TYPOLOGY OF FARMS IN CENTRAL FINLAND

K. R i k k i n e n
Helsinki University
/Finland/

PURPOSE OF STUDY

Agriculture in any district is the result of complicated influences deriving from nature and humanity. Each district and each individual farm is inuque, and the difficulties of comparing regional agriculture at international level are therefore understandable. A worldwide classification of agriculture is an aim which meets with general approval, however, and the Commision on Agricultural Typology of the I.G.U. has been working to that end since 1964.

As the President of the Commission noted in 1964, "a number of quantitative methods elaborated by mathematicians and non-mathematicians are in use in various disciplines to measure similarities or average differences between various phenomena... None of them, however, have been checked in the typological investigations of agriculture." /Kostrowicki 1964, p. 166/. This challenge has since been accepted by several investigators. Some geographers have found factor analytic techniques a useful aid to problem solving in agricultural studies /Henshall 1966; Munton 1970; Munton and Norris 1969; Aitchison 1970; Momsen 1970/. In the latest studies results have been promising: "The factor analysis methodology provides a simple framework for an understanding of the internal variation between small farms within each territory." /Momsen 1970, p. 8./ and "principal component analysis has simplified and redefined the complex data matrices required to describe large numbers of farm systems, and it is possible to group the farms as a result ... Most important of all, by grouping the farm systems in terms in terms of their similarities of function, principal component analysis has provided a valuable basis for their future investigation." /Munton 1970, p. 11-13/.

The purpose of the present study is to examine the suitability of multivariate analysis for the typology of farms in central Finland. The complete study has been published in other journal[15/]. In this paper only the general features of the methods and study results are presented.

MATERIAL

The study material consists of two sample districts in central Finland, one /Kalmari/ containing 172 and the other /Häkkilä/ 146 farms, each with a field area of at least one hectare. The farming census afforded information on many variables illustrating basic features of agriculture. The agricultural census questionnaire contained 136 question groups, some of which included several sub-questions.

15/ Kalevi Rikkinen: Typology of farms in central Finland. Fennia 106. 44 p. Helsinki 1971.

TYPOLOGY OF FARMS ON THE BASIS OF INDIVIDUAL VARIABLES

Typologically the farms are divided into two main types with regard to separate variables. First there exists a continuum- type divison which appears with many central variables describing farms. To be regarded as a second main type is the dichotomous distribution which occurs, for instance, in variables describing the cultivation of most crops. Variations of these main types also naturally occur.

From the typological standpoint the continuum-type division is difficult. Dichotomous division, on the other hand, is in one sense easy, as there are then two distinct classes. On the other hand, the placing of farms in tow classes only may be too rough a division. In any case the drawing of class boundaries and the formation of groups will be highly subjective if individual variables alone are used for division. For this reason the multivariable method was used in the present study.

CORRELATION ANALYSIS

By calculation of correlation coefficients between different variables the formation of homogeneous groups may be attempted /Hagood 1943; Hagood and Price 1952 et. al./. If the interdependences of variables are discovered, a "linkage tree" of variables may be composed.

From the material available 44 variables were formed for correlation analysis. Variables were formed by including at least one variable from each section of the questionnaire which showed a characteristic as well as possible.

Correlation coefficients enable homogeneous groups to be composed with the aid of many different principles. Haggett /1966, p 283-286/ presents there basic concepts: 1/ basic pairs; 2/ p-clusters; 3/ F-groups. As a means of graphic illustration variables in correlation with each other at a different level of significance have been connected by different lines /e.g. op. cit. p. 284-285/.

In the present study correlation analysis was used in an attempt to disclose complete sets of characteristics by changing the order of variables in the correlation matrix in various ways. This was perhaps best achieved by grouping the variables according to the number of other variables with which they are statistically in highly significant positive correlation. The appended matrix /Fig. 1/ were made with reference to Kalmari farms in observance of this principle. First in order is the field area of farms, which is in very strong positive correlation with 25 other variables. Last on the list are variables not in very strong positive correlation with any other variable. Between these two groups is a third consisting of characteristics which are not in very strong statistical core relation with any other variables, or with very few. In this case, the groups of characteristics disclosed by correlation analysis are very inexact. The correlation matrix, however, forms an important basis for the study of farm typology by the multivariate method.

