings of 0.80 or more.
Only a selection of the variables is given in the table.

divergences for variables 064 and 021, which have, respectively, correlations of only 0.69 and 0.63 with the factor.

The information obtained in a factor analysis also allows one to consider the behaviour of all variables in terms of all factors by inspecting the rows of the

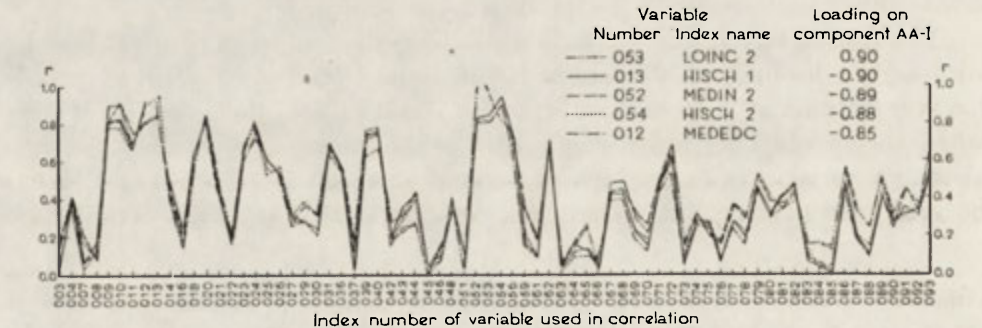


Fig. 2. Correlation coefficient profiles of variables with highest loadings on component AA-I (level-of-living)

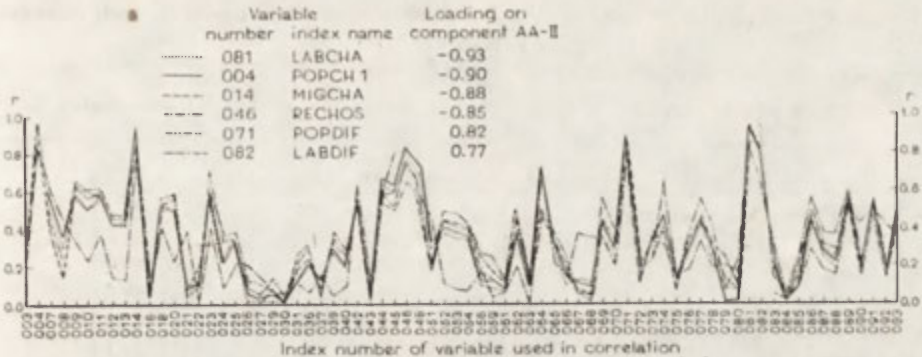


Fig. 3. Correlation coefficient profiles of variables with highest loadings on component AA-II (change)

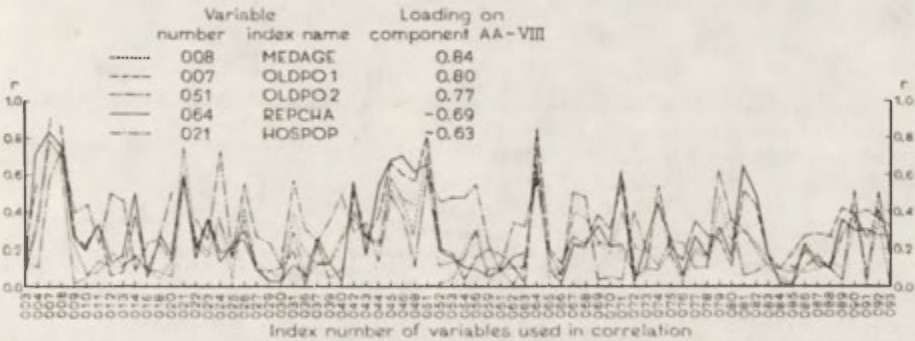


Fig. 4. Correlation coefficient profiles of variables with highest loadings on component AA-VIII (age)

factor matrix. Some variables are highly associated with only one factor, for example those used in Figure 2, whilst others are associated with several. Indices of the latter type are not very useful measures of any single dimension in the environment, as they overlap with several components. In addition, the sum of the squared coefficients of a variable with the extracted factors indicates the amount of variance accounted for. This is called the communality, or h^2 , values of which are given in Table 4 for 12 of the total of 72 variables. The high values for variables 009 and 010 are contrasted with the low communality of index 016 (unemployment), the variance of which amounts to only about half of the total even considering eleven factors. This variable obviously has little in common with the others, even in a multivariable situation. Table 5 provides a summary of the magnitudes of the communalities and reveals a large range from over 90% to under 30%.

TABLE 4. Communalities of some variables after rotation of the components with eigenvalues of at least 1.0

Short title of variable	Index	Communality %
Population density	003	97.32
Population change, 1950—1960	004	95.26
Old population, 1950	007	94.70
Median age, 1960	008	89.73
Median income, 1960	009	94.97
Low income, 1960	010	93.65
High income, 1960	011	79.19
Median education, 1960	012	81.86
High school, 1960	013	88.63
Migration change, 1950—1960	014	92.83
Unemployment, 1960	016	54.24
.	.	.
.	.	.
Rural farm, 1950	093	90.44

Note: Only 12 of the total of 72 variables are given above. The table is for illustration only.

The index number refers to the initial listing of variables from 1 to 93, from which 72 only were selected for analysis.

TABLE 5. Distribution of communalities for all variables

Communality %	Number of variables	Percentage of variables
90.0—99.9	21	29
80.0—89.9	24	33
70.0—79.9	10	14
60.0—69.9	5	7
50.0—59.9	8	11
40.0—49.9	2	3
30.0—39.9	0	0
under 30.0	2	3
total	72	100

In summary, this short account has illustrated the use of factor analysis in exploration of the relationships amongst a variety of environmental characteristics¹⁵. It has been utilised for screening variables, suggesting major dimensions and the original variables most efficiently measuring them, and for indicating the behaviour of the indices in relation to all components.

¹⁵ Factor analysis does have other uses. Factor scores, which represent the measures of the cases on the factors, can be used in classification problems or for input into regression and other analyses. There are several examples of such applications in geography, e.g. Berry [3], Carey [8], McConell [22], and Thompson et al. [32].

EXAMPLE B — CLASSIFICATION

This second example illustrates one of the several classifications problems. This situation involves a test of pre-determined classes, which here are, for convenience, four groups of Standard Regions of England and Wales: the "North", consisting of the "Northern", "East and West Ridings", and "North Western" Standard Regions; the "Midlands", comprising the "North Midlands" and "Midlands"; the "South East", formed by joining the "Eastern" and "London and South Eastern" Regions; and the "South West", consisting of the "Southern", "South Western", and "Wales" Regions.

The individuals in the study are 157 urban areas, that is all County Boroughs, Municipal Boroughs, and Urban Districts, with a population in 1951 of 50,000 or more, and the whole of London Administrative County. They are characterized by sixteen variables, which means we have a truly multivariable situation. The features, listed in an appendix to the paper, were chosen to give variety, and at the same time to provide a composite statement of the nature of towns.

The data matrix for 16 variables and 157 cases arranged in four groups¹⁶ was then subjected to a standard discriminants analysis¹⁷, using a computer for the calculations. The procedure involves the calculation of means for the groups for each variable, and the computation of a matrix of cross products of deviations from the group means and a pooled dispersion matrix. The matrix is then inverted to provide the information used to calculate discriminant functions, of which there is one for each group. The function coefficients and constants are calculated for each group and for each data point. A probability density is then evaluated for each of the latter, and each case is placed in that group for which its probability density is largest. That is, it is assigned to that group to which it is most similar, based on the group's composite character of sixteen variables.

The technique has essentially optimized the classification of individuals by maximizing the ratio of among group deviations to within group deviations. The final output consists of a matrix classifying the individuals as to their correct or incorrect assignment in the actual groups. If the pre-determined groups are sound, then there will be a high proportion of correct assignments.

Before checking the placement of cases it is essential to undertake a test of the null hypothesis that the means of the four groups are identical. Mahalanobis' generalized distance, D^2 , is used in a *chi-square* test with $p(g-1)$ degrees of freedom. In the example the four groups did not have identical means¹⁸, thereby jus-

¹⁶ North, 49; Midland, 23; South East, 64; South West, 21. The technique does not need to have the same number of cases in each group.

¹⁷ Discriminant functions analysis has been used little in geography, but two applications in related fields are those of Miller [24] in climatology and Miller and Kahn [25] in geology. Stone [31] has made use of generalized distance measures.

¹⁸ The obtained *chi-square* value was greater than the value expected for 48 degrees of freedom, and a 0.01 level of significance. The null hypothesis of $\bar{X}_1 = \bar{X}_2 = \bar{X}_3 = \bar{X}_4$ was accordingly rejected.

tifying the calculations of the discriminant functions for the grouping analysis.

The use of the technique can be illustrated geographically by means of Figure 5. This shows the classification number of each of the 157 urban areas as produced by the discriminant functions analysis. The towns are then divided into those which have a function number similar to that of the region in which they are situated, and those with a different number. The latter are regarded as having been mis-classified, and their true location in terms of functional similarity is also indicated on the map. A summary of the classification is provided by Table 6,



Fig. 5. Urban areas mis-classified on the basis of four computed discriminant functions

TABLE 6. Classification matrix for 157 urban areas in England and Wales

Actual group	Discriminant groups					Urban areas classified:	
	1	2	3	4	Total	correctly	incorrectly
1	36	7	3	3	49	36	13
2	3	15	1	4	23	15	8
3	8	5	41	10	64	41	23
4	0	2	1	18	21	18	3
Total	47	29	46	35	157	110	47

which reveals about three-tenths of the 157 urban areas can be regarded as different from the general composite character of the region in which they are located. The misclassification is greatest in groups 2 and 3, but the reasons for this will not be considered here, as the aim is only to illustrate the uses of different techniques.

THE RELEVANCE OF MULTIVARIATE TECHNIQUES TO GEOGRAPHY

The reference to maps conveniently leads into the question of the relevance of the multivariate techniques to geography. Many of our problems do revolve, as in most fields of study, around the two basic features of relationships and classification. Some usual geographical expressions of the former are, for example, the aim of studying the characteristics of places, i.e. their environment, in terms of a group of variables, the need to simplify complexity by identifying major dimensions, and the analysis of the behaviour of the individual elements of the environment. All these are perhaps only preliminary or initial questions, but are nevertheless essential. The lack of well developed theory in geography suggests that we still have many questions of this type.

Beyond such exploratory situations there is a need to examine processes at work within an environment. Multivariate dependence relationships ask questions relevant to such analysis, but one is again faced with the problem of the selection of variables and the overlap amongst different measures. There seems a need to combine factor analysis which provides uncorrelated dimensions, or screens variables, with the usual regression techniques. Even so, any question of explanation of the process relationships awaits interpretation and further experimentation.

Classification problems are well represented in geography, particularly in some of the concepts of regionalization. Basically there are questions concerned with creating groups and testing existing classes, both of which seem to be done largely intuitively at present. There seems considerable scope for the use of multivariate classification techniques to provide more effective answers to our regionalization questions than we have at present. The illustration of discriminant functions analysis serves to point to this in a simple way.

It is also instructive to look again at the three basic features of data points, variables, and large magnitudes, to see if geography handles them any differently from multivariate analysis. There are no great contrasts as to the size of situations or problems, but there does seem to be a strong need for geography to consider more variables than it usually does in order to allow more efficient analyses of both relationships and classifications. The traditional emphasis on only a few indices was partly responsible, it seems, for the long existence of deterministic concepts, which fortunately have now been pushed out by current general systems theory notions of complex interactions and circular connections. Furthermore,

it is unlikely that high levels of explanation can be achieved, particularly in the socio-economic sphere, with only a limited number of "independent" variables. But, whatever the reasons for having many variables, one soon reaches a position of complexity which necessitates the use of multivariate techniques. Perhaps this stage is reached when $p = 4$, because we can no longer visualize such a multi-dimensional situation and require mathematics for its representation.

The variables are themselves not geographic, as they are used in other fields. Nor does geography necessarily use more variables than other disciplines. A difference might lie, however, in the uses of the numerous dimensions, whether original or extracted by factor analysis. Expressions of this might be the comparison of environments of different places, or the creation of multifactor homogeneous regions, problems which are often encountered in geography. Yet all of such uses generally involve the behaviour of individuals for several variables, rather than the examination of indices *per se*.

The characteristic cases of geography are the towns, areas, farms, countries, and so on, conveniently termed "places". We may examine the environments of these, for example the counties of the U.S.A. or the urban areas of Britain, for as many variables as we wish. We may follow this step by some form of generalization of the environments, or by noting the deviant behaviour of individuals¹⁹. Whilst the places are more important than the variables, in effect what we are really accomplishing is an examination of changes in composite (or single-variable) environments over space.

This point leads into the consideration of a crucial difference between multivariate analysis and geography, given the similarities of basic problems and consideration of individuals and variables. What does one do about geographic location, i.e. position on the face of the earth? To return for a moment to our two examples — in the first, the study of relationships considered the individual counties as data points only (except that a later development in the analysis, which was not illustrated, would have considered them); whilst the other considered the characteristics of different places, ignoring their spatial location as opposed to their position in p -dimensional space. The discriminant functions approach provided non-spatial classes, because the urban areas of like nature do not fall into spatially contiguous groups. We have, here, a special geographical concept of preferring "regions" to groups or mere aggregates²⁰. In multivariate terms, we can locate places as data points in p -dimensional space, where the axes are variables, not the familiar "northings" and "eastings".

Some differences exist, therefore, in the methodologies of the two fields. This is most noticeable in the inclusion into geographical concepts of a position

¹⁹ This is often accomplished statistically by calculating residuals from regression, or by comparing the factor scores of individuals.

²⁰ A further point on terminology can usefully be interjected here. A region is regarded as a group where elements possess the property of spatial contiguity.

variable, but also is seen in the emphasis on individuals rather than groups or aggregates of cases, and in the relegation of variables to a secondary place behind the consideration of the individuals. But these are differences in emphasis rather than differences of basic concepts, and can be taken care of by special uses of techniques or by refinements of these to suit geographical purposes.

CONCLUSIONS

Multivariate techniques deal basically with a large number of variables, but their purpose is more fundamental than this. It has been suggested that they are designed to deal with two main types of problem, the study of relationships amongst variables and the classification problem. Furthermore, within each of these two areas the different techniques examine a variety of questions and deal with the basic input of individuals.

On the other hand, the paper has suggested in the last section a few differences between multivariate and geographical methodologies. These perhaps necessitate refinements of the basic methods or their use with special geographical features such as position. Nevertheless, provided they are used carefully, bearing in mind the basic questions they ask, their procedural requirements and assumptions, and the need to use techniques repeatedly and with proper experimental design, they are valuable aids to more efficient research in geography.

University of Sheffield

APPENDIX — LIST OF VARIABLES USED IN DISCRIMINANT FUNCTIONS ANALYSIS

- X₁: population density in 1951 (persons per acre),
- X₂: % change in population, 1931—1951, total change,
- X₃: % change in population, 1931—1951, migration change,
- X₄: % change in population, 1951—1961, total change,
- X₅: males in social classes 1 and 2, as % of total employed males, 1951,
- X₆: males in social classes 4 and 5, as % of total employed males, 1951,
- X₇: occupied dwellings with 1—3 rooms, as % of total occupied dwellings, 1951,
- X₈: households with a density of less than one person per room, as % of all households, 1951,
- X₉: households which have five facilities, as % of all households, 1951,
- X₁₀: employment in primary activities as % of all employment, 1951,
- X₁₁: employment in manufacturing as % of all employment, 1951,
- X₁₂: employment in service activities as % of all employment, 1951,
- X₁₃: balance of commuting as a % of resident population, 1951,
- X₁₄: number employed as % of total resident population, 1951,
- X₁₅: sex ratio — females per 1,000 males, 1951,
- X₁₆: population over 65 years as % of total population, 1951.

The variables were transformed before being used. Numbers 2, 3, 4, and 13 were turned into index numbers by the addition of 100, and X_{10} had a value 1 added to each case. All were then transformed by the logarithm of the amount so as to achieve normal distribution curves.

BIBLIOGRAPHY

- [1] Anderson, T. W., *An introduction to multivariate statistical analysis*, New York, 1958.
- [2] Berry, B. J. L., A note concerning methods of classification. *Annals, A.A.G.*, 48 (1958), 300—303.
- [3] Berry, B. J. L., An inductive approach to the regionalization of economic development. In N. Ginsburg (ed.) *Essays in economic development*, Department of Geography, University of Chicago, Research Paper No. 62, 1960, pp. 78—107.
- [4] Berry, B. J. L., Barnum, H.G., and Tennant, R.J., Retail location and consumer behavior. *Papers, Reg. Sci. Assoc.*, 9, 1962, 65—106.
- [5] Blalock, H. M., *Social Statistics*, New York, 1960.
- [6] Bogue, D. J. and Harris, D. L., *Comparative population and urban research via multiple regression and covariance analysis*, Scripps Foundation for Research in Population Problems, Publication No. 8, Oxford, Ohio, 1954.
- [7] Burt, C., The appropriate uses of factor analysis and analysis of variance, In Cattell, R. B. (ed.), *Handbook of multivariate experimental psychology*, Chicago, 1966, pp. 267—287.
- [8] Carey, G. W., The regional interpretation of Manhattan population and housing patterns through factor analysis, *Geogr. Rev.*, 56 (1966), 551—569.
- [9] Cattell, R. B. (ed.), *Handbook of multivariate experimental psychology*, Chicago, 1966.
- [10] Cattell, R. B., *Factor analysis*, New York, 1952.
- [11] Cooley, W. W. and Lohnes, P. R., *Multivariate procedures for the behavioral sciences*, New York, 1962.
- [12] Ezekial, M. and Fox, K. A., *Methods of correlation and regression analysis*, New York, 1959.
- [13] Fruchter, B., *Introduction to factor analysis*, New York, 1954.
- [14] Haggett, P., Regional and local components in the distribution of forested areas in S. E. Brazil: a multivariate approach, *Geogr. Journ.*, 130 (1964), 365—379.
- [15] Harman, H., *Modern factor analysis*, Chicago, 1960.
- [16] Henshall, J. D., „The demographic factor in the structure of agriculture in Barbados”. *Trans. Inst. Brit. Geogr.*, 38, 1966, 183—195.
- [17] Hart, J. F. and Salisbury, N. E., Population change in middle western villages: a statistical approach. *Annals, A.A.G.*, 55 (1965), 140—160.
- [18] Isard, W., *Methods of regional analysis*, Cambridge, Mass., 1960, chapter 7, Appendix B: Scaling and latent structure techniques, pp. 281—293.
- [19] Kendall, M. G., Factor analysis as a statistical technique. *J. Roy. Stat. Soc.*, B12 (1950), 60—73.
- [20] Kendall, M. G., *A course in multivariate analysis*, London, 1957.
- [21] Lavalley, P., McConnell, H., and Brown, R. G., Certain aspects of the expansion of quantitative methodology in American geography, *Annals, A.A.G.*, 57 (1967), 423—436.
- [22] McConnell, H., A statistical analysis of spatial variability of mean topographic slope on stream dissected glacial materials, *Annals, A.A.G.*, 56 (1966), 712—728.
- [23] Mayfield, R. C., The range of a central good in the Indian Punjab, *Annals, A.A.G.*, 53 (1963), 38—49.
- [24] Miller, R. G., Statistical prediction by discriminant analysis, *Meteorological Monographs*, vol. 4, No. 25, Boston, 1962.
- [25] Miller, R. L., and Kahn, J. S., *Statistical analysis in the geological sciences*, New York, 1962.

- [26] Moser, C. A., and Scott, W., *British towns*, London, 1961.
- [27] Rao, C. R., *Advanced statistical methods in biometric research*, New York, 1952.
- [28] Roberts, M. C. and Rumage, K. W., The spatial variations in urban left-wing voting in England and Wales in 1951, *Annals, A.A.G.*, 55 (1965), 161—178.
- [29] Robinson, A. H., Lindberg, J. H., and Brinkman, L., A correlation and regression analysis applied to rural farm population densities in the Great Plains, *Annals, A.A.G.*, 51 (1961), 211—222.
- [30] Roy, S. N., *Some aspects of multivariate analysis*, New York, 1958.
- [31] Stone, R., A comparison of the economic structure of regions based on the concept of distance, *J. Reg. Sci.*, 2.2 (1960), 1—20.
- [32] Thompson, J. H. et al., Toward a geography of economic health: the case of New York State, *Annals, A.A.G.*, 52 (1962), 1—20.
- [33] Wong, S. T., A multivariate statistical model for predicting mean annual flood in New England, *Annals, A.A.G.*, 53 (1963), 298—311.

Acknowledgements:

The author wishes to thank the Computer Center, Indiana University, U.S.A., for the use of its computing facilities in connection with both examples described in this paper. He also wishes to thank the cartographic staff of the Geography Department, Sheffield University, for their preparation of Figs. 2, 3, 4, and 5.

THE POINT OF CHANGE FOR METHODS IN THEMATIC CARTOGRAPHY

LECH RATAJSKI

During the Third International Cartographic Conference, held in Amsterdam in April, 1967, the author read a paper on the point at which methods of cartographic generalization should change. This is shortly to be published in the International Yearbook of Cartography.

I would here like to outline the problem of the so-called point of changing methods. These points are associated with pairs of related methods. Such pairs of methods can be distinguished, for example, in the case of two cartographic techniques which are graphically similar but differ in the presentation of values, for instance either continuously or at intervals.

The continuous cartodiagram and the interval cartodiagram will serve as an example of such a pair of methods. A cartodiagram is a map on which the values are presented by means of proportionate symbols — circles, squares, triangles, etc. The continuous cartodiagram shows each individual value with a symbol of a specific size and the interval cartodiagram shows symbols grouped into a number of classes of different size. In the first case there is a continuous scale of values and in the second case values are expressed at intervals.

Both these cartographic methods are the subject of the author's investigations when studying the cartometry of industrial maps. In the course of the study it was discovered that when using the continuous cartodiagram errors of measurement began to increase rapidly from a certain point. This error is caused as much by the thickness of the drawn line as by limitations of ocular precision, amounting to 0.2 mm. The magnitude of this error is illustrated on the graph (Fig. 1). It appears that in the case of the most commonly occurring sizes of symbols viz. between 2 and 10 mm, the error of measurement oscillates in the following way:

- from 24.3 to 6.3 per cent for a circle symbol,
- from 40.8 to 12.1 per cent for a square symbol and,
- from 54.4 to 17.3 per cent for an equilateral triangle.

When using the interval diagram the average systematic error of measurement depends on the magnitude of intervals and these in turn depend on the mathematical progression and its index. The magnitude of that error may be illustrated by the following formula:

for the intervals resulting from the geometrical progressions:

$$X = \frac{50(q-1)}{q+1} \quad \text{where } q \text{ is the quotient of the progression.}$$

When q is 2 the error amounts to 16.66 per cent, when q is 3 the error amounts to 25 per cent and so on until it reaches a maximum at 50 per cent. With the use of arithmetical progression the magnitude of this error is constant, amounting to 25 per cent.

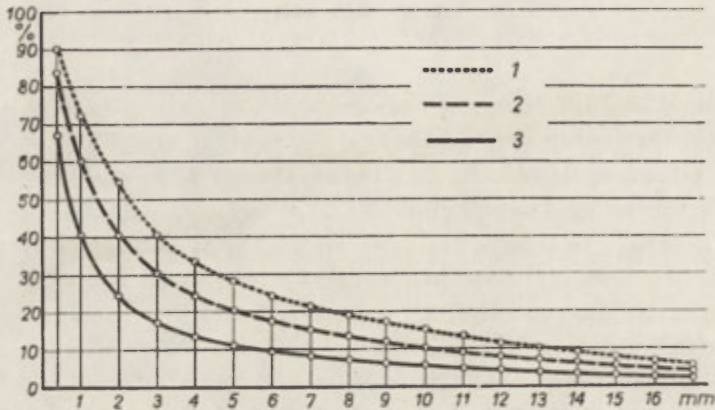


Fig. 1. The error graph

1 — for equilateral triangle (the height), 2 — for square (the side), 3 — for circle (the radius)

In the case of the circle symbol an error of 16.66 is already reached when it has a radius of 3.2 mm, and an error of 25 per cent occurs with a radius of 2 mm. In the square symbol these values are correspondingly 6.3 mm and 3.8 mm for the length of the sides and for an equilateral triangle 9.4 mm and 5.8 mm for its height. In conclusion, for small values the interval cartodiagram is better than the continuous cartodiagram on the grounds that values cannot accurately be measured if the former technique is used. Since the same error occurs with both types of symbol it is more rational to replace the continuous diagram by the interval one and vice versa. This is the point of changing methods (see Fig. 2).

The next pair of methods to be discussed are the continuous choropleth symbol and the interval choropleth symbol. In the continuous case values are represented by the number of lines contained in 1 cm. on the map (Fig. 3). The number of these lines that can be shown is also limited by ocular precision. Assuming the thickness of a single line to be 0.2 mm and the distance between any two lines 0.3 mm, 20 distinguishable lines can be compressed into 1 cm on the map. The precision of the eye also influences one's ability to distinguish the areal units possessing different intensities of blackness or colour. In the first case the Weber-Fechner law operates; in the second we have to allow for the sensitivity of the eye to changes of colour tone. Following Arthur H. Robinson and John S. Keates,

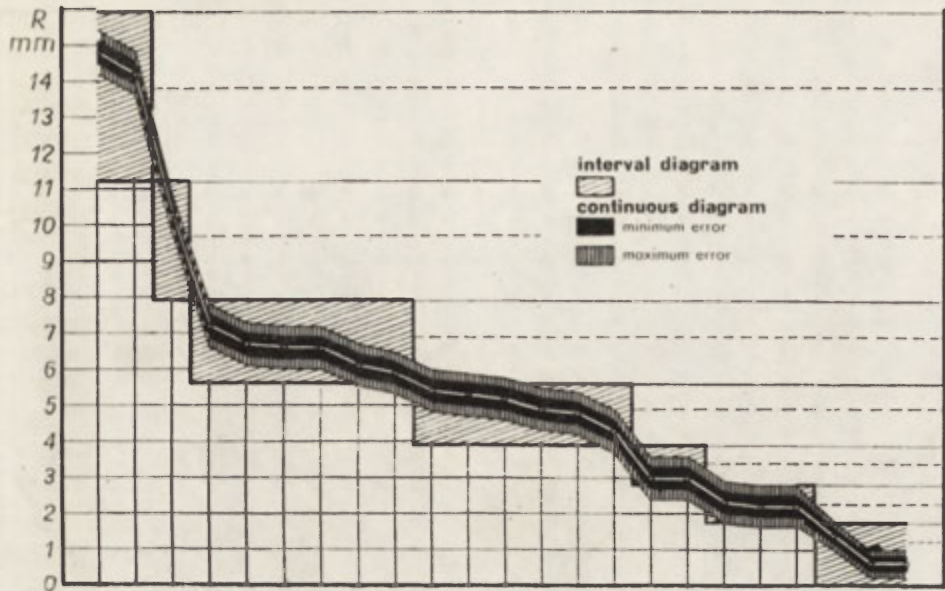


Fig. 2. The graph of error bands

the author thinks practical considerations permit one to distinguish from 10 to 12 gradations. Therefore this number defines the nature of easily legible intervals in the choropleth symbols employed.

