The geographical differences and similarities of radon affected areas in England

Kara Turner

Turner, Kara (1995) The geographical differences and similarities of radon affected areas in England. Masters thesis, Durham University.

Kara Turner
February 1995

THE GEOGRAPHICAL DIFFERENCES AND SIMILARITIES OF RADON AFFECTED AREAS

List of Corrections

Page	Correction
5	premises to house
9	-particle to α-particle
19	Raper *et al.* (1992) reference added to references on pg. 230
40	data ... is to data ... are
48	filed to field
52	reflectivity's to reflectivities
57	later to latter
61	Jonhson to Johnston
72	indirect to irregular
72	known as the to as well as a
75	silled to skilled
80	changed Figure numbers
81	NJ???? to NJ0327
84	McLaren & Tulip (1991) reference added to references on pg. 230
98	enumeration districts to ward boundaries
110	Hampel *et al.* (1986) reference added to references on pg. 230
126	Taylor (1991) reference added to reference list on pg. 230
126	sentence reworded
129	added footnote to explain legend
135	aggregatin to aggregation
151	(10-30%) to (3-30%)
152	redrew Figure 12.4.6
152	better poorer agricultural areas to better agricultural areas
153	changed legend
155	deleted 'the'
156	deleted '-$_5$'
171	proprtion to proportion
176	added footnote to explain legend

The Geographical Differences and Similarities of Radon Affected Areas in England

by

Kara Turner

Thesis submitted to the Department of Geography, University of Durham for the Degree of Master of Science

October 1994

The Geographical Differences and Similarities of Radon Affected Areas in England

by

Kara Turner

Declaration

This MSc thesis is the result of my own work and has not been submitted for consideration in any other examination. Material from the published or unpublished work of others, which is referred to in the thesis, is credited to the author in question in the text. The thesis is approximately 67,000 words in length.

Kara Turner

October 1994

MSc Thesis

The Geographical Differences and Similarities of Radon Affected Areas in England

Abstract

The geographical distribution of radon gas is very uneven. The gas occurs naturally in all buildings at concentrations which can vary from below the United Kingdom national average of 20 Bq m^{-3} to more than 2,000 Bq m^{-3}. Five counties have been identified by the NRPB as 'Affected Areas' where more than 1% of homes have radon levels in excess of the current Action Level of 200 Bq m^{-3} (Miles *et al.*, 1992). These counties are Cornwall, Devon, Somerset, Derbyshire and Northamptonshire.

The level of radon gas in buildings is largely dependent on the underlying geology but geology does not always provide a full answer as to why spatial variations in radon occur. The implication of land capability on indoor radon levels in the five Affected Areas has been assessed using ARC/INFO and in Northamptonshire the influence of social factors (population density, social class and the proportion of households consisting only of pensioners) has been analysed.

There are some similarities in the results for the Affected Areas (especially between the counties located in the south-west of the country) as well as some striking differences (for example, the relationship between urban areas and radon levels differs in all the Affected Areas). Results in Somerset and Northamptonshire are strongly influenced by one or more dominant radon category or land capability grade.

In general, higher radon levels are associated with poor quality agricultural land and, in Northamptonshire, with high population density at ward level. The areas of Northamptonshire which have above average proportions in social classes I and II (1991 Census) are more likely to be associated with low radon levels (at district level), whereas areas with high proportions of households consisting only of pensioners tend to be associated with areas where more than 10% of homes are above the Action Level (at ward level).

The Geographical Differences and Similarities of Radon Affected Areas in England

Part One: Data Availability

Part Two: Methodology

Part Three: Results

APPENDICES

List of Figures

1 INTRODUCTION

2 RADON GAS

3 LAND USE

5 SOURCES OF OTHER DIGITAL DATA

8 HARDWARE AND SOFTWARE

9 DATA CAPTURE AND INTEGRATION IN ARC/INFO

10 ERROR MEASUREMENT AND ESTIMATION

12 RESULTS

Devon and Cornwall

Somerset

Northamptonshire

Derbyshire

13 A COMPARISON BETWEEN THE AFFECTED AREAS

List of Tables

2 RADON GAS

3 LAND USE

4 CENSUS DATA

5 OTHER SOURCES OF SPATIAL DATA

10 Error Measurement and Estimation

12 RESULTS

Devon and Cornwall

Somerset

Northamptonshire

Derbyshire

13 A COMPARISON BETWEEN THE AFFECTED AREAS

List of Abbreviations

ALC	Agricultural Land Classification
AUSLIG	Australian Centre for Remote Sensing
AVHRR	Advanced Very High Resolution Radiometer (USA)
Bq	Bequerel (usage = Bq per m^3)
BRE	Buildings Research Establishment
CD	Compact disc
CDU	Census Dissemination Unit (University of Manchester)
CHEST	Combined Higher Education Software Team
CIS	Countryside Information System
CPD	Central Postcode Directory
CVS	Census Validation Survey
DIME	Dual Independent Map Encoding
DoE	Department of Environment
ED	Enumeration District
EPA	Environmental Protection Agency (USA)
ERS	European Earth Resources Satellite
ESRC	Economic and Social Research Council
ESRI	Environmental Systems Research Institute, Inc.
EU	European Union
GIS	Geographic Information System
GISA	Geochemical Interactive Systems Analysis
GRID	Global Resource Information Database
HRV	High Resolution Visible sensors
IAEA	International Atomic Energy Agency
ICRP	International Commission on Radiological Protection
IRS	Indian Remote Sensing Satellite
ITE	Institute of Terrestrial Ecology
LandIS	Land Information System
LBS	Local Base Statistics
MAFF	Ministry of Agriculture, Fisheries and Food
MAUP	Modifiable areal unit problem
MOS	Marine Observation Satellite
MSS	Multispectral Scanner
NISS	National Information Services and Systems
NOAA	National Oceanographic and Atmospheric Administration
NOMIS	National Online Manpower Information System

NRPB	National Radiological Protection Board
NRSC	National Remote Sensing Centre Ltd.
OPCS	Office of Population Censuses and Surveys
OS	Ordnance Survey
PAF	Postcode Address File
PAT	Polygon attribute table
ppm	Parts per million
SAR	Sample of Anonymised Records (Chapter 4)
SAR	Synthetic Aperture Radar (Chapter 3)
SAS	Small Area Statistics
SINES	Spatial Information Enquiry Service
SPOT	Satellite Pour L'Observation de la Terre
TM	Thematic Mapper
UFC	University Funding Council
UNEP	United Nations Environment Programme

Acknowledgements

I wish to thank the following people for their generosity in terms of time and advice: Sinclair Sutherland (& NOMIS), Gilbert Roberts (formerly of the Computing Centre at Durham University) and Lesley Clarke (Philip's) for her help with the digitising.

Thanks are due to the US Geological Survey for allowing me to use one of their diagrams in Figure 2.3, to Reed International Books and George Philip for permission to reproduce a road map in Figure 3.4, and similarly to Bartholomew's for Figure 5.1, to the National Remote Sensing Centre and the Institute of Terrestrial Ecology for allowing me to use their satellite imagery in Figures 3.6 and 3.7 respectively and to the NRPB for the reproduction of several figures in Chapter 2.

I am very grateful to David Gaylard and George Philip's for allowing me time off work to complete the thesis.

The MSc would not have possible without the support of the National Radiological Protection Board and in particular, Mr. Jon Miles for his help and patience. Without the academic and administrative support of Dr. Danny Donoghue and especially Professor Tooley, the thesis would not have been what it is today and many thanks are owed to them both.

My mother and father gave me the impetus to start the MSc and enabled me to continue it when funds were short. They have offered me continuous encouragement and help and I am eternally grateful for all that they have done.

Extra big thanks go to Matthew for travelling to all ends of the country when I needed him to and for being more helpful than a thousand references could ever be.