FACTOR ANALYSIS AND VARIMAX ROTATION

By means of correlation coefficients conclusions can be drawn as a rule only with regard to the interdependences of individual characteristics. A better notion of the connections between several different variables is afforded by factor analysis. This method enables variables to be grouped in collections relatively independent of each other /Harman 1960; Berry 1961; Steiner 1965/.

In the present study not all variables were chosen for factor analysis which had been subjected to correlation analysis, but on the basis of the latter the number of variables was restricted to 25. In the elimination process care was taken above all that variables in so-called technical correlation were not included. The following were selected for factor analysis:

1/ Field area
2/ Forest area
3/ Presence of otherwise of milking machine
4/ Cows, number
5/ Technical equipment
6/ Presence or otherwise of tractor
7/ Is barley cultivated?
8/ Is there a car?
9/ Is there pasturage?
10/ Are oats cultivated?
11/ Are there horses?
12/ Has the farmed received agricultural training?
13/ Employees, number
14/ Are potatoes or root crops cultivated?
15/ Is hay grown?

16/ Are there pigs?
17/ Is wheat cultivated?
18/ Is there a bull?
19/ Is there a successor in the ownership?
20/ Is rye cultivated?
21/ Working days of owner outside farm.
22/ Are husband and wife both living?
23/ Is the main profession other than farming?
24/ Age of farmer
25/ Is there poultry?

In the present study the correlation matrix was factorized by the principal axis method. It is a natural attribute of this method to include in the first factor as many as possible of the covariance of variables. However, the eigenvalues of the following factors still were high. This indicates that we are not concerned with a one-dimensional body of variables. Thus there was good reason for rotation.

The object of rotation is to remove general factors by reversing factor axes and to obtain the interrelations of variables in a simple, interpretable from. An aim of this kind is in conformity with the attempt to compose a typology of farms. In the present work Varimax rotation was used.

Rotation with 3-5 factors was tried here, and the four factor solution proved most successful. A clear interpretation for four factors was to be found, and the so-called simple structure requirement was realized in the solution. The factors can be interpreted as follows:

1/ This factor gives high loadings to variables indicating farm area, number of cows and technical standard of machinery and equipment. The factor thus indicates size and wealth.

2/ The factor gives fairly high loadings to many variables indicating grain and fodder crops, also to the variable indicating presence of horses. It may be called the factor of traditionalism.

3/ This factor gives high loadings to advanced age but continued ability to work their own land /of farmers/. It may be called the pensionary factor.

4/ This factor gives high loads to journeys to work outside the farm owned /also, fields are often not in fully effective use/. It may be called the work elsewhere factor.

These four factors may also be used as a basis of classification for individual farms.

FARM TYPOLOGY ON BASIS OF FACTOR SCORES

The proportion of individual farms to different factors was obtained by calculating factor scores for each farm. Factor scores were calculated by taking the average score for each factor as 500 and the deviation as 100. Thus the scores are standardized and the division approximates in theory to the normal, which facilitates further treatment.

On the basis of the above interpretation of factors the farms which receive high scores from the first two factors are full-time farms, whose owners gain their principal livelihood from agriculture. Those which receive high scores from the third and fourth factors are mainly part-time farms.

Two methods will now be presented which enable farms to be classified more precisely, taking factor scores as a starting point.

Standard deviation as basis for classification

In the appendeddiagram /Fig. 2/ the first two factor scores of Kalmari in order of rank are taken as examples. It will be seen that the factor scores in general change as continuum types. For this reason it is difficult to define the boundaries between different farm types.

One possibility is to give primary attention to the dominant, factor, i.e. the factor whose score on the farm is highest. Farms can then be classified by division into four groups. In practice, however, the greatest and second greatest factor scores may be almost equal. In such cases the dominant factor gives a poor notion of the farm.

