RAPPORTER 80/31

TWO NOTES ON LAND USE STATISTICS

BY
P. A. GARNÅSJORDET, Ø. LONE AND H. V. SÆBØ

STATISTISK SENTRALBYRÅ
OSLO

RAPPORTER FRA STATISTISK SENTRALBYRÅ 80/31

TWO NOTES ON LAND USE STATISTICS

BY
P. A. GARNÅSJORDET, Ø. LONE AND H. V. SÆBØ

OSLO 1980
ISBN 82-537-1214-6
ISSN 0332-8422

PREFACE

This report contains two invited contributions to meetings held under the auspices of the Statistical Commission and Economic Commission for Europe, Conference of European Statisticians. The first paper, on land use and linkages, was prepared for the Meeting on Land Use Statistics held in Geneva, Switzerland, 17-20 March 1980, and the second paper, on point sampling, was prepared for the Seminar on Environmental Statistics held in Warsaw, Poland, 16-19 September 1980. The Warsaw seminar was arranged in cooperation with the Senior Advisers to ECE Governments on Environmental Statistics.

The papers included in this report deal with issues of central importance in the work the Central Bureau of Statistics is undertaking in relation to environmental statistics, resource accounting and in particular to land accounts and land use statistics. The views expressed are those of the authors and do not necessarily represent those of the ECE or of the Bureau.

Central Bureau of Statistics, Oslo, 20 November 1980

Odd Aukrust

CONTENTS

UNITED NATIONS

ECONOMIC AND SOCIAL COUNCIL

CES/AC.52/5
23 January 1980
Original: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Meeting on Land Use Statistics
(17-20 March 1980)

THE PROVISION FOR LINKAGES TO VARIOUS DATA SYSTEMS IN THE DEVELOPMENT OF LAND USE STATISTICS IN NORWAY 1/

Paper prepared by the Central Bureau of Statistics, Norway.

TABLE OF CONTENTS | Paragraphs

1/ Prepared by Ø. Lone, Unit for Resource Accounting.

TABLE OF CONTENTS (continued)

* * *

Summary

1. In the Norwegian land accounts system, attempts are made to provide for linkages with several different types of data systems. The land accounts system by itself requires linking of a number of rather different data sources; the land accounts are linked to population and production statistics in order to make possible land budgeting, to the accounts for other resource categories, and to environmental data.

2. The problem of determining and identifying the unit of observation is probably particularly acute where land use statistics is concerned. There are no given, natural units such as human individuals or business establishments. Lacking such natural units, the means of geographical identification, holdings, regions, addresses and co-ordinates in space, are all important linkage mechanisms.

3. Examples of the four kinds of data sources on Norwegian land use (censuses, registers, maps and map-derived registers, and point-sampling surveys) are given, with a brief discussion of their main characteristics and in particular their means of geographical identification.

4. The Norwegian land accounts system is based on two main linkage units, the holding (identified by the property/holding-number and in the Economic Map Survey), and the sample points in systems established for this purpose (identified by map co-ordinates). Data is published by administrative regions such as counties, municipalities and urban areas.

5. By way of illustration, the linkages between agricultural censuses and the Land Register, between the National Forest Survey and the Economic Map Survey, between housing censuses and the topographical map, and the linking of land use in urban areas to industrial production are discussed very briefly. The linkages to registers of geocoded data are also mentioned.

6. The land accounts are linked to resource accounts such as those for water resources, wood products and minerals, and to environmental data such as air quality in urban areas. The sample point systems and the watercourse register will be used as a basic geographical reference to observe and monitor environmental data in general, possibly even in Norwegian ocean territory.

I. Introduction

7. In the Norwegian land accounts system attempts are made to provide for linkages with several types of data systems. The land accounts system by itself requires linking of a number of rather different data sources; the land accounts are linked to population and production statistics, to the accounts for other resource categories, and to environmental data.

8. The need to link land use data from different sources is obvious. Very few data sources cover all of the national territory, and those with a high level of coverage are usually specialized and narrow in content in relation to the needs of national and regional land use planning, which presupposes integrated, multi-dimensional land use information. Particularly with regard to potential use of land for different purposes (urban development, agriculture, forestry) and with regard to changes in land use, detailed and integrated data are required.

9. Changes in land use, moreover, are intimately connected to demographic and economic development at the national and regional levels, so that when the land accounts system is to be used as a basis for land budgeting, linkages to population and production data are required. This is also desirable from the point of view of analysing the intensity of land use (population density, economic productivity etc.).

10. The land accounts system is also central to the attempt to integrate the different resource accounts. Several resource categories, such as agricultural products, wildlife, and forest products, are, as biotic resources, directly connected to land as a major factor of production. With the water resource accounts, the land accounts make up the category of geo-accounts, in distinction to the material balances-accounts (as for instance energy accounts, minerals accounts, forest products accounts). This means that the land accounts system provides the geographical framework for integrating several biotic and abiotic (mining, energy production) resource accounts. (Even if more localized in their land requirements, abiotic resources, too, locate in space and may have major land use and environmental implications through the way they are exploited.)

11. Lastly, the land accounts system provides for the linking of environmental data to the land use information produced by the system. That is, different types and categories of land use may be correlated with such indices of environmental quality as concerns air, water and other environmental components.

II. Linking the statistical individual: What is a land use unit ?

12. The idea of linking data from different sources is of course to disaggregate data in order to advance from correlations on a regional or even national level to correlations on the level of the statistical individual. However, the problem of determining and identifying the unit of observation is probably particularly acute where land use statistics is concerned.

13. A land use unit is perhaps best defined as a homogeneous field (or parcel) of land, but the operative word (and the deceptive one) here is "homogeneous". A land use unit is defined and delimited differently according to different dimensions of land use and different types of characteristics. There are no given, natural units (such as human individuals in population statistics or business establishments in industrial statistics), but a plethora of units that multiply in number and get progressively smaller in size as the number of characteristics to take into account increases. With a seven-digit system of classes, the main classification in the Norwegian land accounts system provides for a theoretical maximum of ten million classes and an actual number of four hundred thousand.

14. Most land use statistics are conventionally presented by administrative regions and/or by holdings (particularly for data from agricultural or forestry censuses). The holding is not, strictly speaking, a unit of land use, but an aggregate no different in principle from other types of regions, with no information on the land use units (fields, parcels) as such. The holding is thus a means of geographical identification, along with other regions, geographical coordinates, and adresses. These are all important linkage mechanisms, and as such will be discussed repeatedly throughout the remainder of this paper.

III. Some important Norwegian data sources and their characteristics

15. To provide a unified system of land use statistics, data from several sources are needed. Most of these sources are specialized and cover only part of the territory in question, and even when covering the same (part of the) territory, they often differ on the amount of land in comparable classes.

16. These differences may be due to different statistical populations, to different methods of collection, to different classes and definitions of classes, and to different time references, quite apart from their identifications of the statistical individuals. The sum total of these differences makes it imperative to compare and control land use data at some sort of individual or micro level.

17. Briefly, there may be said to be four kinds of data sources for land use information:

(a) censuses

(b) registers

(c) maps and map-derived registers

(d) point-sampling surveys

III.1 Censuses of agriculture and forestry

18. In recent decades, censuses of agriculture and of forestry have been held at ten-year intervals. The censuses have, however, been held in different years (the one of agriculture two years after the one of forestry), and with non-comparable populations; the agricultural census covering all agricultural holdings in excess of 0,5 ha agricultural land, the forestry census all holdings in excess of 2,5 ha productive forest land. The classifications are not comparable between agricultural censuses (of 1949, 1959, 1969) or between forestry censuses (of 1947, 1957, 1967), nor between the agricultural and forestry censuses, for more than a few main classes of land.

19. However, data on the level of holdings are available within the Central Bureau of Statistics and are particularly valuable because linked to production data. The Census of 1979 was a combined census of agriculture and forestry with all the advantages of such an arrangement and much better classification categories for land use than in the earlier censuses.

III.2 Censuses of population and housing

20. The censuses of population and housing of 1950, 1960, 1970 and the one planned for this year identify households and housing units by holding and adress, and classifies housing units in a way that has been used as a basic reference for the classification of built-up land in the land accounts classification. As is the case for the agricultural and forestry censuses, data is available on the level of the municipality and its subdivisions.

III.3 Censuses of establishments, the Register of Establishments, and other industrial statistics

21. Censuses of (business) establishments are held every ten years, and the Central Register of Establishments and Enterprises in the Bureau is continuously updated. Identification is by way of code numbers internal to the Bureau and by way of adress, which in this case is linked to the geocoded registers discussed in paragraph III.5

22. Industrial and commercial production statistics are identified by administrative regions and by way of ISIC (International Standard Industrial Classification of All Economic Activities) and commodity classification.

III.4 Administrative data routines and registers

23. Ministries and agencies of the central government use a large number of registers and data collection routines, some of which provide land use data or land use related information. Classification of land use categories are seldom according to any common standard, and this will be one of the major challenges in the development of land budgeting.

24. Most information systems with relevance to land use identify units by holding and/or adress, and thus by administrative region as well.

25. One of the particularly important registers with regard to land use is the Register of Roads at the Roads Directorate (agency of the Ministry of Transport and Communications). This identifies units of public roads by coordinates and place-names of end-points. For most categories of public roads, the width of the road and thus its land surface are given or may be estimated, and traffic counts are easily linked to these road units.

III.5 Geocoded data systems

26. Norway has just started implementation of a system of geocoded registers of Ground property, Adresses, and Buildings, the so-called GAB-system. This system of related registers identify its units by holding and by coordinates in space. Hopes are attached to the future use of these registers to provide data on land use changes, and the land accounts system is coordinated with the register systems on the points of land use classification and updating of the registers.

III.6 The Economic Map Survey and the Land Register

27. Maps at the scales of 1: 5000 and 1: 10 000 are made for about 170 000 km^2 of a total land area of 320 000 km^2, and about 130 000 km^2 have been mapped so far. These maps contain property limits as well as a detailed division of the area mapped into land cover and land capability classes, mainly from the point of view of agriculture and forestry.

28. The Land Register, covering so far about fifty out of the more than four hundred and fifty Norwegian municipalities, is based on these maps, and gives data based on digitizing of the map information for holdings as well as ownership units. The Land Register may, through its digitized data, be used to identify individual fields (mapping units) in the coordinate system on these maps.

III.7 Topographical and thematic maps

29. The whole of Norway is covered by topographical maps at the scale of 1: 50 000 containing much valuable information and serving as a basis for thematic maps (e.g. geology, vegetation) with very uneven coverage. The maps are of the Universal Transverse Mercator projection (UTM) and contain information on topography, watercourses, roads, settlement, wood cover and bogs, among other subjects.

30. Some thematic maps are based on the Economic Map Survey maps, which are of a different map projection, but this projection, too, is a transversal Mercator projection. The UTM-coordinates and the so-called NGO-coordinates may be transformed to each other by readily available computer programs.

III.8 The National Forest Survey

31. Some 70-80 per cent of the productive forest land in Norway and all land below the forest line in the counties surveyed is covered by the National Forest Survey in the latest survey of 1964-1976. A large amount of information on land categories and forestry data is collected by fieldwork on sample sites identified on topographical maps and by UTM-coordinates.

IV. Linkages in the land accounts system

32. Two kinds of linkages are important in the land accounts system. Linkages in the sense of connecting data from different sources at the level of the statistical individual have to be supplemented by linkages in a wider sense, by linking different systems at the aggregate as well as the individual level through the use of comparable classes.

IV.1 Linkages through classification

33. The classification system in the land accounts provides for linkages to three main categories of sources.

34. Firstly, the classification system is based on land cover categories as a surrogate for the actual activities making up land use. These land cover categories take into account the classes of land identifiable from remote sensing sources and air photos as well as the categories used in the agricultural and forestry census(es) and those used in the Economic Map Survey. To widen the classification from the sometimes narrow concentration on capability for agriculture and forestry, emphasis is also laid on classification by vegetation types, meant to reflect the wider ecological dimensions of land categories.