According to the range of values to be portrayed the intervals will vary in size and, as a result, there will be a corresponding systematic error defined by the formula cited above. This error could be reduced by the application of the continuous choropleth symbol for the low, but usually common, values of a series of numbers, whilst keeping enough intervals for the higher values, which are usually less numerous. This procedure has been checked by compiling popula-

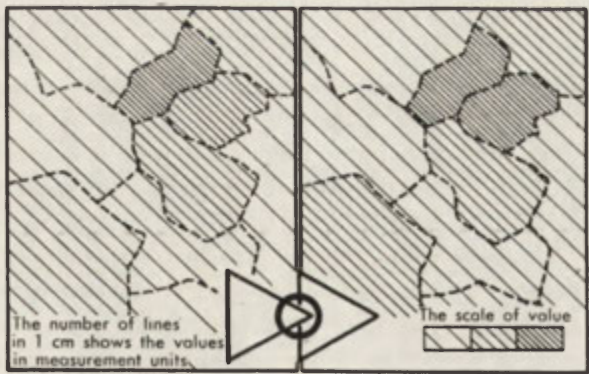


Fig. 3. The continuous (left) and the interval (right) choropleth symbols

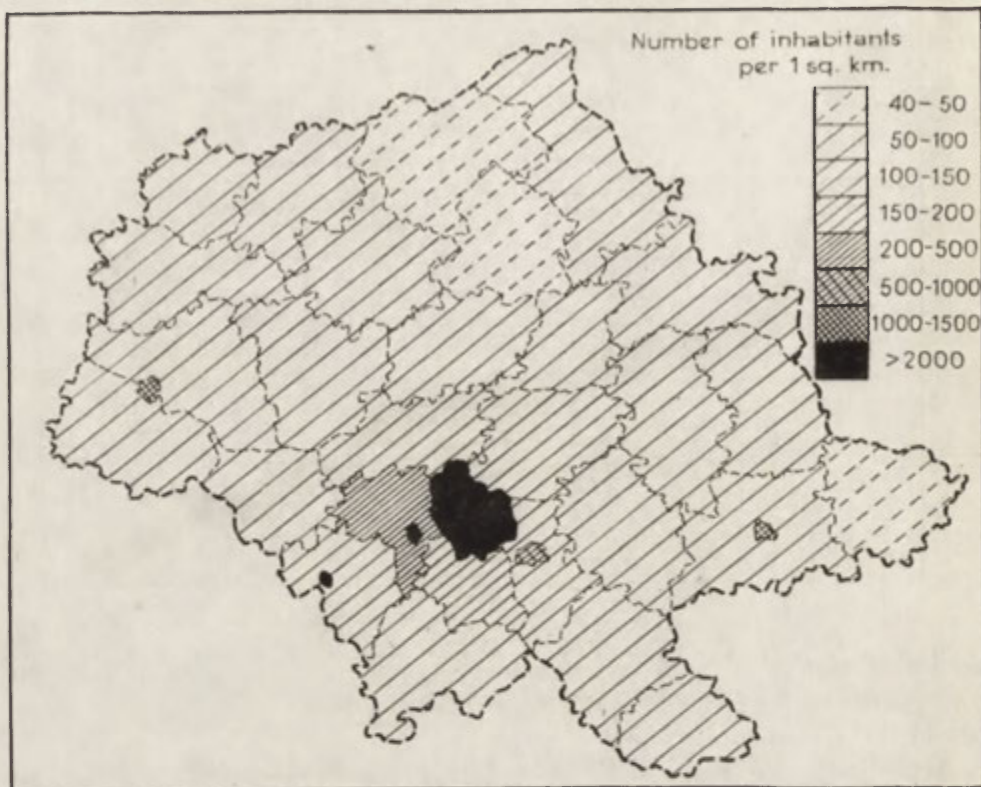


Fig. 4. An interval choropleth map

tion density maps of the Warsaw district. First, an interval choropleth map using only 9 intervals was compiled. Next, the total population for the whole district was read from the map. The value obtained differed from the original number of inhabitants by 10.3 per cent (Fig. 4). Next, another map (Fig. 5) was compiled, using the continuous choropleth for low values and the interval choropleth for higher values. Because the minimum density of population in the enumeration areas is 47 per sq. km. one line was made to correspond to a density of $40 + 2$. Thus, the point of changing methods is met at a value of 80 inhabitants per sq. km. For the remainder of the map the interval choropleth of 8 intervals was employed. The population figure read from the map was very close to the original one, with an error of only 0.3 per cent. The point of changing methods for this pair of choropleth symbols occurs where the density of lines in 1 cm. is at a maximum (20 lines). However, this value is not the same in every case.

Finally, for the third pair of methods the uniform dot method (Fig. 6) and the multi-scale dot method (Fig. 7) the point of changing methods is defined by the limit of legibility. In this case the most important thing is to count dots on

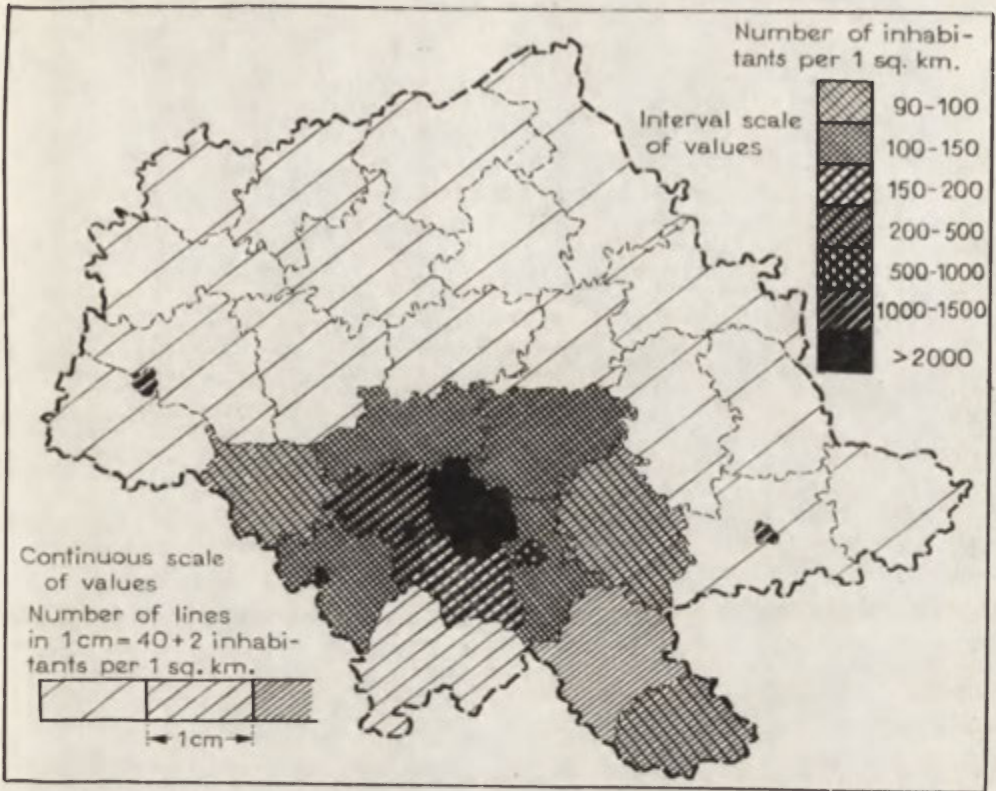


Fig. 5. An interval choropleth map with continuous scale for low values

a sufficiently large area. In practical terms the dot should not be smaller than one having 0.4 mm diameter and the spacing between dots should not be less than 0.3mm. From these assumptions a maximum of 200 legible and distinct dots can be placed in 1 sq. cm. of map area. This density of the dots defines the point of changing methods. If it were necessary to present a value greater than this density for 1 sq. cm. one would be forced to replace the classic uniform dot method by the multi-scale one. By using this idea of the point of changing methods cartometric considerations will be respected and the coalescence of dots into an illegible mass can be avoided.

Unfortunately these investigations on cartographic methods are here limited to those mentioned above. However, from the examples discussed above it may be supposed that other cartographic techniques can also be viewed in terms of related pairs of methods. The realization of the existence of the cartographic pairs of methods conditioned by the point of changing methods should make the construction of economic maps easier by the use of more logical methods.

University of Warsaw



Fig. 6. The uniform dot method. One dot represents 1000 inhabitants



Fig. 7. The multi-scale dot method

PRECISION IN TERMINOLOGY: AN EXAMPLE FROM FLUVIAL MORPHOLOGY

CHRISTINE S. SUNLEY

An examination of the literature concerning all aspects of equilibrium in river systems has indicated the importance of definition of terms in such a way that there is no ambiguity in their interpretation. Almost every term relating to rivers has been used in more than one sense by different workers in the field.

The term "graded river" is a notorious example. "Graded" has been variously used to indicate rivers which possess one, or a combination of the following: an approximate balance between erosion and deposition, permitting only slow change; a smooth longitudinal profile; a worn profile (Makkaveyev, 1955); a mild slope (Scheidegger, 1961); an exponentially diminishing size of bed material and degree of slope in a downstream direction (Sternberg, 1875; Schoklitsch, 1937; Shulits, 1941); a profile of maximum probability (Leopold and Langbein, 1962); a responsiveness to external change; a degree of adjustment between velocity, width, depth and other variables; adjustment to base level; an end profile or a minimum slope. Various other terms which contain some or all of the characteristics included in the term "graded" are in common use. Rivers, or stretches of rivers, may thus be described as stable, balanced, poised (Matthes, 1941; Fisk, 1951), in regime, of regime-type (Blench, 1951), in a steady state, in dynamic equilibrium, in quasi-equilibrium or in silting equilibrium (Crickmay, 1960)¹. The usefulness of the word "grade" has been questioned by J. Challinor (1961) and by G. H. Dury (1966) and its meaning restricted by J. E. Kesseli (1941) and E. W. H. Culling (1955). Yet the term persists in recent geographical, geological and engineering literature. The multiplication of near synonyms has been matched in recent decades by a resurgence of interest in the question of fluvial balance.

The need for a more precise terminology is indicated, but the problem is not overcome simply. The minimum requirement that each author should state his definitions so clearly that his meaning cannot be misunderstood is less important than the need that the distinctive characteristics of rivers may be recognized and expressed in measurable terms. Improvement in methods of measurement and an increase in understanding of the theoretical interrelationships must precede a major improvement in terminology.

Variables such as bedload, dominant discharge, velocity, and many others

¹ References are given above to those terms cited which are not in widespread use.

act as building blocks in the construction of complex theories of grade. The attainable precision in the definition of the graded river will depend in part on the definition of these components.

The charge or load of a stream is an important factor in fluvial equilibrium. Most workers attempt a subdivision of the load when discussing its effects. Thus L. G. Straub (1934) writes: "It is the bedload which certainly must be of primary importance in defining the stability of a channel", and E. Kuiper (1960) writes: "the slope of a graded river is determined by the bed material load and not by the wash load". The subdivision of the load well illustrates a disparity between qualitative recognition and quantitative appreciation. Since the eighteenth century attempts have been made to understand the mechanics of sediment moving near the bed of rivers but the difficulties of defining boundaries which are both measurable and meaningful have not yet been surmounted.

In 1908, W. J. McGee described three methods of transport of load: — "a) in solution, b) in suspension and c) in what may be denoted saltation". G. K. Gilbert (1914) adopted this latter term to indicate not the whole bed load but an intermediate position between movement in close contact with the bed and that in suspension (p. 15): "Streams of water carry forward debris in various ways. The simplest is that in which the particles are slidden or rolled. Sliding rarely takes place except where the bed of the channel is smooth. Pure rolling, in which the particle is continuously in contact with the bed, is also of small relative importance. If the bed is uneven, the particle usually does not retain continuous contact but makes leaps, and the process is then called saltation... With swifter current leaps are extended, and if a particle thus freed from the bed be caught by an ascending portion of a swirling current its excursion may be indefinitely prolonged. Thus borne it is said to be suspended, and the process by which it is transported is called suspension". Although Gilbert was able to make qualitative observations of the various types of movement through the glass walls of his experimental flumes, his definitions have raised objections firstly because of the difficulty of quantification of the actual amount of debris moving in those ways and, secondly, because of theoretical objections to the validity of such a division.

M. P. O'Brien and B. D. Rindlaub (1934) suggest that load can only be defined by reference to the method of measurement (p. 597): — "The term bed load refers to that portion of the material carried by a stream or channel in such a manner as to be caught in a trap of reasonable length and proper design". The design of the trap will influence the amount of material caught. If too large, then both bed load and debris normally travelling in suspension will be deposited. If too short, then the saltating load will be under-represented. This question raises the problem of getting consistent results, either to compare the rate of movement in one stream at different discharges or to compare the rates in different streams. The separation of bed load from the suspended load in rivers

will obviously depend on the types of samplers in use. As new methods of investigation, such as the use of radiosondes, photographic methods, or the use of radioactive pebbles are introduced, the definition of the terms must alter. The type of instrument is important but the length of the records and the frequency of the readings will also influence the definition of each category. For instance, R. Boissier (1916) demonstrated that in the case of rivers with glacial regimes, even daily readings were inadequate to define the material carried in suspension, for the load could double within an hour.

When the total load is measured by the examination of deposits in reservoirs or by turbulence flumes, which transform the whole load into suspension for sampling, or by readings taken in rocky stretches where it appears that the whole load is naturally in suspension, the problem of subdividing the bed load from the rest involves examination of the material composing the bed of a nearby, normal stretch of the river. Whether for this purpose or for collecting data on the transporting ability of the stream directly, the interpretation of material on the bed of a river involves questions of discharge and recognition of lag deposits. During low water, when the material is most easily examined, much material normally in suspension will rest on the bed. The definition must be linked to the discharge in order that consistency, if not precision, may be achieved. H. A. Einstein (1950) proposed the following definitions;

“Bed load: Bed particles moving in the bed layer. This motion occurs by rolling, sliding, and, sometimes, by jumping.

Suspended load: Particles moving outside the bed layer. The weight of suspended particles is continuously supported by the fluid.

Bed layer: A flow layer, two grain diameters thick, immediately above the bed. The thickness of the bed layer varies with the particle size.

Bed material: The sediment mixture of which the moving bed is composed.

Wash load: That part of the sediment load which consists of grain sizes finer than those of the bed.

Bed-material load: That part of the sediment load which consists of grain sizes represented in the bed.”

Although these definitions provided a useful basis for theoretical discussion, R. A. Stein (1965) objects:— “Although this appears precise enough there is no accurate way to measure it”.

The separation of lag deposits from material in motion is no easy matter. J. P. Miller (1958) states: “in high mountain streams little movement of bed material occurs”, and this he tentatively attributes to “residual effects of Pleistocene glaciation and frost action on stream regime.” In contrast to this evidence from the United States, A. Schoklitsch (1937) reports that from “obtainable data on accumulation in reservoirs” (which included material transported as both suspended and bed load) “for Alpine districts, it may be assumed that about $\frac{2}{5}$ to $\frac{4}{5}$ of the deposited material had been transported along the stream bed”.

C. H. Crickmay (1960) suggests that on the bed of streams at grade or "silting equilibrium", "the maximum alluvial fragments, on the one hand, and the bed slope, on the other, have become so adjusted to supermaximum and total discharges that the coarsest fraction of the transported debris lies inertly upon the bed for indefinitely long periods and, by being there, reduces vertical corrasion to an infinitesimal quantity". Whether the material on the bed dates from the Pleistocene period, or from "maximum" flows, presumably in the more recent past, or whether it represents present-day bed load can only be solved by more field research. Until the interpretation of the bed material rests on a firmer footing, the question of the definition of bed material load will be far from precise.

In certain stretches of some streams, mapping of dunes may be used to estimate bed load movement, as was attempted by A. A. Humphreys and H. L. Abbott (1861) near the mouth of the Mississippi. D. B. Simons, E. V. Richardson and C. F. Nordin (1965) state that: — "The bedload in most natural channels cannot be measured economically by presently available (1964) sampling techniques", yet "a simple relation based on the mean height and velocity of the sand waves can be used to determine bedload transport".

This latter method may have some theoretical justification as well as practical potential, although the saltating load and the bed load as defined by Gilbert will be considered together. The division of the load is a theoretical as well as a practical problem.

Kalinske (1942) attacks the term "saltating load" as not only insignificant but inseparable from suspended load in its method of travel. Both the suspended load and the saltating load are picked up by turbulent fluctuations and the difference in the rate of fall is a gradation not justifying a separate classification. Unlike R. A. Bagnold's justifiable separation of surface creep in describing the motion of sand grains in a desert, sand grains in water do not set others in motion by bouncing.

Other workers suggest that the separation of bed load from the "wash load" is the most relevant. In this case much of the suspended load is included in the bed load or bed-materials load, for example, A. G. Anderson (1942). "The sediment load of a natural stream may be considered to be made up of two parts, the wash load and the bed-materials load. The wash load is the fine material washed into and through the reach under consideration by storm runoff. The bed-materials load is derived from the bed of the stream and may be transported either by rolling along the bed or in suspension". This particular definition includes some imprecision by lack of definition of storm runoff and complication by including the derivation of the sediment. A simpler definition was proposed by H. A. Einstein, A. G. Anderson and J. W. Johnson (1940, p. 632): — "*Bed-load* is that part of the total sediment-load composed of all particles greater than a limiting size (42-mesh sieve for the Enoree River) whether moving on the bed or in suspension, and includes all bed-material in movement.

Wash-load is that part of the total sediment-load composed of all particles finer than the limiting size (42-mesh sieve for the Enoree River) which usually is washed into and through the reach under consideration". The use of the 42-mesh sieve size as a criterion for this particular stretch was not entirely arbitrary for "only material coarser than this size is available in the bed in sufficient quantity to be considered abundant" and the "material in these two groups is transported according to different laws". The coarser material "whether moving on the bed or in suspension, is moved at a rate of transportation that seems to be a function of the discharge. On the other hand, the material finer than the 42-mesh sieve is merely "washed through" the section of the river without deposition".

This distinction of size would be difficult to justify in stretches where there is little difference in size between the suspended load and the bed material, when, for instance, "silt and clay on the bed of a slow-moving stream may be motionless at the same time that silt and clay are moved in suspension" (H. W. Menard, 1950).

In the above example the separation was justified on the basis of a different response to discharge. Other reasons have been suggested by R. Sakthivadivel (1965, p. 202):— "the movement in suspension follows principles which are entirely different from those that give bedload transport. While the weight of suspended solid particles is supported by the surrounding fluid and moves with the horizontal velocity component equal to that of surrounding water, the weight of bed load is supported by the moving bed and not by the fluid".

Whilst many attempts have been made to produce bed-load formulae, mainly from flume studies, and methods have been devised of extrapolating the sediment load from a few readings, the development of total load formulae is in its early stages. A. J. Raudkivi (1967) suggests it is for this reason that the separation into suspended and bed load is acceptable: — "This division lacks physical meaning, but so long as there is no general theory explaining all phases of sediment transport by one function it is justified".

As theoretical knowledge progresses new bases for subdivision may emerge. R. A. Bagnold's work with large grains only a little more dense than the transporting fluid suggests that "Fluid turbulence clearly cannot be regarded as an essential requisite in grain transport by fluids. At the mean concentration of 35%, at which the flow becomes laminar, the mean separation distance between the moving grains is still as much as 0.3 of a grain diameter, so no question of permanent rolling contact can arise; yet the grains are maintained indefinitely at this dispersion against the action of gravity. This requires a dispersive force acting between the moving grains large enough at the bed level to support the whole immersed weight of the moving load above". He uses the support of the grain stress and the support by fluid turbulence as a theoretical basis for defining bed load and suspended load. The extension of these experiments to natural channel conditions has not yet been attempted.

The question of bed load has been discussed in some detail as indicating

the demands of precision in terminology. In order to be precise the theoretical background must be well developed to achieve meaningful subdivisions. Qualitative recognition may not be matched by the available methods of measurement. Precision is an unattainable goal with the present methods of measurement and even consistency is difficult to achieve.

Most discussion of grade or equilibrium in rivers involves relating discharge to slope and other parameters. It is agreed that not all discharges are equally significant in the formation of river channels. Although some workers such as L. B. Leopold and T. Maddock (1953) use, for expediency, mean annual discharge, much discussion has taken place to define the most effective discharge in terms of frequency and quantity of work performed.

In 1941, C. C. Inglis introduced the term "dominant discharge" as "the discharge which controls the meander length and breadth. It appears to be slightly in excess of bank-full stage". Despite this latter proviso, L. B. Leopold and M. G. Wolman (1957) consider "bankfull discharge . . . equivalent to "dominant discharge" of the Indian literature". K. M. Bhatia (1946) suggests "a dominant discharge is that discharge which occurs the largest number of times". K. K. Framji (1946) says that "an easier estimate of the dominant discharge was available in the simple arithmetical average of daily discharges within a specified period when discharges were in excess of the discharge at which bed movement began in the river" and he suggests the term "dominant-formative discharge". H. M. Morris (1963) defines dominant discharge as "that discharge above which half the total sediment load is carried. This discharge is also called the bed-building discharge and the corresponding river stage the bed-building stage". T. Blench (1951) suggests that the dominant discharge is an abstraction, a "steady discharge that would produce the same result as the actual varying discharge."

The semantic problem is apparent. At least four terms are being used to describe the same concept. Yet when these definitions are closely examined it is seen that they are not strictly synonymous but are based on different assumptions and sparse data.

The definitions may be classified into those which attempt to define dominance by comparison between discharge frequencies and sediment charge and those which regard the actual channel as a significant indicator.

The first group require records of flow duration and the relationship between the discharge and the sediment load. There are few rivers for which records of flow are available for long periods yet the length of record will very much affect the definition of dominant discharge. This is demonstrated by M. Nixon (1959), who, looking at the frequency of a particular discharge of the River Thames, found that the value varied according to the length of time considered: —

20.6%	1-year period, 1951
4.3%	5-year period, 1951—5
2.9%	72-year period, 1883—1954

This means that for any definition of dominant discharge to be precise, the length of record considered must be included in the definition. H. M. Morris' definition of bed-building discharge and that which A. Schoklitsch attributes to F. Schaffernak, i.e. "The bed-forming discharge is defined as the discharge which, because of its duration, is capable of transporting the most material during the year", require knowledge of the amounts of sediment related to the various discharges. K. K. Framji's definition requires knowledge of the discharge at which bed load movement begins. These relationships are judged either from direct readings or from bed-load or total load formulae. Schoklitsch suggests that his own bed-load formula may be applied, whilst Morris states: — "If most of the load is bed load, a formula of the DuBoys' type may be used. If both bed load and suspended load are present in appreciable amounts, then one of the total load methods must be used." However, the application of bed-load formulae to natural streams has been met with little success and Raudkivi has little confidence in the total load formulae: — "errors of 100 per cent are to be expected". K. K. Framji used Schoklitsch's method and also an examination of the cross-sectional area to record departure from a "datum area". The significant discharge was the one "which caused the datum area to increase markedly; for from that stage onwards the river began to be active and to modify not only the bed but also the sides". M. A. Benson and D. M. Thomas (1966) use Schaffernak's definition, although without reference to his work, but, instead of using formulae to compute bed load or total load, they consider that "because only information on suspended sediment load is generally available, only suspended sediment can be considered and it is necessary therefore to assume that total sediment load is proportional to suspended load". Although the same assumption is made by M. G. Wolman and J. P. Miller (1960) it must be regarded with some scepticism after reports such as that of K. B. Schroeder and C. H. Hembree (1956) which showed that for wide, shallow, sandbed streams, the unmeasured sediment load can often vary from 20 to 60 per cent of the total. The unmeasured load in the last statement lies below the depth to which the DII-48 sampler can reach. Whilst Schoklitsch was satisfied with treating only the bedload in his definition, Benson and Thomas attempt a definition on measured suspended sediment alone. However, they did not meet with much success, for suspended sediment loads "of almost the same amount (per equal discharge intervals) are transported over a large range of discharge. Therefore the concept of dominant discharge may not be too meaningful". Similarly, discussion of Framji's definition showed that the significance of the dominant discharge defined in this way was not apparent. It can be seen that the above attempts at definition are far from precise because of the difficulties of obtaining adequate flow records and the even greater difficulty in estimating the total load of the stream.

The second group do not need to estimate the sediment load but interpret the channel itself as the result of a significant discharge. The significance has

been shown by correlation with other morphological features and by the demonstration that the frequency of attainment of this stage is remarkably constant for streams of all sizes. M. G. Wolman and L. B. Leopold (1957) relate bankfull discharge to meander wavelength and M. G. Wolman and J. P. Miller (1960) argue: — "Because the bankfull stage is equal to the elevation of the flood-plain surface, and overbank flows contribute only a small part of the flood-plain sediment, the bankfull discharge appears to be the "effective" discharge controlling the development of the floodplain". On the other hand, C. W. Carlston (1965) suggests that a stage less than the bankfull is best correlated with meander wavelength and M. G. Wolman (1959) shows from a short study of a stream in Maryland that the "lateral cutting of the cohesive channel bank ... occurs mostly during winter months" at stages less than bankfull. Discharge is not the only factor for the duration of the discharge and the wetting of the banks will influence the amount of erosion performed.

If the bankfull stage is accepted as the dominant discharge, there is still difficulty in its precise definition and determination of frequency. M. Nixon (1959) proposes: — "the bank-full state: a bank-full flood of a river is the highest flood which can be contained within the incised channel of the river and which does not cause spilling of water on to the river's flood plain or washlands. The bank-full discharge is the maximum discharge reached by a bank-full flood at any point in the course of the river." He points out practical difficulties in the application of this definition, that the floodplain may be wet before the river reaches bank-full, that the gauging station sites in stable situations may often be below bank-full after this stage has been reached in other parts of the basin and so on. "Even within the definition used there can be some doubt as to the precise value of the bankfull discharge". To determine the frequency a time scale must be chosen. At present, two studies, by M. Nixon in Britain and Wolman and Leopold in the U.S.A., have produced results which are difficult to compare. Wolman and Leopold base their results on annual flood series and suggest a recurrence interval of about 1.5 years. Nixon takes all floods and states, "the value for an average river in England and Wales is 0.6%, therefore it is equivalent to a little over 2 days per year". Despite the difficulties of definition, Nixon suggests, "rivers will adjust their geometry where they are free to do so until the frequency of their bank-full discharge remains at this value, and therefore it is a condition of regime".

Sediment load and discharge are important variables in the establishment of equilibrium in rivers. Other characteristics, such as velocity, slope, roughness, and base level have also proved elusive to precise definition, e. g. J. H. Mackin (1963) points out some of the caution needed in the interpretation of data on velocity distribution in river courses.

As the problem becomes more complex, as more variables enter into the discussion, the degree of precision to be expected must decline. In discussion of the

concept of grade, the choice cannot be between a precise definition and a vague one but rather whether the utility of the concept outweighs the unavoidable imprecision. It is evident that the degree of tolerance varies widely from author to author. Recently the increasing emphasis on measurable change has led to the rejection by some authors of qualitative definitions based on a long time scale. Thus, G. H. Dury (1966, p. 231) writes; — “the concept of grade . . . is to be recognized as unserviceable both in the study of actual terrains and in the theoretical analysis of landforms generally.” Such conclusions are based on the assumption that the crudely measurable parameters of velocity, discharge, slope and load are the only significant indicators of equilibrium. This attitude is justified on the basis of utility with regard to engineering projects and short-term stability but does not answer the wider problems on which the concept of grade has appeared to throw some light.

