Square Kilometer Grid System: An Efficient Database in Rural Studies

Jarmo Rusanen, Arvo Naukkarinen and Alfred Colpaert
University of Oulu*

Census data have been available in Finland from 1970 onwards based on 1 x 1 km grid squares identified by latitudinal and longitudinal coordinates as the areal units, as well as the various administrative units. Although this statistical unit has been in existence for some time, it is only in the last few years, with increased computer capacity and the development of geographical software, that it has become possible to exploit the spatial information provided by it.

This paper compares results obtained using the new areal unit with those of earlier research and employs GIS techniques to arrive at a definition of a rural area. Experience with the use of this areal unit has been encouraging, especially as it permits more detailed deduction of spatial structures and changes occurring therein. It allows the rural areas of Finland to be studied in a manner that is independent of administrative boundaries, and in this way provides more precise information on rural structure and the future of rural life in Finland. At the same time, interpretation of this material raises problems of a new kind, for the accuracy of the data unit means that deduction of the spatial structure is a more complicated matter than that using traditional units such as the commune.

BACKGROUND

Statutory population censuses have been carried out by the Central Statistical Office of Finland since 1950, initially at 10-year intervals and at 5-year intervals since 1970, the most recent one having taken place in 1990. These censuses have virtually covered 100 percent of the households in the country, and the results have been published as sets of statistics corresponding to administrative divisions such as the provinces or local government districts or functional areas such as regional planning districts. Alongside these traditional areal units, however, statistics have also been compiled since 1970 using a much finer division of the country, into 1 x 1 km map coordinate grid squares. This has been made possible by assigning a set of map coordinates to each building examined in the census, so that all data regarding the population and their places of residence and work can be precisely located.

* Department of Geography, University of Oulu, Linnanmaa, 90570, Oulu, Finland.

Geography Research Forum • Vol. 13 • 1993:129–138

Work has been in progress at the Department of Geography, University of Oulu, since 1987 aimed at testing the suitability of this new statistical unit as a source of data for research into spatial phenomena and processes, as well as identifying sources of error in the material and exploring alternative styles of presentation in graphical or cartographic form (see Rusanen et al., 1991).

The squares in the grid system are defined by their N and E coordinates. The total surface area of Finland in 1985 was 338,145 km^2, of which the land area comprised 304,623 km^2 (STV, 1985), while 107,503 of the resulting 1 x 1 km grid squares proved to be ones in which there was at least one inhabitant. These were taken to form the basic set of grid squares for the statistical material. The implication is, of course, that, calculated in this manner, approximately 65 percent of the land area of Finland may be said to be uninhabited. The number of data units in fact alters according to the variable to be examined, so that an examination of the population by age groups, for example, will show that only 44,351 grid squares contain children of age 0–6.

Grid square census data are now available for the years 1970, 1975, 1980 and 1985, and those for 1990 will be ready in 1992, in addition to which it is possible to obtain employment statistics on a yearly basis from 1987 onwards. The reliability of the statistics has steadily improved over this period, too, so that where the data on the occupational structure of the working population deviated from the official statistics by an average of 0.62 percent in 1970, the discrepancy was only 0.21 percent in 1985. Apart from numerical errors, problems have also arisen from changes in recording methods or criteria from one census to another and from errors in the location coordinates, although a relatively high standard of reliability had been achieved in these respects by 1985 (Naukkarinen et al., 1991a).

Although the variables available include population, age structure, settlement, employment, economic activity and housing, research to date has concentrated on certain sets of population, age structure, occupational structure and employment data.

LINKS WITH GIS

During the last decade geographers have increasingly used GIS (Geographical Information System) techniques (Maguire, 1989; Rhind, Goodchild and Maguire, 1991; Egbert and Slocum, 1992). Due to its brief history, the role of GIS in geography is still a matter of discussion. Is it a mere technology? Is it a discipline of its own? How is it to be defined? (Maguire, 1991). What is its importance for the future of geography? (Abler, 1988; Openshaw, 1991). Is geography the home discipline of GIS? (Morrison, 1991). Is there a GIS crisis in geography? (Openshaw, 1991). What is the place of GIS in the geography curriculum? (Kemp, Goodchild and Dodson, 1992). Although its status in geography cannot be defined precisely, perhaps as a result of constant change and development, we understand GIS to be a crucial part of geography, providing a useful tool for studying the spatial structure of rural areas.