1. Introduction

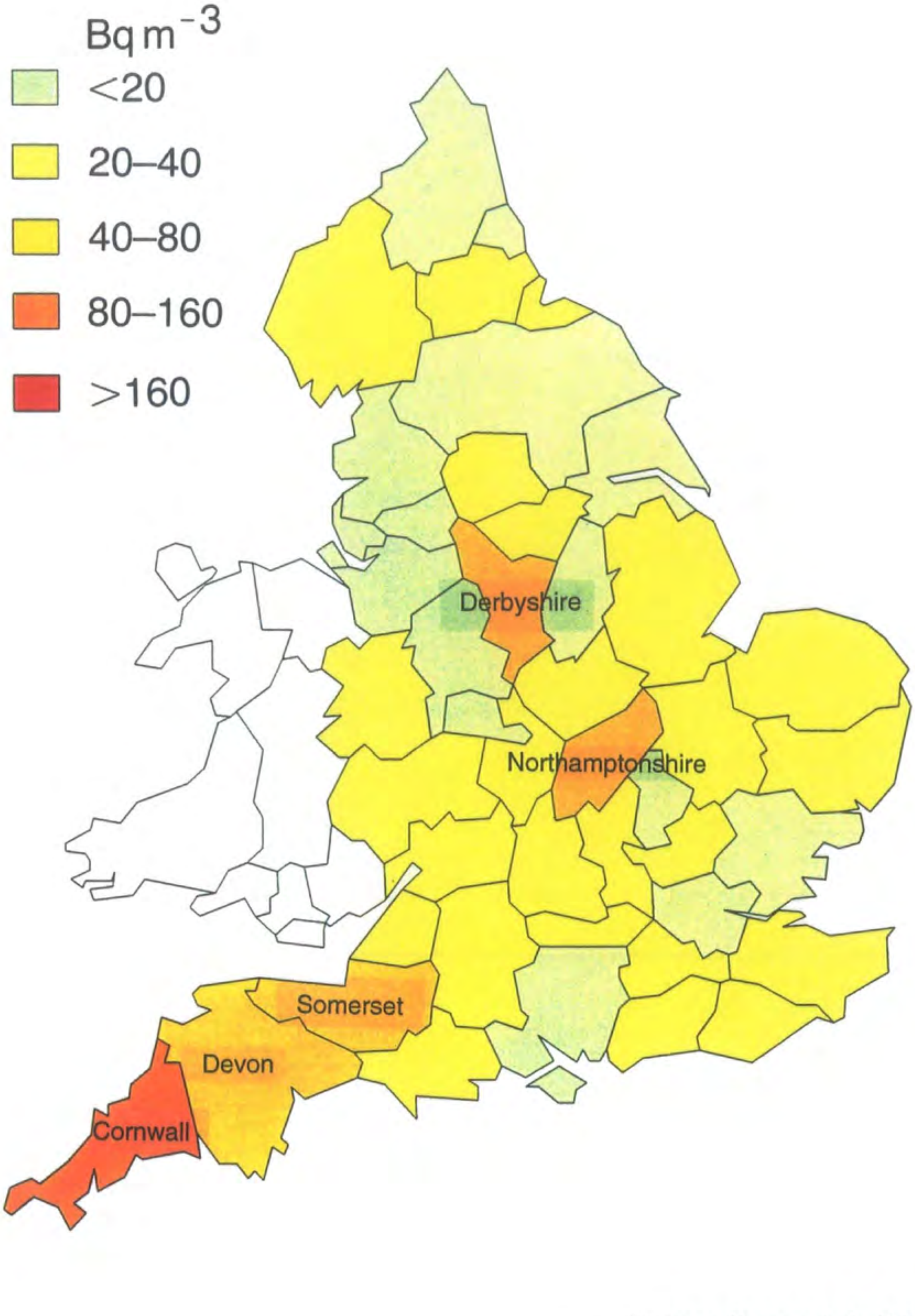
Over the past decade there has been a growing recognition that on average people receive a thousand times more radiation from natural background radiation than from man-made sources (Miles, 1991). Associated with this recognition is the acceptance by most earth scientists and certain sections of the general public, that radon (as the primary source of natural radiation) constitutes an important environmental health hazard (Owen, 1993). The National Radiological Protection Board (NRPB) was established by the Radiological Protection Act of 1970 and is responsible for carrying out research and development and providing information, advice and services to those with responsibilities for radiological protection. From measurements that it has carried out over the last 10 years, the NRPB has shown that at least 50% of the total radiation dose for the average Briton is obtained from combined radon and thoron (Clarke and Southwood, 1989) and this figure is as high as 80% in some parts of the south-west of England.

Radon and thoron are produced entirely by natural radioactive decay processes which have been going on for millions of years. The parent element, uranium, is concentrated in certain rock types. Since it was discovered that the gas builds up in some buildings and other confined spaces such as mines and caves, more and more countries are carrying out large-scale surveys and are introducing guidelines to prevent excessive exposures (Miles *et al.*, 1992).

The main cause of high radon levels in parts of Cornwall and Devon is the presence of granite at or just below the surface. The Hercynian granites of the south-west are known to have a relatively high uranium content which gives off radon as it decays. High radon levels in other parts of England are not so easy to explain. This project arose from NRPB findings that radon 'hot-spots' can occur in areas where there is no clear geological explanation. NRPB surveys carried out (on behalf of the Department of the Environment) between 1987 and 1991 indicated three other areas of the country where radon levels were found to exceed the present safety guidelines set in January 1990. In these three areas more than 1% of the homes that were measured were found to be above the Action Level of 200 Bq m^{-3}. For ease of classification these areas have been delineated by their county boundaries - Somerset, Derbyshire and Northamptonshire (see Figure 1.1). In total the five counties account for 3.5 million people and make up approximately 7% of the total population of England and Wales (OPCS, 1992). Elevated radon levels in Northamptonshire were a somewhat unexpected find as neither the NPRB nor geologists had predicted that the county would be affected. The NRPB national surveys are ongoing, and it may be that other radon 'hot-spots' will be discovered over the next few years.

Geography provides a useful framework for studying the spatial characteristics of radon gas. From recent literature (*e.g.* Johnston, 1986; Gilbert, 1988; Johnston, 1991) there can be little doubt that place specificity is back at the core of academic geography. These debates have

Figure 1.1 Representative radon levels by county in England



(Source: Green *et al.*, 1992)

provided the impetus for important critiques of certain excesses of modernist theorising and have increased sensitivity to space and place (Berg, 1993). In order to study the geographical differences and similarities between Cornwall, Devon, Somerset, Derbyshire and Northamptonshire, aspects of their natural and social environments were assessed. The spatial variations in land capability for the five counties were analysed in terms of the relationship between the different grades of land and the number of homes above the Action Level. A detailed study of the relationship between population density, social class and age structure and the number of homes above the Action Level, was carried out for the county of Northamptonshire.

The physical landscape of the five counties varies greatly from the moorlands and dissected river valleys of Cornwall, through the moorland landscape of Dartmoor and the rolling hills and valleys of Devon, to the flat peaty Levels and sudden rise of the Mendips in Somerset. Northamptonshire is characterised by arable farming and large market towns with little spatial variation in the nature of the land. Derbyshire, on the other hand, lies astride the Peak District at the southern end of the Pennines and is characterised by a variety of karst scenery, escarpments of Millstone Grit and lowland valleys associated with the coal measures. Figure 3.7 provides an aerial view of the variations in land use across the United Kingdom as a whole (see Chapter 3). The similarities that exist between Cornwall and Devon have been maintained in this project by treating the two areas as a single coverage and thus maintaining the affinity between the two counties in terms of geology, land use, social factors and radon levels. Somerset, Derbyshire and Northamptonshire are treated as three separate areas and any association there may be with surrounding regions has not been taken into consideration in the context of this work. Population patterns in the five areas are related to the location and landscape of the counties, with the highest population density in Northamptonshire and parts of Derbyshire and the lowest population densities in rural west Devon and most of Cornwall.

Although this study is concerned with radon gas, which is a known carcinogen, the work is not epidemiological in nature, nor does it aim to predict where future 'hot-spots' may lie. Furthermore, research into the association between radon gas and geology is being undertaken elsewhere (*e.g.* Ball and Miles, 1993) and there is no benefit to be gained from repeating such work here. An outline of the geological factors affecting radon is given as part of the introduction to this research but more comprehensive information is available (*e.g.* Ball, 1991; Ball, 1993; Plant *et al.*, 1983; Reimer, 1990; Schroeder, 1965).