The concept of grade in the widest sense serves two purposes: firstly, it is an expression of the limits to fluvial action and, secondly, it provides a framework for the classification of streams and their valleys.

The idea that there must be some limit or opposing action to the work of running water has excited interest for at least three hundred years. This idea that fluvial processes are working towards an equilibrium condition was as clearly stated by DuBuat (1779) as by C. C. Inglis (1949) to take two examples at random from the voluminous literature. Indeed the stimulus of this concept may perhaps be attributed to its lack of precision. The philosopher R. B. Braithwaite (1953) writes; — “it is only in theories which are not intended to have any function except that of systematizing empirical generalisations already known that the theoretical terms can harmlessly be explicitly defined. A theory which it is hoped may be expanded in the future to explain more generalisations than it was originally designed to explain must allow more freedom to its theoretical terms than would be given them were they to be logical constructions out of observable entities. A scientific theory which, like all good scientific theories, is capable of growth must be more than an alternative way of describing the generalizations upon which it is based, which is all it would be if its theoretical terms were limited by being explicitly defined”. Thus this ambitious idea, the search for balance, for equilibrium or limits has led to the growth of knowledge on river control and the design of irrigation canals. The time scale involved necessarily alters what rate of change may be regarded as a limit. Engineers and many modern geologists may be happy to ignore change which is not sufficiently large to be measured. Thus T. Blench (1951, p. 2) states: — “A river is said to be “in regime” in a reach if its mean measurable behavior during a certain time interval does not differ significantly from its mean measurable behavior during comparable times before or after the given interval . . . A record of a few weeks would be worthless, and to require a record of centuries would probably disqualify every channel on earth from passing the test of regime.” This short-term “limit” may

be contrasted with the value of sea level or base level as a limit in discussion of landscape evolution. Indeed graded and base level were thought to be synonymous by several late nineteenth century geologists until, in 1902, W. M. Davis made the valuable distinction between long-term effect of ultimate base level and the important though less absolute limit of grade. Grade represents a slowing in the rate of change sufficient to be reflected in the landscape. Thus, the graded slope, although adjusted to base level, is also adjusted to climatic conditions, to the sediment supply and size of the drainage basin. A change in these conditions, even without a change in grand base level may produce significant landscape effects, as demonstrated by W. D. Johnson (1901).

Interest in this stage of slowing down of change, of relative stability, inspired the interpretation of the landscape in terms of stages of rapid change and stages when the limits were more closely approximated.

In the short-term, when crude measurements of change are possible, only very rapid change, usually brought on by man's interference, is regarded as a departure from regime or quasi-equilibrium conditions. Even within this period of river records, such a distinction can only be relative and so, because man-induced change is so widespread, the U.S. Geological Survey has found the need for the study of some isolated river basins as hydrological "bench marks" by which to assess the rate of change elsewhere.

The time scale involved influences the cause and effect relationships, as shown by S. A. Schumm and R. W. Lichty (1965). In the short-term, the slope cannot be as radically altered as can the width to meet the changing conditions, but in the long-term the slope becomes a fully dependent variable. The short-term view emphasises that the interrelationships between the characteristics are more important than study of each independently. In the long-term view one characteristic, slope of the river, is given greater emphasis. The justification lies in the inference that, over long periods of time, slope-changes, for example as a mountain stream creates its valley, are likely to be far greater than the width-changes during that time. The two viewpoints should aid each other. The idea of a graded profile as a smooth profile cannot be other than imprecise and detailed study of a stream's profile must take into account the other interrelationships. Small slope changes cannot be analysed with the blunt tool of a geological-time concept of grade. Yet the long-term view may be valuable for understanding of the fundamental causes of the fluvial landscape. C. C. Inglis (1941) points out that an explanation for meandering which suggests that meanders are caused by excess valley slope cannot be fundamental for: — "It is more correct to say that (except where the slope of the country has been affected by tilting due to earthquakes, variations by strata or other disturbing factors) the slope of the valley is determined by the river conditions rather than the river conditions by the slope of the valleys." The short-term view may be myopic, but the long-

-sight of the geological view is not adequate for the close examination of a river's profile.

As a basis for classification of rivers, the concepts of equilibrium relating to both time scales have limitations.

The geological concept is an ambitious one based on indications of the past history of the river, on its present appearance and on prediction of future action. For example, the Lower Mississippi has been termed graded or one of the near-equivalent terms, by several authors, such as H. Gannett (1900, "graded"), H. Baulig (1935, "équilibre mobile"), or H. Fisk (1951, "poised"). Examination of the sediments in the Lower Mississippi basin indicates that rapid aggradation took place after the last glaciation until about 5000 years ago (see Fisk). The contrast between the coarse deposit of this period of aggradation and the relatively fine-grained alluvium of today is one indication that the rate of change has slackened. The present appearance of the river, with its relatively smooth profile leading without break to base level, with a regular diminution of bed material in a downstream direction (see S. Shulits, 1941), has led to a prediction that its future behaviour, in the absence of external change, is likely to show little alteration. In deeming the Shoshone east of Cody "graded", J. H. Mackin (1948) used evidence of terraces indicating maintenance of the present river slope during a considerable period of downcutting, as well as the present smooth profile, and he predicts that this slope would be eventually re-established even were the area to be disturbed by earth movements. However, so much evidence is demanded by such a classification and so many of the criteria can only be assessed subjectively that there has been little agreement about its application. The crude contrast between such a river as the Lower Mississippi and, for example, the irregularity of profile of the Niagara or a mountain torrent with waterfalls and lakes, may be obvious, but the classification cannot be easily applied to the wide range of rivers with histories and characteristics between these extremes.

In the short-term a classification is also difficult to apply: most of the criteria of adjustment are gradations on a scale of values. H. Baulig (1926) and T. Blench (1951) suggest that a continuous alluvial channel is one criterion and they agree that the degree of stability then depends on the measurable rate of change. In the assessment of streams with rock walls or immovable beds the problem becomes more difficult. T. Blench suggests that the stream will approach the regime-type in so far as the flow is able to adjust some characteristics. The only example he gives of a river channel that is completely nonregime is "a rock channel ... when free of deposition". By examining mountain streams, J. P. Miller (1958) confirmed the opinion that these so-called ungraded streams would have characteristics in common with other streams traditionally regarded as graded. He suggests that (p. 49) "The dearth of reliable criteria for recognizing equilibrium

conditions in modern streams is a shortcoming of the concept of grade", yet, he adds (p. 50): — "As a general framework for studying ancient stream features, the concept of grade seems to serve a useful purpose."

The over-expectation of precision has been one of the stumbling blocks in the application of the geological concept of grade. Another has been the equation of the term "grade" with terms such as regime or dynamic equilibrium, which are only based on short-term interactions. So many terms have been linked to the word "grade" that its use can only lead to confusion unless accompanied by long explanation. The concept may have some value, even though imprecise, but the word has lost much of its utility.

W. M. Davis admonished: — "Use no term that you cannot define, make no statements that you cannot defend". This advice is a counsel of perfection. Many terms, such as "gravity", are valuable although precise definition is not possible. All definitions vary in their completeness. In this paper no solutions are offered to the many problems indicated, but an attempt has been made to show the reasons behind the confusion of terminology. The plea is not only for more data, involving both measurement and qualitative observation, but for better communication of the limits of precision of the terms used. To quote M. L. Hill (1963) "The use of names which imply knowledge that does not exist is particularly reprehensible."

The complexity of the phenomena which geographers attempt to describe means that they are dealing with characteristics which cannot all be described quantitatively. The terms used by geographers must, therefore, be defined at different levels of precision.

In this paper some of the problems in the study of rivers have been discussed. Terms such as bed load or dominant discharge lack agreed definition because of difficulty of measurement. The concept of grade, based on the interrelationships of such ill-defined quantities, can only achieve a still lower degree of precision.

The problems of fluvial morphology are of relevance to only a part of geography but the question of terminology and definition is common to all branches. The problem would have been similar if the characteristics discussed had been, instead, the central business district, the city, and the city region.

During the early stages of growth of a subject the introduction of many new terms is essential. The danger comes when the new terms are equated with words already in common use in the misguided attempt to "simplify" the literature. A profusion of terms, free of ambiguity, is preferable to a few words to which so many definitions have been attached that their use leads to confusion. The value of the term "grade", like that of the term "geographic region" has suffered in this respect.

In the engineering literature dealing with rivers a large proportion of space in each article is devoted to the question of definition. It is suggested that this

tedious but essential task has been neglected by geographers and geologists, to the detriment of both communication within the subject and interdisciplinary interchange of ideas.

Queen Mary College, London

BIBLIOGRAPHY

- [1] Anderson, A. G., Distribution of suspended sediment in a natural stream, *Trans. Am. Geophys. Union*, 23 (1942), 678—683.
- [2] Bagnold, R. A., *The physics of blown sands and desert dunes*, London 1941, reprinted 1964.
- [3] Bagnold, R. A., Some flume experiments on large grains but little denser than the transporting fluid and their implications, *Proc. Inst. Civ. Eng.*, 4 (3), Paper 6041, (1955), 174—205.
- [4] Baulig, H., La notion de profil d'équilibre; hisoire et critique, *Cong. Int. Geogr.*, 3 (1925), 51—63 and reprinted in Baulig, H., *Essais de Géomorphologie*, Paris 1950.
- [5] Baulig, H., *Géographie Universelle XIII, Amérique Septentrionale*, Paris 1935, 99.
- [6] Benson, M. A., and Thomas, D. M., A definition of dominant discharge, *Bull. Int. Assoc. Hydrology*, 11 (2), (1966), 76—80.
- [7] Bhatia, K. M., A study of the meanders of the Indus between Sukkur and Rohri, *Ann. Rep. (Techn.) of Central Board of Irrig. India*, Part I (1946), 115—118.
- [8] Blench, T., Regime theory for self-formed sediment-bearing channels, *Proc. Am. Soc. Civ. Eng.*, No. 70 (1951).
- [9] Boissier, R., Le charriage des alluvions en suspension dans l'eau de l'Arve, *Arch. des Sc. phys. et nat.*, 4 periode, 41 (1916), 331—333.
- [10] Braithwaite, R. B., *Scientific explanation*, Cambridge, U.K., 1953.
- [11] du Buat, L. P., *Principes d'hydraulique*, Paris 1779.
- [12] Carlston, C. W., The relation of free meander geometry to stream discharge and its geomorphic implications, *Amer. J. Sci.*, 263 (1965), 864—885.
- [13] Challinor, J., *A dictionary of geology*, Cardiff 1961.
- [14] Crickmay, C. H., Lateral activity in a river of North-western Canada, *J. Geol.*, 68 (1960), 377—391.
- [15] Culling, E. W. H., Longitudinal profiles of Chiltern streams, *Proc. Geol. Ass.*, 67 (1956), 314—345.
- [16] Davis, W. M., Base level, grade and peneplain, *J. Geol.*, 10 (1908), 77—111, reprinted in *Geographical essays by W. M. Davis*, London and Harvard 1954, 381—412.
- [17] Davis, W. M., and Daly, R. A., Geology and geography 1858—1929 in *The development of Harvard University*, Morison, S. E. (Edit.), Harvard 1930, 307—28.
- [18] Dury, G. H., The concept of grade, in Dury, G. H. (Edit.), *Essays in geomorphology*, London 1966, 211—133.
- [19] Einstein, H. A., Anderson, A. G., and Johnson, J. E., A distinction between bed load and suspended load in natural streams, *Trans. Am. Geophys. Union*, 21 (II), (1940), 628—633.
- [20] Einstein, H. A., The bed load function for sediment transportation in open channel flows, *U.S. Dept. Agric. Tech. Bull.* 1026 (1930).
- [21] Fisk, H. N., Mississippi River valley geology relation to river regime, *Trans. Am. Soc. Civ. Eng.*, 117, Paper 2511 (1951), 667—689.
- [22] Framii, K. K., Explanatory note on dominant formative discharge concept in non-boulder rivers, *Ann. Rep. (Techn.) of Central Board of Irrig., India*, Pt. I., (1946), 122—125.

- [23] Gannett, H., *Profiles of river in the United States*, U.S. Geol. Surv. Water-Supply and Irrig., Paper 44 (1901).
- [24] Gilbert, G. K., *The transportation of debris by running water*, U.S. Geol. Surv. Prof. Paper, 86 (1914).
- [25] Hill, M. L., Role of classification in geology, in Albritton, C.C., *The fabric of geology*, Stanford, Calif., 1963, 164—165.
- [26] Humphreys, A. A. and Abbot, H. L., *Physics and hydraulics of the Mississippi River*, U.S. Army Corps. of Eng. Prof. Paper, 13 (1861).
- [27] Inglis, O. C., *Meandering of rivers*, Central Board of Irrig., India, 24 (1941), 98—117.
- [28] Inglis, C. C., *The behaviour and control of rivers and canals*, Central Water Power, Irrig. and Navig. Res. Station, Poona, 13 (1949).
- [29] Johnson, W. D., The High Plains and their utilisation, *U.S. Geol. Surv. 21 st. Ann. Rep.*, Part IV, (1901), 601—674.
- [30] Kalinske, A. A., Criteria for determining sand transport by surface creep and saltation, *Trans. Am. Geophys. Union*, 23 (1942), 639—643.
- [31] Kesseli, J. E., The concept of the graded river, *J. Geol.*, 49 (1941), 561—88.
- [32] Kuiper, E., Sediment transport and delta formation, *Proc. Am. Soc. Civ. Eng.* 86, HY 4, 2371 (1960), 55—68.
- [33] Leopold, L. B. and Maddock, T. J., *The hydraulic geometry of stream channels and some physiographic implications*, U.S. Geol. Surv. Prof. Paper, 252, 1953.
- [34] Leopold, L. B. and Wolman, M. G., *River channel patterns ; braided, meandering and straight*, U.S. Geol. Surv. Prof. Paper, 282—288 (1957).
- [35] Leopold, L. B. and Langbein, W. B., *The concept of entropy in landscape evolution*, U.S. Geol. Surv. Prof. Paper, 500 A (1962).
- [36] Mackin, J. H., Concept of the graded river, *Bull. Geol. Soc. Am.*, 59 (1948), 463—512.
- [37] Mackin, J. H., Rational and empirical methods of investigation in geology, in *The fabric of geology*, C.C. Albritton (Edit.), Stanford, Calif., 1963, 143—149.
- [38] Makkaveyev, N. I., *The river bed and erosion in its basin*, Acad. of Sci. of the U.S.S.R., Moskva 1955.
- [39] Matthes, G. H., Basic aspects of stream meanders, *Trans. Am. Geophys. Union*, 22 (1941), 632—646.
- [40] McGee, W. J., Outlines of hydrology, *Bull. Geol. Soc. Am.*, 19 (1908), 199.
- [41] Menard, H. W., Sediment movement in relation to current velocity, *J. Sed. Pet.*, 20 (3), (1950), 152—153.
- [42] Miller, J. P., *High mountain streams : effects of geology on channel characteristics and bed material*, State Bureau of Mines and Mineral Resources, New Mexico Inst. of Mining and Technology, 4, 1958.
- [43] Morris, H. M., *Applied hydraulics in engineering*, London 1963, 359.
- [44] Nixon, M., A study of the bank-full discharges of rivers in England and Wales, *Proc. Inst. Civ. Eng.*, 12 (1959), 157—173.
- [45] O'Brien, M. P. and Rindlaub, B. D., The transportation of bed-load by streams, *Trans. Am. Geophys. Union*, 15 (1934), 593—602.
- [46] Raudkivi, A. J., *Loose boundary hydraulics*, London and New York 1967, 91.
- [47] Sakthivadivel, R., Discussion of Total bed-material transport by Avery A. Bishop, Daryl B. Simons and Everett V. Richardson, *Proc. Am. Soc. Civ. Eng.*, 91, HY 6, Pt. I, Pap. 4266 (1965), 201—203.
- [48] Scheidegger, A. E., *Theoretical geomorphology*, Berlin 1961, 159.
- [49] Schoklitsch, A., *Hydraulic structures* (trans. by S. Shulits), New York 1937, 141.
- [50] Schroeder, K. B. and Hembree, C. H., Application of the modified Einstein procedure for computation of the total sediment load, *Trans. Am. Geophys. Union*, 37 (1956), 197—212.

- [51] Schumm, S. A. and Lichty, R. W., Time, space and causality in geomorphology, *Am. J. Sci.*, 263 (1965), 110—119.
- [52] Shulits, S., Rational equation of river bed profile, *Trans. Am. Geophys. Union*, 22 (1941), 622—630.
- [53] Simons, D. B., Richardson, E. V., and Nordin, C. F., *Bedload equation for ripples and dunes*, U.S. Geol. Surv. Prof. Pap., 462-H, 1965.
- [54] Stein, R. A., Laboratory studies of total load and apparent bedload, *J. of Geophys. Res.*, 70 (8) (1965), 1836.
- [55] Sternberg, H., Untersuchen über Längen und Querprofil Geschiebeführende Flüsse, *Zeit. Bauwesen*, 25 (1875), 483—506.
- [56] Straub, L. G., Effect of channel contraction works upon the regimen of movable bed streams, *Trans. Am. Geophys. Union*, 15 (1934), 454—463.
- [57] Wolman, M. G. and Leopold, L. B., *River flood plains, some observation on their formation*, U.S. Geol. Surv. Prof. Pap., 282-C (1957).
- [58] Wolman, M. G., Factors influencing erosion of a cohesive river bank, *Am. J. Sci.*, 257 (1959), 204—216.
- [59] Wolman, M. G. and Miller, J. P., Magnitude and frequency of forces in geomorphic processes, *J. Geol.*, 68 (1960), 54—73.

EXAMPLES OF GEOMORPHOLOGICAL INVESTIGATIONS IN POLAND BASED ON QUANTITATIVE METHODS

BOLESŁAW DUMANOWSKI

Quantitative methods are based on measurements. Their usefulness in geomorphological research work has been recognised since the 19th century, but in practice they were seldom employed. It was not until the end of the Second World War that quantitative methods began to be increasingly used in scientific publications. Measurement methods were first applied to forms, next geological structures, and finally processes. The technique and the accuracy of measurements are gradually improving. In spite, however, of considerable progress in this field, methods are still very often rather primitive; moreover these methods are not applied to all phenomena. Notwithstanding all these shortcomings, quantitative methods are obviously very convenient in the description of the forms themselves, as well as their structure and the processes that affect their evolution. In the first place they facilitate the correlation of observations from various areas, hence providing a wider basis for the interpretation of the material available.

Along with the progress made by quantitative methods in other countries, their use in geomorphological investigations is also increasingly extended in Poland. Today most investigations are undertaken on the basis of quantitative methods. Two examples will be here cited to illustrate this work technique, one respecting the investigation of processes, the other of geological structure.

Professor A. Jahn (Cielińska, 1961) was the first to introduce quantitative methods into the investigation of processes in the Karkonosze Mts. They comprised measurements of sheet and rill wash, creep of isolated blocks, dislocation of debris, solifluction and vertical movements due to the freezing of water in the soil.

The importance of the role played in slope formation by the process of sheet and rill wash has often been stressed in geomorphological investigations. The lack was felt, however, of adequate quantitative methods which would allow a more exact determination of the magnitude of these processes. From 1956 to 1960, in the course of slope investigations in the Karkonosze Mts., and also in Spitsbergen, Prof. A. Jahn and his collaborators used the simplest measurement methods, which provided general information on the scale of these processes. The technique used was as follows: Several tin containers, $30 \times 30 \times 40$ cm.,

were placed on variously inclined slopes so that the tops of the containers were raised slightly above the ground. Each container had two openings, flush with the slope surface. The material washed off the slope was introduced into the container through the opening that faced the slope, while the other opening in the opposite side of the container carried off the water. A sector of the slope above the container was fenced off by wooden boards in order to make sure that the material introduced into the container was derived from a strictly defined plot of land. When placing the containers an attempt was made to take into account the differences in local conditions, such as variability of the slope gradient, the plant cover (forest, meadow, absence of plant cover, etc.). Nineteen containers were placed on slopes in the Karkonosze Mts., eight in Spitsbergen. From time to time they were emptied and their contents carefully measured. Part of the material washed off the slope surface, mainly that from the finest fraction, had been carried beyond the containers. Hence, the measurement figures were probably an under-estimate of actual conditions. In spite of this primitive technique the above method allowed the quantitative determination of the magnitude of sheet and rill wash. The data resulting from investigation in the Karkonosze Mts. (Cielińska, 1961) are as follows: the amount of material washed down the slope is controlled not only by the transporting agent, i.e. the character and intensity of melt waters, of streams and of rain water, but also, and first and foremost, by the natural properties of the slope surface affected by the wash process. The down-wash accomplished by rain water is more effective than that by melt waters. The material from sheet wash differs distinctly from the material moved by rill wash. In the former case organic particles represent as much as one third of the transported material, in the latter only about 12 per cent. The amount of material moved by concentrated wash (in rills) is about 40 times that washed off by sheet wash.

The strongest down-wash has been observed in the Bialy Jar ravine, where erosional incisions occur on slopes without a plant cover, having an inclination of from 36 to 39°, and which are composed of strongly weathered granite. It has been found here that 26 grams of material are annually carried down from 1 sq. metre of the surface by concentrated wash flow. This is 1,000 times the amount carried down by sheet wash on a grass covered slope with an inclination of 23°. The computed lowering of the slope surface would be 1 mm per 10 years in the case of rill wash on the steep slope or 1 mm per 10,000 years in the case of sheet wash on the 23° slope. Observations of very strongly eroded slopes in Spitsbergen (A. Jahn 1961) have shown that a period of 150—170 years is needed to accomplish a lowering of 1 mm. Observations made by a similar method by T. Gerlach (1966) in the Carpathians suggest that the amount of the down-washed material is controlled by climatic factors, the plant cover, the shape of the slope and, in the case of cultivated areas, by the lay-out of fields and the cultivation methods. On slopes with a perennial plant cover the amount of

down-washed material is from several thousand to several hundred thousand times less than that on ploughed slopes.

These measurements indicate considerable differences in the amount of material removed under the same climatic conditions (the Sudety Mts., the Carpathians). The amounts depend mainly on such factors as the plant cover and the inclination of slopes. The absence of plant cover may many thousand times increase the amount of the down-washed material. No data are available as regards the magnitude of down-wash on solid rock; it seems probable, however, that, depending on the kind of material, the differences will be much greater.

The creep of detached blocks has long been reported on from the area of the Karkonosze Mts. For the purpose of measuring the rate of the creep movement, steel rods were cemented into the bedrock, along the side of 7 selected blocks. The rods were used as stable points, their respective distances from the blocks being accurately registered. Measurements of the creep of the blocks have shown that their annual advance ranges from 0.0 to 32.65 cm. The measurements were made on slopes inclined from 18 to 37 degrees, and indicate a close dependence of the rate of the creep on the plant cover and the inclination of the slope. It is also reasonably supposed that the weight of the blocks considerably affects the rate of creep.

Measurements of the solifluction phenomena observed in Spitsbergen were very accurately taken by A. Jahn (1961). Pegs were used, inserted at a depth of 10, 30 and 60 cm. respectively, also squares of wood, 5×5 cm in size, and boulders marked with paint. This technique was used in 1957—1959 on the gently inclined ($5-17^\circ$) slope of Bogstranda to measure the rate at which material moved on solifluction tongues. In one of the profiles the maximum yearly travel of the pegs was 12 cm, in another approximately 5—6 cm. The fact that the pegs were set at various depths disclosed that the rate of movement of material decreased with increasing depth and that the surface layers were more uniformly moved. In four peg profiles inserted in the Revdalen valley into a heavy-clay slope having an inclination of $3-4^\circ$, very little movement was observed over a period of two years. Its maximum attained 3 cm. On evidence of the pegs having been upheaved by 10—15 cm owing to the work of the ice. A. Jahn supposes that the solifluction movement is closely connected with the vertical movements of the ground. On this basis A. Jahn supposes that on a very gently inclined slope (3°) solifluction movements occur: a) as a frost-caused thrust (indispensable requisite), b) resultant viscous flowage. In Spitsbergen the vertical soil movements caused by the freezing-thawing process were also measured, with the help of Prof. S. Baca's sliding scale (motometer). This consists of an immobile iron frame and metal rods with small plates attached thereon. The rods, with the plates attached on them, were sunk into the ground so that the top parts of the rods passed through the frame, with a scale marked on it to show

the vertical movement of the plates, i.e. that of the ground. Meteorological and hydrological observations were made together with those concerning the ground movements. The yearly records show that the maximum vertical movements on a surface of heavy-clays attained the figure of 15.5 cm, compared with 1.5 cm for a sandy-gravelly surface. The total of 15.5 cm for the clay consisted of an upward winter movement of 8.7 cm and a downward summer movement of 6.8 cm below the starting line.

These measurements of processes, though lacking precision, made it possible not only to determine their magnitude but also to determine the conditions that favoured or impeded the particular processes. Thus, it is now easier to define the extent to which the development of forms is affected by such factors as the kind of material, the plant cover, meteorological conditions etc.

The geomorphological investigations in the region of the Karkonosze Mts. have shown that a full understanding of the origin of certain forms cannot be attained without a detailed knowledge of the geological structure. In 1925 three types of granite were differentiated in this area by H. Cloos. The first type — that most widely distributed — is a porphyritic granite containing orthoclase crystals up to 8 cm in diameter. Equigranular granite, medium- or fine grained, represents the second type: the third is an aplite granite, almost entirely devoid of micas. The above classification has proved inadequate to determine the connections between the geological structure and the relief forms. Moreover, because of the small scale used in geological maps, the picture they provide is too generalised to be an adequate basis for detailed geological investigations. Observations have shown that the three granitoid features largely affecting the development of forms are: the quantitative mineral composition, density of the net of fissures, and the grain size. Specimens of resistant and unresistant rocks were collected and micrometric analyses were made to determine the mineral composition. The results proved most interesting. It was found that the proportions of the various minerals varied at the sampling sites. Plagioclase ranges from 20 to 45 per cent of the rock volume, orthoclases from 18 to over 42 per cent, quartz from 25 to about 42 per cent, biotite from below 1 per cent to about 10 per cent. The few remaining minerals occur in such small quantities that they probably have no significant effect on the development of forms. Along with the quantitative mineral composition analyses, grain size and density of fissures were also measured. So far as the grain size is concerned the following types of granite were differentiated: fine grained (below 2 mm), medium grained (2—5 mm), coarse grained (above 5 mm) and unequigranular porphyritic granite with large orthoclases. In order of frequency of occurrence the porphyritic type comes first, followed by the fine grained and the medium grained types. The coarse grained granite is the least common one. The rock is also strongly differentiated as regards the density of the fissure net. In most cases the spacing is 0.5—1 m. In some places, however, the concentration of fissures is very strong, the spacing

being about 1 cm. These concentrations have been called "bundles" (fascicles).