This view is based on the fact that vast amounts of census data are available for the grid above mentioned. These data sets are called themes, overlays, coverages, maps or data elements in GIS systems (see Tomlin, 1990). Coverages in GIS systems are usually elements of the physical earth surface, but this method can be used just as well with data concerning human geography. When we compare the availability and extent of geo-coded census data internationally (see Ottoson and Rystedt, 1991; Taylor, 1991; Dahmann, 1992; Eweg, 1992; Rhind, 1991) we find that the situation in Finland is exceptionally good. In addition to the general census data produced at five-year intervals since 1970, data are now available for each year from 1987 onwards.

USE OF THE GRID SQUARE MATERIAL FOR DEFINING RURAL AREAS AT THE MACRO-LEVEL

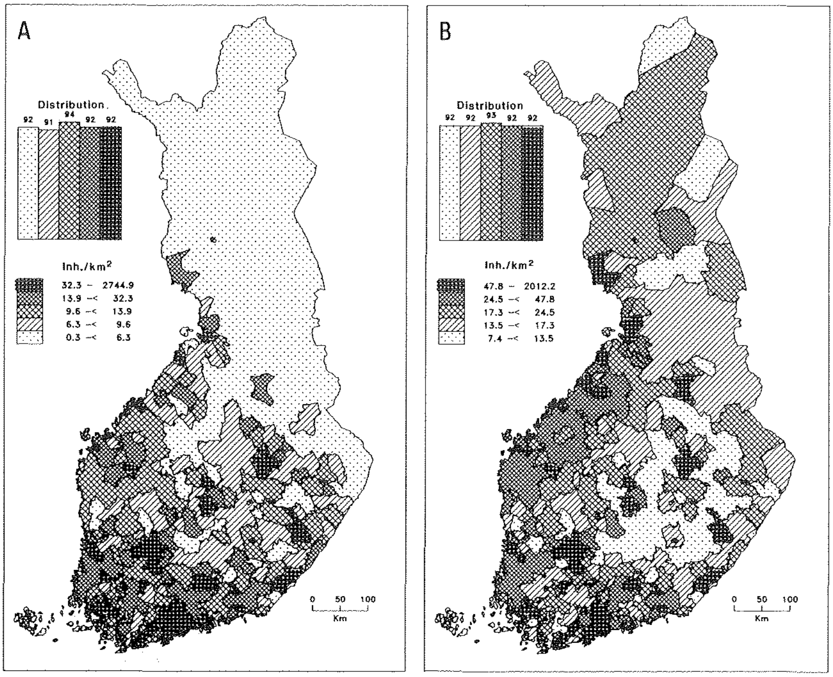
As a data unit, the 1 x 1 km grid square is well suited in theory for use as the basic unit for spatial analysis, and as such provides essentially new information relative to the aggregate statistics compiled at various levels from the same raw data, and enables spatial information to be handled in a manner which is independent of administrative boundaries.

As an example of the differences that arise when data are analyzed in this manner, instead of using the traditional statistical unit, the local government district, or commune, we can consider variations in the density of population in Finland in 1980 (Fig. 1a and 1b).

The traditional model shows the whole of Northern Finland to be very sparsely populated, normally with less than $6/km^2$, whereas the grid square analysis places the most sparsely inhabited areas mainly in the Lake Region, while assigning Northern Finland a near-average density. The difference is explained by the fact that this new unit allows one to study only those areas which possess some population, whereas the traditional model also included uninhabited areas, which are particularly common in Northern Finland.