35. Secondly, all built-up land is classified by physical characteristics as well as by the International Standard of Industrial Classification of All Economic Activities (ISIC). This is to provide for linkages to economic data, and in particular for land budgeting purposes.

36. Thirdly, land is classified according to climate, i.e. actually by heighth above sea level, a major factor of climatic differentiation at the county level. Perhaps even more important, legal restrictions on land designated to the exclusive use for either recreation or conservation purposes are taken into account.

IV.2 Linkage units

37. The land accounts will be published at county level, with data on main classes published at the level of the municipalities and urban areas. The analyses and estimates necessary in order to integrate data from different sources will be based partly on linkages at the level of the holding. Most sources identify holdings through the property/holding number, and even if the holding is not a totally satisfactory unit for all purposes, it is possible to compare and link data from different sources in a controlled way far better than data on the level of administrative regions could make possible.

38. The most important linkage in the land accounts system, however, is the point sample systems established for this purpose.

39. These systems consist of nets of regularly (quadratically) spaced sample points with distances between points varying from 100 m up to 12 km depending on the intensity and heterogeneity of land use in the sampled area.

40. Within urban areas, distance between sample points is 100, 200 or 300 m, depending on the size of the urban area. This is to minimize work load in sampling and keep relative standard deviations at the same level for classes of comparable relative importance.

41. Within the remainder of the urban region (that is, in the commuting zones), distance between points is 600 m, in order to provide for extension of the 100, 200 or 300 m net in case of expansion of the urban area.

42. These nets of sample points are positioned and identified by NGO-coordinates, that is, in the coordinate system and the map projection used as a basis for the Economic Map Survey. The nets of sample points have been established for all urban areas with more than 1000 inhabitants, and, as of January, 1980, the registration from air photos of land use in sample points in 1955, 1965 and 1975 is completed for some 150 urban areas out of a total of 250.

43. The remaining nets are constructed on the basis of 1, 3, 6 and 12 km distances between sample points in the UTM-coordinate system, that is, with the topographical maps as cartographical basis.

44. To provide land use data on a national basis, a net of some 7500 points have so far been established, with nets of a finer mesh for some 2-5 pilot counties under preparation. These nets of sample points will be integrated at the 1 and 3 km distances with the sample sites of the National Forest Survey, coded according to the information available in the topographical maps at scale 1: 50 000, and the UTM-coordinates are transformed to NGO-coordinates and identified in the Economic Map Survey for sample points within the areas covered by these maps. In addition, other land use data and land use related information is linked to sample points (such as geology, vegetation, etc.) from thematic maps and air photos.

45. The advantages of linking land use information through sample points are in our experience considerable. Compared to the alternative one often has, of mapping all relevant land category units, identifying and measuring each minimum mapping unit by planimetrification or digitized data systems, there are obvious gains in speed and economy. The decisive question is whether this sort of detailed correlations is needed in a cartographically precise form or mainly for analytic and planning purposes.

46. The sample point systems also make it possible to resolve some of the conceptual and practical difficulties of deciding which geographical level one is working with when collecting land use statistics. In particular, urban land use is often confused by not differentiating clearly between land use at the field level (each residential, commercial or manufacturing unit) and at the area level (residential areas, which may contain schools, shops, hospitals, and even small-scale manufacturing industries; commercial areas, very often with mixed use at the field level; and manufacturing areas, etc.).

47. In our sample points in urban areas, we classify points according to physical structure (point level), and to land use at the field level and at the area level. We are thus able to provide data on the proportion of residential (and other) fields within the residential (and other) areas, as well as building densities and road densities etc. within classes of fields and within areas.

V. Linkages between land use data and linkages to population and production data: Some examples

V.1 The Agricultural Census and the Land Register

48. For the municipalities which have established land registers on the basis of the Economic Map Survey, the amount of agricultural land in comparable classes given by the land registers differs considerably (with a mean area some 10-15 % higher) from the data given by the agricultural census and the annual sample census of agriculture. These discrepancies are serious, and the Central Bureau of Statistics is cooperating with the Institute for Land Registration to clear up the reasons for the large and seemingly unsystematic differences. This is done through the comparison of data from the two sources at the level of the holding, using the property/holding number, which is a means of identification in both sources.

V.2 The National Forest Survey and the Economic Map Survey

49. These two sources both give productivity classes for forestry production and land categories that theoretically should be comparable. Nonetheless, a comparison between the two sources for one of the pilot counties, Østfold, based on the sample sites of the NFS and some five hundred sample points in the national net mentioned earlier, resulted in a consistently higher productivity evaluation for the data from the EMS. In the very near future, we will compare these productivity evaluations at the level of the sample points, as the five hundred sample points in this county are a sub-sample of the about nine thousand NFS sample sites in the county, both systems identifying sample points by UTM coordinates.

50. Linking these sources also makes it possible to give detailed data on the permutations of productivity, age of the forest growth, cubic mass, vegetation type, trees etc. (all from the NFS) with data on land capability for agriculture, ownership, tenancy etc. (from the EMS).

V.3 Land use and production in urban areas

51. Through physical maps at the scales of 1: 1000 to 1: 5000 land use units with industrial or commercial establishments in the five largest urban areas in Østfold are delimited and measured by digitization. These land use units are then linked to the establishments in the Central Register of Establishments and Enterprises in the Bureau by way of adresses and maps. Though a rather labour-intensive procedure, this linking results in very detailed and reliable data on the correlations between land use, production, and employment, and other correlations highly relevant to the analysis of land use intensity and to land use planning and budgeting.

V.4 Topographical maps and housing censuses

52. The category of dispersed settlement is an important land use class in Norway, some researchers speculating that it may consume twice as much land as the urban areas. There is, however, no reliable data on the amount of land used for dispersed settlement. In connection with the land accounts system, dispersed settlement as shown on topographical maps for the two pilot counties, Østfold and Sør-Trøndelag, is being counted and registered within cells of 1x1 km in the UTM-coordinate system.

53. In order to control the reliability of these data, however, a few sample municipalities are compared to the data given in the Agricultural Census and in the Population and Housing Census, the linkage being the property/holding number, which for this purpose is an adequate unit.

54. Actually, this procedure is a two step linkage. As the topographical maps do not contain any property/holding limits or identification, it is necessary to go first from the topographical map to the economic map, and then identify the property/holding number and look this up in the population and housing file from the census.

V.5 Geocoded data (the GAB-system)

55. Data on land use, buildings, and particularly on changes in land use and construction of new buildings, from the registers earlier mentioned, may be linked to each other (building register units to ground property units) by the adress register, by property/holding number and by coordinates in space. These units may also be linked to the sources mentioned above by the same linkages.

56. There is so far not much practical experience with using the registers for such purposes, as they are hardly implemented for more than a few pilot municipalities. However, in so far as they consist of rationalizing earlier registers and administrative routines (particularly regarding the statistics on bulidings and construction), all signs point to the great value of the geocoded registers. The work with the land accounts system is closely coordinated with the development and implementation of the GAB-system.

VI. Linkages to other resource accounts and to environmental data

VI.1 The water resource accounts

57. The resource accounts for land and water make up the category of geo-accounts, for which geographical space is a major dimension and location an important characteristic. The water resource accounts will be based on a Watercourse Register, established by the Norwegian Water Resources and Electricity Board and implemented by the Unit for Resource Accounting of the Central Bureau of Statistics, as part of the work with the resource accounts for energy. In order to provide data on the environmental consequences of hydro-electric power construction, it became necessary to identify watercourse units unambiguously. This was done through the implementation of the Watercourse Register, for which a proposed scheme had already been put forward by the Water Resources and Electricity Board.

58. The Watercourse Register provides for linking the land accounts to the energy accounts by way of area flooded through hydro-electric power construction. The sample points in the national net are linked to the watercourse units and their characteristics through UTM-coordinates, which is the chief identification of the units.

VI.2 The forest products accounts

59. As forests cover some thirty per cent of the Norwegian land area, the linkage between the land accounts and the resource accounts for forest products is an important one. The forest products accounts are linked to the land accounts particularly as far as stocks and reserves of wood are concerned, through the sample sites of the National Forest Survey.

60. A model for estimating cutting (removal) and transport costs of forested sample sites with different characteristics concerning topography and distance to roads is under preparation, and, when implemented, will provide data based on the National Forest Survey on both wood reserves and productive forest land area by economic cost categories.

VI.3 The minerals accounts

61. An important land use category in many Norwegian counties is land taken for the purposes of providing sand and gravel for building and road construction. Such areas, where this purpose competes with other land uses such as residential buildings, agriculture, forestry, water supply from ground water, and recreation, are frequently the sites of complicated land use conflicts. Integration of resource accounts for sand and gravel with the land accounts may help resolve some of these conflicts through better information on alternative sites and deposits within a region such as the county. Resource accounts for these resource categories are under preparation and will obviously link to the land accounts through land use classes and through geographical coordinates as well as the property/holding number.

62. The national net of sample points in the land accounts system will be used to infer mineral reserves in Norway through the use of geological classes and categories linked to the sample points and empirically-based probabilities correlating geological structures and mineral reserves. The sample points are coded and these probabilities estimated by the Geological Survey of Norway (NGU).

VI.4 Land use and air quality in urban areas

63. In cooperation with the Norwegian Institute for Air Research (NILU), the Central Bureau of Statistics is working to establish a model for the correlations between land use (based on the sample points in the urban areas) and air quality, with the case of Oslo, the capital, as a pilot study. With information on wind and point-based observations on air quality, a simulation model will correlate air quality with characteristics of different areas and land use categories in this city.

VI.5 Geo-accounts and environmental monitoring

64 The basic units of the land accounts system and the water resource accounts, the sample points system and the watercourse register, will be used as a geographical frame of reference to observe, monitor, and link data on environmental quality in general, that is, for air, land and water. As most of the territory under Norwegian jurisdiction is ocean territory, it has even been suggested that the national net of sample points be extended to observe and monitor environmental data in these areas (from meteorological data through sea-water characteristics to the geology of the sea-bed).

NATIONS UNIES | ОБЪЕДИНЕННЫЕ НАЦИИ | UNITED NATIONS

COMMISSION ÉCONOMIQUE POUR L'EUROPE | ЭКОНОМИЧЕСКАЯ КОМИССИЯ ДЛЯ ЕВРОПЫ | ECONOMIC COMMISSION FOR EUROPE

SEMINAIRE | СЕМИНАР | SEMINAR

CES/SEM.12/R.4
ENV/SEM.14/R.4
July 1980

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

ECONOMIC COMMISSION FOR EUROPE

SENIOR ADVISERS TO ECE GOVERNMENTS ON ENVIRONMENTAL PROBLEMS

Seminar on Environmental Statistics
Warsaw (Poland)
(16-19 September 1980)

POINT SAMPLING IN NORWEGIAN LAND USE AND ENVIRONMENTAL STATISTICS

Prepared by the Central Bureau of Statistics of Norway*)

I. INTRODUCTION

1. An important part of the Norwegian work with resource accounting is to establish land accounts.**) The purpose of the land accounts is to present information about actual land use, as well as planned land use and land capability (potential land use). The land accounts will supply data in terms of land balances and statistics of annual changes in land use. The accounting system itself defines the relationship between different types of land information. It also specifies how new balances can be produced by employing different sources of land use statistics. This is a complicated matter because of the various classifications and data collecting methods that exist today. Agricultural censuses, for instance, differ considerably from forestry surveys and censuses of population and housing in terms of classification methods adopted.***)

2. In order to establish a statistically sound and firm base for the land accounts, we had to develop a separate system of data collection. The system consists of a national grid, in which land use and other related information is classified or measured at each point of intersection. Within those parts of the country composing an intensive and complicated pattern of land use, the grid is closely spaced, whereas in other parts, such as mountainous areas, the distance between the points is larger. This approach is similar to a standard statistical stratification procedure, being applied in different types of statistical surveys on population, employment, etc.