A method of forming class boundaries is to use standard deviation as a criterion. This technique has been used, for instance, by Nelson /1955/ in classifying American cities on the basis of their occupational structure. According to Nelson's classification a city can be specialized in more than one factor and to varying degrees. If the average factor score is taken as 500 and the standard deviation as 100, farms which are one, two or more standard deviations away from the average are easy to assemble /cf. Fig. 2/.

Factor scores may deviate from the average both upwards and downwards. From the classification standpoint, what a farm contains is perhaps more significant than what it lacks. A s a first step in farm classification it may therefore be best to consider those whose standard deviation is one or more upwards. We now present a simple classification in which scores below 600 are marked with the symbol 0, and scores of 600 or over are marked +. In this way each farm receives a foursymbol index. For instance, the index 0 + 00 means that the farm's scores by factors I, III and IV are below 600, but by factor II above 600, i.e. at least one standard deviation away from the average.

Table 1.
Farm classification based on standard deviations uppwards from average, at Kalmari.

Type	Number of farms
0000	103
+000	23
0+00	7
00+0	8
000+	21
00++	7
+0+0	1
+00+	1
++0+	1
	172

Type 0000 is clearly the most common both at Kalmari and Häkkilä /Table 1/. It can naturally be divided into subtypes in accordance with the factor by which scores may have a standard deviation of one or more downwards /below 400/.

In a typology of farms obtained entirely in this manner there is emphasis on some special characteristic. In other words, farms belong to the same type because of features they share and which differentiate them from others. On the other hand, farms belonging to the same type on the strength of many other characteristics may differ markedly.

Grouping analysis

By the former method class boundaries were drawn to one standard deviation. This subjective method may be avoided by the use of grouping analysis. There are several grouping methods /Harvey 1969, p. 345-346/. The general principle is that groups should be formed in such a way that their within-groups variance is as small as possible. The present study employed the method evolved at the Computing Centre of Helsinki University /HYLPS/GA, version H/.

In analysis the number of groups desired must first be estimated and starting values shosen. Grouping of observations is then tested. An observation is considered to belog to the group in which it differs least from the group average. In the solution which is mathematically best the total withingroups distance /D/ is smallest. Naturally, the higher the grade of homogeneity demanded within the group, the greater the number of groups which must be chosen.

In grouping, therefore, the observation values chosen as starting values in group formation are a matter of central importance. There are many possible combinations, and the mathematically best solution is not necessarily best in a typological sense. For this reason several starting values and group numbers were experimented with in the present study.

Factor scores for farms by four different factors are taken as a starting point for grouping in this study. No factor scores are weighted. Here is the essential difference between this and the standard deviation method shown earlier, which laid stress on special differentiating characteristics.

Grouping was performed in 4-7 groups, and three different starting values were used used for each group number. As a subject for closer examination we shall now take a grouping of farms in four groups only.

From the typologycal standpoint it is essential to compare how different grouping cause farms to be placed in different groups. Farms whose factor scores by all factors are almost the same are naturally placed often in the same group. Their opposites are "solitary" farms, which are associated with different farms in different groupings. Table 10 shows by three figures to which group a farm belongs according to different groupings. The first figure of the distinguishing number signifies the group in order of size to which a farm belongs according to Grouping I, and the second and third figures the group to which it belongs according to Groupings II and III. Thus, for instance, the distinguishing number 124 signifies a farm which according to Grouping I belongs to the first or biggest group, but according to Grouping III to the fourth or smallest group.

The interrelation of farm types is illustrated in Fig. 3. The size of the symbols shows the number of farms belonging to each type. The types resembling each other most are those whose farms belong to the same group according to two groupings /e.g. 111 and 131/. Such cases are connected by a line in Fig. 3. This provides a good notion of the similarity of various farm types and also reasons for the possible combination of types. As an example, farm types in Fig. 3. are also divided into combined types A-C.