This study aims to identify, and offer suggestions for, the relationships between the five radon Affected Areas with respect to radon levels, land capability and various aspects of the population. The value of the radiological data, collected by the NRPB, is enhanced by integration with other relevant spatial data sets, creating a new information source which can be used in the planning and implementation of future radiological protection strategies. The geographic information system ARC/INFO was used as a tool for integrating and analysing the data. There are few examples in British or European literature citing the results of integrating socio-economic data with physical databases using GIS. This thesis hopes to redress the balance slightly by tackling some of the problems associated with the geo-referencing of human and physical databases. The availability of the relevant data sets has been assessed as an integral part of the project and the reasons for the choice of specific data sets have been discussed. Possible openings for future work have been addressed.

The key questions analysed for each of the counties under investigation are as follows:

1. What spatial patterns exist in the radon data set?
2. What spatial patterns exist in the land capability data set?
3. What spatial patterns and relationships exist between radon gas and land capability?

In Northamptonshire, three additional questions have been posed:

4. What is the relationship between population density and radon levels?
5. What is the relationship between social class and radon levels?
6. What is the relationship between the proportion of households containing only pensioners and radon levels?

A résumé of the findings on the availability of the data sets used in this project can be found at the end of Part One and similarly a résumé of the methods and the software that were used can be found at the end of Part Two. A summary of the overall findings is found at the end of Part Three.

Part One: Data Availability

PART ONE

2. Radon Gas

2.1 History of Radioactivity

Radioactivity was discovered by Antoine-Henry Becquerel in 1896 when he discovered that a uranium salt caused the blackening of a photographic plate. He concluded that the invisible radiation was emitted due to an intrinsic property of the element uranium and that uranium shared several properties with x-rays, such as the ability to discharge an electroscope (Draganic *et al.*, 1991). Radioactivity is a phenomenon in which energy is released by certain substances in the form of invisible radiation. It is a natural process that is universal and already existed when the chemical elements began to form (Draganic *et al.*, 1991). The radioactive gas radon, was first discovered by a German named Dorn in 1900.

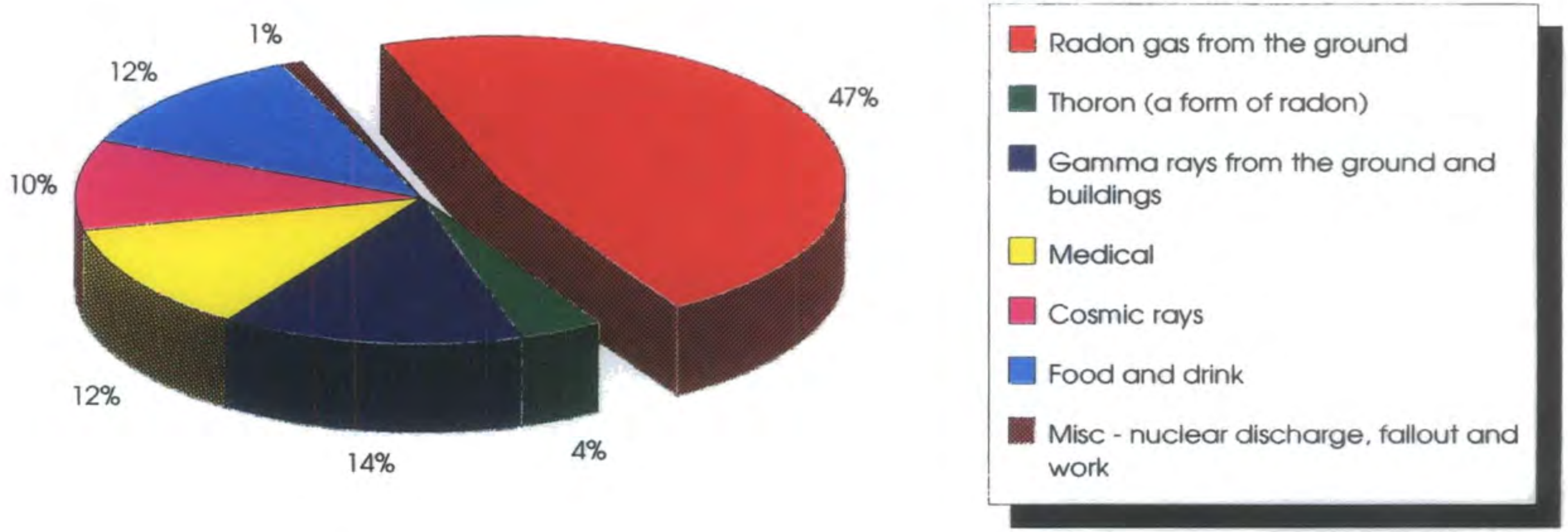
Radon is a naturally occurring, chemically inert radioactive gas with the atomic number 86. It has three naturally occurring isotopes, none of which is stable; radon-219, radon-220 and radon-222. Only radon-220 and radon-222 are of radiological consequence. Radon-220 is called thoron and is produced by the decay of Thorium-232 but has a shorter half-life (54.7 seconds) and is therefore of less concern than radon-222 (half-life of 3.8 days), known commonly as radon. Radon is about eight times denser than air but its scarcity in the atmosphere (about one part in 10^{18} in volume) prevents the gas from stratifying (Clarke and O'Riordan, 1990). This study concentrates solely on radon-222 and is referred to here simply as radon.

Radon is odourless, colourless and tasteless and cannot therefore be detected by the human senses. The activity of radon is measured in Becquerel (Bq)[1] with the activity concentration given in Bq m^{-3}. Socially acceptable levels have had to be progressively reduced in order to protect the public from the effects of the gas, which is known to cause cancer. In 1990 the NRPB revised its advice to the government on the action to be taken to limit exposure to high levels of radon in the home and a new 'Action Level' of 200 Bq m^{-3} (mean indoor activity concentration, where the mean is taken over one year) was set for existing dwellings and 100 Bq m^{-3} for future dwellings. If a house is found to have a level which exceeds the Action Level the government recommends that remedial action be taken to reduce the radon level to 200 Bq m^{-3} or below. The 200 Bq m^{-3} level is an arbitrary figure however, and as long as radon is present there is some degree of risk of lung cancer, albeit a very small one. Table 2.1 lists the estimated risk of

---

[1] Where one Bequerel is equal to one atomic disintegration per second. One picocurie equals 0.037 Bequerel.

Figure 2.1 Sources of daily exposure to radon



Source: NRPB

contracting lung cancer in the United Kingdom for various radon concentrations. In the United States the current action level is 4 picocuries per litre of air (c. 148 Bq m^{-3} - Brookins, 1990) and it is possible that the level in the United Kingdom will be further reduced in the future.

Radon is a health hazard because it emits alpha particles, which once ingested or inhaled in air or water, can damage tissue as they are not penetrative and therefore give up their energy to a relatively small volume of tissue (Ball, 1993). The assessment of risk is a major area of study in itself and there is some disagreement as to the estimated risk to human health. The NRPB has assessed the risk to people in the United Kingdom from differing radon levels (see Table 2.1) but it is still not known with certainty exactly how radiation causes cancer. The lifelong exposure of the average person to the average radon level in Britain, implies an estimated lifetime risk of only around 0.3% of contracting lung-cancer, assuming mean life expectancy of 75 years. For people who smoke the risk is slightly more than the average and for those who do not the risk is slightly less (Clarke and O'Riordan, 1990). Dosimetry standards are periodically revised, modified and supplemented as more measurements are carried out and progress is made. A new survey carried out in Sweden has confirmed that radon in homes causes lung cancer (NRPB, 1993). In Sweden, radon levels in the homes of 1,400 people with lung cancer were compared with the levels in 2,800 people without lung cancer. The results are consistent with the NRPB estimate that the lifetime risk of lung cancer (in the United Kingdom) at the Action Level for radon is 1 in 30.