---

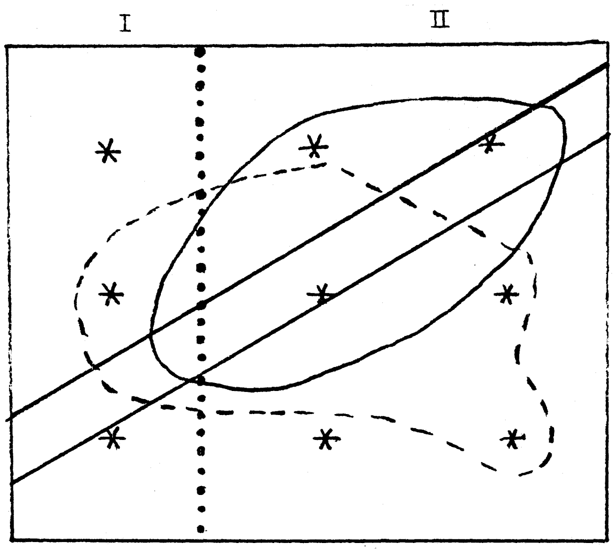
*) Prepared by P.A. Garnåsjordet and H.V. Sæbø, Unit for Resource Accounting, Central Bureau of Statistics of Norway.

**) An overall description is given in: 'A system of resource accounts. The Norwegian experience', OECD, Paris 1980.

***) For a more thourough discussion, see: 'The Provision for Linkages to various Data Systems in the Development of Land Use Statistics in Norway', CES AC. 52/5, Paper prepared by the Central Bureau of Statistics, Norway, to the ECE/CES meeting on land use statistics, in Geneva 17 - 20 March 1980.

3. Reduction of cost is one reason for adopting the method of point sampling. Another important reason is that point sampling makes it possible to link various information on land use, such as information on soil, vegetation, altitude and physical structure. This is of special importance in order to present statistical tables and analyses. Figure 1 below may illustrate how easy this is done by point sampling.

Figure 1. An example to demonstrate point sampling.



——— Boundary for an agricultural field
-------- Boundary for a sand and gravel deposit
.... Administrative boundary
I, II Different administrative units

═══ Agricultural Road
* Point of observation, grid system.

4. By obtaining information about a sufficiently large number of sample points, we will, for instance, be able to state the extent of agricultural land within an area, as well as to say something about the physical structure of the land. Examples are: Land being used for technical purposes (roads, fences, farm-yards a.s.o.) and land use that will prevent other types of resource utilization, such as sand and gravel pits. In the example above the amount of land in the first category within the administrative area marked II, would be calculated as 2/6 = 1/3, while the amout of land that is used for agricultural purposes and which could be used for sand and gravel pits can be calculated to 1/6 in area II. Given many points it is obvious that point sampling is much faster than measuring each small land use configuration.

5. The following tasks have been given priority in the continous work on land accounting in Norway:

A. To produce a comprehensive land use statistics for urban areas (settlements).

B. To carry out a survey of total national land use.

C. To produce land account for pilot counties.

In order to achieve the two former aims, we have employed point sampling.

6. The first part of this paper describes how the method is being used in practice, in order to demonstrate the potentialities of the point sampling method. The second part of the paper will discuss the sampling method from a statistical and an analytical point of view.

II POINT SAMPLING IN URBAN AREAS

7. The land use patterns in urban areas are rather complicated. Various different classification systems have been developed. The Norwegian system of land use classification is fairly close to the new 'Draft international classification of land use', suggested by the ECE in March 1980. The purpose of the system is to classify 'homogenous' areas down to a minimum of approximately 0.1 hectar.

8. Physical planning is not consentrated at one single geographical level. Planners often deal with rather detailed information, such as, for example, building densities, as well as the actual physical structure (roads, parking lots etc.). At the regional level, planners will use information of the character of an area, neighbouring fields, etc. Examples of such data on different geographical levels will be discussed later on.

9. The Sarpsborg/Fredrikstad settlement is situated in Østfold county in the south-eastern part of Norway. In 1970 the settlement counted approximately 80 000 inhabitants. Figure 2 shows the Sarpsborg/Fredrikstad conurbation. The settlement is covered by a grid 100 m spacing, which gives 6 898 sample points.

10. In order to demonstrate the classification procedure two small areas are used as examples. The location of those areas is shown in figure 3.

11. Figure 4 is an air photograph of one of these areas (in Sarpsborg), around a factory which consists of various elements: buildings, roads, railroad tracks, storage area, and as a matter of fact, agricultural land being situated in between the roads to the factory. This land is under cultivation, illustrating how complicated the land use pattern is in practice.

12. Figure 6 shows an area in Fredrikstad covering:

- the elder town with an old fort
- post-war residential areas south of the elder town
- post-war industrial areas " " " " "

Important matters are how the different elements within the area ought to be classified, as for the walls of the elder town, the moat surrounding it, or the wood situated close to the industrial sites. The example illustrates parts of the reason why ordinary land use statistics lacks rather detailed information. It would be a tremendous job to delimit all the small parcels (with homogenous land use), and to measure them either manually or to digitize them on maps. Such problems are solved by adopting point sampling methods.

Figure 2. The location of the Sarpsborg/Fredrikstad settlement. Scale 1 : 1000 000.

Urban settlements.

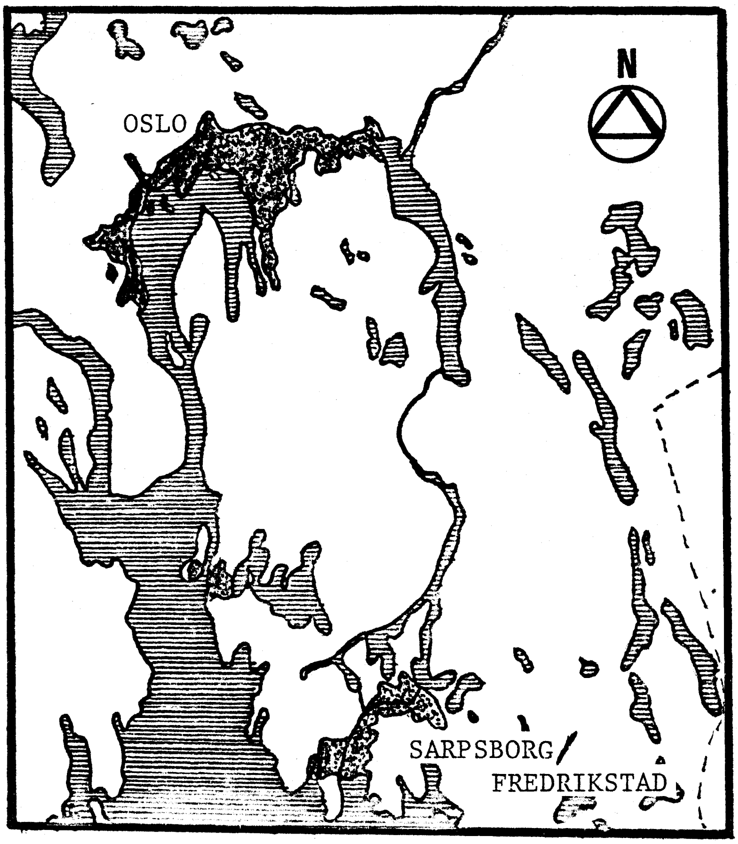
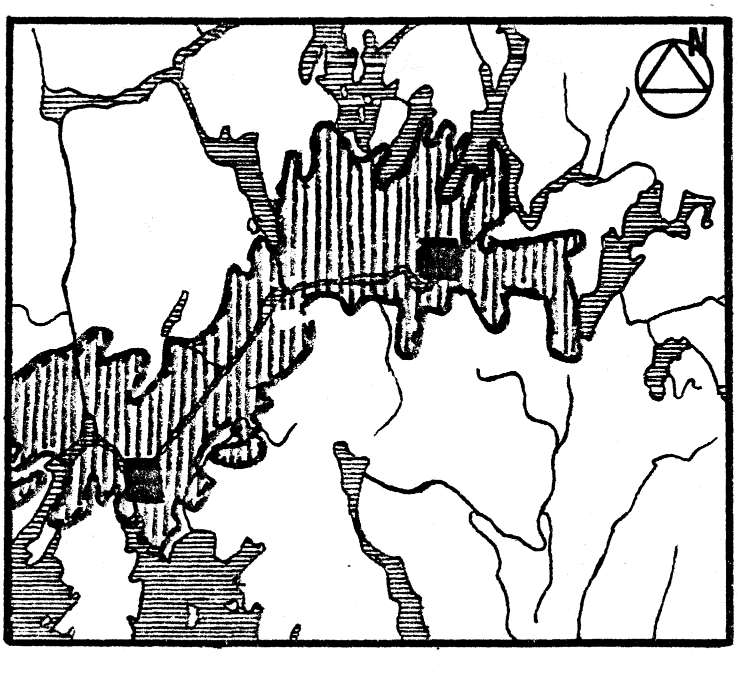


Figure 3. Location of small areas used as examples in this paper. Scale 1 : 250 000.

Boundary for the urban settlement.



13. The corresponding maps (figures 5 and 7) show how the grid is placed. One will see the actual points in which land use and other interesting characteristics are classified. (A brief description of the classification system is presented in appendix 1). The classification is carried out on air photographs in the same scale as the maps. However, air photos will have slightly different scales in different parts of the picture. This is accounted for by placing a transparent map foil on the air photo, adjusting manually in order to match the map and the photo. The interpretators use illuminated tables to ease the work. In the classification process, information from maps and other sources is used as a supplement.

14. As an example, we will try to go trough the classification at four points, two on each of the figures 5 and 7. In figure 5 we classify:

Point 1A: 03 021 03

Classification level 1 (area level):
Manufacturing and warehousing area (code 03).

Classification level 2 (field level):
Manufacturing and/or warehousing field (code 021).

Classification level 3 (physical structure and surface):
Land with improved surface and open space for other economic activities than agriculture (code 03).

Point 1B: 07 081 07

Classification level 1:
Agricultural area (code 07)

Classification level 2:
Agricultural field (code 081)

Classification level 3:
Cultivated land (code 07)

In figure 7 we classify the two example points as:

Point 2A: 06 051 09

Classification level 1:
Area of institutions and improved open space (code 06)

Classified level 2:
Public park (code 051)

Classification level 3:
Other land with soil cover (code 09).

Point 2B: 01 066 04

Classification level 1:
Residential area with mainly one- and two-storey buildings (code 01)

Classification level 2:
Local road (code 066)

Classification level 3:
Elongated land, improved or stabilized (line-shaped) (code 04).

Figure 4. Air photo of the study area in Sarpsborg. Scale 1 : 10 000. 1A and 1B are sample points used as examples.