Another example would be the concentration of population in Finland, in which we can set out to study the minimum area which accommodates a half of Finland's population. It has been customary to look at the concentration of population from the perspective of Southern Finland (Fig. 2a), employing a zonal model. Hustich (1977) claims, for example, that the southern half of Finland was shrinking up to the mid-1970s, implying a concentration of population in the south. This southerner's view, based on the use of the commune as the statistical unit, continues to prevail in Finnish geographical circles. Following the classification proposed by Alestalo (1983), however, half of the Finnish population live in the major cities (Fig. 2b), generating an image which departs very markedly from that fostered by the Hustich model.

Figure 1: A: Density of population in Finland in 1980, conventional model.
B: Density of population in Finland in 1980, grid square model.



Using the new unit, the inhabited squares are arranged in order of size, allowing the distribution of population to be depicted cumulatively by quartiles, beginning with the most populous squares. Each of the quartiles contains the same proportion of the population in Finland (sec Alestalo 1983). This method using the new unit enables us to look at the location of the most densely populated grid squares quite irrespective of administrative boundaries, providing us with a still more accurate result (Fig. 2c). We now see that half of Finland's population is dispersed among a set of agglomerations comprising towns and cities of varying size, as well as the centers of the largest rural communes. The pattern of these most densely inhabited squares reflects fairly accurately the hierarchical nature of the regional structure in Finland.

Statistically speaking, the situation in 1985 was that a half of the Finnish population, approximately 2,420,000 persons, were living in an area amounting to 1,249 km^2 (Table 1), i.e., only 0.41 percent of the country's total area. Correspondingly, it may be said that three out of every four Finns live in an area of 5,579 km^2, or 1.83 percent of the total.

Figure 2: A: Hustich's map depicting the shrinking of the minimum area occupied by one half of the population of Finland from 1880 to 1975 (Hustich, 1975:215). B: Minimum area occupied by a half of the population of Finland in 1980 (adapted from the map of Alestalo, 1983). C: Minimum area occupied by a half of the population of Finland in 1985, based on cumulative grid-square population density data (Naukkarinen, 1991:31).

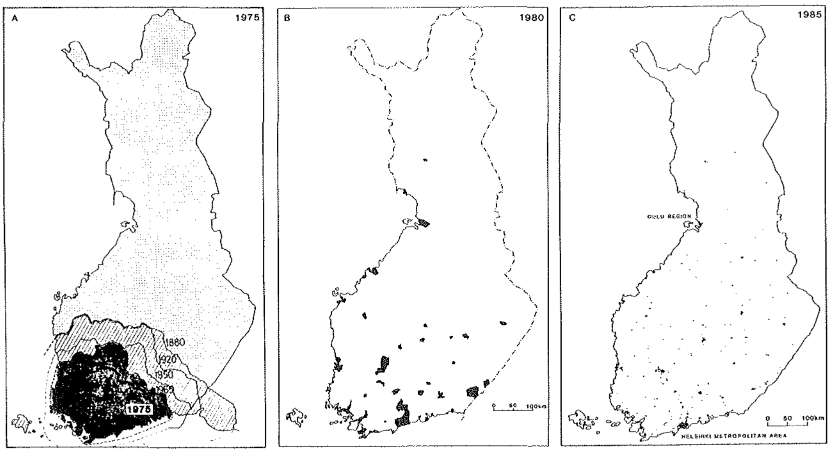


Table 1: Concentration of population in Finland and mean population density in 1980 and 1985 (Statistics Finland: grid square data).

	No. of inhabited squares		*Population density inh./km^2*	
	1980	*1985*	*1980*	*1985*
1. quartile	292	324	3,951	3,749
2. quartile	891	925	1,294	1,311
3. quartile	5,270	4,330	220	282
4. quartile	100,945	101,923	11	12
TOTAL	107,398	107,502		
MEAN			42.90	45.08

The remaining quarter of the population is dispersed over about 102,000 km^2, comprising approximately 95 percent of the inhabited area of the country and representing the really sparsely-populated districts, with an average population density of less than 12 inhabitants /km^2. It is this last quarter that can be referred to at the major aggregational level as the rural population, and these districts as 'rural areas', lying for the most part beyond the officially defined built-up areas.