CONCLUDING REMARKS

This study has been purely taxonomical in the sense that the central problem has been the typology of farms, with little attention paid to the explanation of causal relationships between the groups obtained. Also, no clear criteria were adopted in advance for the merits of the grouping, nor was the number of groups pre-established. The interpretability of groups was considered the most important criterion, indefinite though it is. The sole purpose was to arrive at a reasonably objective classification of farms by the multivariate method.

How was the criterion of objectivity fulfilled? The multivariable methods employed are in themselves technically objective. But the methods yield results in accordance with the variables, which are included in the analysis. "The results of the factor analysis are only as good as the choice of the original variables." /Momsen 1970, p. 3/. And indeed, the researcher's sebjective notions were revealed by the choice of variables in the present sutdy.

Because no clear criteria were adopted for the classification of farms, we may note what has generally been noted with regard to classification. "It is generally agreed by logicians that there can be many valid classifications of a given universe of individuals... The property chosen as the differentiating characteristic depends primarily upon the purpose of the classifications." /Grigg 1965, p. 470./. Thus are "good". Such a question might be answered, however, when some practical requirement is at issue. It might be asked, for instance: What is the nature of the farms which have a successor in the ownership and possibilities of continued existence in future? The typology which reveals such farms in a group of their own to the researcher is good from the standpoint of this practical problem.

Althoughthis study deals with farm typology as a purely taxonomical problem, the classification methods employed are serviceable also for practical requirements in which criteria for the number and content of groups are precisely defined.

R e f e r e n c e s

1. Aitchison, J.W. /1970/: The farming system of Wales: a study of spatial variability. Paper presented to the Fourth Meeting of the IGU Commission on Agricultural Typology, Verona, September 1970.

2. Berry, Brian J.L./1961/: A method for deriving multi-factor uniform regions. Przeglad Geograficzny XXXIII, 263-275.

3. Grigg, David /1965/: The logic of regional systems. Annals of the Association of American Geographers 55, 465-491.

4. Haggett, Peter /1966/: Locational analysis in human geography. London

5. Hagood, M.J. /1943/: Statistical methods for delineation of regions applied to data on agriculture and population. Social Forces 21, 288-297.

6. Hagood, M.J. and Price, D.O. /1952/: Statistics for sociologists. New. York

7. Harman, Harry H./1960/: Modern factor analysis. Chicago

8. Harvey, David /1969/: Explanation in geography. London

9. Henshall, J.D. /1966/: The demographic factor in the structure of agriculture in Barbados. Transactions and Papers of the Institute of British Geographers 38, 183-195.

10 Kostrowicki, Jerzy /1964/: Geographical typology of agriculture. Principles and methods. An invitation to discussion. Geographia Polnica 2, 159-167. Beziehung zur Landwirtschaft. Fennia 93: 2, 1-171.

11 Momsen, Janet D. /1970/: Classification of agriculture: a case study from the Caribbean. Paper presented to the Fourth Meeting of the IGU Commission on Agricultural Typology, Verona, September 1970.

12 Munton, R.J.C. /1970/: Farm systems classification: a use of multivariate analysis. Paper presented to the Fourth Meeting of the IGU Commission on Agricultural Typology, Verona, September 1970.

13 Munton, R.J.C. /and Norris, J.M. /1969/: The analysis of farm organisation: an approach to the classification of agricultural land in Britain. Geografiska Annaler, Ser. B, 52, 95-103.

14 Nelson, Howard, J. /1955/: A service classification of American cities. Economic Geography 31, 189-210.

15 Rikkinen, Kalevi /1971/: Typology of farms in central Finland. Fennia 106, 44 p. Helsinki.

16 Steiner, Dieter /1965/: Die Faktoranalyse - ein modernes statistisches Hilfsmittel des Geographen für objektive Raumgliederung und Typenbildung. Geographica Helvetica 20, 20-34.