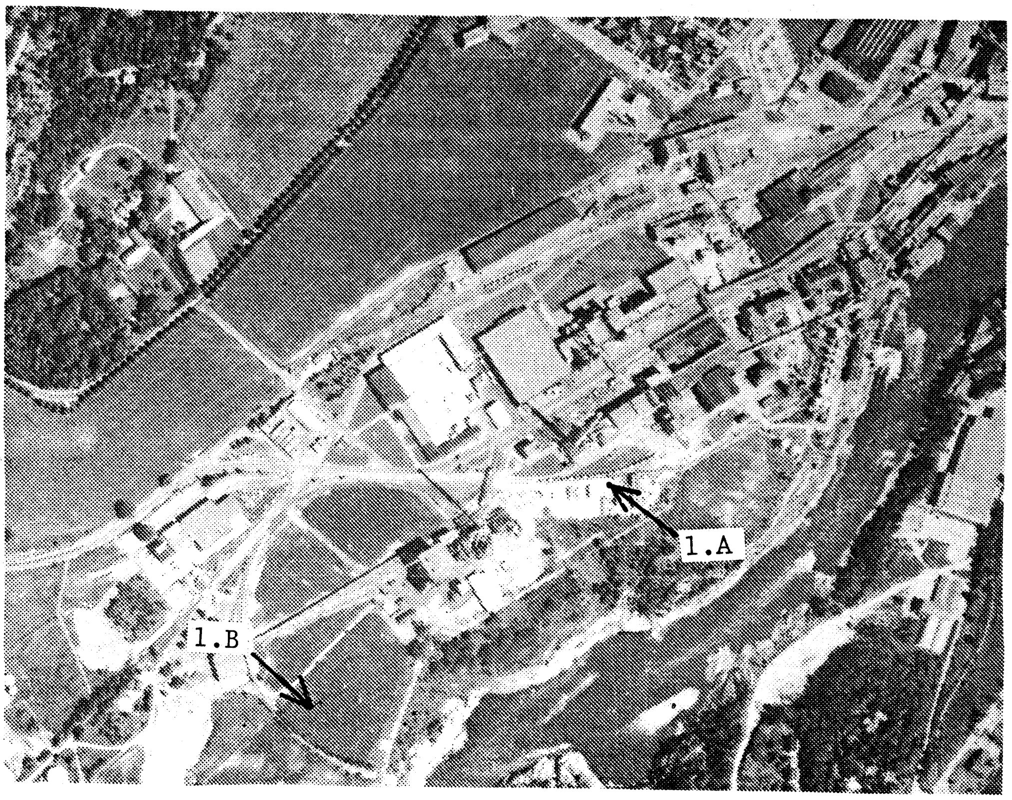


Figure 5. Map of the study are in Sarpsborg. Scale 1 : 10 000. 1A and 1B are sample points used as examples.

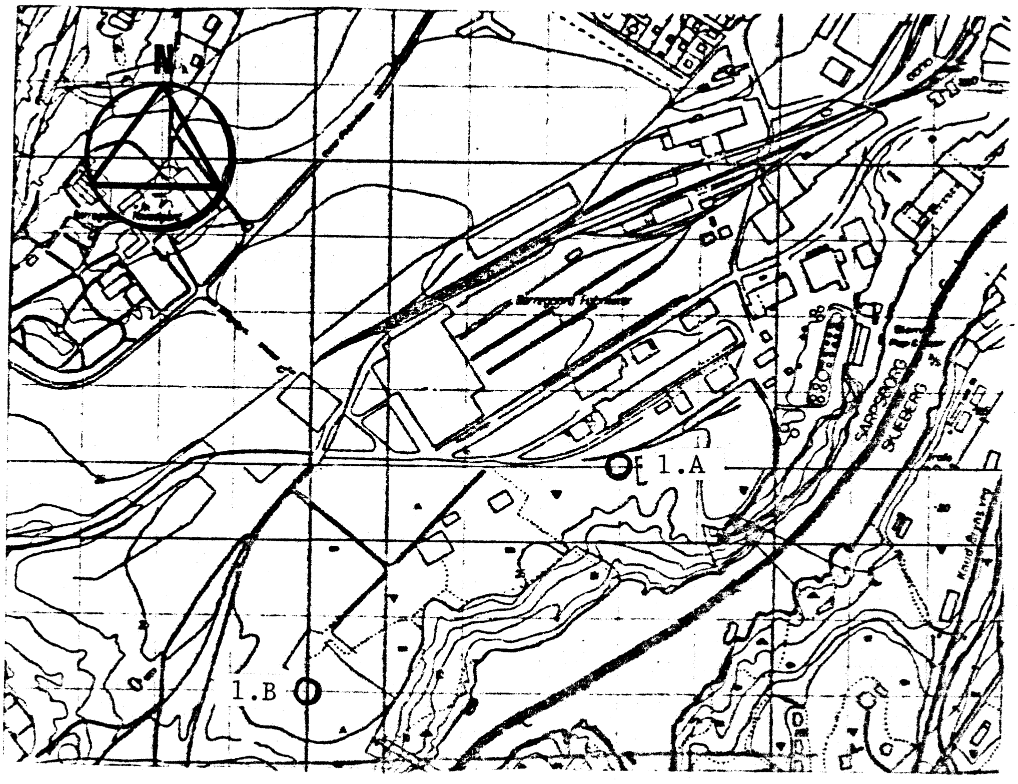


Figure 6. Air photo of the study area in Fredrikstad.
2A and 2B are sample points used as examples.
Scale 1 : 10 000.

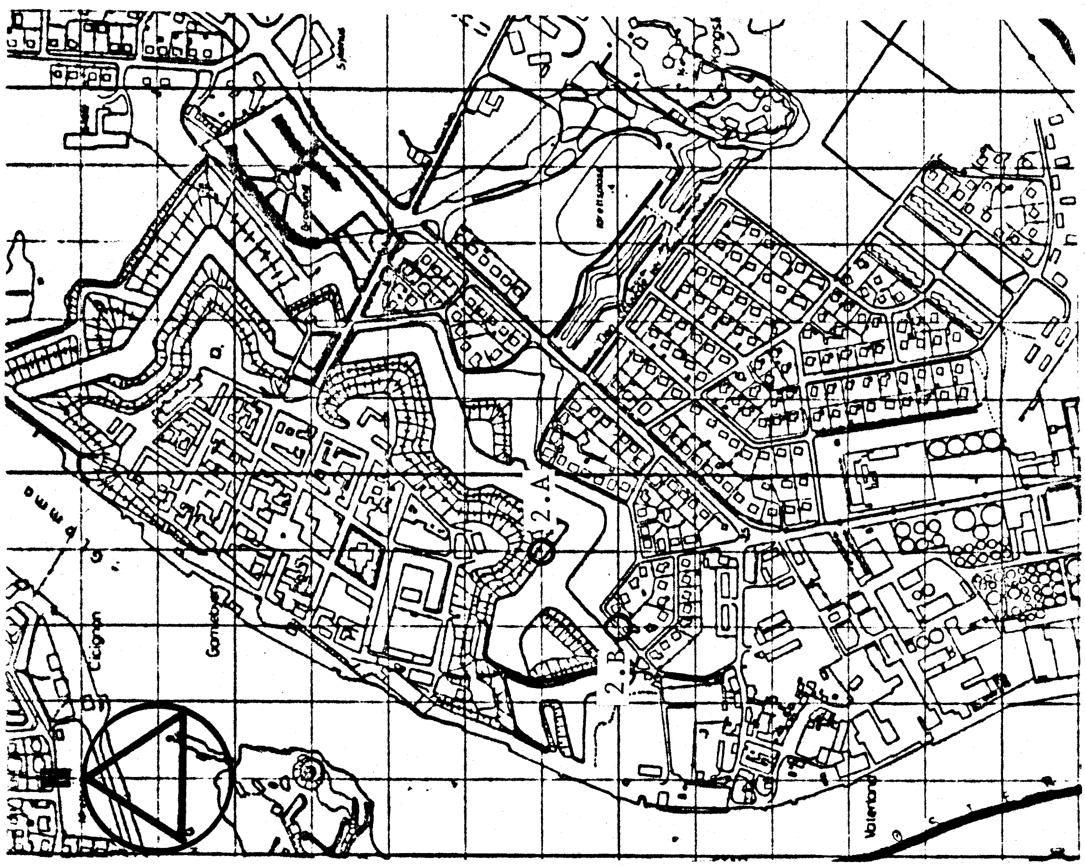


Figure 7. Map of the study area in Fredrikstad.
2A and 2B are sample points used as examples.
Scale 1 : 10 000.

15. A brief review on the low altitude photographs will show that the classification seems correct. In some places, however, as for instance in the central parts of the urban areas, the classification must be verified by direct observation. The confirmation is carried out by the local municipalities on request from the Central Bureau of Statistics. (Applies to approximately 20 larger urban areas.) The municipalities of these areas have the manpower to carry out this work, as well as being interested in obtaining relevant land use statistics for the urban areas in question.

16. In all, approximately 250 urban areas will be classified for the years 1955, 1965 and 1975 (or years as close as possible if air photos for the years in question do not exist). The classification at all three points of time have been carried out simultaneously in the same points. This reduces classification errors and gives reliable information on changes in land use. (See paragraph IV.)

17. The data will make possible to analyze the changes in land use in Norwegian urban areas on a detailed level, for instance by grouping according to size, physical properties, geographical location etc. This will enable us to develop a pattern of causes and effects in order to explain land use development.

18. We wish to present land use statistics for individual urban areas. For the smaller urban areas, however, it will be possible to present statistically reliable information only for the largest land use classes. For conurbations as Sarpsborg/Fredrikstad (6 898 points), it is possible to present rather detailed information. We shall present some main results. Table 1 shows the land use in 1955, 1963 and 1975 in Sarpsborg/Fredrikstad (within the 1975 borderlines).

19. Maps will be produçed to present supplementary information on large urban areas. The maps are not accurate, as they are produced using point sampling data. They will, nevertheless, give an impression of land use patterns. An example of such a map is shown in the next paragraph (figure 8).

20. The table shows that the growth of built up land in the region has declined, which is the result of a stagnation in the manufacturing industries and in population growth. It is, however, interesting to notice that the main cause to the decrease is the reduced growth of residential land, whereas the other categories of built-up land seem to have almost identical growth within the two periods.

Table 1. Land use in Sarpsborg/Fredrikstad. 1955, 1963 and 1975.

	Land use 1955	Annual growth 1955-1963	Land use 1963	Annual growth 1963-75	Land use 1975
	ha	per cent	ha	per cent	ha
Land within the settlement boundary in 1975	6 898		6 898		6 898
Built-up land (011-073)	2 159	3.0	2 732	2.2	3 564
Residential (011-013)	1 156	3.7	1 546	2.3	2 028
Manufacturing (031-032)	396	1.8	457	1.9	571
Services and city-center (021-022)	74	2.2	88	2.3	115
Institutions (041-053)	217	2.9	272	2.0	345
Technical and communications (061-073)	316	2.0	369	2.0	505
Agricultural land (081)	1 969		1 713		1 314
Forest land (091)	1 138		984		712
Other non-built-up land (101)	1 116		967		824
Water (111)	516		502		484

21. In order to understand and to explain such reduction in growth, it is necessary to examine the changes in land use. Table 2 shows changes in land use from non-built-up categories of land use into various built-up categories for two time-periods. Such tables can be regarded as parts of more detailed 'input-output' tables of land use changes.

Table 2. Urban development (changes from non-built-up to built-up land use classes at the field level) 1955 to 1963 and 1963 to 1975 in Sarpsborg/Fredrikstad. ha

To (1963) From (1955)	Total	Residential	Manufacturing etc.	Services and city-center	Institutions	Technical and communications
Total	594	403	72	13	62	44
Agricultural land	176	105	25	6	25	15
Forest land	156	138	3	-	7	8
Other non-built-up land	262	160	44	7	30	21

To (1975) From (1963)	Total	Residential	Manufacturing etc.	Services and city-center	Institutions	Technical and communications
Total	872	523	147	16	76	110
Agricultural land	302	125	75	10	33	59
Forest land	278	241	5	-	11	21
Other non-built-up land	292	157	67	6	32	30

22. The table illustrates the fact that Norwegian land use policy of conserving agricultural land has had little success in this region. To some extent residential development has been forced to take place on forest land as opposed to agricultural land. Other built-up categories, however, have the same 'consumption pattern' in the last period as in the first one, taking mainly agricultural land and other non-built-up land.

23. Land use policy in Norway during the 1970's has been forcused upon agricultural land being used for built-up purposes. Therefore it seems rather strange that the reduction in agricultural land within the boundaries of the settlements has been larger during the last period compared to the first.