Table 2.1 The Lifetime Risk of Lung Cancer From Lifelong Exposure to Radon

Average concentration (Bq m^{-3})	Lifetime risk %
20	0.3
100	1.5
200	3.0
400	6.0

(Source: Green *et al.*, 1992)

The first steps towards radiation protection regulations were taken in Germany in 1913 and then in Britain in 1915, in the main to protect patients and medical staff from radiation given off by equipment at work, for example from medical x-rays. National advisory committees on radiation were formed and these issued the first recommendations for protection. The International Committee on Radiation Protection (ICRP) evolved in order to keep definitions of units and methods of measurements up to date and to provide recommended action for radiation protection (see section 2.3 below).

2.2 Radon Levels in England and the Affected Areas

Every second of the day, human beings are bombarded with approximately 30,000 radioactive atoms and 200 million gamma rays from the soil and from building materials (Goldsmith and Hildyard, 1990). Radiation of natural origin accounts for 87% of the effective dose equivalent to the population of the United Kingdom. At least half of the natural radiation dose emanates from short-lived decay products of radon and this percentage is much higher (around 80%) in certain parts of the country, namely the south-west. Figure 2.1 shows the relative importance of radon within the context of the average radiation dose received in the United Kingdom. The population-weighted annual average radon level for England (for indoor air) is 21 Bq m^{-3}, five times greater than outdoors, with the distribution positively skewed (Clarke and O'Riordan, 1990). According to Hughes and O'Riordan (1993) this figure is low compared to most other European countries.

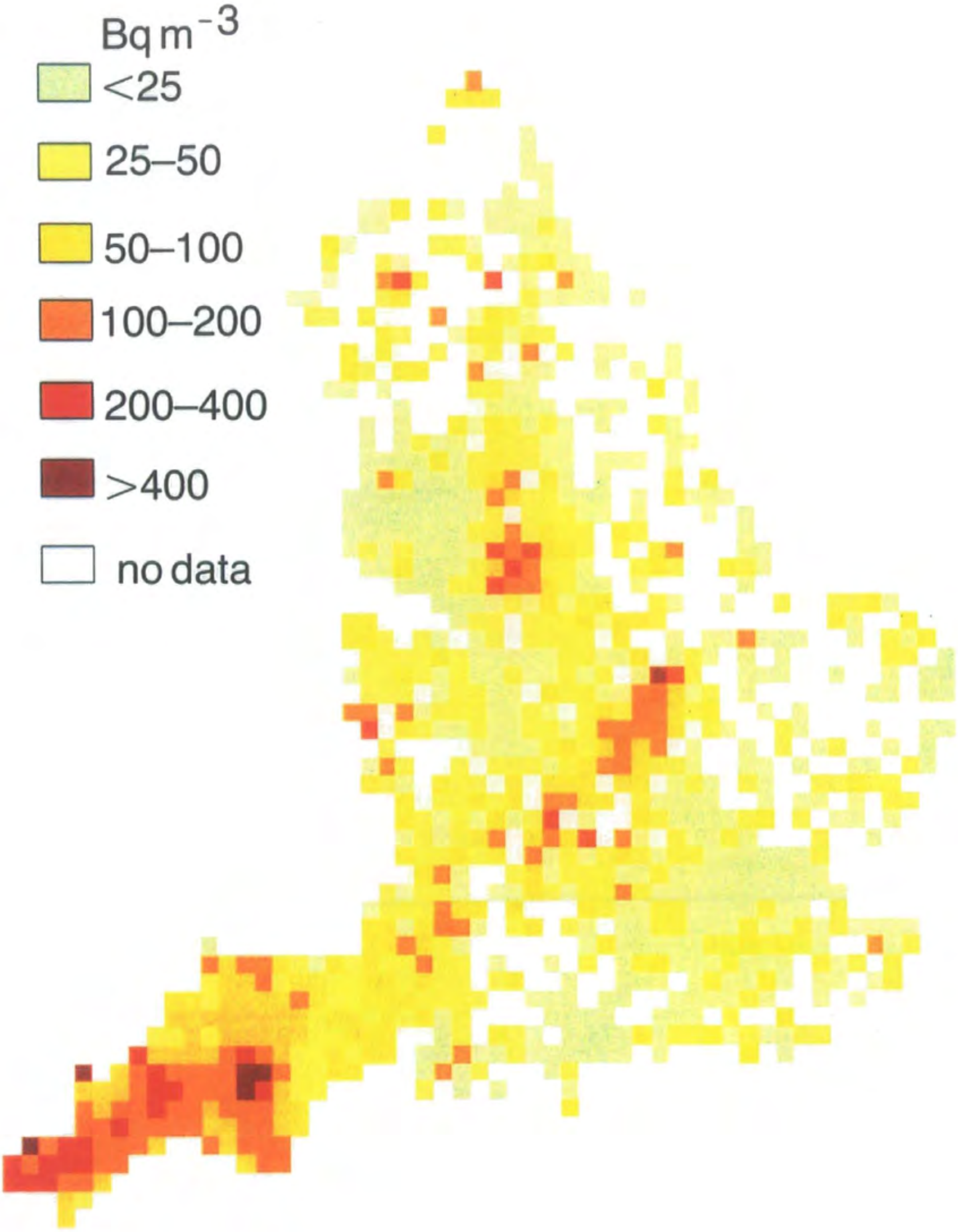
The geographical distribution of radon is very uneven (see Figure 1.1). The majority of homes that exceed the Action Level are situated in the two counties of Cornwall and Devon. Cornwall is the most affected county, with an average radon concentration of about 110 Bq m^{-3} (Dixon and Gregory, 1992). O'Riordan *et al.* (1987) yielded results in Devon and Cornwall where between 24% and 19% of the dwellings sampled exceeded the Action Level (from a total sample number of 62 for the two counties). The variations in radon concentration were virtually identical in the two counties and could be represented approximately by a log normal distribution with a median of 160 Bq m^{-3} and a standard deviation of about three (Cliff and Haggett, 1988). The NRPB have carried out various surveys (some of which are on-going, see section 2.3 below) and the results of the surveys (up to mid-1991) are shown in Figure 2.2 for 10 kilometre grid squares. In order to delimit areas in the United Kingdom where radon poses a threat to human health, the NRPB issued its first assessment of a radon 'Affected Area' in 1990. This specifies that 'areas of the United Kingdom where 1% or more of homes exceed the Action Level of 200 Bq m^{-3} should be regarded as Affected Areas' (Miles *et al.*, 1990), such areas are identified from radiological evidence and are periodically reviewed. Following this assessment Cornwall and Devon were designated Affected Areas with approximately 12% (60,000) of homes above 200 Bq m^{-3} and an estimated 5% above 1,000 Bq m^{-3}. Somerset, Northamptonshire and Derbyshire were added to the list in 1992 with between 1% and 2% (10,000) of homes above the Action Level. Figure 1.1 shows the representative radon levels for all English counties, the Affected Areas clearly standing out from the rest of the country. The average radon levels for the five Affected Areas are given in Table 2.2. Some Scottish, Welsh and other English counties also contain areas where radon levels exceed the Action Level and it is estimated that these areas contain a total of 5,000 homes. There are as yet insufficient data to satisfy the necessary criteria for designating them Affected Areas. In Scotland, these areas include parts of the Grampians and the north-west Highlands. The large urban areas of England such as London, Manchester and Birmingham by comparison, generally have very low radon readings.

In spite of the fact that elevated radon levels transcend administrative boundaries, Affected Areas are delimited by county boundaries to facilitate policy implementation and aid planning and organisation. In Affected Areas preventative action against radon may be required during the construction of buildings and homes. It is recommended that homes should be designed and built with the threat of radon in mind, so that concentrations of the gas within the home can be kept as low as possible. The construction industry is kept informed of research into radon (and carries out its own research) and has developed a variety of measures to keep radon out of the home. Such measures include fans and sumps and other methods of increasing ventilation in the home, as well as plastic membranes which are laid in the foundations of a building or beneath the floorboards and present an impermeable barrier to radon gas. The Building Research Establishment operates a telephone hotline for members of the public who need advice concerning remedial measures.