F i g u r e s

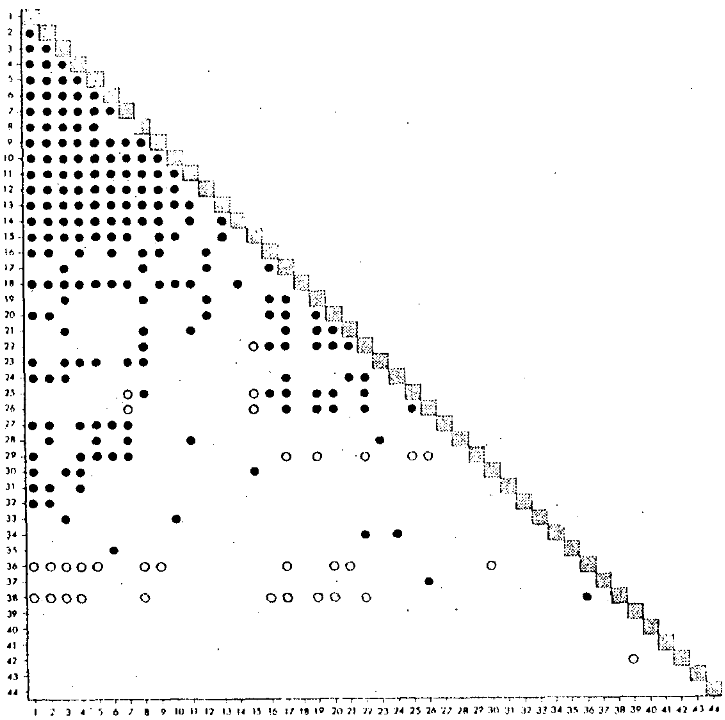

Fig. 12. Correlation matrix of variables in Kalmari.
● = positive correlation at 0.1 % level.
O = negative correlation at 0.1 % level.

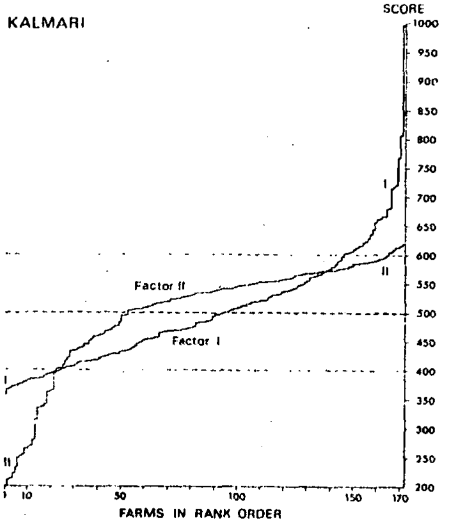

Fig. 2. Factor scores of farms in rank order after
Factors I and II.

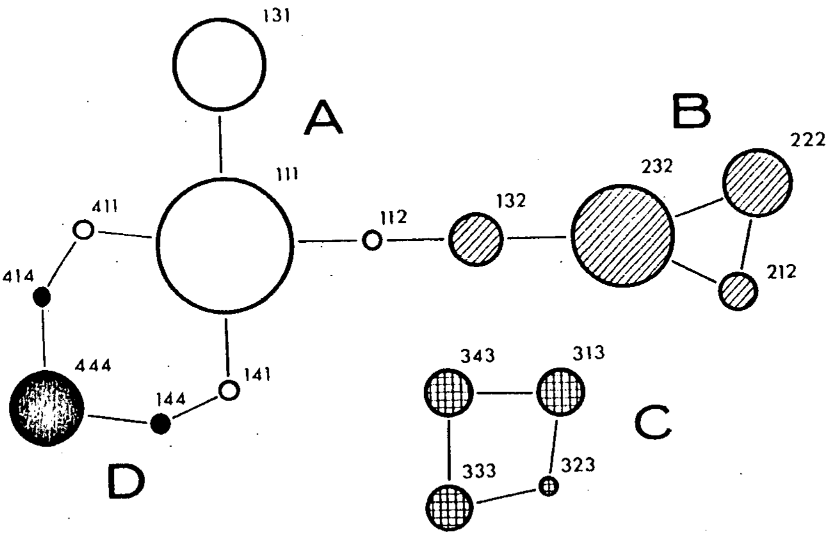

Fig. 3. Types of farms in Kalmari based on grouping analysis.