24. The registration of land use in urban areas from air photoes for the three years 1955, 1965 and 1975 is done for three separate levels, the area level (residential areas, industrial areas, business and commercial areas etc.), field (or parcel) level, and a lowest level within fields, which we have called physical structure. Each sample point is thus classified according to the physical surface or structure, to the field, and to the area class surrounding the sample point. The classification system is given in appendix 1. This procedure makes it possible to combine data on different dimensions and geographical levels of land use.

25. Land use statistics based upon point sampling makes it possible to calculate, for instance, physical structure within residential land use. Table 3 below shows building densities and land stabilized for traffic purposes etc.

26. Table 3 shows that building densities as well as 'traffic' desities have increased 1-2% within the residential fields during the last two periods. The increase in 'traffic space' is of special interest to the environmental authorities. To some extent we can explain this by comparing the figures with the trends within residential areas. Table 4 below shows the development of land use for communication purposes within residential areas.

Table 3. Buildings and artificial stabilized ground at the point level as percentages of residential land at the field level. Sarpsborg/Fredrikstad.

	Buildings	Artificially stabilized ground
1955	8.7	8.3
1963	9.0	8.6
1975	10.0	10.3

27. Table 4 shows no specific trend as regards communications. The conclusion is simple: The residential areas have the same need for traffic lines, that is traffic to and from the regions as before, but within each parcel which is classified as residential land, more space is used as parking ground and smaller roads! Data from other Norwegian urban areas also indicate that this conclusion is a correct one.

Table 4. Land for communication purposes at the field level as a percentage of residential land at the area level. Sarpsborg/Fredrikstad.

1955	11.0
1963	10.7
1975	11.0

28. It is difficult to estimate the production costs of this type of land use statistics, including:

- interpretation
- supervision and control
- acquisition of maps
- reproduction of air photos
- presentation of results

Some of the work has been done free of charge by the Geographical Survey and the municipalities. The total costs seem, however, to be about 1 mill. norwegian kroner, or approximately 5 kroner or 1 US$ per point. For this rather low cost, we can produce land use statistics of high quality covering approximately 1 200 km^2 (460 sq.miles) of built-up land with a population of 2.4 mill. people.

29. The classification will be completed this autumn, and the first results will be published March 1981. Preliminary analyses will be available at the same time.

30. In our view, production of this sort of land use statistics should be repeated every 5 or 10 years. One solution may be to produce some statistics for the largest urban areas every 5 years, whereas statistics for the smaller areas could be revised every 10 years.

31. The accuracy of such statistics seems to correspond to that of ordinary statistics concerning population, occupation and employment. With reasonably skilled assistants and a fair degree of supervision, the classification error will be approximately 10 per cent. The sampling error, which comes in addition, is treated in paragraph IV. In medium sized urban areas, the classification error and the sampling error will be of the same order of magnitude.

III NATIONAL LAND USE STATISTICS

32. In order to provide national land use statistics within a very short period of time, 1½ year, the point sampling method has been employed to collect numerous data describing land use, geology, vegetation, terrain, landscape and geographical location. The data collected at each point are described in appendix 2. The sampling procedure is based on a national grid, consisting of approximately 7 000 points. We work with two grid systems. Above the forest line*), the mountainous area, the distance between sample points is 12 km. Below the forest line the grid spacing is 6 x 6 km.

33. Sources of information are maps in different scales and coverage, and to some extent air photographs. Various institutions working with vegetation mapping, geological mapping etc. have cooperated in terms of producing data for various areas for which maps have not yet been produced. Thus we will have a complete set of data for almost all sample points, for which ordinary mapping would have been too expensive.

34. For one county, Østfold, situated in the south-eastern part of Norway, some results are available. This is a pilot county, and the grid is 3 x 3 km in this case (there are no mountain areas in the county), which gives 464 points. Different tables can be produced with data derived from this base. Table 5 shows the main land use classes in Østfold in 1975.

35. As mentioned in connection with the data on the urban settlement, maps can be uses to illustrate the data and give som idea of the land use pattern (figure 8).

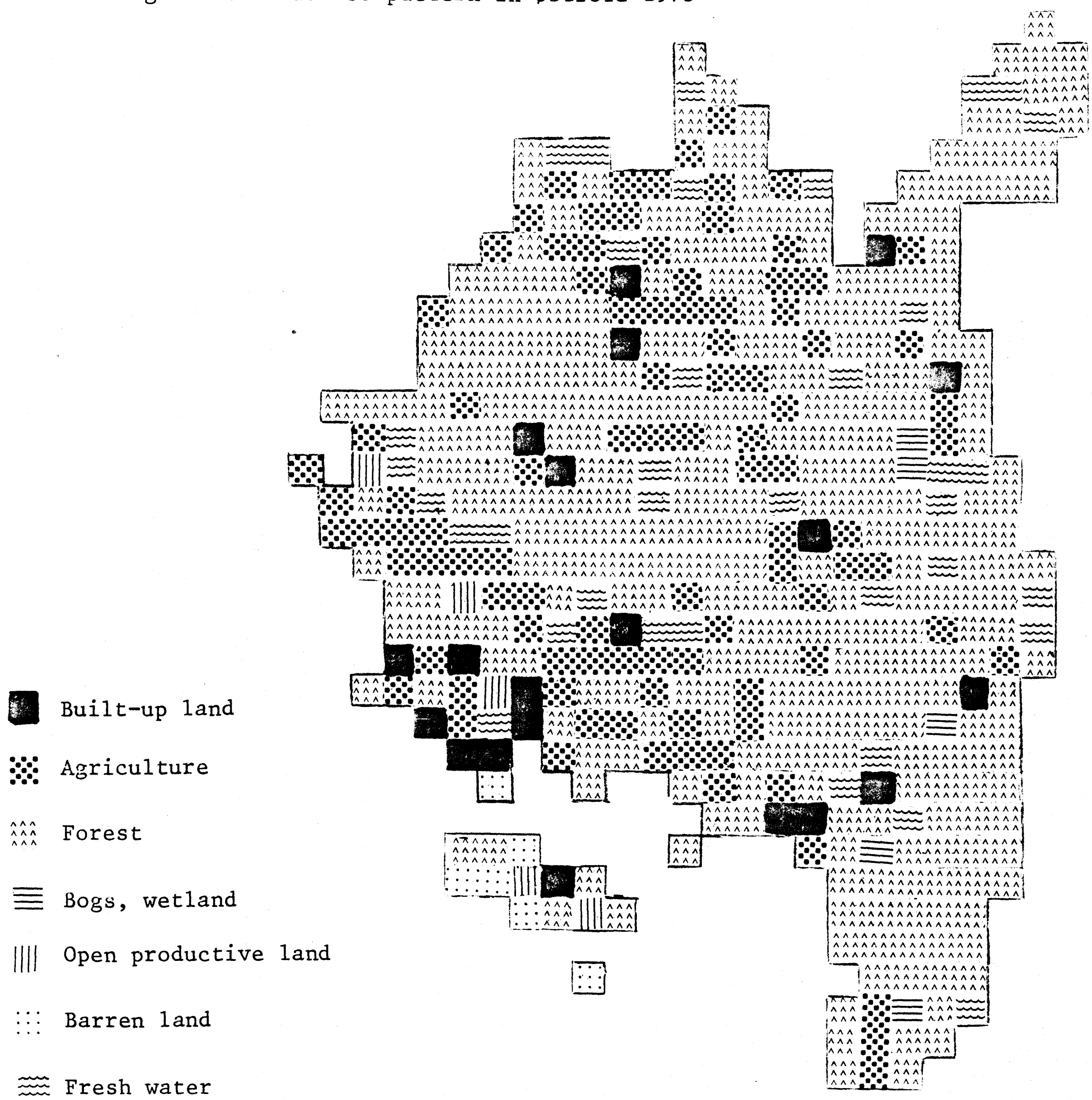
---

*) In practice this is chosen as an area for which the agricultural authorities have not found it necessary to produce detailed maps (Economic maps in scale 1 : 5 000).

Table 5. Land use in Østfold 1975.

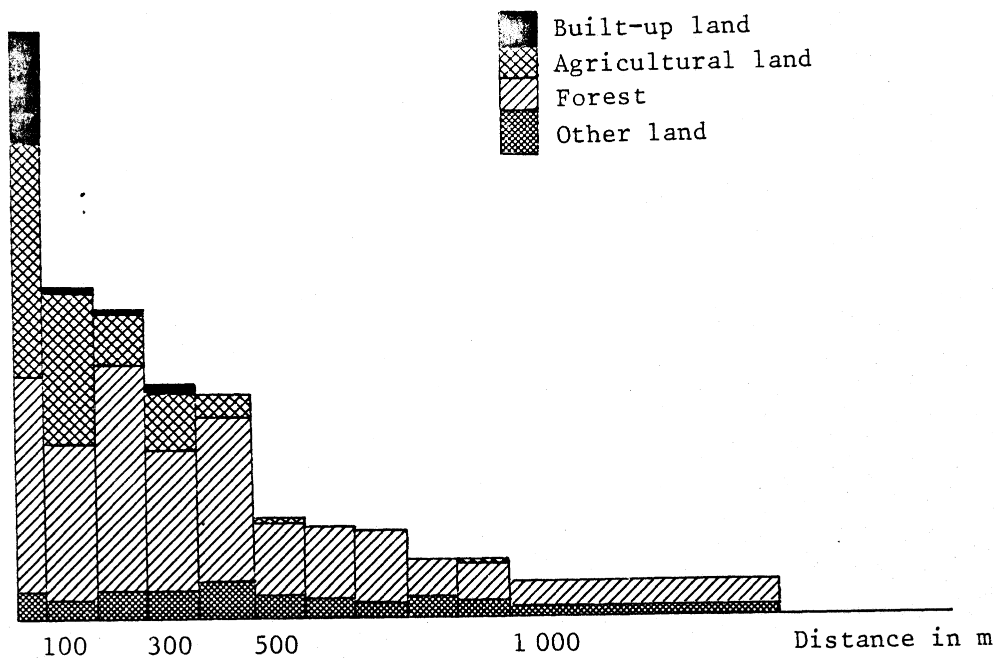
	Number of points	Km^2	Per cent
Built-up land	16	144	3.4
Agricultural land	96	864	20.7
Forest	297	2 673	64.0
Bogs, wetland	5	45	1.1
Open productive land	8	72	1.7
Barren land	6	54	1.3
Fresh water	36	324	7.8
T o t a l	464	4 176	100.0

Figure 8. Land use pattern in Østfold 1975



36. An interesting point is that he registrations will enable us to estimate the extent of wilderness being left in the different parts of Norway. Up to now, it has been difficult to produce an exact and operational definition of wilderness. From a philosophical point of view, we would like to define wilderness as free nature, which means that the rythms should be exclusively governed by nature itself. The crucical question is the extent of impacts in terms of distances to roads, railroads, dams, power-lines, or settlements will have upon the natural cycles. We think it is almost impossible to give a general definition for different landscapes, vegetation, altitudes etc. for the whole country. By employing point sampling, we can solve this problem: Instead of applying only one definition based upon a set of distance measures, we can collect data on the distances from each sample point to different types of man-made installations. The figure below shows how distance to roads vary within the county for different categories of land use.