The mountain peaks within the main ridge of the Karkonosze range are built of granite or of monzonite granite. The grain size is rather small and the granites are fine- or medium grained. The variety having a porphyritic structure is here less common. The net of fissures is rather dense, the average spacing being about 0.5 m. The bulk of the ridge consists of a similar rock. The individual peaks rise above the more or less even line of the ridge where the rocks they are built of have a lower plagioclase content or are finer grained. It is characteristic that the convexity in the upper of the slope forming the ridge zone is more gentle in those places where the rock has a porphyritic structure. In the case of fine-grained varieties this curve is much sharper. A distinct break of slope occurs if the boundary between the two varieties is sharp. The break of slope is especially sharp when the change in the quantitative mineral composition occurs with a change in grain size.

Passes in the main ridge are first and foremost associated with the differentiation of the quantitative mineral composition. As a rule they correspond to granodiorite zones. Small creeks descending from the main ridge in some places produce microforms of shaped valleys, elsewhere they flow in shallow channels on the surface of the slope. The appearance and disappearance of valley forms here is closely connected with the mineral composition of the rock. The presence of valley forms is observable in granodiorite areas, while on monzonite granite the streams flow as a rule in shallow channels. A step (threshold) often occurs at the granodiorite-monzonite granite boundary in the longitudinal section of the streams. Differentiation in the quantitative mineral composition also affects the shape of larger valleys. Granodiorite areas are associated with the widening of valleys, monzonite granite areas with their narrowing.

As well as these landforms, the distribution of chemically altered regolith corresponds to the differentiation in the quantitative mineral composition of granitoids. These regoliths are encountered almost exclusively on granodiorites. The monzonite granite does not, as a rule, display any important chemical alteration.

The above description shows the great importance of investigations of the quantitative mineral composition when clarifying the development of forms on granitoids. The plagioclase content plays an important role in the area under consideration. When it is low, positive forms are produced by the rocks, and negative forms are produced when the plagioclase content is high. Variations in the biotite content, however, do not seem to affect forms to any great extent. The highest biotite content has been observed in tors. This inability of biotite to exert any influence may probably be explained by the fact that the amount in the rock is too small to affect in any large measure the development of forms.

In 1966 my investigations in the Karkonosze Mts. covered, in addition to the solid rocks, the chemically altered weathering residue. The differen-

tial-thermal and X-ray analyses of 21 samples made it possible to determine what clay minerals occur in the weathering residue and in what approximate amounts. The samples were collected at depths ranging from 0.9 to 8 m below the surface, from the main ridge, the slope and the depression of Jelenia Góra. The investigations were concerned with an examination of the fraction less than 5 microns in diameter. The per cent content of material from this fraction was relatively very low, since in more than 60 per cent of the samples it represented only 1 to 3 per cent of the weight of the sample. In only one sample the 5 micron fraction represented 54.5 per cent of the weight. Hence the weathered material represents the so-called "granite grit" i.e. fine granite debris, from 1 to 10 mm in diameter, bearing distinct signs of chemical alteration. Only locally, along clay fissures, is it clay-like in character. Illite, kaolinite and montmorillonite are among the clay minerals whose occurrence has been noted. Illite is the most common constituent. Its absence was noted from only one sample; in the other samples its per cent content varied from 10 to 100 per cent of the whole volume. In 40 per cent of samples it was the predominant constituent among the clay minerals. In 50 per cent of the samples kaolinite was the dominant mineral, while montmorillonite represented only 10 per cent. It is characteristic that whenever kaolinite was present in the sample it always dominated over the remaining clay minerals (40—80 per cent). Montmorillonite was found in 5 samples; in two of them it displayed a distinct predominance over the remaining minerals (its amount being from 60 to 100 per cent).

A distinct connection has been observed between the distribution and the thickness of the weathered material on the one hand and the type of rock on the other hand. The rock mantle occurs foremost on granodiorites. This connection seems a matter of course, since many investigations (e.g. by S.S. Goldich 1938) have shown that plagioclases yield to weathering much quicker than the remaining more important rock-building minerals. On the basis of investigations in SW Asia, W. N. Razumova (1956) concluded that the initial thickness of the mantle rock depends on such features of the rock as its petrographic composition, structure, the extent of fissuring and the grain size. A close connection also exists between the distribution of the rock mantle and the forms. The former occurs mainly in the various depressions.

So far, research studies (J. P. Bakker, 1960) have suggested that a distinct dominance of kaolinite in samples indicates a rather warm and moist climate. Hence, it might be supposed that material is Tertiary in age. I. I. Gincburg (1963) is of the same opinion but, in addition, he stresses the importance of the parent rock as a factor in the mineral composition of the rock mantle. Zones characterised by a definite clay mineral composition have been differentiated in the rock mantle by many Russian authors (W. M. Razumova, 1956; K. K. Nikitin, 1956; I. I. Gincburg et al., 1963). In most cases 4 zones were differentiated. Zone 1 is that of seepage or disintegration, displaying the work of the weathering

process — mainly physical — along the fissures. Zone 2 is characterised by leaching, with hydromicas (illite) as the dominant constituent, and kaolinite and montmorillonite present, too. Zone 3 is that of disintegration; kaolinite is predominant, and montmorillonite and illite are present. Zone 4 represents the final products of the weathering processes, with kaolinite as the dominant constituent. The thickness of the particular zones varies from 1 to about 15 metres.

The above described zones are not indicated within the region of the Karkonosze Mts. though the thickness of the rock mantle there sometimes attains 20 metres. So far as the mechanical composition is concerned, the Karkonosze rock mantle would correspond to Zone 1, because it consists mainly of debris with a clay admixture. By reference to the mineral composition in various places, the Karkonosze rock mantle shows features characteristic of the first 3 zones. For example in samples taken at similar depths (0.9—2 m) from various points of the same outcrop, illite dominated in one sample (90—100 per cent), montmorillonite in another (60—80 per cent) and kaolinite in the two remaining ones (60—80 per cent).

In view of the distribution of the mantle rock weathering residue, its thickness and mineral composition, the writer supposes that it is at least of Tertiary age, when the prevailing conditions were a warm and moist climate. Only the lowest parts have survived so far. In particular the marked variation in the relative proportions of weathering minerals present suggests that the weathering processes attacked the rock not only from the surface downwards but also along the fissures. Hence the zonal vertical pattern distinguished by Russian geologists is most probably repeated on a smaller scale horizontally, from the fissures into the interior of the rock. Moreover, the variation in mineral composition of the weathering residue may also be due to the original mineral composition of the rock.

University of Warsaw

BIBLIOGRAPHY

- [1] Bakker, J. P., Some observations in connection with recent Dutch investigations about granite weathering and slope development in different climate and climate changes, *Zeitschrift f. Geomorphologie*, 1960, Suppl. Bd. 1. Morphologie des Versants.
- [2] Cielńska, M., Comment on quantitative investigations of some processes of modern denudation in the Karkonosze Mountains, *Zeszyty Naukowe Univ. Wrocł., Seria B*, 8, 1961.
- [3] Cloos, H., *Einführung in die tektonische Behandlung magnetischer Erscheinungen*, B. 1, Das Riesengebirge in Schlesien, Berlin 1925.
- [4] Gerlach, T., *Współczesny rozwój stoków w dorzeczu Górnego Grajcarza (Beskid Wysoki — Karpaty Zachodnie)*, Inst. Geogr. PAN, Prace Geograficzne, 52, Warszawa 1966.
- [5] Gincburg, I. I., *Typy drewnich kor vyvietryvaniya, formy ich proyavleniya i klassyfikaciya, Kora Vyvietryvaniya*, Vyp. 6, 1963.
- [6] Goldich, S. S., A study in rock weathering, *Journ. Geol.*, 46 (1938).

- [7] Jahn, A., Quantitative analysis of some periglacial processes in Spitsbergen, *Zeszyty Naukowe Ułiw. Wrocl., Nauki Przyrodnicze, Seria B*, 5, 1961.
- [8] Nikitin, K. K., Kora vyvietrivaniya na dioritach juzhnogo Urala, *Kora Vyvietrivaniya*, Wyp. 2, 1956.
- [9] Razumova, W. N., Kora vyvietrivaniya syevyero-zapadnoy chasti Kazachskogo Nagorya, *Kora Vyvietrivaniya*, Wyp. 2, 1956.

SOME METHODS OF MORPHODYNAMIC RESEARCH USED IN CRACOW

ZDZISŁAW CZEPPE

A group of geomorphologists in the Cracow geographical centre are carrying out research on the contemporary geomorphological processes active in southern Poland. Some work has also been done in Spitsbergen. They are trying to measure the intensity of selected processes under known environmental conditions. The final results will lead to the recognition of certain types of contemporary remodelling of relief and to an estimation of the rate at which the changes take place.

The methods employed may conveniently be divided into a group of direct methods and a group of indirect methods. The first group aim at direct measurement of the intensity of a given process. Measurements are carried out continually or repeatedly, with greater or lesser frequency, for at least a year and often for many years. They are linked with meteorological, climatological, hydrological and other supplementary observations. We may also place in this group the measurement of landforms created rapidly by an unusual storm or flood. Indirect methods are based on the measurement of young forms of relief created during longer periods of time and representing the total effect of the activity of one process or of a set of processes. The methods of recording the contemporary morphogenetic processes have been developed, among others, by a group of research workers under the leadership of L. Starkel [5]. Limited space permits to mention only some of the direct methods.

Any morphodynamic research begins with geomorphological mapping of the region. Often a general geomorphological map is prepared first and then detailed plans are made on such a scale as to make possible the insertion of all the observed changes in relief of the chosen portions of the terrain.

METHODS OF OBSERVATION AND MEASUREMENT

1. *Fall of weathered material.* Weathered material falling off small rock faces was caught on canvas, covering a surface of several square metres, pinned on to the ground at the foot of the rock. The quantity of fallen material, when related to the area of the face supplying the material, gives the rate of slope retreat. With

this method M. Niemirowski [7] calculated that, on Babia Góra (1725 m) in the High Beskids, 1—2 cm³ of material falls off annually from an area of 1 m² of perpendicular face of Magura sandstone, which corresponds to a retreat of 0,001—0,002 mm/per year. Seasonal changes of intensity of the process were also noted.

2. *Soil-creep*. Soil-creep has been measured in the Carpathians with the help of wooden pegs forced into soil both along the slope contour and down the maximum slope. The exact position of each peg was determined in relation to main and auxiliary fixed points. T. Gerlach [3, 4] installed in Jaworki, in the eastern part of the High Beskids, a net of 3 lines running down slope and 3 along the contour lines. The fixed points, made of concrete and placed between the pegs, permitted him to make a detailed slope profile. Levelling along this profile was repeated every year. Comparing the consecutive profiles one can notice degradation or aggradation going on in particular sections of the slope. The position of pegs is also checked every year. Their displacement shows the intensity of soil-creep.

3. *Displacement caused by needle-ice*. On slopes needle-ice displaces surface material which moves down slope. To measure the amount of displaced material and the distance of this displacement T. Gerlach used, in the Tatra Mts., a wire stretched on the ground between two geodetic pins. The surface below the wire was cleaned and smoothed. After a period of activity of needle-ice the area completely covered by displaced material below the wire was measured, as well as the position of single rock fragments further below. Material then was collected, weighed and analysed [3]. In Spitsbergen 4 ceramic tiles were placed on the more or less flat surface in the middle of the earthy central part of a sorted ring. Their displacement by needle-ice was noted for 3 years; no directional preference of their movements was observed, which shows that an inclination of surface is necessary for such a preference to become established [2].

4. *Vertical frost movements of soil*. Freezing of moisture in waste composed of a certain combination of size fractions causes marked changes of its volume, which results in frost heave, sorting of fractions and similar phenomena. The annual run and intensity of the vertical movements of soil due to freezing and thawing are observed with motometers. These consist of a steel frame, the posts of which are cemented below the depth of frost penetration. The cross-bar of the frame has holes through which the tops of measuring rods stick out. The rods are made of brass alloy and have small plates at their lower ends. These plates are pushed into an undisturbed face of regolith at various depths. The hole is then filled with earth. During the period of frost action the length of the rods above the cross-bar is measured every day. Then curves of movement are drawn. The curves from Spitsbergen and Labrador are similar, but those from Poland are different. In the Arctic there is an autumnal period of up-and-down movements due to waves of freezing and thawing. Then comes a period of upward movement

ending with the full winter immobility. Spring brings a continuous downward movement caused by thawing penetrating down the active zone. The unfrozen layer is soaked with meltwater and subsides, achieving its normal density. The excess of water is squeezed to the surface. This happens because an impermeable permafrost underlies the active zone. In Poland meltwater coming from snow mostly flows off and evaporates before the frozen ground melts. There is not much water from the ground ice and it partly evaporates and partly penetrates down into the regolith. Therefore soil is not soaked, does not subside and is left in a swollen, spongy state. Motometers do not show the springtime downward movement. However, frost activity in our country is complicated and it is necessary to collect more observations before a fuller interpretation is possible [1].

5. *Surface wash*. T. Gerlach has developed an interesting method of research on the intensity of surface wash in relation to the form of a slope and its vegetation cover. He uses tin troughs, or gutters as we call them, 50 cm long and 10 cm broad, closed at the sides and with a moveable cover. There is a little pipe in the bottom and a flat tin strip in front. Gutters are placed on slopes so that the flat strip is pressed into the soil and water flows into the trough. The water is led by means of both the pipe in the bottom and a rubber pipe into a glass container taking 5 or 10 litres. This container is easily exchangeable, and water collected in it is filtered, and the material washed in is then weighed. Troughs are set two or three side by side where the angle of slope changes. Hence, if the area and declivity of that part of the slope supplying washed-down material is known, the amount of surface wash per unit area can be calculated. The other factors, such as vegetation, soil and meteorological conditions, are known. Observations conducted with this method for three years in the upper part of Grajcarek river basin in the eastern part of the High Beskids enabled T. Gerlach to show which parts of the slope are degraded and which aggraded by surface wash and how intensity changes with the form of the slope and the season of the year. He has calculated also that intensity is greatest on arable land, from which $2530 \text{ m}^3/\text{km}^2/\text{per year}$ of material is carried away. From pasture the rate is $3.3 \text{ m}^3/\text{km}^2/\text{per year}$ and from meadow $0.4 \text{ m}^3/\text{km}^2/\text{per year}$. It is least from a forested slope, the rate here being $0.03 \text{ m}^3/\text{km}^2/\text{per year}$ [4].

6. *Stream bed processes*. L. Kaszowski and M. Niemirowski are undertaking research on the seasonal flow and the intensity of processes in mountain stream beds. One of the streams being studied is Potok Bialy in the Tatra Mts. Its basin of 4 km^2 stretches from 900 to 1700 m above sea level. The other two are the twin streams Jaszczce and Jamne in the High Beskids. Their basins cover 20 km^2 and reach from 600 to 1200 m above sea level. The Potok Bialy basin is developed mainly on dolomites and the other two mainly on flysch sandstones. There are water-gauging stations on these streams.

Erosion and sedimentation in the beds are investigated mainly by indirect methods. The streams are frequently visited and all changes in their beds are

plotted on a plan with a scale of 1 : 400. Some chosen segments of the beds are repeatedly photographed for later study. Transport of coarse material is observed with the help of marked gravels. This method can show the relationship between the form of stream bed, hydrologic parameters and the character of gravel transportation. Sample sets of gravels of certain size classes, namely 4—6, 6—8, 8—16 and over 16 cm, are prepared. In each set there are gravels of varied form: elongated, flat and spherical. In sets comprising 30 to 100 gravels every one is measured and its form is noted. Then they are painted with resistant enamel. Colour and design are changed for every site of their primary deposition.

The painted gravels were deposited in chosen segments of stream beds differing from each other as regards width, length, form, steepness and amount of water. Each point of primary deposition received a fixed mark. There are 8 points of deposition on each of the Beskid streams, and about 300 gravels were put at each site. After every high water the painted gravels have to be found, their position plotted and their characteristics, orientation and inclination of the longest axis noted. All the changes occurring on their surface, e.g. scratches, cracks and breaks are also noted.

The investigation is not yet finished, but the dependence of the character of transportation and the distance between the points of temporary deposition of gravels on the form and size of material, steepness of channel, and volume and velocity of water is very well demonstrated. The frequency of high water stages is similar in all streams, but, owing to the geological differences, development of their beds and dynamics of the bed processes are markedly different. The erosional transformation of the stream beds cut in sandstone is much more intense than that of the bed cut in dolomite.

Retreat of the undercut stream banks built of loose material is measured with iron sticks 50 cm long driven right into the bank so that their ends do not show at the surface. They are placed in the middle of the undercut faces at intervals of 4—5 m. On curved faces the intervals are smaller but consistent on any one site. After every high water which causes a retreat of the bank the protruding ends of the sticks are measured. This gives the basis for valuation of retreat of the whole face and of the volume of lost material. The sticks are then driven right in again. This simple method used for 3 years supplied the data for quantitative evaluation of the rate of retreat of eroded banks and for finding the relationship between the flow and lateral erosion.

In addition to the methods mentioned here there are others which have not been described and these also will be further developed and checked. Some of them have been already accepted abroad and are in wider use.

Jagellonian University, Cracow

BIBLIOGRAPHY

- [1] Czeppe, Z., Roczny przebieg mrozowych ruchów gruntu w Hornsundzie (Spitsbergen) 1957—1958 (Annual course of frost ground movements at Hornsund) Spitsbergen (1957—1958), *Zeszyty Naukowe Uniw. Jagiell., Prace geogr.*, 3, Kraków 1961.
- [2] Czeppe, Z., Przebieg głównych procesów morfogenetycznych w południowo-zachodnim Spitsbergenie (The course of the main morphogenetic processes in south-west Spitsbergen), *Zeszyty Naukowe Uniw. Jagiell., Prace geogr.*, 13, Kraków 1966.
- [3] Gerlach, T., Metódy a terajší stav vyskumu morfogenetických procesov v polskich Karpatach, *Geograf. Časopis*, 16, Bratislava 1964, 256—270.
- [4] Gerlach, T., *Współczesny rozwój stoków w dorzeczu górnego Grajcarek (Beskid Wysoki — Karpaty Zachodnie)* (Res.: Developpement actuel des versants dans le bassin du haut Grajcarek (les Hautes Beskides — les Carpates Occidentales), Polish Ac. Sc., Inst. of Geogr., *Prace Geograficzne* 52, Warszawa 1966.
- [5] Kaszowski, L., Kotarba, A., Niemirowski, M., and Starkel, L., Maps of contemporaneous morphogenetic processes in Southern Poland, *Bull. de l'Acad. Polon. des Sciences, ser. géol. et géogr.*, 14, (2), (1966), 113—118.
- [6] Kłapa, M., Prace Stacji Badawczej Instytutu Geografii PAN na Hali Gąsienicowej w latach 1960 i 1961 (Report on research work carried out at the Scientific Station of the Inst. of Geogr. of the Pol. Acad. of Sc. in the Tatra Mts (Hala Gąsienicowa), in 1960—1961), *Przegl. Geogr.*, 35 (1963), 221—237.
- [7] Niemirowski, M., Rola współczesnych procesów morfogenetycznych w kształtowaniu rzeźby podszczytowej partii Babiej Góry (The role of contemporaneous morphogenetic processes in the formation of the relief in the highest parts of Babia Góra), *Zeszyty Naukowe Uniw. Jagiell., Prace Geogr.*, 10, Kraków 1964, 45—75.

SOME RECENT FIELD AND LABORATORY
EXPERIMENTS IN GEOMORPHOLOGY

V. BRUCE PROUDFOOT

The present paper summarises some recent experimental work in geomorphology that has been carried out in Britain or published in British journals. In addition there is a brief discussion on some experimental work with which the writer is associated, not all of which has been previously published. This paper should be regarded as additional to the chapter on experiments in C.A.M. King's recently published *Techniques in Geomorphology* [1], for experiments included there are mentioned here only incidentally, if at all. It is not meant as a comprehensive survey, but rather as an indication of some of the experimental work being carried out, and the problems arising from this.

The essential character of experimental work is that the results should be reproducible and that at least some of the parameters which are thought to govern the processes being studied should be measurable. The distinction between quantitative field observation and experiment is not always clear. Strictly the placing of markers in particular positions and the subsequent measurement of their new positions after the lapse of intervals of time does not constitute an experiment, for neither of the conditions of experimentation is strictly satisfied. The insertion of the markers modifies the original situation, although careful replication of experiments could go some way towards mitigating such disturbance [2]. Young [3] has noted some of the problems associated with the measurement of soil movement on slopes. To study such movement he used metal pegs inserted into the surface of the ground, as well as pegs inserted at various depths in the sides of pits dug to insert the pegs and then re-filled. Observations were made by excavating the pits after intervals of 6—12 months over 2 years, and again after 4 years. A number of pegs had to be rejected since they showed inconsistent movements which it was assumed had been caused by disturbance during re-excavation. As part of the same research programme Young also carried out experiments in which blocks of soil were alternately wetted and dried and the consequent movements measured. Such experiments recall the long-term investigations into soil movement initiated by Charles Darwin and his son Horace in 1877, the results of which were communicated to the Royal Society of London in 1901 [4]. These experiments showed clearly the way in which the vertical

movement of a stone laid on the surface of the ground was affected by changes in the moisture content of the soil, and by alternate freezing and thawing. The importance of frost-heaving has also been clearly demonstrated in experiments carried out by Edwards as part of an investigation into various factors affecting the growth of cereals on artificial seed-beds prepared by dry-sieving of the soil [5].

Edwards and Young both had difficulties in measuring small movements, partly because of the relatively large size of their markers. In experiments where the markers are small in relation to the features being observed, as with survey pegs in glaciers, the effect of the markers can probably be neglected. Other difficulties arise when measurements must be "once and for all" because of the nature of the features being studied, as is the case with landslides. Similarly the laboratory determination of the characteristics of materials is scarcely experimental in the strict sense, although the behaviour of materials under test conditions may provide some guide as to their behaviour under natural conditions.

Among the limited number of experiments designed to test the properties of markers used in quantitative observations, reference should be made to a paper by Joliffe [6], which describes two simple experiments designed by the Hydraulics Research Station to compare the relative rates of movement of several sizes of pebble. Natural beach pebbles marked with marine paint and colour pigments, and artificial shingle made of concrete incorporating fluorescent particles were employed. The pebbles and shingle were traced almost daily for up to three weeks and the centroids of movement were calculated for each particle size used. Differential movement took place according to particle size and it appeared that for a particular size of wave there was a particular size of material which moved at a greater rate. Larger sizes on average travelled further than smaller sizes, so that it seems likely that the larger waves moved the largest material most rapidly, whilst very small waves moved only the smallest beach particles, the larger particles remaining more or less stationary. These experimental findings not only assist in interpreting the movement of other marker material but help to explain the grading on many natural beaches and provide evidence for interpreting variations on beaches.

In most experimental work which seeks to reproduce natural conditions or provide analogues to them the problem of scale is of fundamental importance. Furthermore, the physical or dynamical similarities between the experimental model and reality are as necessary as morphological similarities. With regard to scale, the laws of dynamical similarity and the practices of dimensional analysis provide a generally adequate framework for model use [7]. In some circumstance it is possible to utilise successively smaller experimental scales, or to use small natural features as models for larger, in such a way that one has a continuous sequence of scales, for example, from the large river, through the small stream to the laboratory flume.

On the other hand some experiments are scarcely ever likely to reproduce

very closely natural conditions. Included in this category are experiments designed to elucidate such major problems in structural geomorphology as the effects of mountain building. In well known experiments, Griggs in 1939 had induced flows in water glass beneath a layer of oil loaded with sand and had obtained patterns similar to the root formation inferred by geophysicists beneath surface mountain folds [8]. Similar experiments were initiated by the late Professor S. W. Wooldridge and M. J. Kenn, who used lubricating oil above glycerol [9]. Convection currents in the latter were simulated by two contra-rotating perspex drums. In these latter experiments it was found impossible to reproduce the single root below surface mountain folds: instead, extended or duplicated roots were formed, associated with a broad gentle uplift in the upper layer.

There has been considerable discussion of the mechanisms involved in folding, much less of those involved in faulting. Wilson [10] has described again the experimental fracture patterns produced by Riedel in 1929 [11] in his discussion of the tectonics of the Great Ice Chasm on the Filchner Ice Shelf, Antarctica, an east-west zone of brecciated ice which he interprets as the result of a wrench-or tear-fault, in which the northern side moved relative to the southern.

Especially within the last decade the results of laboratory experiments have increasingly been used in the interpretation of folds and sedimentary structures. Experiments to elucidate such structures were carried out last century by Gilbert [12], but much of the recent work derives its impetus from the experiments of Ph. H. Kuenen, conveniently summarised by him in 1958 [13]. Experimental results combined with the critical examination of modern sedimentary patterns has led to a much fuller understanding of many sedimentary sequences recorded in the hard rock geology of the British Isles [14], for example, studies of the Torridonian by A. D. Stewart, [15] and R. C. Selley and others [16]. As well as sedimentary sequences, numerous individual features have also been described and explained on the same basis, for example W. D. I. Rolfe has described an air heave structure from the probable playa sediments in the Old Red Sandstone of Lanarkshire, Scotland [17].

Williams [18] has discussed theoretical aspects of the mechanism of convolute folding, paying special attention to the property of liquefaction of confined water-logged unconsolidated sedimentary layers. He has suggested alternative mechanisms to those proposed by Sanders who, in discussing the origin of convolute laminae, appealed for further experimental work to continue that already carried out by Kuenen [19, 20]. Of the few recent papers that actually include experimental work mention should be made of one by Ramberg in which is discussed the use of cemented multi-layers of rubber with unlike rigidity, in experiments designed to simulate the evolution of drag folds [21]. This same paper provides a theoretical analysis of the results in terms of buckling theory as well as experimental tests of the elastic case. The alternation of competent and incompetent layers such as Ramberg used in his experiments (Fig. 1) is characteristic of many

glacial deposits and it is probable that the arched and contorted structures found in many sub-glacial disturbances and push-moraines [22] have been formed by compressive forces, with results analogous to those produced experimentally. Ramberg himself has also studied the flow-pattern of moraine-carrying glaciers using models whirled in a centrifuge to simulate ice-flow [23]. Substances with high viscosity, such as bouncing putty were used as glacier-imitation materials. It was possible to produce experimentally realistic-looking fold structures of the kind seen in such piedmont glaciers as the Malaspina Glacier in south Alaska.

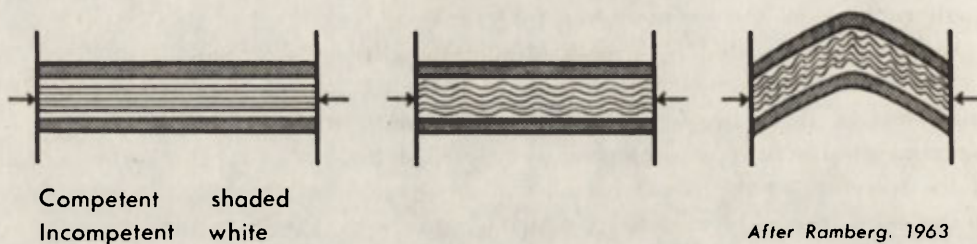


Fig. 1. Laboratory compression of competent and incompetent layers to produce chevron folds

In a series of experiments carried out by Dzulynski and Walton suspensions of plaster of paris and water were poured into tanks containing settled clay or gelatine to produce artificial turbidity currents [24]. They were able to reproduce experimentally a wide range of features similar to those recorded in natural sedimentary rocks. Load structures were also produced by carefully sieving a mixture of fine sand and plaster of paris so that it settled through water on to the surface of already settled muds. Resulting from these and other experiments they suggested that two general principles were applicable: 1. Turbid, sediment-laden flow under water and in sub-aerial conditions is similar; 2. Such flows may move easily whenever they enter into a quiescent water basin regardless of its depth.