The official mean population density in Finland is indicated as being 16 inhabitants/km^2, i.e., the total population relative to the total land area. The grid square material nevertheless allows us to quote a figure which applies only to those areas which are actually inhabited, which comes closer to the true situation. The mean figure for 1985 would then be 45 inh./km^2. This illustrates the way in which use of the new statistical unit enables the notion of rural area to be redefined with greater precision so that the image of the density of settlement which it provides approximates better to the reality as experienced by the people themselves.

USE OF THE GRID SQUARE MATERIAL TO DEFINE FUNCTIONAL RURAL AREAS

The grid square material is most powerful when two or more variables are examined for each square and their alterations under different sets of conditions are examined mathematically, using principles of GIS. Thus it is also possible to approach the definition of a rural area from the bottom upwards, working from data applying to individual grid squares.

This methodological example makes use of the fourth quartile distinguished above, that with 1–92 inh./km^2, together with primary sector jobs as a proportion of total employment in the square, with those squares lacking any primary sector activity excluded from the material. The emphasis is thus on definition of the functional structure of rural areas.

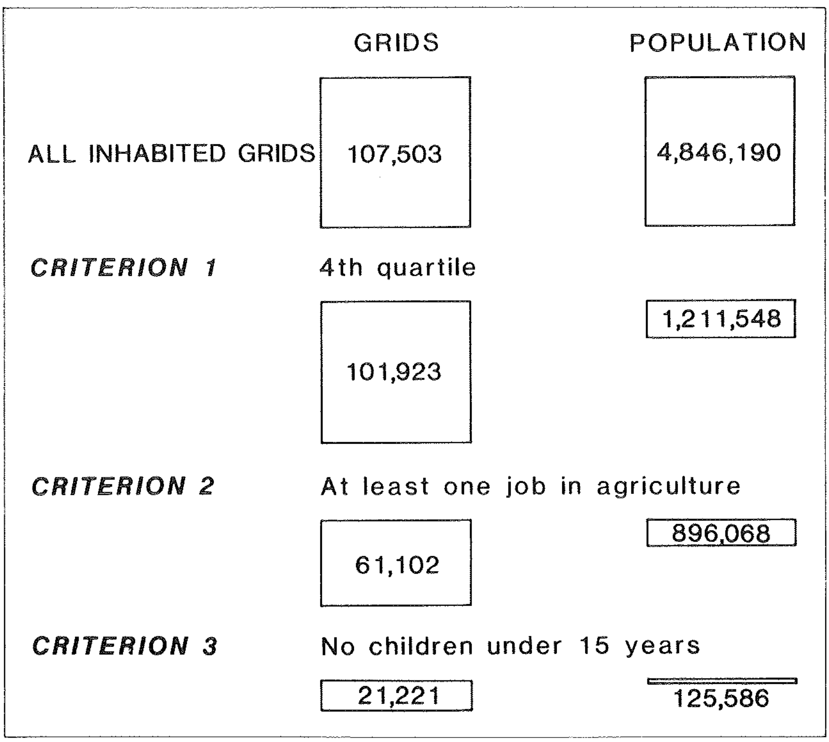
As seen in Table 2, the vast majority of the grid squares are ones in which employment is exclusively in the primary sector, among which just under 50,000 squares form an area that could be referred to as the 'basic countryside', constituting the core of rural Finland (see also Naukkarinen et al., 1991b). The combined population of the grid squares with only primary sector employment amounts to some 11 percent of the national population, and that of all the squares possessing some primary sector employment 18.5 percent. This gives us a minimum estimate for the rural population, since there are also some people engaged in primary production in the larger built-up areas.

Table 2: Jobs in agriculture as a proportion of total employment in the grid squares of the fourth quartile and their population 1985 (Statistics Finland: Grid square data).