Figure 9. Distance to roads for different categories of land use in Østfold 1965



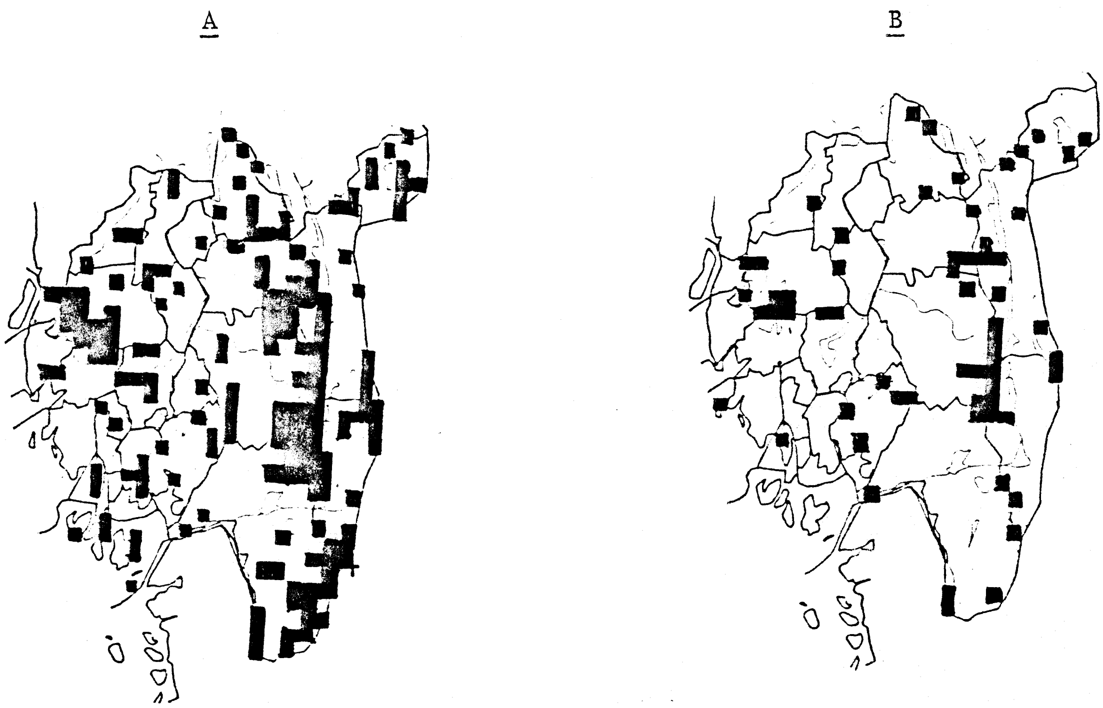
37. The two maps below show the extent of wilderness in Østfold county given two possible definitions:

A. That the area is situated more than 500 m away from roads, railways, power-lines and regulated lakes or rivers.

B. That the area is situated more than 1 km away from roads, railways, power-lines and regulated lakes or rivers.

This illustrates how a number of various definitions can be made operational by using data from the point sampling.

Figure 10. Two possible definitions of wilderness in Østfold 1965.



IV THE PRECISION OF POINT SAMPLING

38. Point sampling may be described in the light of the mathematical theory of sampling. In fact, point sampling on maps and air photos is analogous to sampling people or households in surveys presenting information on social and economical conditions. However, one issue turns out to be important concerning sampling in space, namely the correlation between adjacent points. This kind of correlation is important for the precision of the sampling techniques to be used. It implies that the presision of systematic sampling usually is better than in a random sampling procedure. Unfortunately, it is difficult to take the correlation into account in a simple and exact way when making and comparing sampling designs.

39. In this paper we will first refer to some results about the precision of point sampling , which may be deduced mathematically without considering correlations in space. These remarks concern random sampling procedures (simple random sampling and stratified random sampling). As regards simple random sampling, we have as well given some formulas expressing the uncertainity in estimations of changes.

40. The correlations in space will be taken into account in the next place, and we shall see how they affect the uncertainity in estimation. In addition, estimates of these correlations may be used in analyzing geographical structure. Examples of this will be examined in the next paragraph.

41. Let us assume that we want to estimate the share of the total area A that belongs to a certain land use class (for instance agricultural fields). Let n represent the total number of points in the sample. The points are numbered $j = 1, 2, \ldots, n$ as they are sampled. We define a random variable X_j which follows a binomial distribution:

$$X_j = \begin{cases} 1 & \text{if point } j \text{ belongs to the class we want to estimate,} \\ 0 & \text{otherwise.} \end{cases}$$

42. In simple random sampling $EX_j = p$ and $\text{var } X_j = p(1-p)$. An unbiased estimator for p is given by

$$\hat{p} = \frac{1}{n} \Sigma X_j. \qquad (1)$$

The variance of this estimator is

$$\text{var } \hat{p} = \frac{p(1-p)}{n}. \qquad (2)$$

This gives a standard deviation

$$s = \sqrt{\frac{p(1-p)}{n}}. \qquad (3)$$

Variance and standard deviation usually have to be estimated, the values for p are then replaced with $\hat{p}$ and those for n with (n-1) in (2) and (3).

43. We may now estimate a confidence interval for p, that is an interval including the correct calue with a certain probability. If np(1-p) > 10, $\hat{p}$ is approximately normally distributed, and an approximate 95 per cent confidence interval is given by

$$\left[\hat{p} - 2\hat{s},\ \hat{p} + 2\hat{s}\right], \qquad (4)$$

where $\hat{s}$ is the estimate for s. Table 6 shows the confidence interval for different values of $\hat{p}$ and sample size n = 100, 1 000 and 10 000.

Table 6. 95 per cent confidence interval of area shares in simple random point sampling. Per cent

Estimated class	Confidence interval at n points		
	n=100	n=1 000	n=10 000
50	40-60	46.9-53.1	49.0-51.0
20	12-28	17.5-22.5	19.2-20.8
10	4-16	8.1-11.9	9.4-10.6
5	1.6-10.8	3.6- 6.4	4.5- 5.5
3	0.6-8.3	1.9- 4.1	2.6- 3.4
2	0.2-6.9	1.1- 2.9	1.7- 2.3
1	0-5.4	0.4- 1.6	0.8- 1.2
0	0-3.6	0- 0.4	0- 0.1

44. In practice, the points are placed in a systematic, regular grid. In this case the confidence intervals in the table usually provide upper limits for the real confidence intervals, and the variance formula (2) and the table must be used for this purpose. The reason for this is the correlation between adjacent points in space (and on the map). Since land use classes have a certain size, sampling of two adjacent points would be of little value, each of them hardly representing new information compared to the other. Sampling of adjacent points is as far as possible avoided by systematic sampling in a regular grid.

45. There is, however, a more efficient random procedure, too, namely random stratified sampling. Let us assume A divided into r known strata $A_1, A_2, \ldots, A_r$. In each stratum, s, we estimate the share of the actual class p_s with the help of n_s points. The estimator for p_s, $\hat{p}_s$, is the share of points of the class in question in stratum s. We have

$$p = \frac{1}{A} \sum_s p_s A_s . \qquad (5)$$

An unbiased estimator for p is

$$p^* = \frac{1}{A} \sum_s \hat{p}_s A_s. \qquad (6)$$

It may be shown that

$$\text{var } p^* = \frac{1}{A^2} \sum_s A_s^2 \frac{p_s(1-p_s)}{n_s} . \qquad (7)$$

46. Assume n_s is selected proportionally to the area A_s,

$$n_s = \frac{A_s}{A} n. \qquad (8)$$

It may be shown that proportionally stratified random sampling is consequently more efficient than simple random sampling. We have such a kind of stratification if we, for instance, divide A into equal squares and sample one random point within each square (stratum). In this case the variance is

$$\text{var } p^* = \frac{1}{n^2} \sum_s p_s(1-p_s). \qquad (9)$$

This expression may be used to give an approximation for variance in systematic sampling with a square grid of points. Unfortunately, this formula may not be estimated without knowledge of the structure of area which goes beyond the data from the point sampling (p_s is unknown).

47. We will note that the variance (7) is minimized if

$$\frac{n_s}{n} = \frac{A_s\sqrt{p_s(1-p_s)}}{\sum_s A\sqrt{p_s(1-p_s)}} . \qquad (10)$$

To obtain the best estimator of a certain class, we consequently have to increase the number of sampling points in strata where

i) the area A_s is large,

ii) $p_s(1-p_s)$ is large.

48. Suppose now we have the estimates $\hat{p}_1$ and $\hat{p}_2$ for the same area class, but measured at two different points of time. The estimates of the change Δ then is:

$$\hat{\Delta} = \hat{p}_2 - \hat{p}_1,$$

$$\text{var } \hat{\Delta} = \text{var } \hat{p}_2 + \text{var } \hat{p}_1 - 2 \text{ cov } (\hat{p}_1, \hat{p}_2). \tag{11}$$

If the points are sampled independently of each other at the two points of time, the covariance disappears, and the variance of the change is the sum of the variance of the shares. If we use the same points, we will usually have a positive covariance, and the variance of the change estimate is reduced.

49. If the points are sampled randomly once and for all, it may be shown that

$$\text{var } \hat{\Delta} = \frac{p_2(1-p_2)}{n} + \frac{p_1(1-p_1)}{n} - \frac{2}{n} (q-p_1p_2), \tag{12}$$

where q is the share of the area belonging to the class at both times of measurement. If the points are placed in a systematic, regular grid, the formula usually provides an upper limit for the real variance.

50. If the changes are small ($q>p_1p_2$), the variance of the change is reduced when using a fixed net. If $q >> p_1p_2$ and in addition, $p_2 >> p_2^2$ and $p_1 >> p_1^2$, we have

$$\text{var } \hat{\Delta} = \frac{1}{n} (p_1+p_2-2q). \tag{13}$$

51. Land use class changes are often one-way changes, thus making $q = \min. \{p_1,p_2\}$. An example is usually the class of built-up land. This class is increasing at the expence of most other land use classes, and virtually no built-up land is transferred to other classes. In such a case $q = p_1$ and we of course have

$$\text{var } \hat{\Delta} = \frac{\Delta(1-\Delta)}{n}. \tag{14}$$

52. In the example from Sarpsborg/Fredrikstad, the built-up land in 1963 was estimated to 39.6 per cent, and in 1975 to 51.6 per cent. The change was 12 per cent. The formula gives a standard deviation for this change of 0.4 per cent and the confidence interval is 11.2 - 12.8 per cent.

53. Table 7-8 shows the standard deviation $s(\hat{\Delta}) = \sqrt{\text{var}\hat{\Delta}}$ in estimating different changes. Number of points n = 500. The formula (11) is symetrical in p_1 and p_2, and we choose $p_2 > p_1$ in the tables. The standard deviations are calculated for a fixed point net and compared to the standard deviations we would have found if the points had been sampled independently of each other at the two points of time. The standard deviation for other values of n can be calculated by employing $s(\hat{\Delta}) \sim 1/\sqrt{n}$.

Table 7. Standard deviation of estimated changes with n = 500, p_1 = 10 per cent. Per cent

	Standard deviation with fixed point net				Standard deviation with independently sampled point nets
p_1 = 10	q = 10	9	5	0	
p_2 = 10	0	0.6	1.4	2.0	1.9
11	0.4	0.8	1.5	2.0	1.9
15	1.0	1.2	1.7	2.2	2.1
20	1.3	1.5	1.9	2.4	2.2
50	2.2	2.3	2.6	3.0	2.6

Table 8. Standard deviation for estimated changes with n = 500, p_1 = 50 per cent. Per cent

	Standard deviation with fixed point net						Standard deviation with independently sampled point nets
p_1 = 50	q = 50	49	45	40	20	0	
p_2 = 50	0	0.6	1.4	2.0	3.5	4.5	3.2
51	0.4	0.8	1.5	2.0	3.5	.	3.2
55	1.0	1.2	1.7	2.2	3.6	.	3.2
60	1.3	1.5	1.9	2.4	3.7	.	3.1
80	2.0	2.1	2.5	2.9	.	.	2.9
100	2.2	.	.	.	.	.	2.2

54. In systematic sampling the relative positions of the points are given. Since the random procedure were discussed on the basis of randomness only in the selection of points, we now have to use another model. The correlation between points in space must be taken into account.