Selley and Shearman [25] have reported other experiments. In the first series very fine grained sand was allowed to settle and accumulate with a relatively loose packing. Thin layers of black sand were added at intervals to give a visible bedding in the deposit. After deposition had ceased the container was given a severe shock which caused the sand grains to fall into a tighter packing. The bedding in the lower parts of the sand remained virtually undisturbed but the pore-water expelled from these lower parts produced local intense disturbance of the bedding in the upper layers. Where current bedding had originally been simulated the upper layers of this were completely obliterated. In a second series of experiments, a thin bed of clay was interposed with the regular bedding of sand. When the tank was shaken after deposition had ceased, excess pore water from the lower sand collected in a thin layer below the clay, then burst through

at a number of points and flowed upwards through the overlying sand disrupting its original bedding.

Although such experimental evidence has largely been used to interpret hard rock structures, it has considerable validity in the interpretation of the relatively unconsolidated glacial and fluvial deposits which are the main concern of geomorphologists in Britain [26, 27]. Jardine has applied the findings of Selley and Shearman and Dzulynski and Walton in a study of small-scale structures in a Quaternary fine sand-clay series, which he has attributed to loading followed by thrusting and injection during temporary liquefaction [28].

Amongst other Quaternary deposits to be studied experimentally have been varves, the palaeomagnetism of which has been examined by A. I. Rees [29]. Flume experiments were used to determine the phenomena occurring during the deposition of a fine silt because of an apparent anomaly in the magnetic properties of a laboratory-deposited silt. Basing his theoretical analysis on the classic work of Bagnold [30], Rees outlined the conditions under which sediment movement would occur and the bed would become rippled. In the range of stress investigated three regions divided by clearly established critical stress values were observed. At lowest stress the plane bed was stable without bed movement under all conditions and at highest stress the plane bed was stable with a recognizable transportable load. In the intermediate stress range the situation was more complex, but the bed could become rippled at any stress in the range, provided there was a source of sediment.

Experimental work to evaluate the changes occurring after deposition and during compaction has been reported by Smart [31]. He also provided a theoretical note on the consolidation of sediments arguing that during consolidation, if slipping or breaking of particles is more important than re-arrangement of the water-ionic complex, then it can be postulated that there are a number of locations at each of which it is possible for an irreversible volume-reducing event to occur. On average each such event allows the volume to reduce by the same amount. However, since the weaker locations will tend to collapse first there will be a tendency for the rate of collapse to decrease as consolidation proceeds. At the same time there may be an increase in the ability of the structure to bridge points of weakness. It can therefore be postulated that volume-reducing events occur at a rate with stress proportional to their number. The validity of this theoretical model was substantiated by the experimental compaction of a calcareous ooze in the laboratory.

Laboratory evidence relating to the physical characteristics of consolidated sediments has continued to accumulate although it is not yet possible to use many laboratory-determined physical properties of rocks to account for their geomorphic significance. As part of the Research Programme of the Mining Research Establishment D. W. Hobbs has described the strength and strain characteristics of a particular coal under triaxial compression [32] and N. J. Price has discussed the

influence of geological factors on the strength of Coal Measure rocks. [33]. Strength is controlled by three main factors: composition, porosity and degree of compaction. These same factors are presumably responsible for lithological control of landforms; for example, from the experimental results a close relationship would be expected between the spacing of joints and such factors as porosity that control strength. Price has also dealt at some length with theories of fault and joint development in rocks when they behave as brittle or semibrittle materials [34]. Combining field and laboratory investigations, Chorley is investigating the parameters which are assumed to control the present height variations of the Lower Greensand ridge. He has already published data relating to factors assumed to exercise a control over the shearing resistance of the sandy soils associated with the Lower Greensand [35].

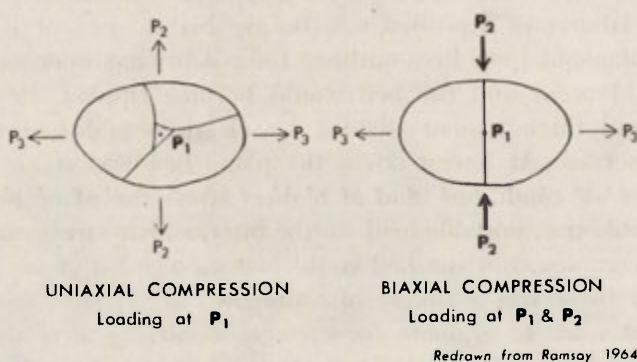


Fig. 2. Diagrammatic representation of laboratory crushing of pebbles in uniaxial and biaxial compression

The mechanical disintegration of rocks has been studied by D. M. Ramsay, who has described the different results of the laboratory crushing of pebbles in uniaxial and biaxial compression [36] (Fig 2). With uniaxial compression the pebbles were deformed until cracking appeared, and rupture in most cases occurred with explosive force on three radially oriented cracks somewhat oblique to the long and intermediate axes of the pebbles. At the point of loading on each pebble there was a circular area of crushing and flattening and in the more brittle rocks a roughly cylindrical zone of splintering parallel to the short dimensional axis. When compression was continued the cracks gaped and the margins of the wedge-like opening crashed in as slivers, and there was increased crushing in the circular zone where pressure was applied. With biaxial compression the fracturing of the specimen generally occurred as a single plane containing both the maximum and intermediate stress directions and normal to the unconfined direction. In a few cases cracking was a combination of vertical cracks with others inclined at 10—15 degrees. Whereas Ramsay was concerned with applying his

experimental observations to the observed deformation of pebbles in the Lower Old Red Sandstone conglomerate adjacent to the Highland Boundary Fault in Scotland, his work has wider implications for geomorphologists, for example in the study of pebbles shattered under periglacial conditions.

J. D. Clark has made a special study of the artificial fracture of pebbles in various different materials as part of a study into the development of early pebble tool cultures in Africa [37]. When pebbles are dropped or thrown, the proportion split in the same plane as the major or minor axis is as high as the proportion split at right angles to the axis. In naturally fractured pebbles from the Batoka Gorge of the river Zambezi in northern Rhodesia the fractures are usually parallel to the plane of the major axis. The simplest Kafuan pebble tools seem to be transversely split, that is, at right angles to the plane of the major axis.

Also, in connection with archaeological research, Schmalz has reported on the experimental production of flint patination, which, it seems, may proceed very rapidly under favourable conditions [38]. The flint examined by him was composed of a dense fine-grained aggregate of anhedral quartz crystals without any amorphous or opaline silica. Alkali solutions, even when very dilute, attacked the grains exposed to this action, etching the grain surfaces and in extreme cases perceptibly reducing the grain size. The modified aggregate scatters light to a much higher degree than does the unmodified aggregate, causing the white or blue-white surface effect known as patination. The increased porosity of the altered surface layer makes it more susceptible to mineral absorption from ground water, resulting in the commonly observed coloured patinas.

This study by Schmalz of the solubility of silica is only one of several carried out recently. A wide range of evidence concerning silica solubility was presented at a symposium on Experimental Pedology held at Nottingham in 1964 [39]. Data derived from drainage water analyses and lysimeter and laboratory experiments indicated the effect of organic matter in enhancing loss of silica during weathering and soil formation. Presence of calcium in the parent material likewise leads to the loss of silica. On the other hand the leaching of silica is restricted in the presence of iron and aluminium.

Whereas it is easy to establish weathering experiments in which the starting materials and end products are the same as those occurring in natural conditions, it is much more difficult to simulate experimentally the exact biological and physico-chemical environments found in nature. At best it is generally only possible to argue by analogy from the experimental results. This is well brought out in a discussion of experimental work carried out by Henderson and Duff on the effects of fungal action on minerals, rocks and soils [40]. They were able to show that strains of fungi which produced citric and/or oxalic acid were effective in decomposing certain natural silicates, while an oxalic acid-producing strain released metallic ions and silica from rocks and soils. It was argued that, under favourable circumstances in the field, the type of decomposition demonstrated

experimentally could occur. Such circumstances would arise, possibly only in a micro-habitat, when there was an accumulation of organic matter which could be readily metabolized. The cumulative effect of small-scale reactions over a long period of time could be an important factor in the weathering of minerals.

Opportunity to study the weathering of fresh rock debris has recently been provided by the erection of an experimental earthwork at Overton Down, Wiltshire [41] (Fig. 3). This is part of an experimental programme initiated by a research committee of the British Association for the Advancement of Science.

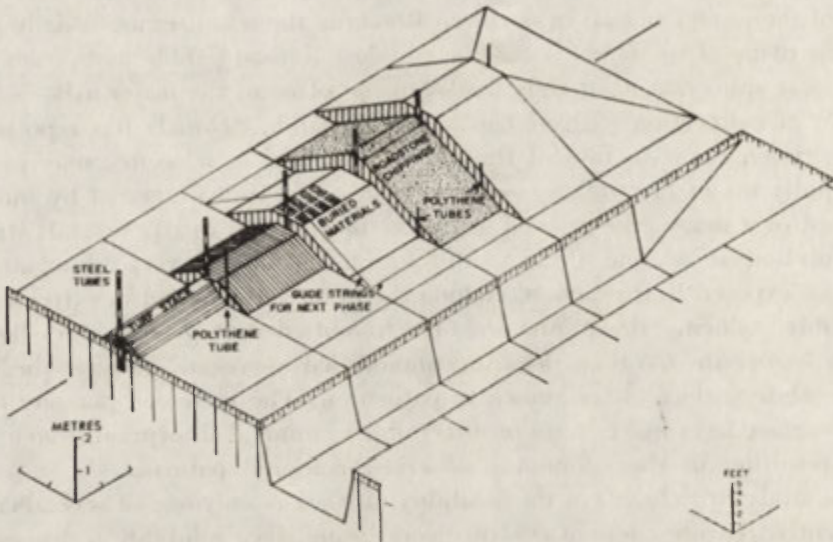


Fig. 3. Isometric drawing of experimental earthwork on Overton Down, Wiltshire, England

The primary purpose of the experiments, which are designed to last for more than a century, is to study the erection and denudation of earthworks so that quantitative information can be obtained that will be useful in the interpretation of archaeological excavations. Incidentally, however, considerable information of geomorphic significance is being obtained and it is this which will be discussed here.

Since experiments are being carried out over a relatively long period of time and the results must be easy to interpret and to compare with those from other experiments and from archaeological excavations, the experiments have been kept relatively simple. Each consists essentially of a linear bank and ditch built to uniform specification so that sections cut through the bank and ditch-fill after successive intervals of time will indicate changes through time, parent material and other factors remaining essentially constant. The first earthwork was built on chalk downland at Overton in Wiltshire in 1960, the second on sandy heathland in Dorset in 1963. Basically the earthworks consist of a ditch

some 90 feet (27.5 m) long, cut 5 feet (1.5 m) deep in "unweathered" parent material, 10 feet (3 m) wide at the top and 8 feet (2.4 m) wide at the bottom, the sides being evenly sloped. Each bank was built of material dug from its ditch, and was separated from it by a berm or platform 4 feet (1.2 m) wide. The bank on the chalk site at Overton was 23 feet (7 m) wide at the base and sloped towards an apex at 6 feet 7 1/2 inches (2 m). The second bank on the heathland site in Dorset was 19 feet 8 inches (6 m) wide, and sloped towards an apex at 5 feet 8 inches (1.7 m). The different dimensions resulted from the different expansion factors of chalk and sand.

The central cores of the banks were built of stacked turves cut from the area of the ditch. On top of these was then heaped the remainder of the soil. The rest of each bank was built in three distinct phases, material for the several phases coming from arbitrary but levelled layers in the ditch. Each phase was carefully built to a predetermined profile and between successive layers distinctive chip-pings of different rocks were spread. Two series of vertical polythene tubes fixed at ground level were also incorporated in the banks. These together provide data on movement within the banks as they settle. Two other types of markers were also incorporated in the earthworks, one type spread randomly over the site before any construction work began; the other type, numbered pottery discs, placed in precisely measured positions in the banks, on the berms and on the edges of the ditches. To ensure the accurate cutting of the final stages of the ditch a template having the shape of the required cross-section was used.

From the point of view of breakdown of rock debris and the development of soils, the earthwork on chalk at Overton is more suitable than the other, since the chalk is an extremely homogeneous parent material and normal soil development had proceeded to a relatively shallow uniform depth on the site before the experiment began. In 1962 and again in 1964 (two and four years after the construction of the earthwork) three large samples were collected from the base, middle and top of the sloping surface of the bank on the side away from the ditch. The size composition of the surface material was determined by slightly different methods in each year so that results are not strictly comparable. However, in both years there were marked differences between the samples from the top of the bank slope and those from the middle and base. Those from the middle and base had a very considerable proportion of material in the coarser grades, whereas those from the top had a greater amount of fine material present. In both years the analyses have reinforced the visual impression of a noticeable concentration of coarser material towards the base of the slope, amongst which there was a high proportion of flints. A crude estimate of this is given by the percentage of acid insoluble material in the basal samples — 26% of the basal samples in 1962, 20% in 1964, compared with between 2% and 9% in middle and top samples. Excavations through the bank in 1962 indicated that the crest of the bank had sunk by 4 inches (10 cm.) and that there had been lateral spreading

of about 9 inches (22.5 cm.) at front and back. By 1964 the bank had sunk a further 2 1/2 inches (6 cm.) but there had been no further lateral spread. The buried polythene tubes showed that there had been a downslope movement of the upper layers amounting during the four years to about 7 or 8 inches (18 or 20 cm.) at the surface and progressively less with increasing depth.

Field observations and laboratory data enable an outline to be given of suggested processes in the weathering of the bank surface. After construction the effect of weathering would be to break up the material on the surface of the bank, largely by alternate freezing and thawing, by wetting and drying and by solution of calcium carbonate from the chalk, leading to disaggregation. Although initially the flinty material in the bank was probably generally larger in size than the chalky material the greater percentage of flint present in the larger grades does suggest that the flint blocks have been more resistant than the chalk to disintegration. Under the conditions of creep, wash and consolidation, leading to the slow downhill movement of material at the bank surface, as shown by the bent polythene tubes, there would be a tendency for the larger material to roll, or slide to the lower parts of the slope and there accumulate. If the movement of the larger material occurs sporadically rather than continuously, then we would expect to find a patchy distribution of larger material on the surface of the bank. On occasions, particular areas of bank surface would lose a considerable part of their larger material and there would be a more even distribution of the various grades of material, as in the top sample collected in 1962. Downslope movement has progressively uncovered a small but significant supply of large unweathered rubble which, spreading down the two sides of the bank, has provided as it broke up, a supply of rubble dominantly rather less than 2 inches (5 cm.) in diameter.

As well as disintegration of material, some aggregation has also taken place, with the "cementing" of chalk particles by redeposition of calcium carbonate. This was noted in all the samples collected at the surface in 1964, and was especially common with material between 1/4-inch (0.6 cm.) and 1/8-inch (0.3 cm.) in diameter, although it did occur on both larger and smaller particles. Similar aggregation occurred within the bank. Usually this "cemented" chalk could be disaggregated by gently crushing between the fingers, while its porous nature could be clearly seen. Presumably it is this material which holds water that is available for plant growth, and which allows the colonisation of weathering chalk surfaces by higher plants [42].

No such plants had colonised the bank surfaces by 1964. However, at a depth of about 1 inch (2.5 cm.) below the surface, a distinct layer of blue-green algae was noted in 1962 and 1964. These algae have the ability to fix atmospheric nitrogen, but the failure of higher plants to colonise the surface is more likely to be due to drought than to nitrogen shortage. In all the 1964 samples from the bank surface other organic matter was also found. This consisted of remains of higher plants, fragments of leaves, stems, seed-cases and roots and fruits;

fragments of mosses and other plants; and crumbs of humus. With the exception of the moss fragments, which might have been derived from plants growing on the bank, although none had been seen in spite of careful search, this organic material had obviously been blown on to the bank from the surrounding area. There was concentration of a considerable amount of organic matter in the fine fraction ($< 2\text{ mm.}$) of all the samples from the surface of the bank. Values in all size grades are higher for base (maximum 3.1% in the fine fraction) and top samples than for middle samples.

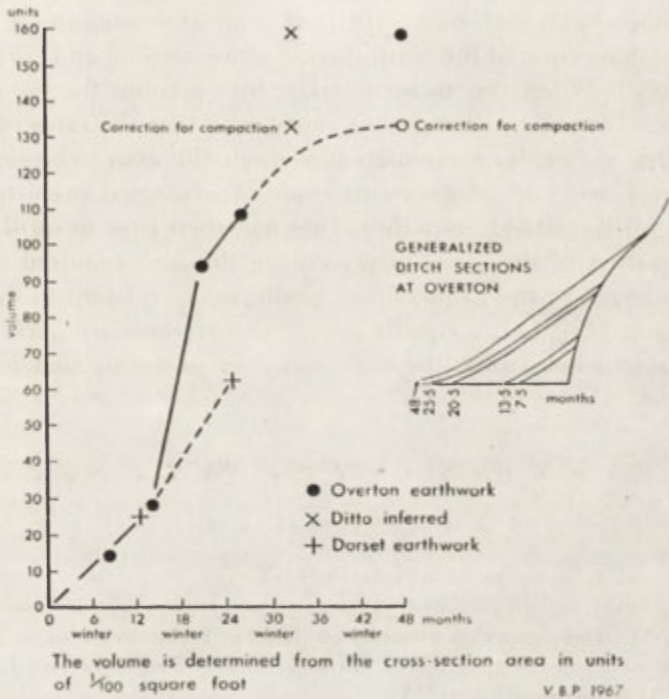


Fig. 4. Rate of filling of ditches of experimental earthworks, England

Virtually all this organic matter represents material blown on to the surface of the bank or deposited during snow fall. The outermost layer of the bank initially contained some organic matter derived from the ditch trimmings but this was so insignificant in amount that it was not recorded in the samples collected from the bank surface in 1962.

The silting of the ditch as a result of the weathering of the sides provides opportunity on a small scale to examine one of the classical problems in geomorphology — that of slope development. As expected according to theory, the material falling on the ditch floor has protected the base of the ditch side from further weathering and left a “nose” of unweathered rock beneath the ditch-fill

[43]. As the ditch-fill has mounted up the side of the ditch the original straight slope has been weathered into a curved one which is successively protected from further weathering by the accumulation of more ditch-fill. Within two years of the start of the experiment sufficient material had weathered from the side of the ditch to indicate that its final form would be very close to that predicted on the basis of the equations developed by Scheidegger [44].

When the volume of ditch-fill that has accumulated since the beginning of both experiments is plotted against time the effects of a single winter on the rates of accumulation are well shown (Fig. 4). In winter there is greater fluctuation in temperatures and moisture, thus a greater amount of weathering. After the initial deposition of ditch-fill there is some settling and surface redistribution of material. When this factor is taken into account for the more recent ditch-fill of the Overton earthwork it is easily seen that the rates of weathering of the ditch sides and of the accumulation of ditch-fill seem to be slowing down. This is in accordance with observations made of excavated prehistoric ditches, which seem to infill relatively rapidly at first and then later to infill much more slowly. Continuation of the experiments during the next hundred years should provide data relevant to the formation of landforms in relation to the frequency of climatic events [45]. Other equally simple experiments are likely to produce quantitative data to help solve the wide range of problems that confronts the geomorphologist.

Formerly at University of Durham, now at University of Alberta

BIBLIOGRAPHY

- [1] King, C. A. M., *Techniques in Geomorphology*, London 1966, 185—230.
- [2] Compare the comments by Selby, M. J. Methods of measuring soil creep, *J. Hydrol. (New Zealand)*, 5 (1966), especially pp. 54—57.
- [3] Young, A., Soil movement by denudational processes on slopes, *Nature*, 188 (1960), 120—122; Young, A., Soil movement on slopes, *Nature*, 200 (1962), 129—130.
- [4] Darwin, H., On the small vertical movements of a stone laid on the surface of the ground, *Proc. Roy. Soc. London*, 68 (1901), 253—261.
- [5] Edwards, R. S., Studies of the growth of cereals on seed-beds prepared by dry sieving of the soil. I, *Empire J. Exp. Agric.*, 25 (1957), 167—184; II, *ibid.*, 304—320. III, *Empire J. Exp. Agric.*, 26 (1958), 323—343.
- [6] Joliffe, I. P., An experiment designed to compare the relative rates of movement of different sizes of beach pebble *Proc. Geol. Assoc.*, 75 (1964), 67—86.
- [7] Allen, J. *Scale models in hydraulic engineering*, London 1947.
- [8] Griggs, D., A theory of mountain building *Amer. J. Sci.*, 237 (1939), 611—50.
- [9] Kenn, M. J., Experiments simulating the effects of mountain-building, initiated by the late Professor S. W. Wooldridge, *Proc. Geol. Assoc.*, 67 (1965), 21—28.
- [10] Wilson, G., The tectonics of the 'Great Ice Chasm', Filchner Ice Shelf, Antarctica, *Proc. Geol. Assoc.*, 71 (1960), 130—138.

- [11] Riedel, W., Zur Mechanik geologischer Brucherscheinungen, *Zbl. Miner. Geol. Palaont. Abt. B.*, 1929, pp. 354.
- [12] Gilbert, G. K., *The transportation of debris by running water*, U. S. Geol. Survey. Prof. Paper 86, 1914.
- [13] Kuenen, Ph. H. Experiments in geology, *Trans. Geol. Soc. Glasgow*, 23 (1958), 1—28; For more recent experiments see: Kuenen, Ph. H., Matrix of turbidites: experimental approach, *Sedimentology*, 7 (1966), 267—297.
- [14] For a broad survey utilising a variety of data, observational and experimental, see; Allen, J. R. L., On bed forms and palaeocurrents, *Sedimentology*, 6 (1966), 153—190.
- [15] Stewart, A. D., On certain slump structures in the Torridonian sandstones of Applecross, *Geol. Mag.*, 100 (1963), 205—218.
- [16] Selley, R. C. et al., Some underwater disturbances in the Torridonian of Skye and Raasay, *Geol. Mag.*, 100 (1963), 224—243.
- [17] Rolfe, W. D. I., A fine air-heave structure from the Old Red Sandstone of Lanarkshire, Scotland, *Geol. Mag.*, 97 (1960), 133—136.
- [18] Williams, E., Intra-stratal flow and convolute folding *Geol. Mag.*, 97 (1960), 208—214.
- [19] Sanders, J. E., Origin of convoluted laminae *Geol. Mag.*, 97 (1960), 409—421.
- [20] Williams, E., Origin of convoluted laminae — correspondence, *Geol. Mag.*, 98 (1961), 168—170.
- [21] Ramberg, H., Evolution of drag folds, *Geol. Mag.*, 100 (1963), 97—106; see also Ramberg, H., Fluid dynamics of viscous buckling applicable to folding of layered rocks, *Bull. Amer. Assoc. Petroleum Geol.*, 47 (1963), 484—505.
- [22] Charlesworth, J. K., *The Quaternary Era*, London 1957, e.g. 255—259, 410—411.
- [23] Ramberg, H., Note on model studies of folding of moraines in piedmont glaciers, *J. Glaciol.*, 5 (1964—5), 207—218.
- [24] Dzulynski, St., and Walton, E. K., Experimental production of sole markings, *Trans. Geol. Soc. Edinburgh*, 19 (1962—3), 279—305.
- [25] Selley, R. C., and Shearman, D. J., The experimental production of sedimentary structures in quicksands, *Proc. Geol. Soc. London*, 1599 (1961—2), 101—102.
- [26] Compare the experimental work carried out by Köster to reproduce features seen in glacial deposits: Köster, R., Zur experimentellen Darstellung glazialer Stauchungszonen, *Freiburger Forschungshefte C*, 80 (1960), 58—86.
- [27] More than one quarter of all geomorphic work carried out in Britain is specifically concerned with Quaternary deposits, landforms and chronology. Proudfoot, V. B., Geomorphological Research in Britain and Ireland, 1966—7, *Geogr. J.*, 133 (1967), 591—593.
- [28] Jardine, W. G., Note on a temporary exposure in central Glasgow of Quaternary sediments with slump and load structures. *Scot. J. Geol.*, 1 (1965), 221—224.
- [29] Rees, A. I., Some flume experiments with a fine silt, *Sedimentology*, 6 (1966), 209—240.
- [30] Bagnold, R. A., The flow of cohesionless grains in fluids, *Phil. Trans. Roy. Soc. London, Ser. A.*, 249 (1956), 235—297.
- [31] Smart, P. Consolidation of sediments, *Geol. Mag.*, 101 (1964), 565—6.
- [32] Hobbs, D. W., The strength and stress — strain characteristics of Oakdale Coal under triaxial compression, *Geol. Mag.*, 97 (1960), 422—435.
- [33] Price, N. J., The influence of geological factors on the strength of Coal Measure rocks, *Geol. Mag.*, 100 (1963), 428—443.
- [34] Price, N. J., *Fault and joint development in brittle and semi-brittle rock*, London 1966.
- [35] Chorley, R. J., An analysis of the areal distribution of soil size facies on the Lower Greensand rocks of East-Central England by the use of trend surface analysis, *Geol. Mag.*, 101 (1964), 314—321; Chorley, R. J., Geomorphic evaluation of factors controlling shearing resistance of surface soils in sandstone, *J. Geophys. Res.* 69 (1964), 1507—1516.

- [36] Ramsay, D. M. Deformation of pebbles in Lower Old Red Sandstone conglomerates adjacent to the Highland Boundary Fault, *Geol. Mag.*, 101 (1964), 228—248.
- [37] Clark, J. D., The natural fracture of pebbles from the Batoka Gorge, Northern Rhodesia, and its bearing on the Kafuan Industries of Africa, *Proc. Prehist. Soc.*, 24 (1958), 24—42.
- [38] Schmalz, R. E., Flint and the patination of flint artifacts, *Proc. Prehist. Soc.*, 26 (1960), 44—49.
- [39] Hallsworth, E. G., and Crawford, D. V., eds., *Experimental Pedology*, London 1965.
- [40] Henderson, M. E. K. and Duff, R. B., The release of metallic and silicate ions from minerals, rocks, and soils by fungal activity, *J. Soil Sci.*, 14 (1963), 236—246.
- [41] Jewell, P. A., ed., *The Experimental Earthwork on Overton Down, Wiltshire, 1960*. British Association for the Advancement of Science, London 1963, pp. 100. Jewell, P. A. and Dimbleby, G. W., eds., The experimental earthwork on Overton Down, Wiltshire, England: the first four years, *Proc. Prehist. Soc.*, 32 (1962), 313—342. Proudfoot, V. B., The study of soil development from the construction and excavation of experimental earthworks, in Hallsworth, E. G., and Crawford, D. V., eds., (1965), as footnote 39, 282—294.
- [42] Compare Locket, G. H., Observations on the colonization of bare chalk, *J. Ecol.*, 33 (1945—6), 305—9 and Locket, G. H., A preliminary investigation of the availability to plants of the water in chalk, *J. Ecol.*, 33 (1945—6), 222—9.
- [43] Fisher, O., On the disintegration of a chalk cliff, *Geol. Mag.*, 3 (1966), 354—356. Wood, A., The development of hillside slopes, *Proc. Geol. Assoc.*, 53 (1942), 128—138. Scheidegger, A. E., *Theoretical Geomorphology*, Springer-Verlag, Berlin 1961, 88—117, especially 88—98.
- [44] Scheidegger, A. E., as footnote 43 and see Jewell P. A., ed., (1963) as footnote 41, 86—9.
- [45] Leopold, L. B. et al. *Fluvial processes in geomorphology*, San Francisco and London 1964, especially 80—94.