Radon is formed in the decay series of ^{238}Uranium and Thorium which are common, naturally occurring elements found in low concentrations in rocks and soil, though uranium is seldom distributed homogeneously but is concentrated in discrete uranium bearing minerals. Uranium is the ancestor of a family of 14 radioactive elements, see Appendix 1. The concentration of uranium in rocks is usually low, average crustal abundance being between 2.7 ppm and 2.5 ppm (Plant, 1978) and, although radon is associated with elevated uranium levels, it is frequently found at high levels in areas where there is no uranium mineralisation at all. ^{238}Uranium has a half-life of 4,510 million years and decays with the emission of an α-particle. The inertia of radon means that it is relatively unaffected by the chemical buffering reactions, which often control the generation of other gases, and radon is left free to travel as a gas (EPA, 1987). Although it is a very mobile gas, the movement of radon is limited by its radioactive half-life of 3.8 days.

The U^{+4} ion is incompatible with crystal lattice sites in major rock forming minerals and during igneous processes it is enriched in highly evolved granitic and alkaline rocks, such as those found in the south-west of England and in parts of Scotland, in preference to intermediate, basic or ultrabasic lithologies (Plant *et al.*, 1983). It has been shown that the greatest proportion of uranium in granite occurs in the mineral uraninite (Ball and Basham, 1979; Ball and Miles, 1993). Uranium may also occur in accessory minerals such as zircon, allanite, pyrochlore and also in ore minerals such as pitchblende and coffinite and tends to be concentrated in black shales up to ore grade (350 ppm). Most

Figure 2.2 Radon concentrations in England by 10 km squares of the Ordnance Survey grid



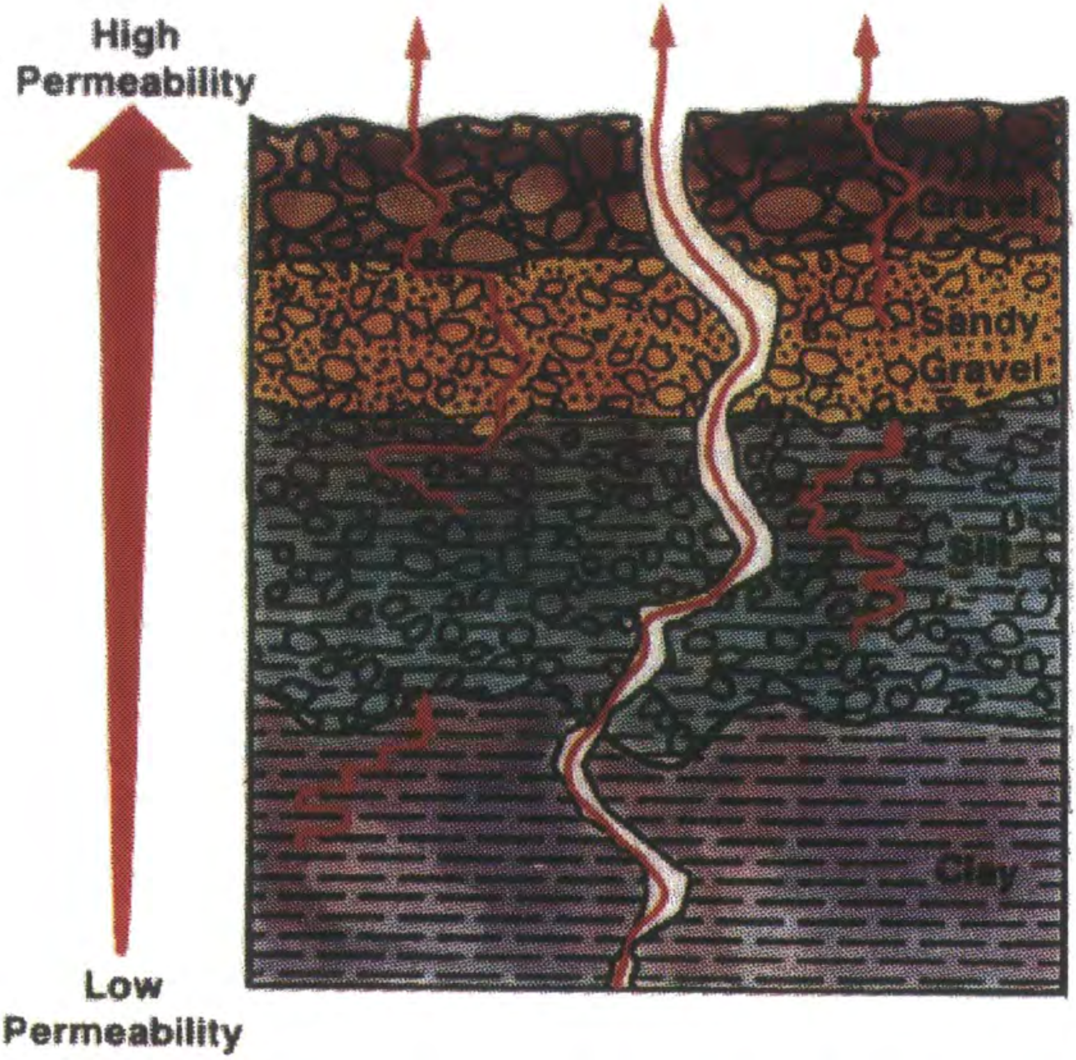
(Source: Green *et al.*, 1992)

economic uranium deposits occur as hydrothermal veins in sandstones and conglomerates containing 0.03 to 0.1% U_3O_8 (Bowie, 1978). Radon isotopes are the only alpha emitting gases, so their concentration in soil gas may be determined accurately using relatively simple technology (Ball *et al.*, 1991) and radon is often used by geologists as a tool for the detection of uranium deposits at depth.

There are many features of rocks and soil which influence the release and diffusion of radon. Uranium and thoron concentration and mineralogy are the main ones, but also important are the permeability of the host rocks, the nature and the extent of carrier fluid transport mechanisms and the weather (Ball *et al.*, 1991). Gilletti and Kulp (1955) cited in Ball *et al.* (1991), noted that the emanation coefficients for uranium minerals increased with temperature but decreased with pressure - with soils having the highest coefficients, followed by rocks and lastly, minerals. Figure 2.3 diagrammatically shows the permeabilities of several types of soil and rocks. Shroeder *et al.* (1965) gives a range of diffusion values for radon from 0.02 to 0.10 cm^2 per second, with an average diffusion rate of 0.03 cm^2 per second for moderately dry, sandy soils. Tanner (1964) cited in Ball *et al.* (1991), has shown that in ideal conditions of dry gravel, the maximum expected diffusion of radon is 10 metres and in soils this distance would be substantially less. Radon tends to be dispersed therefore partly as a result of diffusion but the short half life means that it is only widely dispersed where there is a high rate of groundwater movement (Plant *et al.*, 1983) or where there are deep fault structures, natural passageways or underground mine workings. The radon content of surface and groundwaters in the granites of south-west England has been the subject of intensive studies and radon concentrations between 110 and 740 Bq per litre have been recorded in the Carmenelis Granite area (Ball *et al.*, 1991). The Upper Palaeozoic mudstone and limestone lithologies in the Pennines have been identified as having high uranium levels (Ball *et al.*, 1991) and this may be a contributing factor towards the elevated radon levels in parts of Derbyshire. Digital data for uranium in stream sediments are available from the Geochemical Interactive Systems Analysis (GISA) service of the Geochemical Survey Programme (Ball *et al.*, 1991).