55. Let Z(x) be a random variable with values depending on the position x in space*). Z may for instance denote temperature or height above sea level. In the case of considering land use shares, we may define Z in the plane:

$$Z(x) = \begin{cases} 1 \text{ if the point x belongs to the class,} \\ 0 \text{ otherwise.} \end{cases}$$

*) Such a variable is called a regionalized variable. The theory of regionalized variables is mathematically formalized by Matheron: 'The Theory of Regionalized Variables and its Applications'. Ecole Nationale Superieure des Mines de Paris (1971).

The share p to be estimated is

$$p = \frac{1}{A} \int_A Z(x)dx \tag{15}$$

and the estimator to be used is

$$\hat{p} = \frac{1}{n} \sum_j Z(\mathbf{x}_j). \tag{16}$$

$Z(x_j)$ denote the value of the variable in point number j.

56. Let us consider the variable Z in points x and x + h (x and h denote vectors in two dimensions).

Since Z(x) is defined as a binomial variable, we have

$EZ(x) = Pr(Z(x)=1)$, and the covariance

$$cov(Z(x),\ Z(x+h)) = E(Z(x)Z(x+h)) - EZ(x)EZ(x+h)$$

$$= Pr(Z(x)=1 \cap Z(x+h)=1) - Pr(Z(x)=1)Pr(Z(x+h)=1). \tag{17}$$

Here $Pr(Z(x)=1 \cap Z(x+h)=1)$ is the probability for both x and x+h to be in the class.

57. This probability may be expressed by the <u>conditional probability</u> for x+h to be in the class, given that x is in the class:

$$Pr(Z(x)=1 \cap Z(x+h)=1) = Pr(Z(x+h)=1|Z(x)=1)\ Pr(Z(x)=1). \tag{18}$$

58. We now assume:

1) $Pr(Z(x)=1)$ is constant for all $x \in A$.
2) The covariance between Z(x) and Z(x+h) and the conditional probability $Pr(Z(x+h)=1|Z(x)=1)$ does only depend on the relative position h of the two points (or the distance $|h|$ in the case we are considering as isotropic phenomena).

These assumptions enable us to estimate the probabilities mentioned. In fact, $\hat{p}$ is an unbiased estimator for the share p and $Pr(Z(x)=1)$.*)

59. Both the covariance and conditional probability are measuring the correlations in space. Covariance is the most convenient parameter considering sampling variances. As we shall see in the next paragraph, the conditional probability is adequate in analyzing geographical structure. The combining of (17), (18) and the assumptions gives the covariance

*) In this model p itself is a random variable, but we have $E\hat{p} = Ep = Pr(Z(x)=1)$.

$$K(h) = \mathrm{cov}(Z(x),\ Z(x+h))$$

$$= \Pr(Z(x)=1)\ [\Pr(Z(x+h)=1 \mid Z(x)=1) - \Pr(Z(x)=1)]. \qquad (19)$$

If we assume $K(h) = K|h|)$ (isotropic), the distance $|h|$ where the conditional probability equals the unconditional probability $\Pr(Z(x)=1)$ is called the range of the actual land use class. At this distance the covariance above is zero. Beyond this distance, one point does usually not give information about other points.

60. In the case of sampling in a regular grid it can be shown that the variance of the sampling error approximately is given by *)

$$\mathrm{var}(\hat{p} - p) \approx \frac{p(1-p)}{n} + \frac{1}{n} \sum_{j\neq i} K(x_i - x_j) - \frac{1}{A} \int_A K(x_i - x)dx. \qquad (20)$$

This formula indicates that the variance will be smaller than $\frac{p(1-p)}{n}$ if the term $\frac{1}{n} \sum_{j\neq i} K(x_i-x_j)$ denoting 'mean covariance' between the points in the sample, is smaller than the 'mean covariance' $\frac{1}{A} \int_A K(x_i-x)dx$ between all points in the area A.

61. In a square grid the points are placed as far from each other as possible, and the variance is usually less than $\frac{p(1-p)}{n}$. The exception from the rule appears when the grid coincidies with a regular structure on the ground. If the range ε where the covariance is greater than zero is less than the minimum distance between the points and the covariance is zero elsewhere, the second term in (20) vanishes, and the variance will always be less than $\frac{p(1-p)}{n}$:

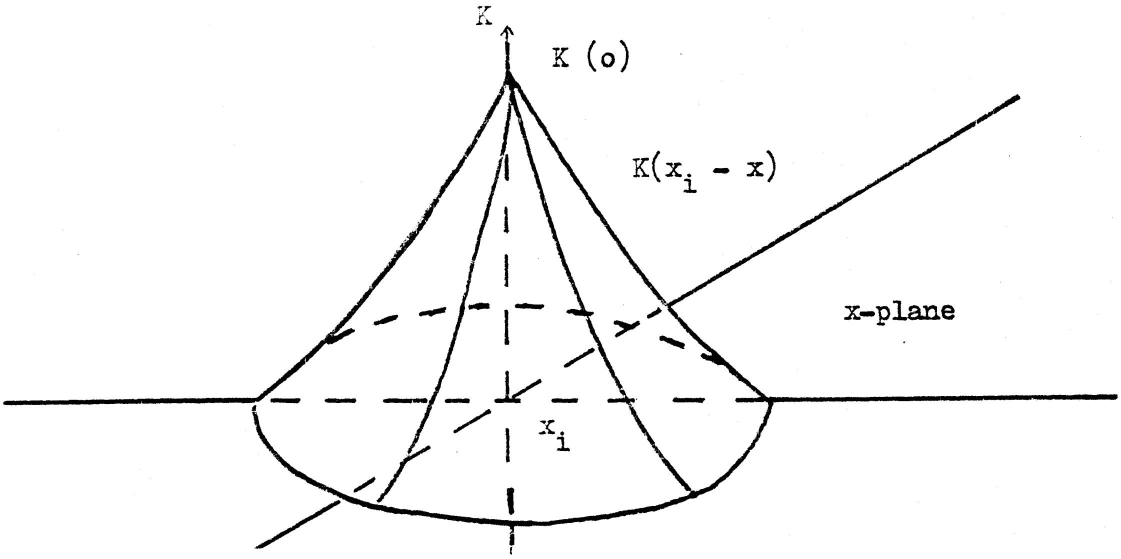
$$\mathrm{var}\ (\hat{p}-p) \approx \frac{p(1-p)}{n} - \frac{1}{A} \int K(x_i-x)dx, \quad |x_i-x| \leq \varepsilon. \qquad (21)$$

---

*) We have assumed that there are no 'border effects', that is $\int_A K(x_i-x)dx$ is independent of x_i.

62. The last term in the formula is proportional to the volume below the covariance function, as shown in figure 11.

Figure 11. The covariance cone about a random point x_i.



63. Unfortunately, the integral in (20) cannot be estimated by the point observations alone. To estimate the variance in systematic point sampling, we have to use a model in addition to point observations. So far we only have few experiences in modelling curves expressing correlations in space for land use classes in Norway. However, simple experiments indicate that the variance reduction in systematic grid sampling compared to simple random sampling may be considerable (more than 50 per cent).

V. ANALYZING GEOGRAPHICAL STRUCTURE

64. The conditional probability of finding a specified land use class in a distance h from a random point in some specified class, is a convenient measure for the correlation between the classes. So far, we have only considered the conditional probability of finding a class starting in the same class. In this paragraph we will examine some examples where we have estimated conditional probabilities of finding different classes starting in one specified class. Here we will assume that the conditional probabilities only depend on the distance between the points. For $h = 0$ the probability of finding the same class is one, the probability of finding any other class is zero. As h increases, the probability of finding the same class will decrease until the range is reached, where this probability equals the overall share p for the actual land use class.

66. In figure 12 we have estimated the conditional probability of finding different land use classes as a function of distance from built-up land in Østfold. The range of built-up land is more than 600 m. We may note, however, that at 300 m it is almost as probable

to find agricultural land as finding built-up land. At this distance it is considerably less probable to find forest land. At distances longer than 600 m from built-up land it is most probable to find forest. This shows that in Østfold built-up land is surrounded by agricultural land. The large forests in the county are located further away from built-up land.

66. Figure 13 shows the conditional probability of finding the different land use classes as a function of distance from residential land in the Sarpsborg/Fredrikstad urban area. The range of residential land is about 500 m. It is relatively probable to find land used for communications nearby the residential land. Other land use classes are situated further away.

Figure 12. The conditional probability of finding land use classes as a function of distance from built-up land in Østfold.

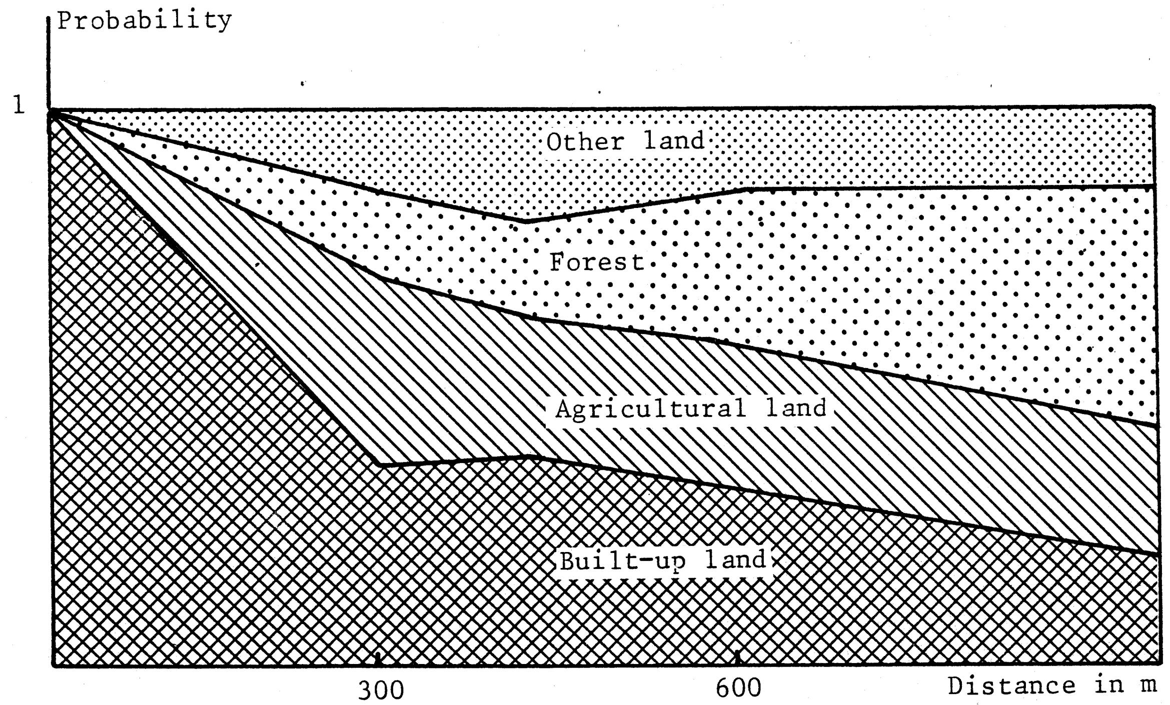
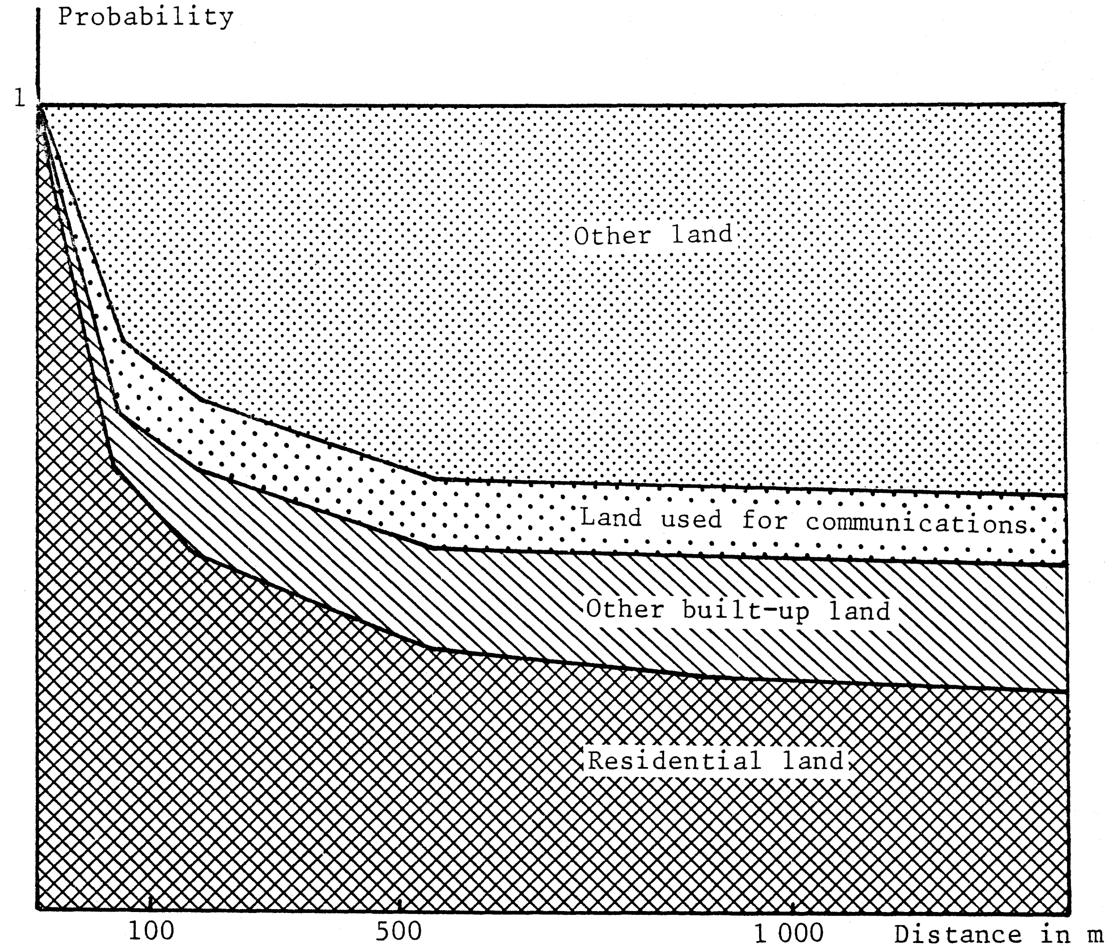