Acknowledgment

Figures 1 and 2 are reproduced with the permission of the authors and the editors of *The Geological Magazine* and *The Bulletin of the Association of American Petroleum Geologists*. Figure 3 is by courtesy of the British Association for the Advancement of Science.

RECENT DEVELOPMENTS AND TECHNIQUES IN THE STUDY OF KARST LANDFORMS IN THE BRITISH ISLES

MARJORIE M. SWEETING

The study of karst landforms has developed rapidly in Britain in the last few years. The aim of this paper is to discuss some of these recent developments particularly in relation to the development of the techniques used. Two main developments will be discussed, 1) the study of the solution of limestones and of karst waters; 2) advances in the study of karst landforms in Britain.

THE SOLUTION OF LIMESTONES AND STUDIES OF KARST WATER

A number of workers are now engaged on problems connected with the composition of limestone waters. The most usual methods of chemical analysis include conventional titration using E.D.T.A.¹ However, various rapid methods have been tried, involving the use of tablets, which when added to the limestone waters give a colour change. The kinds of tablets which have proved most helpful under British conditions are the Durognost and the Palintest varieties, whose merits have been discussed by Douglas [1]. Recently Douglas has compared the Durognost, Palintest and E.D.T.A. titration methods of hardness determination. He finds that on average Durognost determinations are accurate to within 6.5% of the titrated hardness and the Palintest back-titration to within 7.5%. However in 6 out of 10 determinations, the Palintest method was more accurate than the Durognost method. Table 1 shows a comparison of 10 determinations [2].

An account of certain British work on the calcium and magnesium content of karst waters in the British Isles was given in the Symposium of 1965 [3]. Since then much more has been done and the problems, instead of becoming more simple, have become more complex. The variations of natural waters on a local scale appear to be much greater than was apparent at that time; this can be illustrated by a consideration of the work in many areas. For instance, in the Mendip Hills in S. W. England, D. I. Smith claimed an average of about 250 p.p.m. calcium in the cave and spring waters of that area, and maintained that a model for the solution of the limestone in the Mendips could be made [4]. More recent work, by Ford, suggests that the picture is in fact complex, and that waters from areas in quite close proximity have considerable variations in their calcium content. Thus Ford's values for nearby and closely similar caves vary

¹ See also Stenner, R. D., The measurement of the aggressiveness of water towards calcium carbonate, *Trans. Cave Res. Group*, 11 (1969), 175—200.

TABLE 1. Comparison of the Durognost, Palintest and E.D.T.A. titration methods of hardness determination

	Durognost	Palintest	E.D.T.A. Titration
Derbyshire Samples	249	230	248
	142	144	147
	258	231	259
	205	181	189
	249	250	254
Norwegian Samples (collected by A. F. Pitty)	36	41	39
	71	70	68
	45	43	46
	37	36	31
	53	51	63

(from I. Douglas [2])

form 103 p.p.m. to over 320 p.p.m. calcium hardness. Ford believes that variations in the fabric and lithology of the limestones have not been taken into consideration sufficiently [5].

The solubility of limestones is influenced by many factors, of which the amount of carbon dioxide in the soil and the nature of the lithology of the limestones are some of the most important. Work along two main lines illustrates these conclusions.

First, in the Peak district of Derbyshire, Pitty has conducted a careful study of the karst water in a sample cave, Pooles Cavern [6]. The variations in the quantities of dissolved calcium carbonate are shown in Fig. 1 for throughout the year, from May 1963 to July 1964. The difficulty of recognising any pattern in the changes from month to month can be seen from this figure. Pitty has, however, subjected these results to statistical analysis and this has yielded correlation coefficients between the climatic data and hardness fluctuations which appear to be statistically highly significant. The most reliably established and the most interesting positive association established by Pitty is that between the calcium hardness of the water and the temperature at an interval of about 84 days before sampling. This association is unlikely to result from a direct chemical relationship, because the speed of the reaction is unimportant if the water takes so long to move through the limestone. This correlation suggests that biochemically generated carbon dioxide is involved, controlled in its fluctuations by the temperature. From a study of the importance of carbon dioxide in the soil Pitty concludes that the positive correlation with hardness shown in most of the sampling stations is a complex physical and biochemical chain. He goes on to say (p. 38): "In view of the complexity of the processes involved perhaps the term rhizosphere carbon dioxide could be useful and more precise, by implying the

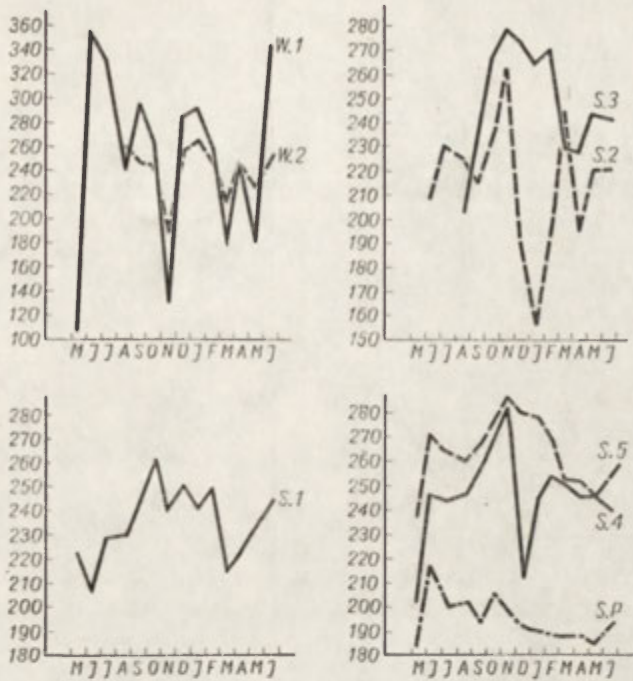


Fig. 1. Calcium carbonate content in parts per million
(from A. F. Pitty, 1966)

manner in which surface vegetation is involved in the production of soil carbon dioxide.” The pattern shown in Figure 2 shows the rapid increase in hardness caused as a result of carbon dioxide increase in February and March, and which Pitty likens to the “spring burst” of activity frequently noted in the studies of micro-organisms. Furthermore there is a positive association with an increased calcium content in November: this characteristic can be attributed to the influence of maximum production of rhizosphere carbon dioxide in mid-summer, the effects delayed by the slow movement of the karst water. Hence the conclusions from this piece of work are that there is an appreciable correlation of the content of the karst water with the temperature 84 days or so earlier: rhizosphere carbon dioxide may explain the high hardness and the fluctuations, even if other modifications mask the rhythm of the main factor. Moreover, the results also show that there can be marked contrasts between amounts and rhythms of solution in the cave waters between sampling points which are less than a metre apart, and from year to year. Thus it is concluded that any results drawn from one locality must be carefully viewed, particularly from the morphoclimatic point of view. This has also been stated by Gams in his work on karst waters in Slovenia [7].

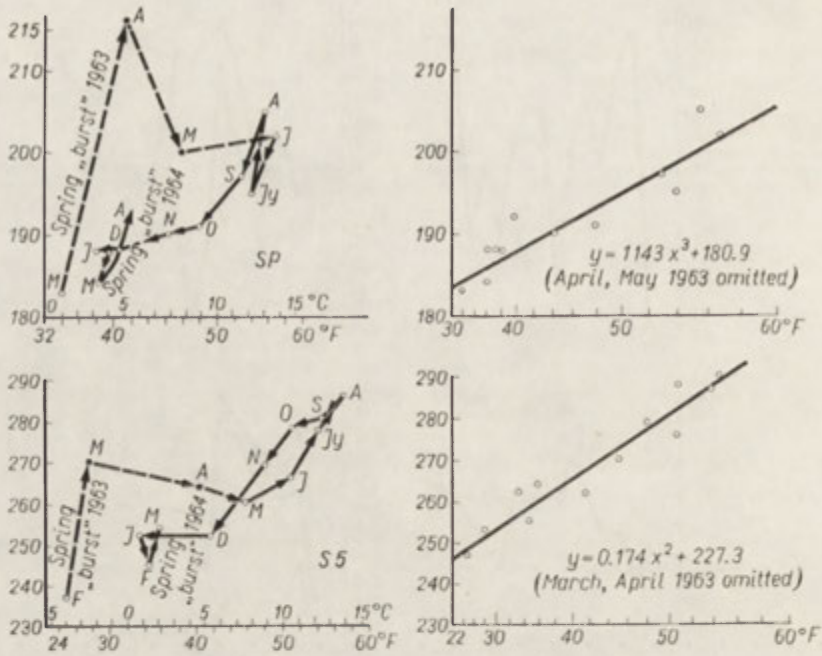


Fig. 2. Calcium carbonate content in parts per million

Air temperatures are mean figures for periods of 9—12 weeks (SP) and 15—18 weeks (S5) prior to sampling (from A. F. Pitty, 1966)



Fig. 3. A divisional scheme for the textural characters of carbonate rocks. For N.W. England

1 — Micrites; 1—10% calcite. Dense groundmass (cement) with occasional dust-like particles and micro fragments of fossils

2 — Biomicrites

a. Sporadic; 10—20% calcite. Dense to dusty groundmass (cement) with small and shred-like flakes of calcite; and with small and occasional determinate fossils, all dispersed and scattered through matrix

b. Packed; 20—50% calcite. Fine grained groundmass (cement) made up of closely packed particles and angular fragments of calcite with fossils occas oolites (Oomicrites)

3 — Sparry limestones

a. 50—75% calcite. Medium to coarse-grained groundmass. Well formed crystal plates of calcite. Also fragmental and granules of calcite as a cement. Fossils common

b. Over 75% calcite. Coarse—grained groundmass, with large plates and mosaic-forms of calcite. Numerous fossils can occur and these are often preserved in cryst. calcite. Cement very sparry with little pore-space (modified from R.C. Folk, 1962)

Secondly, in studies of karst waters in Northern England the author has become interested in the differential solution rates of limestones, particularly in the Carboniferous Limestones. In addition to chemical analyses of the waters, the rocks have been examined petrologically. Preliminary results suggest that there are considerable variations in the solution rates between one type of limestone and another. There are many types of limestone, possibly over 500, which vary widely in porosity and in texture [8]. Porosities in the Carboniferous Limestones examined vary from about 2% or less to over 20%. The table (Fig. 3) illustrates the main limestone types in N. England. The sparry limestones (sparites) have an estimated porosity of about 5—8%; the crystalline calcite reaches 90% in amount and over, and occurs in large united plates and in mosaic aggregates. Textures such as this will tend greatly to restrict the formation of the cement and make the pore space reduced. Micrites have a very low porosity probably 2% or less; they consist of a very dense compact groundmass or cement made up of microscopic calcite particles and clay-paste carbonate. The limestones which have been found to have the highest porosity, of at least 15%—25%, are the biomicrites (especially the packed biomicrites); in these rocks the groundmass is mainly composed of finely granular and fragmental matter allowing for many interstices or spaces to occur and for the formation of cement of varying composition and amount.

Preliminary results in Northern England suggest that the biomicrites are as much as twice as soluble in a given time and under the same conditions as the sparites or the micrites. Analyses of karst waters draining from areas of biomicrites show calcium hardness values of about 160—180 p.p.m.; those draining from areas of sparites values of about 60—80 p.p.m., for conditions of equivalent time and rainfall (Fig. 4).

Detailed analyses of limestone waters in the central English scarplands likewise illustrate the effects of differing lithologies, within a small area and one of general climatic uniformity. Three limestones of varying lithology — the Jurassic oolitic limestones of the Cotswold Hills, the Jurassic Corallian limestones, and the Cretaceous chalky limestones, outcrop within a few kilometres of Oxford. The waters and springs of these limestones have been analysed systematically by Mr. K. Paterson during the past year; some thousands of analyses have been made but have not yet been statistically analysed. The most consistently high values are found in the springs on the Corallian limestones, with between 400—500 p.p.m. calcium; these limestones are to some extent dolomitic, and it has been found by Gams that karst water from dolomitic limestones tends to have the greatest calcium hardness [7]. Springs from the Jurassic oolitic limestones show variations within the range 300—360 p.p.m. Springs from the chalk have rather lower values, with figures ranging from about 180 to about 280 p.p.m. Studies in the fabric and lithology of these rocks as they affect their solubilities are now being undertaken.

A programme of water sampling in the north west of Scotland in the area of the Durness limestone of Cambrian age, has now been started. Here the limestones are partly dolomitic. It has been pointed out by D. I. Smith that the proportion of magnesium in karst waters is approximately in proportion of magnesium in karst waters is approximately in proportion to the amount of magnesium carbonate in the limestones [9]. Other aspects of magnesium solution have been dealt with by Douglas and by the author [10, 11].

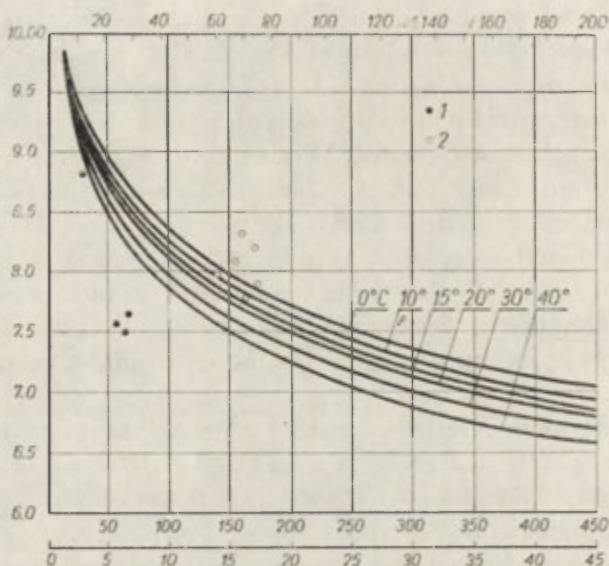


Fig. 4. Differential solution of limestones. N.W. Yorkshire

1 — Sparry limestones; W. part, 2 — Biomicrites; E. part

Attempts to measure the rate of limestone solution directly are also being made. These are being done by the University of Bristol Spelaeological Society in their work in county Clare, Ireland. The instrument for measuring the rate of erosion consists essentially of a tripod with a depth measuring meter on it. On this tripod is mounted a 'clock-gauge' precision meter which indicates on a scale the linear movement of its probe. The particular clock-gauge used indicates the movements' increments of 0.0025—4 mm. Several of these instruments have been placed in caves in county Clare and measurements are being taken. Under good conditions it is claimed that it is possible to measure the depth of limestone removed over periods as short as two days. The repeatability of the readings is ± 0.00025 mm [12]. The first results confirm the relatively rapid solution of the limestones by cave waters in this area.

Karst waters are also being studied along the line that Jakučs has developed in Hungary, by means of flood wave and pulse phenomena [13, 14]. In this

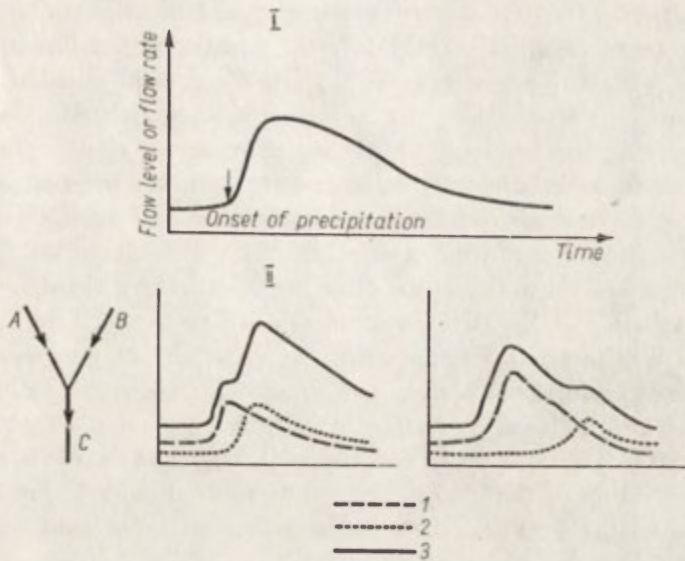


Fig. 5. I. General form of a flood wave. II. Addition of pulses
1 — Form of flow at A, 2 — Form of flow at B, 3 — Form of resultant flow at C
(from K. Ashton, 1966)

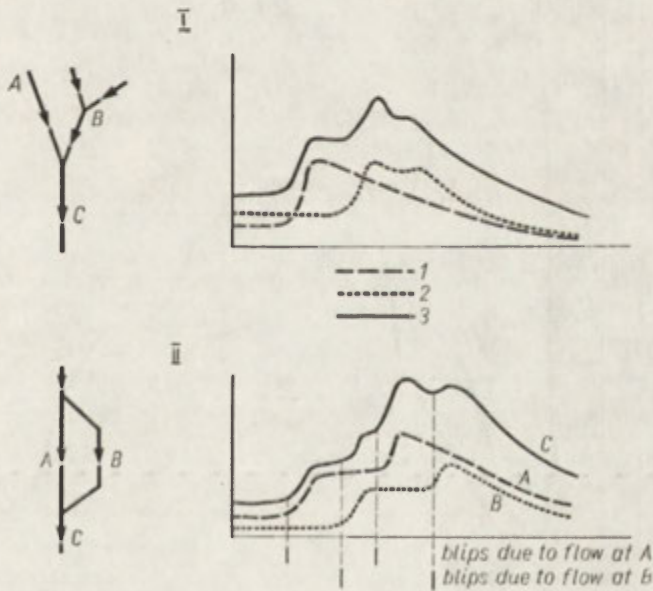


Fig. 6. I. Addition of complex pulses. II. Flow characteristics of a loop
1 — Form of flow at A, 2 — Form of flow at B, 3 — Form of resultant flow at C
(from K. Ashton, 1966)

method the waters of a cave or rising are assumed to originate from varying sources and to be made up of several different types of water. For instance they contain waters from both the streams which originate on the non-karstic ground, and waters from the general percolation which have percolated slowly through the limestones only; the temperature, pH and hardness of these differing waters are normally distinctive. The methods used by Jakuĉs represent some of the first to attempt to separate out these waters in a cave or spring system. With the help of computers it is now possible to analyse the results of the flood pulses. The analysis is based on the fact that after heavy rains or a flood, the waters in a cave or spring will rise, but that the water that is first expelled through the cave or spring will be the ground water which is forced out of its passages by the pressure of the incoming flood water. This ground or phreatic water is recognised by its high pH and high concentration in calcium carbonate. Later the ground water is replaced by the flood water as it comes through and the pH becomes lower and the concentration of the calcium carbonate more dilute. A common pattern is indicated in Figure 5. The length and duration of the ground water charac-

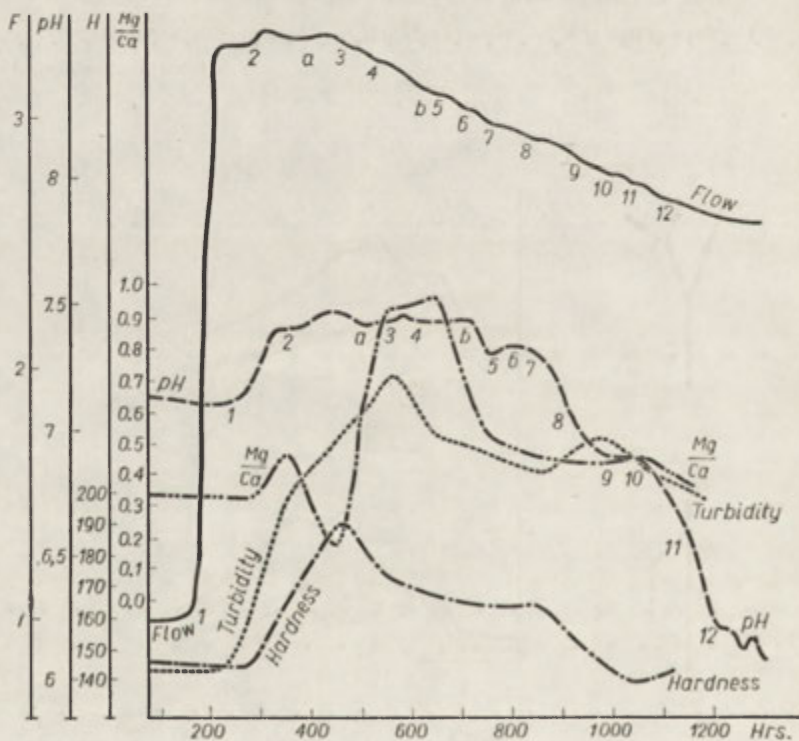


Fig. 7. Black River, St. Catherine's, Jamaica. 9.9.1963

(from K. Ashton, 1966)

teristics and the length of time needed for both it and the following flood water to pass through a system, gives us some idea of the extent of catchment of both the ground water and the flood water. If for instance a cave or a spring is simple, one flood will produce one set of phenomena. More often caves or springs have many contributory inlets each producing flood water; the graphs will then consist of many such pulses, as seen in Figure 6 and such graphs need to be analysed arithmetically. Ashton has applied this kind of technique to the analysis of some karst waters in north west England. It has also been tried with success in some bigger caves in the White Limestone cave areas in Jamaica, and an example is given, in Figure 7, of the Black River. This type of water analysis is of undoubted help in tracing the underground waters of an area, and will also throw light upon the nature of the passages supplying the springs.

ADVANCES IN THE STUDY OF THE LANDFORMS

Much progress has also been made in Britain on the study of actual karst landforms. In particular, progress has taken place in the inter-relationships between lithology and landforms; in the study of caves; in the studies of tropical karstlands; and in the elucidation of fossil karst in Britain.

Much petrological work has been done on certain of the British limestones. In northern England, in addition to their differential solubilities already mentioned, the different limestone types produce different landforms. For instance, jointing in the more sparry types of the Carboniferous Limestones tends to be less frequent (at about $\frac{1}{2}$ to 1 metre intervals) and the beds are massive. Solution is directed along the widely-spaced joint planes and relatively large and isolated blocks of the rock become detached. In the biomicrites and in the micrites, the jointing is much more frequent (at about $\frac{1}{4}$ — $\frac{1}{2}$ m. intervals) and the beds are much less massive; solution is more widely diffused throughout the rock and a much more comminuted and weathered rock debris or scree is formed. Such variations in joint frequency was the subject of a study by Doughty [15].

The styles of weathering of the different limestone types are also variable. Thus the sparites tend to form the more massive cliffs. Streams, both surface and underground, are concentrated along the joint planes, and narrow canyon-like valleys and narrow cave-passages are formed. Large-scale collapse of the massive blocks is rare, and the sparites are less susceptible to frost than are the other types of limestones. Slipping of the massive blocks does however take place. Enclosed hollows formed in these massive beds tend to be deep and narrow and to weather back in steep walls by parallel retreat. The impressive steep walled chasms or potholes so characteristic of the upper beds of the Carboniferous Limestones



Plate 1. Valley cut into sparry limestones (N.W. Yorkshire)



Plate 2. Valley cut into biomicrite (N.W. Yorkshire)

in the Ingleborough district of Yorkshire tend to originate predominantly in these highly sparitic and massive beds. These sparry beds also give rise to the most important limestone pavements [16, 17].

The biomicrites, on the other hand, in N. England give rise to lower cliffs. These rocks are more porous, and this together with the greater frequency of the jointing means that they tend to be more soluble and easily weathered. Vallies are wider and shallower. The biomicrites are more susceptible to joint weathering; overhanging rock shelters (abris) formed by periglacial action in the Pleistocene commonly occur in these rocks. Collapse is also important, and many of the larger collapsed caverns in N. W. Yorkshire are associated with the micrites and biomicrites [18]. Plates I and II illustrate the style of weathering.

The same points can be made by a comparison of the Carboniferous Limestones in the Burren region of county Clare in western Ireland. In general the limestones there are much more sparry, averaging over 84% compared with 73% for N. W. England. This has resulted in a series of rather more massive and less soluble limestones in western Ireland. Collapse landforms are less frequent but a greater percentage of the beds form limestone pavements. Absolute solution rates, despite a higher rainfall, are also lower [19].

In southern and central England studies in the areas underlain by the chalk illustrate the importance of pore-space, both in its effect on solution and on the landforms. It can be shown that the maximum solution, and possibly also maximum reprecipitation, occur in the top two metres of the ground. Because the chalk has such a large porosity — nearly 50% — its water holding capacity is much greater than in the more crystalline limestones. Not only is its solution rate higher, as shown already, but the nature and locus of the solution also differs. Whereas in the sparry limestones solution is directed along the joints and fractures, solution of a porous limestone like chalk is more greatly diffused throughout the rock.

Such solution causes the uniform lowering and denudation of the chalk, almost grain by grain, compared with the much more selective solution that takes place in massive limestones. In massive limestones such selective solution gives rise to local accelerated corrosion, as discussed by Gams [20], whereas accelerated corrosion is unimportant in the chalk. This difference partly accounts for the absence of karst hollows and caves in the chalk.

In the study of caves in Britain the Cave Research Group has expanded its activities and has been able to publish many of the results of cave studies in Britain, both from the point of view of exploration and of the more scientific aspect. In the south west of England, the first British school of speleology has been established in the William Pengelly Cave Research Centre at Buckfastleigh, in S. Devon. Work is now well ahead in the establishment of this centre, which is situated close by a group of caves in the Devonian Limestones. A museum and workshop are planned and it is hoped to be able to accommodate, albeit in a simple way, research workers who wish to work on the varied aspects of the

caves of south west England. Though the size of the caves and the areas of the limestone are small, the caves are of great intrinsic interest in themselves in that they are some of the most southerly occurring in the British Isles, and have better stalactites and a richer fauna than the other English caves and are much more comparable with caves on the continent. With the establishment of this centre there has appeared another English journal devoted to the problems in speleology, called "Studies in Speleology" [21].

In tropical areas many British workers have been active, particularly in the West Indies and in central America. In Jamaica, much speleological exploration has been carried out and many of the techniques first devised in Europe are now being adapted to the problems of this tropical area. In particular the methods of water tracing by using spores and first introduced by Zotl [22] are now being used in Jamaica² to solve problems of underground waterflow. In central America the types of limestone landscape found in British Honduras have been examined by the author. Variations in the limestone landforms of that area can be shown to depend both upon climatic and lithological factors, cone or kegel karst being formed in areas where the rainfall is at least 30" (760 mm.) per year, and where the limestones are crystalline or micritic in texture. In areas of lower rainfall or where the limestones are porous, rolling limestone hills much more like Yucatan rather than cone karst are formed.