Jobs in agriculture as a proportion of total employment	*Grid square* N	%	*Population of grid squares*	
0.1–25.0	2,378	3.9	87,829	9.8
25.1–50.0	3,700	6.1	102,949	11.5
50.1–75.0	3,304	5.4	94,997	10.6
75.1–99.0	2,036	3.3	68,556	7.7
100	49,684	81.3	541,737	60.4
TOTAL	61,102	100.0	896,068	100.0

The next step was to evaluate the future development of rural areas using three coverages. The first is the previously mentioned fourth quartile (1–92 inh./km^2), the second comprises all grid cells having at least one job, and the third consists of those grid cells having no children under the age of 15. The future prospects for the rural areas and the structure of settlement therein will be crucially affected not only by the existing numbers of jobs but also by the age structure of the population. Greater depth can be added to the above table by examining only those grid squares in which there are no children under 15 years of age. These make up more than a third of all those containing primary sector jobs (Fig. 3), and represent the areas in the greatest danger of depopulation, which are also those where most of the employment available is in agriculture. They account for 14 percent of the total population of the squares containing agricultural employment, but only 2.6 percent of the population of Finland as a whole. Thus extensive areas may be said to be threatened with depopulation, but they are areas with little demographic potential on a national scale.

Figure 3: Numbers of grid squares and population figures included in the definition of rural areas according to three alternative criteria.



EXPERIENCE WITH USE OF THE GRID SQUARE DATA

Experience to date with the use of the coordinate-based grid square data for spatial research has for the most part been encouraging, and it is evident that this material provides an entirely new perspective on functional and spatial questions concerning rural areas.

A new problem arises, however, with the interpretation of the data. The traditional data units for studying employment opportunities, for example, have been either individual people, communes, regional planning boards, provinces or other administrative divisions, whereas with the new grid square data the basic units are individual 1 x 1 km squares. It will undoubtedly take time before practitioners of this research become accustomed to this new unit, the use of which for planning and research purposes will inevitably increase in the future. The best data unit of all would of course be the individual person and the coordinates indicating his location, but this prospect will remain theoretical because of the demands of personal privacy. It is now common to extract data for 500 x 500 m grid squares in built-up areas, and the size may well be reduced in the future to 250 x 250 m.

Most experiments with the presentation of grid square data have so far employed conventional cartographic methods, but the material will also provide greater flexibility for analysis at various aggregate levels, filling out the results obtained by the methods of spatial statistics. On the other hand, the great number of data points restricts all processing of the results to quantitative, computerized methods.

The experiments carried out so far encourage the continuation of this exploration. Can the grid square data be used to ascertain the numbers of commuters moving into rural areas or to predict the depopulation of certain sparsely inhabited areas? How can the method be used to delimit other functional or homogeneous areas? Would such grid square material be adaptable to spatial research based on the drainage basin as the areal unit? These are just some of the questions on which the grid square data could provide a new perspective arising out of the preliminary evaluation of the material.

Fields that are still entirely unexplored at present include the combining of remote sensing data with the grid square system in order to study factors affecting the location of settlement, as well as factors linking the state of the environment with changes in the settlement pattern.

CONCLUSIONS

On the basis of the results obtained with the cumulative method we can divide Finland into a densely populated region (quartiles 1–3) and a sparsely populated region (4th quartile), as is consistent with the center-periphery theory. Analysis of the age structure of rural areas shows that large parts of the countryside lack the most important resource for the future, namely children under the age of 15. These are the areas most likely to become depopulated.

GIS methods with multiple coverages can be used to study rural areas and their future prospects in a new and more accurate manner. Previously the study of rural areas was based primarily on the different types of administrative unit, namely urban and rural communes. The new spatial unit enables rural areas to be studied without the restrictions of administrative boundaries and provides a tool for more accurately defining the concept of rural areas.

REFERENCES

Abler, R.F. (1988) Awards, rewards, and excellence: Keeping geography alive and well. *Professional Geographer,* 40:135–140.

Alestalo, J. (1983) The concentration of population in Finland between 1880 and 1980. *Fennia,* 161:263–288.

Dahmann, D.C. (1992) Geographic entities to represent the U.S. settlement system. Paper presented at the session on 'Analysis and Representation of Spatial Structures' of the 27th International Geographical Congress, Washington, D.C. 14 August 1992.