In the south-west, the high uranium content of the granites provides the initial source of radon. The overall picture is of a radioactive granite batholith, intruded into folded sedimentary rocks (dominantly argillaceous and arenaceous sediments of Devonian and Carboniferous age), uniform along its length and stretching over 200 kilometres from Dartmoor, west-south-west through Bodmin Moor, St. Austell, Carmenellis and Lands End and plutons to the Scilly Isles (see Figure 2.4). From work done on gravity anomalies the batholith is estimated to be 10-12 kilometres thick, with the roof between outcrops being no more than 2-3 kilometres thick (Tammemagi and Smith, 1975). Most of the uranium vein mineralisation occurs within about 2 kilometres either side of the granite contact (Ball and Miles, 1993). There is a significant correlation between the

Figure 2.3 Contrasting permeabilities of rock and soil types



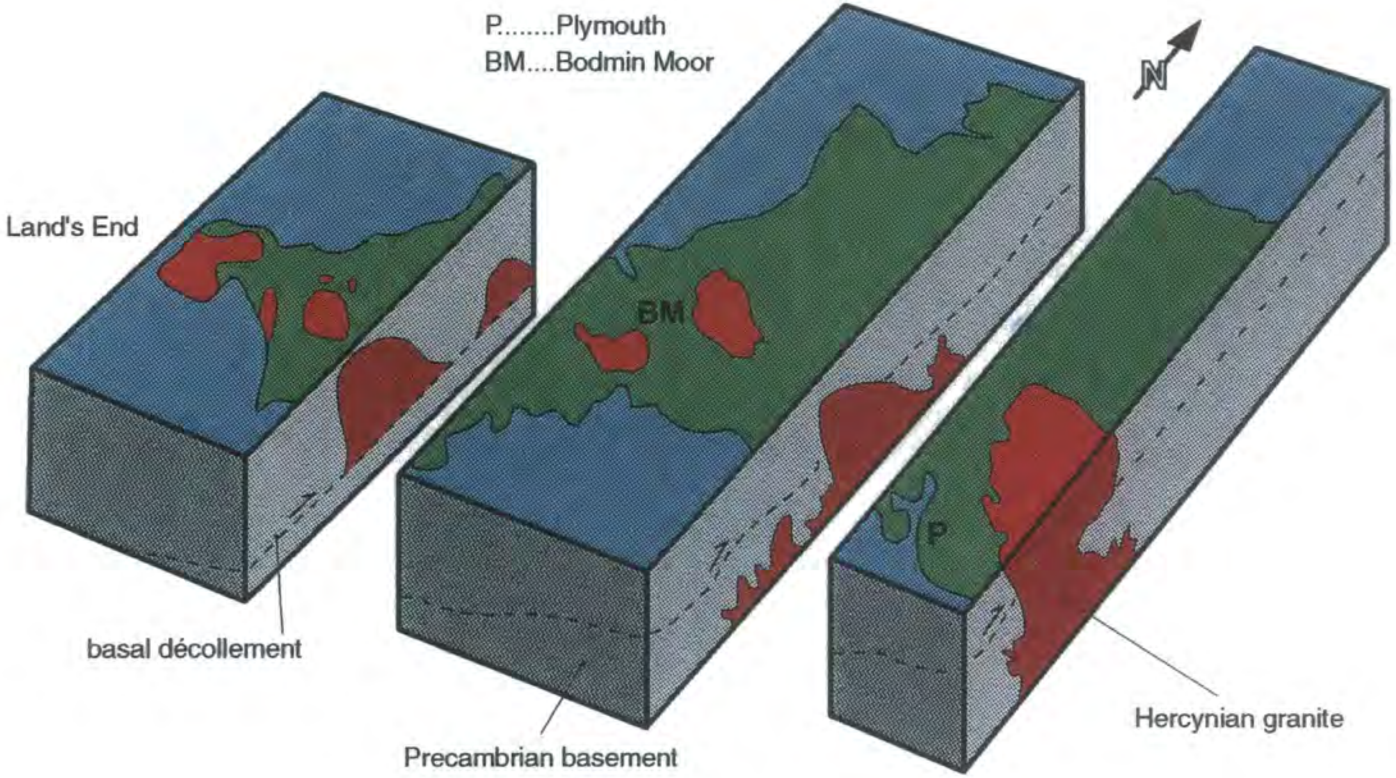
(Source: USGS)

surface granite masses of Dartmoor, Bodmin Moor, Blackmoor, Wendron Moors and West Penwith and areas of Cornwall and Devon with the highest proportion of homes exceeding the Action Level (between 10% and 30% of homes - see Figure 2.2). Southwest England was south of the Pleistocene ice front in the last Ice Age and hence retained Tertiary erosion surfaces, characterised by their deep weathering profiles and consequent disaggregation of the granite. The granites of the south-west are also substantially mineralised and the presence of lodes and other fractures provides potential pathways to the surface. Extensive mining of the various metals and minerals has created a labyrinth of passageways through which the radon can diffuse. These factors have combined to give high radon readings in many parts of Cornwall and Devon. The varying geology of England and the differing concentration of uranium in granites has resulted in the variety of radon concentrations that have been measured in different areas of granitic intrusion.

Radon is not solely correlated with granitic areas however. In a paper by Ball *et al.* (1991) the distribution of radon in soil gas was analysed in relation to the geology of the Ashover anticline in Derbyshire. The results show significantly elevated levels over the Eyam limestone and to a lesser extent in association with the Longstone Mudstone. The mudstones contain much more uranium (8-35 ppm) than the limestones (2-10 ppm) but because of their waterlogged nature contribute much less to the soil gas phase. High radon levels were also shown to occur over fault lines. In Carboniferous shales in and around the Peak District, the highest uranium values occur in the basal Namurian, with some values over 100ppm (Sutherland, 1994). In Derbyshire as a whole there is a clear relationship between the high indoor radon levels and the Carboniferous Limestones found predominantly in the west of the county. Anomalous concentrations of radon far from the original uranium source can often be explained by sub-surface water carrying radon in solution through the limestone cavities. Elevated radon levels in Northamptonshire, mainly correspond with the Northampton Sand Formation, although high values have also been obtained on Middle Lias Marlstone Rock Bed (Sutherland, 1994). The rocks comprise phosphatic ironstones, sandstones and limestones which are thought to contain slightly elevated levels of uranium. The slightly elevated radon levels in Exmoor are more difficult to explain in terms of the geology of the area. Jones *et al.* (1987) cited in Ball and Miles (1993) found only one anomalous concentration of uranium in the whole area, all other values were very close to the crustal average. The high permeability of the underlying limestone in some parts of Somerset however, enables the radon to diffuse greater distances and to be carried in solution in sub-surface water. In the north-east of the county there is an association between high radon levels and Carboniferous and Mesozoic Limestones and in the west with Devonian sandstones and shales.

Figure 2.4 The approximate extent of granite in the south-west of England

The diagram shows that granite bodies have considerable subsurface extent.



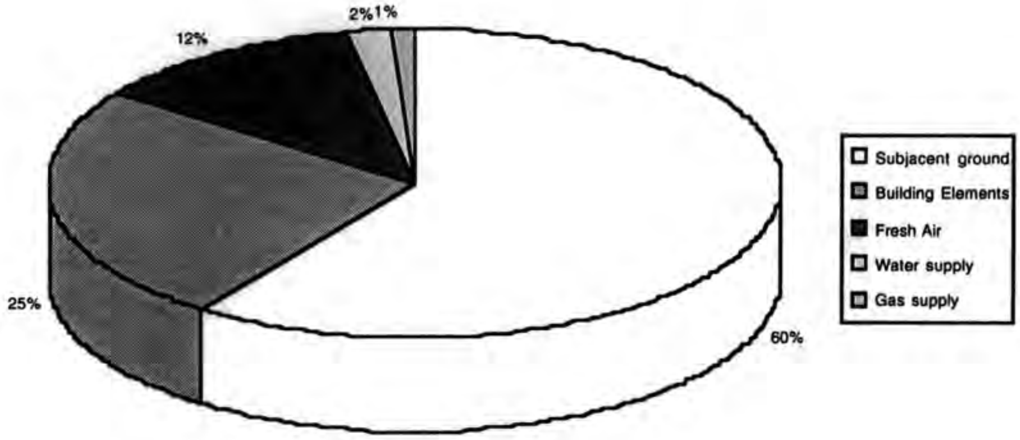
(Source: Freshney & Taylor, 1989)

Radon is not harmful to man until it builds up to high concentrations in buildings or confined spaces such as caves and tunnels. In the second half of the twentieth century the drive towards better home insulation and double glazing has inhibited the escape of radon into the atmosphere by causing a decrease in the rate of change of air in buildings and an increase in the pressure differential in relation to outdoors, although it continues to be the case that the main cause of elevated radon levels lies in the uranium content of subjacent rocks and soil. Results from the NRPB National Survey (Wrixon *et al.*, 1988) found that the presence of either double glazing or draught-proofing around doors and windows typically increases the radon concentration by about 30%. In confined spaces such as caves, mines and buildings, radon gas can build up to high levels by seeping through cracks and diffusing from building materials and from the water and gas supply. A report by the Building Research Establishment (1991) lists the following as the main ingress routes of radon gas into buildings (Ball, 1993):

- cracks in solid floors
- construction joints
- cracks in walls below ground level
- gaps in suspended concrete or timber floors
- gaps around service pipes
- cavities in walls

Figure 2.5 below, shows the relationship between the main sources of radon in the home.