A main advance in British landforms studies in the last few years has been in the applications of the ideas of climatic morphology to the British Isles — a subject which the British have been slow to adopt. However, past climatic phases can be shown to have played an important part in the development of the limestone relief in Britain. This work is just beginning in Britain. Ford, in his studies of the Mendip caves in south west England, has suggested that certain erosional and depositional phases can be related to oscillations in climate during the Pleistocene [23]. This work together with our new ideas on the relatively rapid limestone solution rates indicates the recent origin of many surface landforms and caves both in north west England and in county Clare, Ireland (Fig. 8). We are therefore much more able to differentiate between the landforms that owe their origin to the Pleistocene glaciations and those that are more ancient. We can speak of distinct phases of karstification.

In north west Yorkshire, an older pre-glacial karst can be shown to exist in the eastern part of the area where there are a few large *dolines* and *uvalas* and a marginal *polje* [24]. The western part is an area of glacial and post-glacial karst — largely a glacially-stripped karst or *schichttreppen* karst [25]. Vertical potholes and caves, and limestone pavements are important; true *dolines* are not properly developed, but remnants of hollows formed by solution under snow patches (*kotlici*) are common.

² By D. I. Smith.

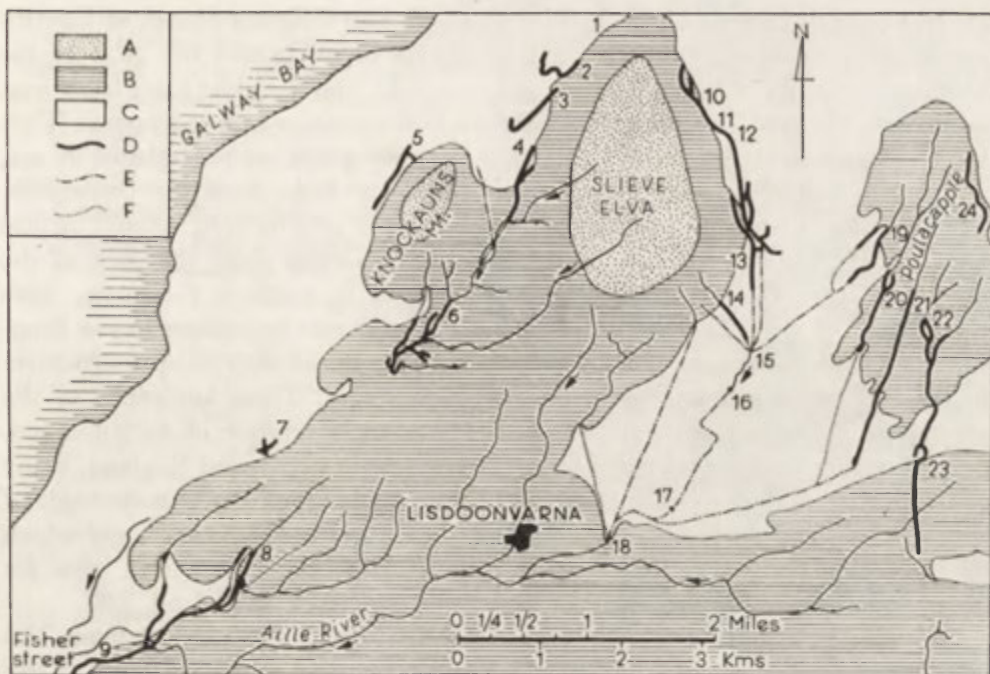


Fig. 8. Caves of post-glacial age in the Burren of County Clare, Ireland

A. Flagstone, B. Shale, C. Limestone, D. Cave passages, E. Drainage known, F. Drainage postulated
 1 — Pollapooka, 2 — Faunarooska, 3 — Pollballiny, 4 — Polldubh, 5 — Poulmagree, 6 — Coolagh River 7 — Pol-an-lonain, 8 — St. Catherine's I, 9 — Doolin Cave (Fisherstreet Pothole), 10 — Pollnua, 11 — Pollbinn, 12 — Pollnagollum, 13 — Poll Elva, 14 — Pollcragreagh, 15 — Killeany Rising, 16 — Owenterbolea, 17 — Upper St. Brendan's Rising, 18 — St. Brendan's Well, 19 — Cullaun I, 20 — Cullaun II, 21 — Cullaun III, 22 — Cullaun IV, 23 — Pollcahermaan, 24 — Gragan West Cave

(From *British Caving*: Ed. by C. D. Cullingford, 1962, and prepared by the Univ. of Bristol Speleological Society)

Several karst phases can be recognised in the region surrounding Morecambe Bay, in N. W. England. This area consists of many low limestone hills (from 100—300 metres high). The slopes of these hills are largely covered with limestone screes at an angle of 30—35° which are banked up against steeper cliffs of limestone. Scree formation is not at all active in this part of Britain today [24], and it must be supposed that these large formations of scree occurred during colder phases of the Pleistocene. The steeper slopes which occur underneath the screes are smoothed and potholed by water erosion and solution; at the foot of the steeper cliffs there are occasional foot caves or *fusshöhle*. These cliffs have clearly been subject to quite a different climate from that which formed the limestone scree slopes. This climate may have been a warmer phase during the great Interglacial in the Pleistocene or, more likely, a warmer climate during the Tertiary period³.

³ Ashmead notes also the development of horizontal caves in this area, which may be related to Pleistocene sea-levels. Ashmead, P., The origin and development of caves in the Morecambe Bay area, *Trans. Cave Research Group*, 11 (1969), 3, 201—208.

In the Peak District of Derbyshire there are also different phases of karstification. In general the younger karst landforms occur around the edge of the limestone, near its contact with the overlying rocks, and the older karst landforms occur near the centre of the area where the limestone was exposed earlier. Thus the landforms in the peripheral areas are generally glacial or post-glacial in age, and result from the recent stripping of the cover rocks from the limestones. The older karst landforms towards the centre of the area include isolated conical hills, called "lows", and these are particularly important along the crest of the limestone area. They resemble the Kuppen Alb in southern Germany. This central part of the Peak District also possesses large pits or hollows in the limestones now filled with sand [26, 27]. Such pits are associated with collapse structures in the limestones and appear to be of a karst origin. These landforms of the central Peak District are probably part of an ancient phase of karstification, formed during the later part of Tertiary period. As in north west England, many of the limestone slopes in the Peak District are made up of scree at an angle of about 30° formed during periods of periglaciation in the Quaternary, and which modified the Tertiary relief. No true tropical fossil karst landforms have yet been described in Britain, such as have been described in Poland [28].

Because limestones are such sensitive indicators of the varying intensities of solution rates, studies of limestone landforms form an important facet of any work on rates of erosion both at the present time and in palaeoclimatology. Thus, with improved chemical techniques and improved diagnosis of the origin of the landforms, the studies of karst landforms in Britain in the next few years should make great strides. Though this paper does not mention all the present work in karst morphology in Britain, it is hoped that some idea of the varied activities has been given.

School of Geography, Oxford University

BIBLIOGRAPHY

- [1] Douglas, I., Field methods for hardness determinations of cave and river waters, Cave Research Group, *Newsletter*, 88 (1963), 3—6.
- [2] Douglas, I., *Field methods for water hardness determinations*, British Geomorphological Research Group Technical Publication, To be published.
- [3] Sweeting, M. M. et al., Denudation in limestone regions: a symposium, *Geographical Journal*, 131 (1965), 34—56.
- [4] Smith, D. I. and Mead, D. G., The solution of limestone, with special reference to Mendip, *Proc. Univ. Bristol Spelaeo. Soc.*, 9 (1962), 188—211.
- [5] Ford, D. C. Calcium carbonate solution in some Central Mendip caves, Somerset. *Proc. Univ. Bristol Spelaeo. Soc.*, 11 (1966), 46—53.
- [6] Pitty, A. F., An approach to the study of karst water, *Univ. of Hull Occasional Papers in Geography*, 5 (1966).

- [7] Gams, I., Faktorji in dinamika korozije na karbonatnih kameninah slovenskega dinarskega in alpskega krasa, *Geog. Vestnik*, 38 (1966), 11—68.
- [8] Folk, R. L., Spectral subdivision of limestone types, classification of carbonate rocks: a symposium, *Amer. Assoc. Petrol. Geologists, Mem.*, 1 (1962), 62—85.
- [9] Smith, D. I., Paper presented to the 4th International Congress of Speleology, Ljubljana, (1965). *Proceedings*, Vol. 3 (1968), 213—218.
- [10] Douglas, I., Calcium and magnesium in karst waters, *Helictite*, 3 (1965) (2), 23—37.
- [11] Sweeting, M. M., Some variations in the types of limestones and their relations to cave formation, Paper presented to the 4th International Congress of Speleology, Ljubljana, (1965), *Proceedings*, Vol. 3 (1968), 227—232.
- [12] Hanna, F. K., A technique for measuring the rate of erosion of cave passages, *Proc. Univ. Bristol Spelaeo. Soc.*, 11 (1966), 83—86.
- [13] Jakucs, L., Neue Methoden der Höhlenforschung in Ungarn, *Die Höhle*, 1959 Part 4.
- [14] Ashton, K., The analysis of flow data from karst drainage systems, *Trans. Cave Research Group*, 7 (1966), 161—203.
- [15] Doughty, P. S., Joint densities and their relation to lithology in the Great Scar Limestone, *Proc. Yorks. Geol. Soc.*, (1967—8), 479—512.
- [16] Sweeting, M. M., The weathering of limestones, in *Essays in Geomorphology*, Ed. by G. H. Dury, Heinemann, London, 1966, 177—210.
- [17] Williams, P. W., Limestone pavements with special reference to Western Ireland, *Trans. Inst. British Geographers*, 40 (1966), 155—172.
- [18] Sweeting, M. M. and Sweeting, G. S., *Lithological variations and their effects on limestone landforms*, Unpublished paper presented to Section E of the British Association in Leeds, 1967.
- [19] Williams, P. W., An initial estimate of the speed of limestone solution in County Clare, *Irish Geogr.*, 4 (1963), 432—441.
- [20] Gams, I., *Types of accelerated corrosion, Problems of the Speleological Research*, Prague, 1965.
- [21] Sutcliffe, A. J., Planning England's first Cave Studies Centre, *Studies in Speleology*, I (1965), 106—124.
- [22] Zötl, J., Die Hydrographie des nordostalpinen Karstes, *Steierische Beiträge zur Hydrogeologie*, Graz, 1960/1961.
- [23] Ford, D. C., The origin of limestone caverns: a model from the Central Mendip Hills, England, *Bull. Nat. Speleo. Soc. America*, 27 (1965), 109—132.
- [24] Sweeting, M. M., *Karst features of the Ingleborough-Settle area*, Paper presented to the Karst Symposium, Settle, International Geographical Union, 20th Congress, 1964.
- [25] Bögli, A., Le Schichttreppen Karst, *Rev. Belge Geogr.*, 88 (1964), 63—82.
- [26] York, C., *The pocket deposits of Derbyshire*, Birkenhead, 1961.
- [27] Ford, T. D. and King, R. J., The Golconda caverns, Brassington, Derbyshire. *Trans. Cave Res. Group*, 7 (1966), 91—114.
- [28] Gilewska, S., Fossil karst in Poland, *Erdkunde*, 18 (1964), 124—135.

SATELLITES IN GEOGRAPHICAL RESEARCH

ERIC C. BARRETT

The impact of two major post-war technological revolutions upon the scope and practice of Geography is growing daily. First, the computer revolution, which has already permitted existing data to be analysed in more detail than before. And second, the development of an advancing Space technology which has made, and will make, available to the student of the Earth and its atmosphere complete global sets of data unobtainable by other means.

In a real sense the second revolution depends upon the first. Computers of various kinds provide the only efficient means whereby spacecraft can be located accurately and tracked continuously, and computers alone have the necessary capacity to store and process the incredibly numerous observations that satellites can make and report each day. It has been estimated that a single weather satellite transmits of the order of 10^9 data points during an average day's flight, and the handling and presentation of data as numerous as these clearly must be geared to the more important uses to which they will be put. At present the materials are being processed for use in pre-existing operational programmes, e.g. cartography and weather forecasting; this by no means precludes the possibility of their use in original research. Indeed, the problem will be to find adequate research workers, rather than important new fields of study to be explored.

The types of data acquired from satellites depend on the design characteristics of those satellites, some of which fulfil a number of geographical purposes although they were not designed with the geographer's requirements specifically in mind. The next decade will be one of great opportunity for the geographer to capitalise upon the current state of Space exploration. Many remote-sensing devices are being developed for the ultimate purpose of the investigation of surface features on other planets, and most will be tested firstly with the Earth as their object of enquiry. Meanwhile, in the U.S.A. and U.S.S.R., families of purpose-built geographical satellites are being designed to study the Earth and its atmosphere by means of camera and radiometric equipment. Aerial photography from aircraft altitudes has wide acceptance as a useful medium for studying visible features on the Earth across a certain range of scale. Photography from spacecraft is potentially more useful since it can range across a wider range of scale, and photographic satellites, once in flight, provide a repetitive photo-coverage for no more than the cost of operating appropriate tracking and data-receiving facilities.

Likewise, spacecraft provide a superior coverage by many other remote environmental sensors designed to measure radiation in other wavebands. American design teams are seeking suggestions for the classes of data that will be the most highly valuable in geography. Opportunities to influence satellite design should not be neglected, nor the chance to catalogue the environment of man — and even man himself — from the Space standpoint may be lost.

Three types of satellites have proved their worth especially in the wide context of geography, and each will be outlined in turn. The data they have provided lend themselves to profitable examination by geographers, interested as ever in distribution studies, and now aided by expanding computer facilities and developing statistical techniques. In this paper the greatest stress will be placed upon American Space projects, since their details and data have been the most readily available to the author; Russian work has been along broadly parallel lines. The American weather satellites will be discussed in more detail than manned spacecraft and geodetic satellites since the author is currently engaged in studies of climatology and meteorology based on weather satellite data.

Finally, some details of the proposals and prospects for the forthcoming purpose-built geographic satellites will be discussed.

GEODETIC SATELLITES

Bearing in mind the traditional relationships between geography and the compilation of maps, it is appropriate that priority of place should be given here to satellite geodesy. The objectives may be summarised as follows:

- 1) To investigate more thoroughly than hitherto the size and shape of the planet Earth;
- 2) To establish accurately the relative positions of land masses, and control new maps series which are completely compatible with each other. These objectives are dependent on the success of the first, since they can only be achieved when the Earth is represented by a precise universal mathematical model.
- 3) To map in detail the global pattern of gravitational anomalies, and thus provide the necessary quantitative results to substantiate geophysical theories of the solid Earth. (Fig. 1).

Satellites are used as tools in geodetic research in two ways. In the first, the precise position of the satellite is determined by measuring the angle it makes in the sky with two or more tracking stations of known position. In the second, the altitude of the satellite is compared with its expected altitude assuming that the Earth is a perfect, homogeneous sphere. Perturbations of the satellite orbits are produced either by variations in the average density of the terrain below, or by departures of the Earth's surface from the form of a perfect sphere.

The satellites themselves may be "active", "passive", or "co-operative".



Fig. 1. Gravitational anomalies revealed by perturbations in a satellite orbit

The four arrowed regions have excessively high gravitational pulls

(From O. W. Williams, *Surveying the Earth by satellite*, *Science Journal* (1967), 7)

The first, including ANNA and GEOS satellites, emit signals using an on-board power source; the second, including the balloon-type satellites ECHO and PAGEOS (Passive Geodetic Satellite) must be illuminated by the sun and photographed against a stellar background: hence a need for clear skies; and the third, including the SECOR family, carry reflectors to co-operate with ground-based emissions such as radar and laser beams.

Eight years' results are described in detail elsewhere (e.g. Dulberger 1965), and only a very brief summary can be included in the present paper. Relating to the objectives outlined above, significant results include the following: 1) In crude, non-mathematical terms, the Earth has been described as pear-shaped, with an elliptical equator, and a number of irregularly-distributed depressions and protuberances. 2) Current measurement techniques can relate and can map positions on the Earth's surface to an accuracy of 10–30 metres within distances of 1000–3000 kilometres. High precision maps are facilitating new techniques

in communications and navigation. 3) Gravitational anomalies have been satisfactorily mapped across the whole world.

Geographers stand to benefit from these developments not only through the provision of more accurate maps, but also from the new geographical data. For example, questions relating to the distributions of land and sea should be answered more adequately. Long-standing problems of continental drift and the nature of large-scale sub-crustal convection currents can also be approached in these new ways.

MERCURY AND GEMINI SPACECRAFT

Manned spacecraft of the American Project Mercury and Gemini programmes have provided useful, high-quality colour photographs of many tropical and sub-tropical regions. Whilst radio and television methods of photography are necessary in the more permanent weather satellites, the temporary recoverable manned spacecraft permit the return to Earth of colour roll film. Black and white photographs are quite acceptable for meteorological purposes, but for the even more complex purposes of terrestrial studies the use of colour is highly desirable. Colour has already proved its added worth in studies of vegetation associations in ocean water masses, and many other terrestrial phenomena.

Alexander (1964) pointed out the fact that Space photography achieves, with a single integration of the radiation reflected from the Earth, a degree of generalisation only achieved before through considerable cartographic labour. Small-scale maps need no longer be the result of laborious survey on the ground, requiring months or even years to complete. To this virtue must be added the ability of satellites to make repeated observations of the same area, at any pre-selected time interval from once each day to every few minutes. Dynamic considerations can, and will be, injected into many geographic studies which have had to be relatively static in the past. At last there is the opportunity for geography to become a synoptic science in the fullest sense of that term — involving relationships over wide areas and deduced from simultaneous observations.

Individual manned spacecraft photographs each show areas several thousand square kilometres in extent, with a resolution of about one kilometre. (See Plate 1). Technically, it is a simple matter to vary the optics of the cameras to photograph objects down to the sizes of individual fields, streams, buildings, surface transportation facilities, and even man himself. The one important pre-requisite for Earth surface photography is an absence of cloud. Repetitive photography over a season or a year would lessen this difficulty, but manned spaceflights of these lengths are still in the future. The problem of an opaque atmosphere could also be reduced by the use of sensors in wavebands other than that visible light: these possibilities are enlarged upon in the final section of this paper.



Plate 1. A Gemini-IV manned spacecraft photograph, taken on 4th June 1965, over the Arabian Peninsula. The Hadhramaut Plateau is dissected by the dendritic tributaries of the Hadhramaut Wadi. The Gulf of Aden is towards the top of the photograph

(Courtesy of N.A.S.A.)

WEATHER SATELLITES

Whilst cloudiness is undesirable during photographic investigations of distributions on the Earth's surface, it is inherently valuable in atmospheric studies. In the ten years that have elapsed since the first specialised weather satellite TIROS I was launched on April 1st 1960, operational and research usage of weather satellite photographs has increasingly underlined the fact that clouds are effective tracers of atmospheric motion within the layer of the atmosphere through which they extend. The various families of weather satellites have all been described in adequate technical detail in various publications (Widger 1966, Barrett 1967), so that only brief notes need be given to that end here.

1) TIROS ("Television and Infra-Red Observation Satellites")

These included 8 of the Space-orientated type, 4 in orbits of 48° to the equator, and 4 at 58° . TIROS IX and X were of the "cartwheel" type, viewing the Earth, by means of cameras and infra-red radiometers, through the side-walls and not through the base plate as in TIROS I—VIII. The average orbital altitude was about 800 km. above the Earth.

2) NIMBUS

These more advanced satellites were flown in 1964, 1966 and 1969. They were Earth-orientated in orbits of 80° to the equator, and at altitudes about 1200 km. The resolution of the NIMBUS infra-red radiometer was much higher than those of the TIROS radiometers, and, at about 3—5 km., compared very favourably with the normal resolution achieved by the television photographic systems.

TIROS VIII and the NIMBUS satellites have been equipped with systems for Automatic Picture Transmission (A.P.T.), permitting simple tracking stations to acquire radio pictures direct from the satellites when they are within radio range. The pictures, of the regions local to these tracking stations, are particularly useful in short-term forecast preparation, and have been used for this purpose for over 2 years by the U.K. Meteorological Office. The Polish Meteorological Office is currently evaluating their usefulness for the same purpose. The other 9 TIROS, and the 3 NIMBUS satellites, all had storage facilities for T.V. pictures, so that swaths, covering the entire sub-orbital track across the daylight hemisphere of the Earth, could be displayed in sequence on a T.V. receiver as the satellites returned to within radio range of the more complex Data Acquisition stations. The T.V. pictures are made permanent by 35 mm photography of the receiver screen. (See Plate 2).

3) ESSA ("Environmental Survey Satellites")

The ESSA weather satellites are operational cartwheel-type TIROS. Two are kept in commission at any one time, one of the T.V. variety for global use, the other an A.P.T. satellite to aid regional forecasters throughout the world. The orbits are near-polar.

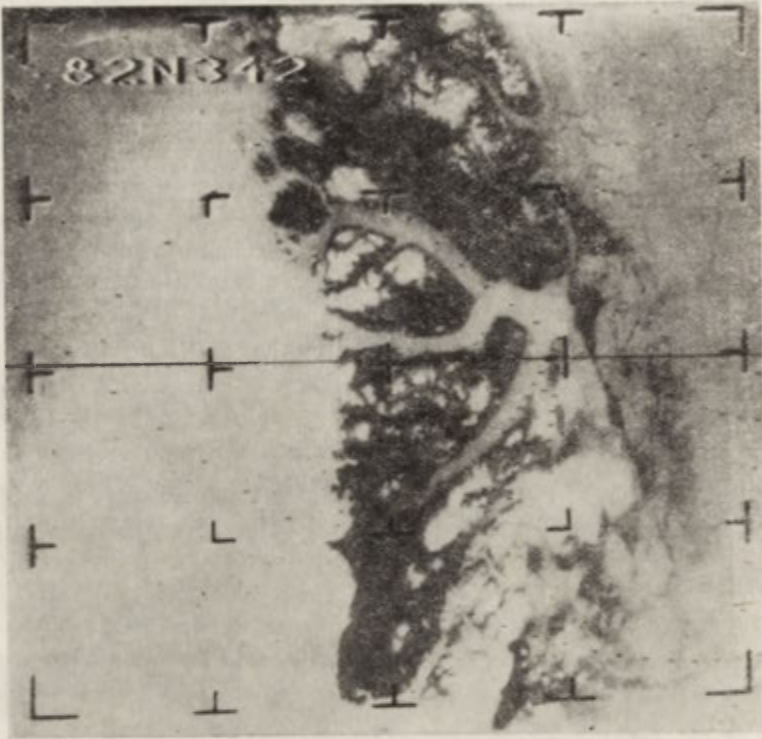


Plate 2. Fjords and glaciers in northern Greenland, seen by NIMBUS I, 31st August, 1964. The grid co-ordinates of the arrow tip are 82°N , 18°W . The continental ice-sheet is to the west, and relatively snow-free coastal mountains in the centre, across which drain Independence, Hagen, and Danmark Fjords to the frozen Arctic Ocean in the east

(Courtesy of N.A.S.A.)

4) COSMOS

Although most of the long series of Russian COSMOS satellites have been designed for very brief "lives" of 1—7 days, to investigate, by their re-entry drag, aspects of the structure of the upper atmosphere, a few have been akin to the American NIMBUS family. COSMOS 122 and 144, flown in 1966 and 1967, investigated weather through similar photographic and infra-red means. Their orbital angle, 65° to the equator, gave specially good coverage of the U.S.S.R.

5) A.T.S. ("Application Technology Satellites")

The first of this new family of multi-purpose satellites was injected into a geo-synchronous orbit 36,000 km. above the North Pacific during December 1966. Although its active life was short, it demonstrated the usefulness of cloud photography of the same area at 20 minute intervals throughout the day. Suomi (1967) has sequenced A.T.S. photographs on daily and monthly bases: the results are exciting time-lapse films of cloud changes over almost one-third of

the Earth's surface. These will aid research into the general circulation of the atmosphere, and a wide variety of other atmospheric and terrestrial studies may be based on data from hyper-altitude satellites.

Most research studies founded on specialised weather satellite data have, of course, been of atmospheric phenomena. It is the conviction of the author that climatology, the atmospheric science that is, in practice and purpose, most definitely the concern of geography, has yet to benefit as fully as it ought from these satellite developments. Meteorological analyses have been very numerous; climatological analyses few and far between. Satellites contribute significantly to climatological data pools through the global breadth and uniformity of their observations. It has been stated recently by the World Meteorological Organisation that only 10% of the Earth's surface is adequately covered, for synoptic weather study, by conventional surface recording stations. Satellite data fill the gaps, bringing into view the previously inhospitable and data-sparse regions, for example in tropical, polar and oceanic areas of the world. Many climatic problems, from that relating to the global-scale classification of climates, down to those on the synoptic and meso-scale can be approached through Space media.

A brief review of the possibilities can be presented in three sections: —

1. *The usefulness of weather satellite photography in atmospheric science*

The usefulness is limited in part by photo-resolution, and in part by the frequency of photography. In the case of the earlier Space-orientated TIROS, the increasing obliqueness of view away from the one point on each orbit at which the satellite observed the Earth from a vertical angle, was a significant factor limiting the usefulness of the photographs. Cartwheel and Earth-orientated satellites do not suffer from this restriction. The most significant recent advance has been the commencement of a routine daily preparation of hemispheric montages from ESSA T.V. photographs, with applications to general circulation studies. (See Plates 3 and 4).

Whilst the best resolution was about 0.3 km., achieved by narrow-angle cameras in TIROS I and II, most satellite photographs have resolved objects down to about 2.8 km. Hence, the finer details of a cloud field cannot be detected, but meso- and macro-characteristics are well displayed. Thus the photographs are *complementary* to conventionally-made observations of the state of the sky, which are by no means rendered redundant. Perhaps the most serious restriction still affecting the use of cloudiness as a tracer of atmospheric motion exists in those rather common situations where different, largely discrete, layers of cloud overlay one another. Wind shear with altitude is normal rather than exceptional through the atmosphere, and different cloud layers are often aligned in correspondingly different directions. These alignments are often difficult, if not impossible, to separate in photo analysis, although where they can be separated satisfactorily, the three-dimensional structure of wind and pressure fields can be deduced usefully from the clouds.