Figure 13. The conditional probability of finding land use classes as a function of distance from residential land in Sarpsborg/Fredrikstad.



APPENDIX 1

LAND USE CLASSIFICATION SYSTEM FOR URBAN SETTLEMENTS

Classification level 1: area level

Code	Area class
01	Residential area with mainly one and two storey-buildings
02	Residential area with mainly multi-storey buildings
03	Manufacturing and warehousing area
04	Central business district and other business areas
05	Communications and other infrastructure areas
06	Area of institutions and improved open space
07	Agricultural area
08	Forest area
09	Area not used for any of the above specified purposes
10	Fresh water area

Classification level 2: field (or parcel) level

Code	Field class
011	Non-built-up parts of residential field with one or two-storey building
012	Built-up parts of residential field with one or two-storey building
013	Residential field with multi-storey building
021	Manufacturing and/or warehousing field
022	Mining and quarrying field
031	Central business district or other business field
032	Mixed residential/business field
041	Institutional field
051	Public park
052	Churchyard
053	Sports facilities field
061	Railway, tramway lines
062	Railway terminal and station field
063	Airports
064	Ports and harbours
065	Throughfare road (Major road)
066	Local road
067	Transport terminals other than above
068	Parking fields
071	Waste disposal and treatment
072	Sewage treatment, discharge, and water supply facilities
073	Energy supply facilities
081	Agricultural field
091	Forest field
101	Not elsewhere specified
111	Fresh water field

Classification level 3: physical structure and surface

Code	Physcial structure class
01	Building
02	Physical and technical constructions other than buildings
03	Land with improved surface and open space for other economic activities than agriculture
04	Elongated land, improved or stabilized (line-shaped)
05	Other improved or stabilized (field-shaped)
06	Gardens and lawns, cultivable parts of park land
07	Cultivated land
08	Steeply sloping land with soil cover
09	Other land with soil cover
10	Steeply sloping land without soil cover
11	Other land without soil cover
12	Sea water
13	Fresh water lake
14	River, 5 m wide or more
15	River, less than 5 m wide

APPENDIX 2

DATA COLLECTED IN THE SAMPLE POINTS IN THE NATIONAL LAND USE STATISTICS

- geograpichal coordinates
- local authority, county
- altitude
- steepness of slope
- aspect
- terrain
- land surface (at the point level)
- distance to the nearest road
- distance to the nearest railroad
- distance to the nearest power-line
- distance to the nearest regulated lake or river
- land use at the field level
- land use in points 300 m north, south, west and east to the registration point
- mainland, island, coast (200 m limit)
- landscape (registered in a circle with radius = 300 m)
- quaternary geology
- geology
- vegetation
- ownership
- protection/conservation

Issued in the series Reports from the Central Bureau of Statistics (REP) ISSN 0332-8422

Printed 1980

No. 80/1 Svein Longva, Lorents Lorentsen and Øystein Olsen: Energy in a Multi-Sectoral Growth Model Energi i en flersektors vekstmodell Sidetall 22 Pris kr 9,00 ISBN 82-537-1082-8

- 80/2 Viggo Jean-Hansen: Totalregnskap for fiske- og fangstnæringen 1975 - 1978 Sidetall 33 Pris kr 9,00 ISBN 82-537-1080-1

- 80/3 Erik Biørn og Hans Erik Fosby: Kvartalsserier for brukerpriser på realkapital i norske produksjonssektorer Sidetall 60 Pris kr 11,00 ISBN 82-537-1087-9

- 80/4 Erik Biørn and Eilev S. Jansen: Consumer Demand in Norwegian Households 1973 - 1977 A Data Base for Micro-Econometrics Sidetall 130 Pris kr 13,00 ISBN 82-537-1086-0

- 80/5 Ole K. Hovland: Skattemodellen LOTTE Testing av framskrivingsmetoder Sidetall 30 Pris kr 9,00 ISBN 82-537-1088-7

- 80/6 Fylkesvise elektrisitetsprognoser for 1985 og 1990 En metodestudie Sidetall 56 Pris kr 11,00 ISBN 82-537-1091-7

- 80/7 Analyse av utviklingen i elektrisitetsforbruket 1978 og første halvår 1979 Sidetall 22 Pris kr 7,00 ISBN 82-537-1129-8

- 80/8 Øyvind Lone: Hovedklassifiseringa i arealregnskapet Sidetall 50 Pris kr 9,00 ISBN 82-537-1104-2

- 80/9 Tor Bjerkedal: Yrke og fødsel En undersøkelse over betydningen av kvinners yrkesaktivitet for opptreden av fosterskader Occupation and Outcome of Pregnancy Sidetall 93 Pris kr 11,00 ISBN 82-537-1111-5

- 80/10 Statistikk fra det økonomiske og medisinske informasjonssystem Alminnelige somatiske sykehus 1978 Sidetall 65 Pris kr 11,00 ISBN 82-537-1119-0

- 80/11 John Dagsvik: A Dynamic Model for Qualitative Choice Behaviour Implications for the Analysis of Labour Force Participation when the Total Supply of Labour is latent En dynamisk teori for kvalitativ valghandling Implikasjoner for analyse av yrkesdeltaking når det totale tilbud av arbeid er latent Sidetall 25 Pris kr 9,00 ISBN 82-537-1152-2

- 80/12 Torgeir Melien: Ressursregnskap for jern Sidetall 56 Pris kr 9,00 ISBN 82-537-1138-7

- 80/13 Eystein Glattre og Ellen Blix: En vurdering av dødsårsaksstatistikken Feil på dødsmeldingene Evaluation of the Cause-of-Death-Statistics Sidetall 64 Pris kr 11,00 ISBN 82-537-1136-0

- 80/14 Petter Frenger: Import Share Functions in Input - Output Analysis Importandelsfunksjoner i kryssløpsmodeller Sidetall 41 Pris kr 9,00 ISBN 82-537-1143-3

- 80/15 Den statistiske behandlingen av oljevirksomheten Sidetall 56 Pris kr 11,00 ISBN 82-537-1150-6

- 80/16 Ådne Cappelen, Eva Ivås og Paal Sand: MODIS IV Detaljerte virkningstabeller for 1978 Sidetall 261 Pris kr 15,00 ISBN 82-537-1142-5

- 80/18 Susan Lingsom: Dagbøker med og uten faste tidsintervaller: En sammenlikning basert på prøveundersøkelse om tidsnytting 1979 Open and Fixed Interval Time Diaries: A Comparison Based on a Pilot Study on Time Use 1979 Sidetall 31 Pris kr 9,00 ISBN 82-537-1158-1

- 80/19 Sigurd Høst og Trygve Solheim: Radio- og fjernsynsundersøkelsen januar - februar 1980 Sidetall 101 Pris kr 13,00 ISBN 82-537-1155-7

- 80/20 Skatter og overføringer til private Historisk oversikt over satser mv. Årene 1969 - 1980 Sidetall 53 Pris kr 11,00 ISBN 82-537-1151-4

- 80/21 Olav Bjerkholt og Øystein Olsen: Optimal kapasitet og fastkraftpotensial i et vannkraftsystem Sidetall 36 Pris kr 9,00 ISBN 82-537-1154-9

- 80/22 Rolf Aaberge: Eksakte metodar for analyse av to-vegstabellar Sidetall 80 Pris kr 11,00 ISBN 82-537-1161-1

Issued in the series Reports from the Central Bureau of Statistics (REP) (cont.) - ISSN 0332-8422

Printed 1980 (cont.)

No. 80/23 P. Frenger, E. S. Jansen and M. Reymert: Tariffs in a World Trade Model An Analysis of Changing Competitiveness due to Tariff Reductions in the 1960's and 1970's Sidetall 47 Pris kr 9,00 ISBN 82-537-1163-8

- 80/24 Jan Mønnesland: Bestandsuavhengige giftermålsrater Sidetall 50 Pris kr 9,00 ISBN 82-537-1167-0

- 80/25 Kari Lotsberg: Virkninger for norsk økonomi av endringer i samhandel Norge - utviklingslandene Sidetall 67 Pris kr 11,00 ISBN 82-537-1170-0

- 80/26 Lasse Fridstrøm: Lineære og log-lineære modeller for kvalitative avhengige variable Linear and Log-Linear Qualitative Response Models Sidetall 122 Pris kr 13,00 ISBN 82-537-1184-0

- 80/27 Aktuelle skattetall 1980 Current Tax Data Sidetall 43 Pris kr 9,00 ISBN 82-537-1194-8

- 80/28 Forbruksundersøkelse blant soldater 1979 Sidetall 45 Pris kr 11,00 ISBN 82-537-1199-9

- 80/30 Totalregnskap for fiske- og fangstnæringen 1976 - 1979 Sidetall 37 Pris kr 9,00 ISBN 82-537-1205-7

- 80/31 P. A. Garnåsjordet, Ø. Lone and H. V. Sæbø: Two Notes on Land Use Statistics Sidetall 47 Pris kr 9,00 ISBN 82-537-1214-6

- 80/32 Knut Ø. Sørensen: Glatting av flytterater i Statistisk Sentralbyrås befolkningsframskrivinger Sidetall 26 Pris kr 7,00 ISBN 82-537-1216-2

Pris kr 9,00

Publikasjonen utgis i kommisjon hos H. Aschehoug & Co. og Universitetsforlaget, Oslo, og er til salgs hos alle bokhandlere.

ISBN 82-537-1214-6
ISSN 0332-8422