Egbert, S.L. and Slocum, T.A. (1992) EXPLOREMAP: An exploration system for choropleth maps. *Annals of the Association of American Geographer,* 88: 275–288.

Eweg, R. (1992) Exploring possible directions of development in a rural area with GIS. *EGIS '92. Conference Proceedings. Third European Conference and Exhibition on Geographical Information Systems.* Munich, Germany. March 1992, 1:19–28.

Hustich, I. (1977) Suomen kutistumisen vauhti hiljenee (The speed of the shrinkage of Finland is slackening). *Terra,* 89:214–221.

Kemp, K.K., Goodchild, M.F. and Dodson, R.F. (1992) Teaching GIS in Geography. *Professional Geographer,* 44:181–191.

Maguire, D.J. (1991) An overview and definition of GIS. In D.J. Maguire, M.F. Goodchild and D.F. Rhind (eds.) *Geographical Information Systems: Principles and Applications.* London: Longman, 1:9–20.

Maguire, D.J. (1989) *Computers in Geography.* London: Longman.

Morrison, J.L. (1991) The organizational home for GIS in the scientific professional community. In D.J. Maguire, M.F. Goodchild and D.F. Rhind (eds.) *Geographical Information Systems: Principles and Applications.* London: Longman, 1:91–100.

Naukkarinen, A. (1991) Suomen aluerakenteen muutos (Changes in regional structure in Finland). In Malinen, P. (ed.) *The Future of Kainuu—Alternatives and Strategies.* Research Institute of Northern Finland, University of Oulu. Research report 106, pp. 27–41.

Naukkarinen, A., Rusanen, J. and Colpaert, A. (1991a) Mitä koordinaattipohjaiset ruututiedot kertóvat väestön ja toimintojen keskittymisestä Suomessa 1980-luvun alkupuoliskolla (What do coordinate-based grid square data tell us of the concentration of population and functions in Finland in the early 1980s?). In M. Löytönen and J. Pietala, (eds.) *Reijo Helle 60 vuotta.* Helsinki: The Helsinki School of Economics and Business Administration. Acta Academiae Oeconomicae Helsingiensis. Series A, 77:157–168.

Naukkarinen, A., Rusanen, J. and Colpaert, A. (1991b) Maaseutu karttaruututietojen perusteella 1980-luvun ensimmäisellä puoliskolla (The Finnish countryside in the early 1980s on the basis of map grid data). *Terra*, 103 (in press).

Openshaw, S. (1991) Commentary: A view on the GIS crisis in geography, or, using GIS to put Humpty-Dumpty back together again. *Environment and Planning A*, 23:621–628.

Ottoson, L. and Rystedt, B. (1991) National GIS programmes in Sweden. In D.J. Maguire, M.F. Goodchild and D.F. Rhind (eds.) *Geographical Information Systems: Principles and Applications*. London: Longman, 2:39–46.

Rhind, D.W., Goodchild, M.F. and Maguire, D.J. (1991) Epilogue. In D.J. Maguire, M.F. Goodchild and D.F. Rhind (eds.) *Geographical Information Systems: Principles and Applications*. London: Longman, 2: 313–327.

Rusanen, J., Naukkarinen, A. and Colpaert, A. (1991) Ammatissa toimivan västön keskittymisestä Suomessa vuosina 1970–1985 koordinaattipohjaisten ruututietojen mukaan (Concentration of the employed population in Finland in 1970–1985 as shown by coordinate-based grid square data). *Lompolo*, 1/91:5–18.

Statistical Yearbook of Finland (STV) 1985/85. Helsinki: Central Statistical Office.

Taylor, D.R.F. (1991) GIS and developing nations. In D.J. Maguire, M.F. Goodchild and D.F. Rhind (eds.) *Geographical Information Systems: Principles and Applications*. London: Longman, 2:71–84.

Tomlin, D.S. (1990) *Geographic Information Systems and Cartographic Modeling.* Englewood Cliffs, N.J.: Prentice Hall.