Figure 2.5 Sources of Radon in Buildings



(Source: Green *et al.*, 1992)

High radon levels in buildings are due primarily to the bulk flow of soil gas carrying radon, with the diffusion of radon through and from building materials playing a more minor role (Ball and Miles, 1993). Molecular diffusion (the movement of a gas from an area of high concentration to an area of low concentration at constant pressure) was originally believed to be the main transport mechanism that transported radon particles from the source rock to areas with little or no intrinsic uranium content, but scientific investigations have indicated that diffusion alone can not account for the high levels of indoor radon discovered in some homes. Radon is drawn into buildings by the slightly negative pressure differences (about 0.1% atmospheres or a few Pascals) that exist between indoor and outdoor atmospheres. These pressure differences cause soil air and atmospheric air to be drawn inwards, either through gaps in the walls and around doors, or through cracks and holes in the floor (Clarke and O'Riordan, 1990). The pressure differential is affected by the warmer air inside a building which creates a tendency for warm indoor air to be displaced by cooler outdoor air - a method of entry known as 'pressure-driven flow'. Variations in temperatures throughout the year cause fluctuating radon levels, with the highest levels in the winter months when the temperature differences are greatest. The relationship between the indoor and outdoor air is affected by the wind, which in turn affects the pressure differential and can drive the flow of radon into a building, the effect of chimneys and fans, the ventilation habits of the occupants and the specific characteristics of the building (Dixon and Gregory, 1992). Other factors such as barometric pressure and precipitation play a lesser role. Fresh air entering the home can also carry low levels of radon.

There has been considerable coverage in the local press in all of the five counties. The issue has been covered most rigorously in Cornwall, especially by the local papers covering the parts of the county where radon levels are highest. Press coverage first began around seven years ago and was most noticeable in the period from 1989 - 1992 when speculation about the medical consequences of elevated indoor radon levels was rife. Press coverage has since decreased in the south-west although articles do still appear in the local press from time to time, but articles on radon continue to be covered in papers in Northamptonshire and Derbyshire where the issue is more fresh. Plans to extend public awareness in Somerset, Derbyshire and Northamptonshire are being evaluated at present by the NRPB and the Department of Environment.

2.3 Radon Measurement and Surveys - National and Regional

Systematic data collection of the natural levels of uranium in the United Kingdom was restricted until the late 1960s when the Institute of Geological Sciences (now the British Geological Survey) began work on a series of regional geochemical atlases of the land area (Plant *et al.*, 1983). Contrary to popular belief, the scientific basis of radiation protection is now reasonably well established. It is fair to say that radiation hazards, although far from being understood, are nowadays better assessed than most of the other dangerous agents and pollutants of the environment (Draganic *et al.*, 1991). The occurrence of uranium throughout the country is well documented and the NRPB are carrying out ongoing research in conjunction with the British Geological Survey to determine the effect of geology on radon levels.

The International Commission on Radiological Protection (ICRP) and the International Atomic Energy Agency (IAEA) have identified four basic objectives that should underpin all monitoring programmes (Heywood and Cornelius, 1988; ICRP, 1977; IEAE, 1976):

1. Assessment of the effectiveness of release controls in preventing the emissions of radioactive materials into the terrestrial and aquatic environment;
2. **Evaluation of the threat to man from environmental radiation;**
3. Gauging compliance with national and international safety standards and
4. Investigation of possible long term trends in the environmental accumulation of radioisotopes due to the operation of nuclear installations.

Only the second point is of direct relevance to this study, but it is useful to note the role of similar research in the wider context of man-made radiation emissions. In the United Kingdom, the NRPB is concerned with setting national safety levels for both natural and man-made radiation levels and advises the government on all issues relating to radon, such as the Action Level.

Following the advice of the NRPB the government launched a publicity campaign in 1991 to educate the public on aspects of radon gas and to encourage householders in Affected Areas to reduce elevated radon levels in their homes. A leaflet drop was carried out in Cornwall and Devon in association with the Post Office, outlining the basic facts about radon and advertising the free radon home measuring kit. Cornwall County Council recognised its statutory obligations at an early stage and undertook an extensive programme of radon measurements in the premises under its control. This led to a considerable programme of radon remedial work which set a precedent for such work in the United Kingdom (Dixon and Gregory, 1992). The most effective method of reducing high radon levels is to install a fan and sump to withdraw air from beneath the floorboards and expel it into the atmosphere. In some cases grants for remedial work are available from the local county council.

Radon in the home is measured using a small plastic track etch detector provided, in this country, by the NRPB. Two detectors are issued, one for the main living area and one for the master bedroom and are left in position for three months (to ensure an accurate annual average and to reduce the effect of seasonal variation), after which time they are sent to the NRPB laboratories to be measured. The detector is disposable, tamper-resistant and indelibly numbered. The detecting element is a rectangular piece of polyallyl diglycol carbonate which is damaged when alpha particles emitted by radon strike its surface, leaving pits. The number of pits is proportional to the exposure of the detector to radon and these are counted automatically. Information on the duration of exposure of the detector in the home allows the average radon concentration to be calculated (Green *et al.*, 1992). In the United Kingdom over 160,000 householders have so far taken up the offer to have their home tested and over 16,000 homes have been identified as exceeding the Action Level, although the NRPB estimate that the national figure is around 100,000 (NRPB, 1993). Over the past five to 10 years however, the rate of discovery of homes above the Action Level has been low - around 1% per annum (Clarke and O'Riordan, 1990). The number of radon measurements (up to the end of 1991) for each 10 kilometre grid square is shown in Figure 2.6.

The average concentration of radon is not constant over time or space and variations of an order of magnitude can be found in contiguous dwellings. People living in a one storey house (compared to those in homes with two or more storeys and especially to those living apartment blocks) are most prone to the radon hazard since its migration is reduced in successively higher floors: Radon readings for upstairs bedrooms are comparatively lower than those for the main living room in the same house. This variation is a function of the principal radon source being the adjacent and subjacent soil. Indoor levels of the gas in homes show marked diurnal variation with the highest daily levels occurring in the morning and peaking at the time the occupants arise. The levels of the gas are affected by the opening and closing of doors and windows, which create a through-flow of fresh air. There is also a marked annual variation, with the highest radon concentrations occurring in January and the lowest in July.