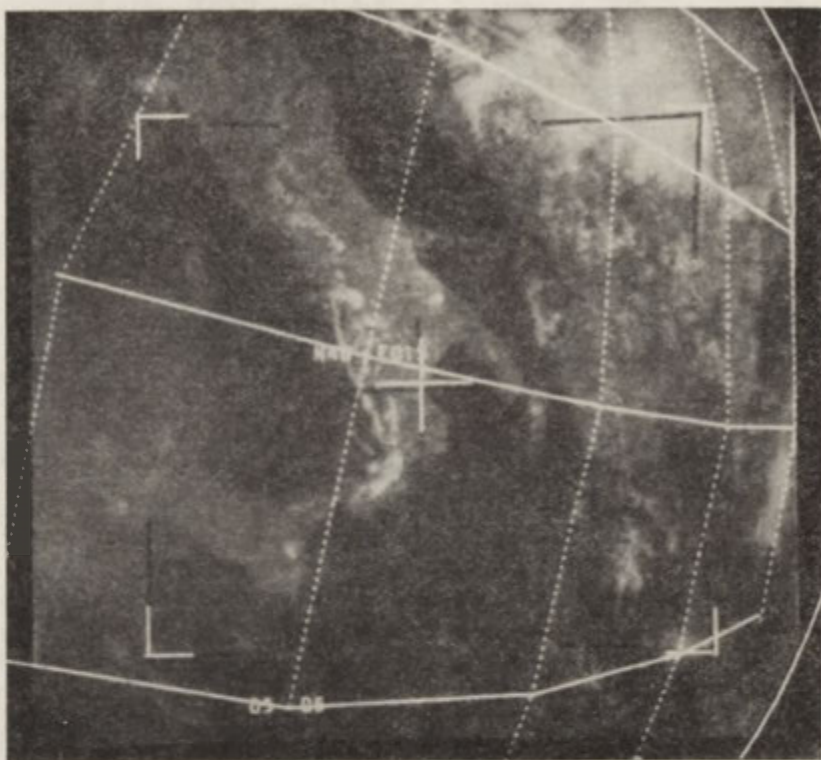


Plate 3. An individual television photograph received from ESSA 1, 28th June 1966. The grid-lines were computer drawn. The area portrayed is the central Mediterranean, with Italy in the centre of the picture

(Courtesy of E.S.S.A.)

The patterns appearing in the cloud fields may be related to a number of key factors, including: —

a) The predominance of stable or unstable conditions.

Relatively cloud-free skies indicate subsidence under anticyclonic conditions, or very dry continental air, whilst deep cloudiness, whether disorganised, or organised over a wide area, indicates instability. Cloudiness building upwards from near the Earth's surface relates closely to the depth of instability.

b) The relative importance of vertical or horizontal motion in the troposphere.

Where horizontal motion is small, cloudiness tends to organise in mottled, cellular patterns; where it strengthens, and takes precedence over vertical motion, the cloud cells become elongated, or even blown into cloud lines. This factor applies best within air streams, especially in ocean anticyclonic cloud fields, and in cold sector extra-tropical cyclone cumuliform cloudiness. It can be applied also to air flow along jet stream axes in the upper troposphere.



Plate 4. An ESSA 5 computer-compiled photo-mosaic, showing the entire northern hemisphere on 8th June 1967. The individual photographs have been rectified to conform to a polar stereographic projection. The major weather belts are readily identifiable

(Courtesy of E.S.S.A.)

c) Contact between air streams contrasting in their temperature and/or humidity characteristics.

Strong bands of cloud are frequently generated along these zones of contact and mixing. The bands tend to be normal to the direction of air flow where one air mass is displacing another, or between converging air-streams, e.g. in the equatorial zone. These are the cases where cloudiness is most frequently layered in a complex manner, with sequences of cloud strata overlapping in the vertical along characteristically inclined air-mass interfaces.

d) Variations in the direction and speed of air flow at different levels in the troposphere.

These can be discerned frequently provided that the clouds at different levels can be distinguished. The most critical problems have been mentioned already. On the debit side, upper cloud is generally thin and fibrous compared with lower cloud, which, whether continuous or broken, tends to be opaque. Hence, an overlay of high cloud does not often hide clouds at lower levels or prevent their reasonable interpretation; the problem is most acute where middle and low clouds are concerned.

e) Air-mass and air-stream modifications induced by the underlying topography.

A wide range of interesting and informative cloud patterns may be related to land and sea surface patterns of topography and temperature. Lee-wave clouds, eddy, crest and cap clouds, and cloud rifts by ocean currents, are some examples.

These relationships, although essentially meteorological in nature, are all significant in modern climatology, which is becoming increasingly dynamic, decreasingly purely descriptive. The application of geographical analytic techniques can only result in a fuller understanding of weather and climate.

From a consideration of the key factors outlined above, it is possible to move towards a classification of cloudiness portrayed by satellite photographs which involves not only the physical appearance of clouds, but also the mechanism involved in their development. The first problem which faced early satellite photoanalysts was the identification of cloud types from the new viewpoint, as far as possible employing the pre-existing schemes of nomenclature. The definitive work here was done by Conover (1962), who listed six characteristics of clouds that could be used in cloud recognition. These were *cloud brightness*, relating to the depth and composition of clouds, the angle of the sun's illumination, and various technical factors in the satellite and on the ground; then the *texture* of the clouds, whether smooth or fibrous (cirriform), smoothly opaque (low stratus and fog), mottled or irregular (cumuliform and strato-cumuliform), and smooth but ragged, with various interdigitating degrees of brightness (stratiform), or various hybrid forms; *vertical structure* can sometimes be deduced from cloud shadows and/or from layers of different textures; *the form of cloud elements*, whether regular or irregular; the *patterns* of those elements, relating to such things as air flow, vertical wind shear and underlying topography; and finally, the *sizes* of the elements and patterns. Conover recognised a very wide variety of cloud types under reasonable photographic conditions, and there is definitely a need for an essentially generic classification of this kind.

Since the emphasis in climatological circles is now swinging towards the dynamic aspects of climate this is an opportune moment to introduce the outline of a cloud classification involving *atmospheric genetics*. In dynamic climatological studies, important criteria include the prevalence of different pressure systems in selected areas, and preferred regions of cyclonic and anticyclonic growth, move-

ment, and decline. Weather satellite photographs have much to offer here. Various atmospheric systems have been examined through this medium in isolation rather than in their regional contexts. The geographer, in approaching the same photographs may need to interpret a much wider range of cloud patterns. Figures 2 and 3 show the author's initial attempt to devise basic outlines for a genetic classification of meso- and macro-scale arrangements of cloud. The primary division is on the basis of cloud related to undisturbed atmospheric flow (Fig. 2) and that perturbed by abrupt irregularities on the Earth's surface (Fig. 3). Later, electronic computers may be set to search for stereotypes of these and similar kinds. The ability to recognise a broad selection of atmospheric systems, made visible by clouds, is essential to the New Climatologist.

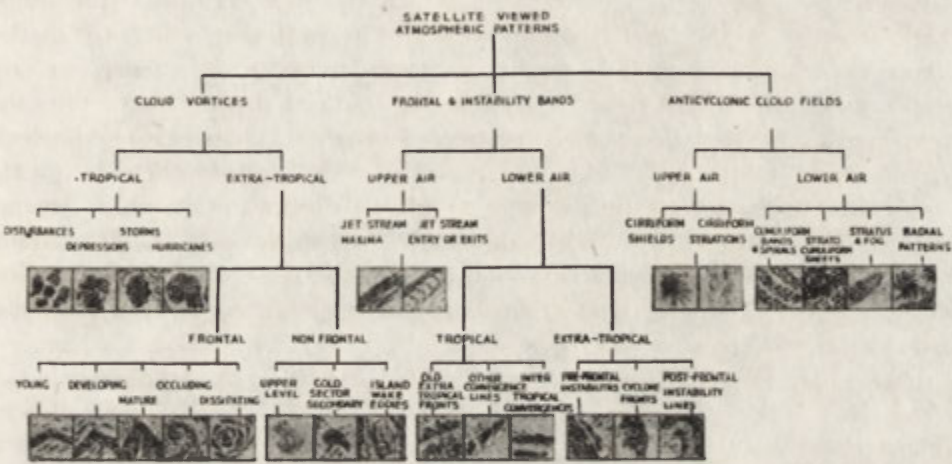


Fig. 2. A basis for the genetic classification of satellite-viewed cloudiness
This scheme is capable of elaborate development in the light of present knowledge

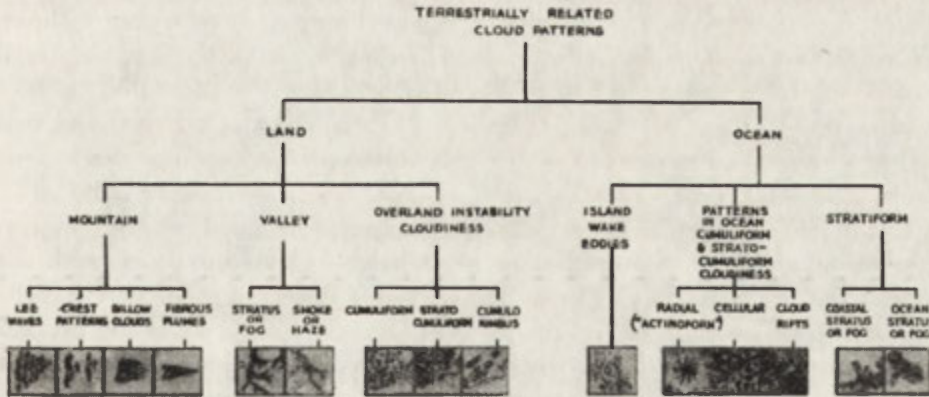


Fig. 3. A basis for the genetic classification of satellite-viewed cloudiness which is related to the underlying terrain

2. *Nephanalyses and the climatology of cloud cover*

Satellite cloud photographs, on receipt at the U.S. Weather Bureau, are analysed routinely by experts who compile cloud charts or *nephanalyses* showing four categories of cloudiness: —

- | | |
|---------------------------|--------------------------|
| i) OPEN (O) | — below 20% cloud cover; |
| ii) MOSTLY OPEN (MOP) | — 20—50% cloud cover; |
| iii) MOSTLY COVERED (MCO) | — 50—80% cloud cover; |
| iv) COVERED (C) | — over 80% cloud cover. |

The nephanalyses are distributed to national forecasting centres throughout the world via the meteorological radio facsimile network, and they have a number of uses in short-term weather forecasting operations.

Whilst the development of weather satellites has been significant for atmospheric scientists in the 1960's, so too has been the continuing search for mathematical models to describe and predict various patterns in the atmosphere, and temperature studies have figured prominently amongst them. One of the most promising thermo-dynamic models has been devised by Adem (1964) who seeks to specify and predict mean seasonal temperatures in the atmosphere and at the earth's surface, employing the amount of cloud cover as the only independently variable parameter. Cloud cover was difficult to assess in the pre-satellite era, largely because of the deficiencies of the conventional meteorological recording network. It is this kind of context that nephanalyses can be of most assistance in distributional research.

Each nephanalysis represents a professional summary of cloud distribution as seen by a weather satellite. Clapp (1964) took all the available nephanalyses for a twelve month period and, by averaging the percentage of cloud cover at 5° intersection of latitude and longitude for months and seasons, prepared acceptable cloud cover maps for much of the world. Since the advent of cartwheel TIROS and ESSA satellites in near-polar orbits, both the estimation of cloud cover and the total global photo-coverage have improved. Figs 4 and 5 illustrate the distribution of individual data obtained by the present author from nephanalyses covering the area from 60°W of Greenwich to 60°E, from and 30° to 60°N, for the months of January and July 1966. The frequency of observations, better for July than January, approaches the immediate optimum of one per day over parts of the subject area. Today the frequency closely approaches 100%.

Figs 6 and 7 show the resulting average patterns of cloudiness for those two high-season months. Comparisons of nephanalysis cloud categories with oktas of cloud cover from contemporary conventional observations over the British Isles show good agreement, and the maps are therefore superior to any pre-satellite attempts to study cloud cover over wide areas and long periods of time. Furthermore, the maps can be applied to the new mathematical predictive models, and also to long-range forecasting whilst this depends upon analogue methods like those currently employed by the British Meteorological Office.

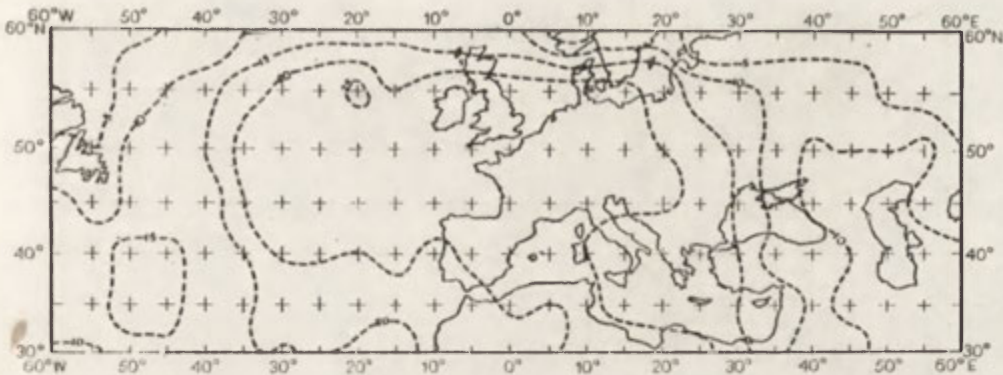


Fig. 4. Distribution of individual data from nephanalyses. January 1966 — number of observations

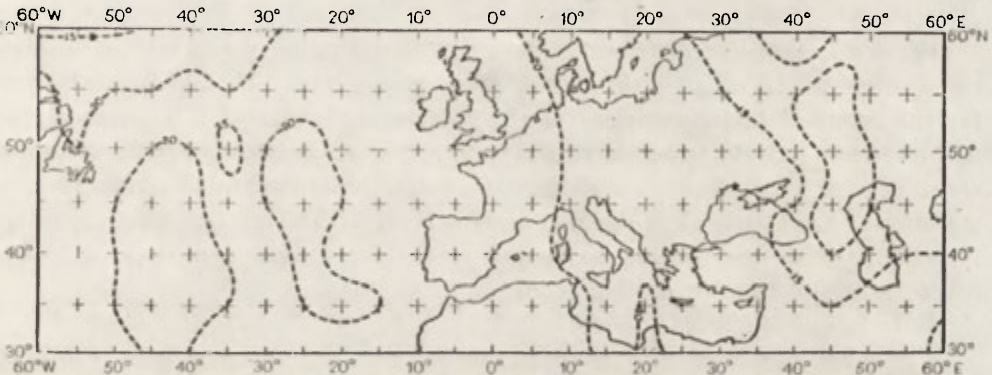


Fig. 5. Distribution of individual data from nephanalyses. July 1966 — number of observations

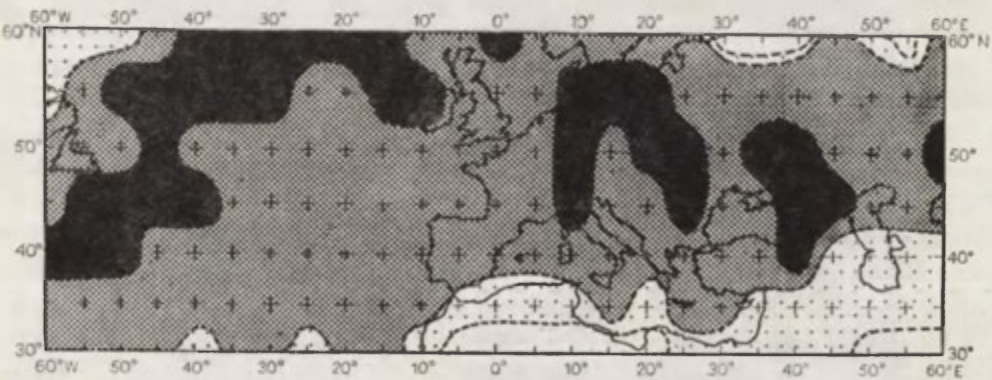


Fig. 6. The distribution of cloudiness in January 1966

Unshaded areas averaged less than 20% cloud cover; lightly stippled areas 20—50%; heavily stippled areas 50—80%; and dark grey shaded areas above 80% cloud cover

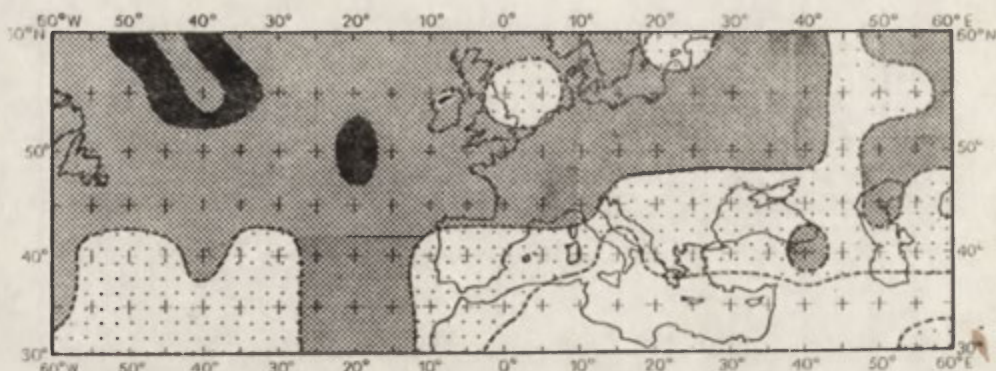


Fig. 7. The distribution of cloudiness in July 1966

Unshaded areas averaged less than 20% cloud cover; lightly stippled areas 20—50%; heavily stippled areas 50—80%; and dark grey shaded areas above 80% cloud cover

For 1967, the data distribution was even more uniform and complete, since single composite nephanalyses are now being issued by the U.S. Weather Bureau for the North Atlantic quadrant of the Northern Hemisphere. The researcher will no longer have to piece together several charts to obtain a partially complete cover of his chosen region each day, and, secondly and even more importantly, with fewer nephanalyses being issued, the pressure on the facsimile recorders at Bracknell is lessened: fewer nephanalyses will be lost because of the unavailability of machine time to receive them.

3. *Investigations in the infra-red*

These are highly significant in that the raw data received from satellites are temperature recordings. Satellite cloud photographs illustrate aspects of the structures of atmospheric systems; infra-red data quantify the energy involved in them. Table 1 summarises the wavebands investigated by TIROS and NIMBUS satellites. The Russian satellite COSMOS 122 recorded within the wavebands from 8—12 μ , 0.3—3 μ , and 3—30 μ , comparable with three of the TIROS medium-resolution channels.

Since the low, medium, and high-resolution radiometers resolved radiation patterns to within 750, 75, and 2.5—5 square kms. respectively, their applicability to atmospheric problems differ accordingly. Whilst the first has provided information relevant to general global heat budget problems, the second has been more valuable in the study of individual weather systems. For example, medium-resolution infra-red recordings have been used to construct temperature maps of hurricanes and the atmospheric fields in which they are embedded. One computer programme has averaged data within grid elements measuring 2.5° square at the equator, for all five of the medium-resolution channels employed in a TIROS satellite (Bandein et al. 1964). Small, intense disturbances of the hurricane variety stand out in stark relief on all the resulting maps: channels 3 and 5 both

TABLE 1. Infra-red wavebands investigated by weather satellites

Type of radiometer and satellite family	Channel where appropriate	Spectral range	Measurement
TIROS Low resolution radiometer	—	0.2—4 μ 7.0—35 μ	Short-wave solar radiation Long-wave Earth radiation
TIROS Medium resolution radiometer	1 (or) 1 2 3 4 5	6.0—6.5 μ 14.5—15.5 μ 8.0—12.0 μ 0.2—6.0 μ 7.0—30.0 μ 0.55—0.75 μ	Absorption by water vapour Absorption by CO ₂ Major “atmospheric window” Reflected solar radiation Long-wave Earth radiation Reflected visible solar radiation
NIMBUS High resolution radiometer	—	3.5—4.1 μ	Surface temperature of earth or cloud tops

show high reflectivity of light from the dense overcast of the vortices, and channels 1, 2 and 4 all indicate low temperatures in the same areas. Estimates of the energy involved in these and other storms can be made from these quantitative data. Procedures are being developed to estimate aspects of the vertical structure of the atmosphere from the same data, including the heights of cloud tops, and vertical temperature profiles up to the tropopause. Here again is an opportunity for the climatologist to approach his subject through factors which can more aptly lay claim to the description of “truly active” than any employed by Thornthwaite. Finally, the high-resolution recordings have permitted detailed studies to be carried out, including those of variations in land and sea-surface temperatures, and will be especially useful in night-time cloud studies. By feeding the data into a facsimile machine, images of night-time cloudiness have been obtained (See Plate 5). In these, the darkest areas are those of highest temperature, namely the oceans; the brightest areas are coldest, namely high cloud tops and snow-covered mountains; the intermediate areas, appearing in shades of grey, include other land areas. Diurnal variations in cloud cover can be examined by comparisons of satellite photographs and the satellite infra-red recordings of sufficiently high resolution.

Thus the energy observations designed to aid the atmospheric physicist in his work again provide material that the climatologist should not neglect.

TERRESTRIAL GEOGRAPHY FROM SATELLITES

In this final section consideration will be given to two separate, yet related, questions. First, the distributions on the Earth’s surface that have been clarified by weather satellite data, and second, the proposals that have been made for



Plate 5. This is a portion of a film strip compiled from NIMBUS I High Resolution Infra-red Radiometer data, 18th September 1964. The high temperatures in the eye of the tropical cyclone show clearly as a dark spot in the centre of the lower temperatures of the cloud tops of the cyclone vortex. The storm is centred near 30°N , 15°W

(Courtesy of N.A.S.A.)

future geographical satellite sensors to improve and extend the range of data of value in geographical research.

1. *Surface patterns from weather satellite data*

The most significant from the applied, or operational viewpoint, have been the photographic infra-red studies of ice and snow, particularly around the North American continent. Research has shown that ice and snow can usually be recognised as such where clouds are well broken, and in mountainous regions and along steep, indented coasts. Across the seasons, the waxing and waning of polar and sub-polar areas of ice and snow can be traced. On a more local scale, both television and A.P.T. pictures can be used as reconnaissance material,

indicating such features as leads or breaks in sea-ice. These features can then be investigated more closely by aircraft, whose services are less widely and more efficiently employed than before. High-resolution infra-red data from NIMBUS satellites have been analysed to show that they permit the recognition of variations in the age and state of ice and snow. Popham and Samuelson (1965) described clear contrasts between infra-red readings from continental shelf ice, continuous and discontinuous pack-ice, and areas of ice-floes. With only a little improvement in resolution, resource estimates of upland snow accumulation for potential hydro-electric usage will become feasible.

Very recent work by Warnecke et al. (1967) demonstrates the use of high-resolution infra-red data, from cloud-free areas, in mapping paths of ocean currents. Earlier studies based on similar data have led to the discovery of former lake beds and abandoned river channels, which are detectable in the infra-red by virtue of variations in moisture retention amongst sorted surface materials.

Many other topographical phenomena have been recognised on weather satellite photography, including vegetation associations, geological structures, soil types, and desert land-forms. Without exception, however, these can be studied better using colour photographs of the Mercury and Gemini types.

TABLE 2. Proposed experiments for remote sensing of the Earth

Applications	Agriculture and forestry	Geology	Hydrology	Oceanography	Geography
Photography	Soils; Veg. growth; Plant disease	Surface structure	Drainage patterns; Soil moisture	State of the sea; Turbidity	Cartography; Land use; Transport; Population; Topography
Infra-red	Terrain composition; Plant condition	Areas of thermal anomaly; Mineral distribution	Areas of cooling	Ocean currents; Sea ice	Land use
Radar imagery	Soil characteristics	Surface roughness; Tectonics	Soil moisture; Run-off slopes	State of the sea; Ice characteristics	Cartography; Geodesy; Vegetation
Microwave radiometry	Thermal state of terrain	Sub-surface layering	Snow and ice		Snow and ice
Remote geochemical testing		Mineral deposits; Trace elements		Surface flora	

2. *Proposals for geographical satellite sensors*

Table II (after Badgley 1966) summarises the lines along which geographical investigations from Space are most likely to move in the foreseeable future. The expected applications listed in Table 2 are very generalised; they could be made more specific by reference to chosen regions of the world; by study at any one of a number of different scales; by examining data acquired during different months or seasons of the year, and by the adoption of either a synthetic or analytical approach. The physical geographer will not monopolise geographical data from Space: current discussions are centring around studies of city structure and transportation from that viewpoint; in some areas it may be possible to conduct population census work by satellite. Very high resolution images are anticipated from the newly-designed EROS (Earth Resources Observation) satellite, combining laser and television techniques.

As in meteorology, ground observations will not be out-dated, but the field work of the future in geography is likely to include many new exercises designed to substantiate the satellite studies. Space data will not create a new kind of geography, but will supplement and reinforce other data sources, thereby greatly extending the scope of geographic observations. Many positive benefits are to be derived from an early application of satellite data to geographical research.

University of Bristol

BIBLIOGRAPHY

- [1] Adem, J., On the physical basis for numerical prediction of monthly and seasonal temperatures in the troposphere-Ocean-Continent System, *Monthly Weather Review*, 92, (1964), 91.
- [2] Alexander, R. H., Geographic data from Space, *Professional Geographer*, 16 (1964), 1.
- [3] Badgley, P. C., Planetary exploration from orbital altitudes, *Photogrammetric Engineering* 32 (1966), 250.
See also: *Spacecraft in geographical research*, Publication 1953, National Academy of Sciences, Washington D.C. 1966.
- [4] Bandeen, W. R. et al., TIROS III meteorological satellite radiation observations of a tropical hurricane, *Tellus*, 16 (1964), 481.
- [5] Barrett, E. C., *Viewing weather from Space*, Longmans, London, 1967.
- [6] Clapp, P. F., Global cloud cover for seasons using TIROS nephel analyses, *Monthly Weather Review*, 92 (1964), 495.
- [7] Conover, J. H., *Cloud interpretation from satellite altitudes*, Air Force Cambridge Research Laboratories, Research Note 81, 1962.
- [8] Dulberger, L. H., Geodetic measurements from Space, *Space Aeronautics*, 12 (1965), 32.
- [9] Popham, R. W. and Samuelson, R. E., Polar exploration with NIMBUS meteorological satellite, *Arctic*, 18 (1965), 246.
- [10] Suomi, V. E., *A.T.S.-B Scanning Camera Experiment*, paper presented to the Tenth Plenary Meeting of COSPAR, London 1967.
- [11] Warnecke, G., Allison, L. J., and Foshee, L. L., *Ocean surface temperature discontinuities as detected from NIMBUS II high resolution radiometric measurements*, Paper presented to the Tenth Plenary Meeting of COSPAR, London 1967.
- [12] Widger, W. K., *Meteorological satellites*, Rheinhard and Holt, New York 1966.

CONTENTS OF VOLUMES
GEOGRAPHIA POLONICA

Vol. 12. Formation et l'Aménagement du Réseau 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, \$ 5.50.

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., 41 Figures, 1968, \$ 3.00.

Vol. 14. Special issue for the 21st International Geographical Congress in New Delhi, 1968. 43 papers prepared by Polish geographers: 24 dealing with physical and 19 with economic and human geography. List of geographical institutions in Poland and Polish geographers, 406 pp., 80 Figures, 1968, \$ 9.15.

Vol. 15. Economic Regionalization and Numerical Methods. The volume contains the final report on the 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, \$ 5.60.

Vol. 16. 11 papers dealing with research problems and techniques in both economic and physical geography, 136 pp., 27 Figures, 1969, \$ 2.70.

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, \$ 8.30.

Subscription orders for the
GEOGRAPHIA POLONICA
should be placed with
FOREIGN TRADE ENTERPRISE
ARS POLONA

Warszawa,
Krakowskie Przedmieście 7,
Poland
Cables, Arspolona, Warszawa

lv