Surveys of radon in homes in the United Kingdom have been carried out since 1976 in order to measure exposure to radon and assess the implications to the health of the British population (Green, 1991). The surveys are funded on the whole by the Department of the Environment and the ongoing nature of the surveys often render them difficult to distinguish from one another. Concurrent surveys in different parts of the country are usually linked and a set of measurements may be stimulated by a leaflet drop or the designation of an area as 'Affected' (Miles, correspondence, 1993). The demand for surveys in England by county to the end of 1991 was 97,168. Of these requested measurements 94,975 (97.7%) were in the five Affected Areas. It is indicative of the location-specific nature of radon gas that elevated levels are in the main, confined to areas with certain geological characteristics and it is in these areas (defined by their county boundaries) that the numbers of homes measured is the highest. A summary of the results of the requested measurements is given below (the percentages do not add up exactly to 97.7% due to the effects of rounding):

Table 2.2 Radon Levels in the Five Affected Counties

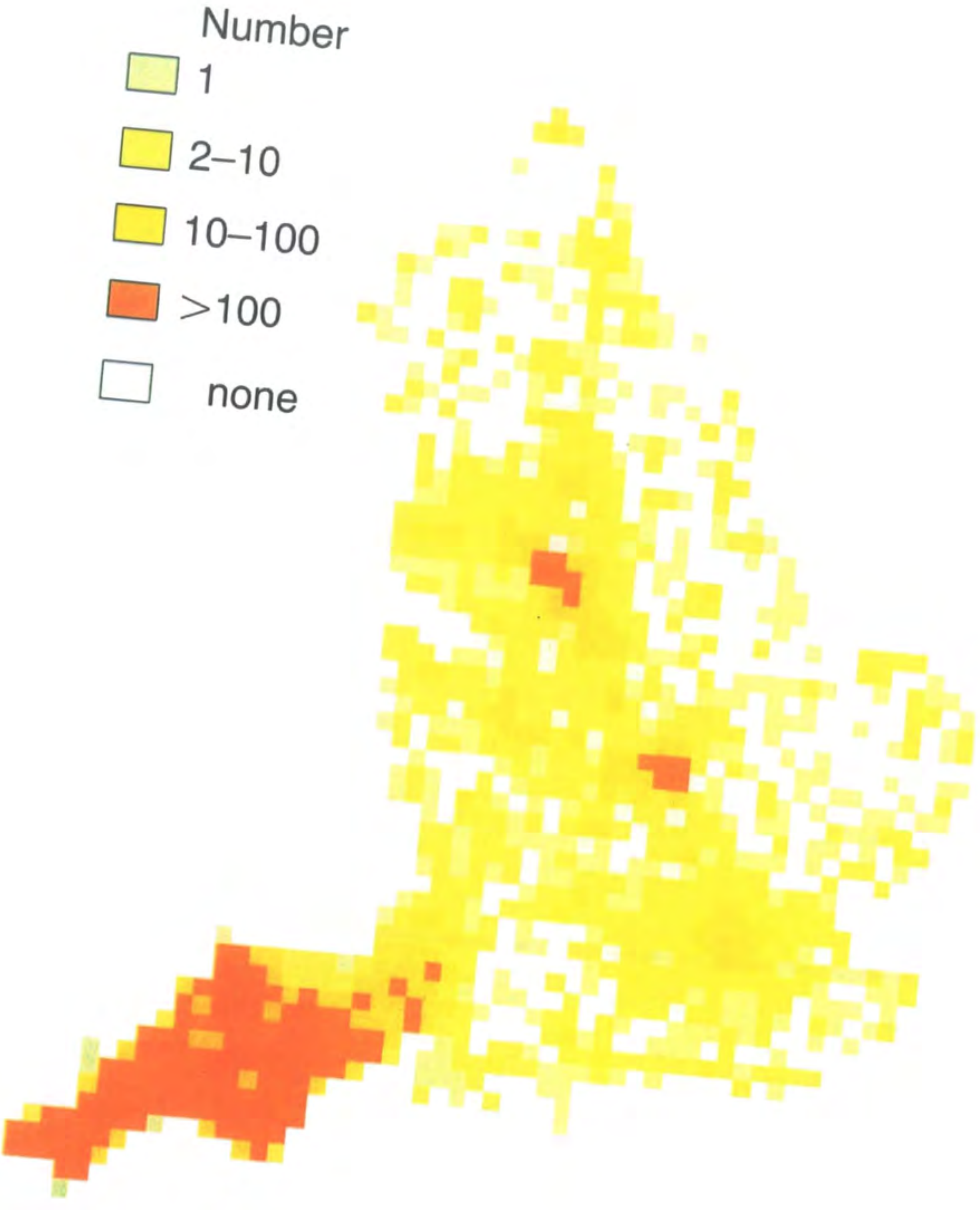
Radon Affected Area	No. of Radon Measurements	% of total measurements	Average Radon Level
Cornwall	31,743	32.7	177 Bq m^{-3}
Devon	53,209	54.8	76 Bq m^{-3}
Somerset	1,872	1.9	69 Bq m^{-3}
Northamptonshire	6,028	6.2	104 Bq m^{-3}
Derbyshire	2,113	2.2	155 Bq m^{-3}

(Source: Green *et al.*, 1992)

The average radon levels in Northamptonshire and Derbyshire are greater than would be expected due to the bias of measurements in the areas of the counties that are affected by elevated radon levels. Indeed, the results from all five Affected Areas may not give an accurate picture of the counties as a whole because of the bias in this particular set of results. The national survey of radon levels for ten and five kilometre grid squares gives a more representative view of regional and county radon levels (see below).

The results of the measurements collected by the NRPB for their national radon database are aggregated for five kilometre grid squares. This sampling method maintains confidentiality and provides a regular framework for analysis which is constant over time. For the British Isles as a whole, there are a total of around 7,807 five kilometre grid squares (Raper *et al.*, 1992) , which can be aggregated for larger areas *e.g.* Figure 2.2. The grid squares are small enough to enable worthwhile spatial analysis at the county and national level, yet are large enough to provide some data in sparsely populated areas. Evans (1991), an exponent of the use of grid squares for analysis, wrote that 'for different grid squares, different aggregations produce similar correlations,' thus the grid squares provide a regular and static structure for the analysis of the results. However, as Pearson (1986 cited in Openshaw, 1986) first pointed out, correlations of averages will not give identical results to correlations of ungrouped data. There is considerable methodological debate concerned with the analysis of grouped data and a summary of the main views is presented in Chapter 10. Some surveys carried out by the NRPB, with funding from the Department of the Environment, have been spatially referenced using postcodes, for example those described by Green *et al.* (1992). Figure 2.7 displays data collected by the NRPB for postcode sectors in Cornwall and Devon. Postcodes are organised hierarchically giving three levels of aggregation, ranging from individual buildings or small groups of houses to large regional units. However, they have the disadvantage of being highly irregular areal units that may vary over time and their

Figure 2.6 Radon measurements in England by 10 km squares of the Ordnance Survey grid



(Source: Green *et al.*, 1992)

location can only be identified using a special-purpose postcode map. These two methods of displaying data highlight the problem with the majority of environmental information, namely the lack of consistency in the use of a basic spatial referencing system. Currently there are two nationally agreed guidelines for standardisation on spatial referencing in the United Kingdom - the Ordnance Survey national grid and the postcode - but data are frequently aggregated for over fifty different spatial units (Openshaw *et al.*, 1987). The use of the postcode is generally reserved for data that relate to households or other properties. The main data collectors and holders may not necessarily wish to exploit the spatial dimension of their data themselves and may not therefore provide the data in a form that will facilitate spatial analysis by others. Alternatively, the data may be aggregated in a way that creates problems for data integration and spatial analysis with other data sets. There is a need for environmental agencies to agree on a United Kingdom standard, or a hierarchy of standards, to ensure future compatibility of environmental information and to avoid the current dilemma faced by those trying to integrate social and economic information, a problem which may be solved by the adoption of new environmentally based zones (Openshaw *et al.*, 1987; Heywood, 1991).

There are a number of organisations that collect and maintain information on various aspects of radon (for example some Universities and County Councils), but the NRPB are the official organisation in charge of collecting information and data and advise the government and the general public on all aspects of the gas. The following list is not fully comprehensive but aims to provide a summary of the main data sets collected by the principal parties interested in radon gas (namely the NRPB, BRE and others), not all of which are necessarily available to the general public or academics. Many of the surveys are ongoing (any figures given are the latest available) and the results from one survey can not necessarily be distinguished from those of another survey. The sponsors of the surveys are varied, but include the NRPB, the Department of the Environment, the Ministry of Defence, the International Environmental Health Organisation and the Department of Health. Some of the data sets have been published and others are currently being assembled:

- A systematic National survey of the gamma-ray dose rates and radon concentrations in 2,000 dwellings in the United Kingdom (1987) (Wrixon *et al.*, 1988).
- Various regional surveys targeting specific areas believed to have elevated radon levels (namely the granite areas of Devon and Cornwall, the uraniferous nodule area of south-east Devon, parts of North Yorkshire and Derbyshire and parts of Somerset), following detailed radiological and geological surveys. These so-called 'directed surveys' were carried out between January 1987 and April 1989. Parts of Scotland, south Wales and Northern Ireland have since been surveyed and

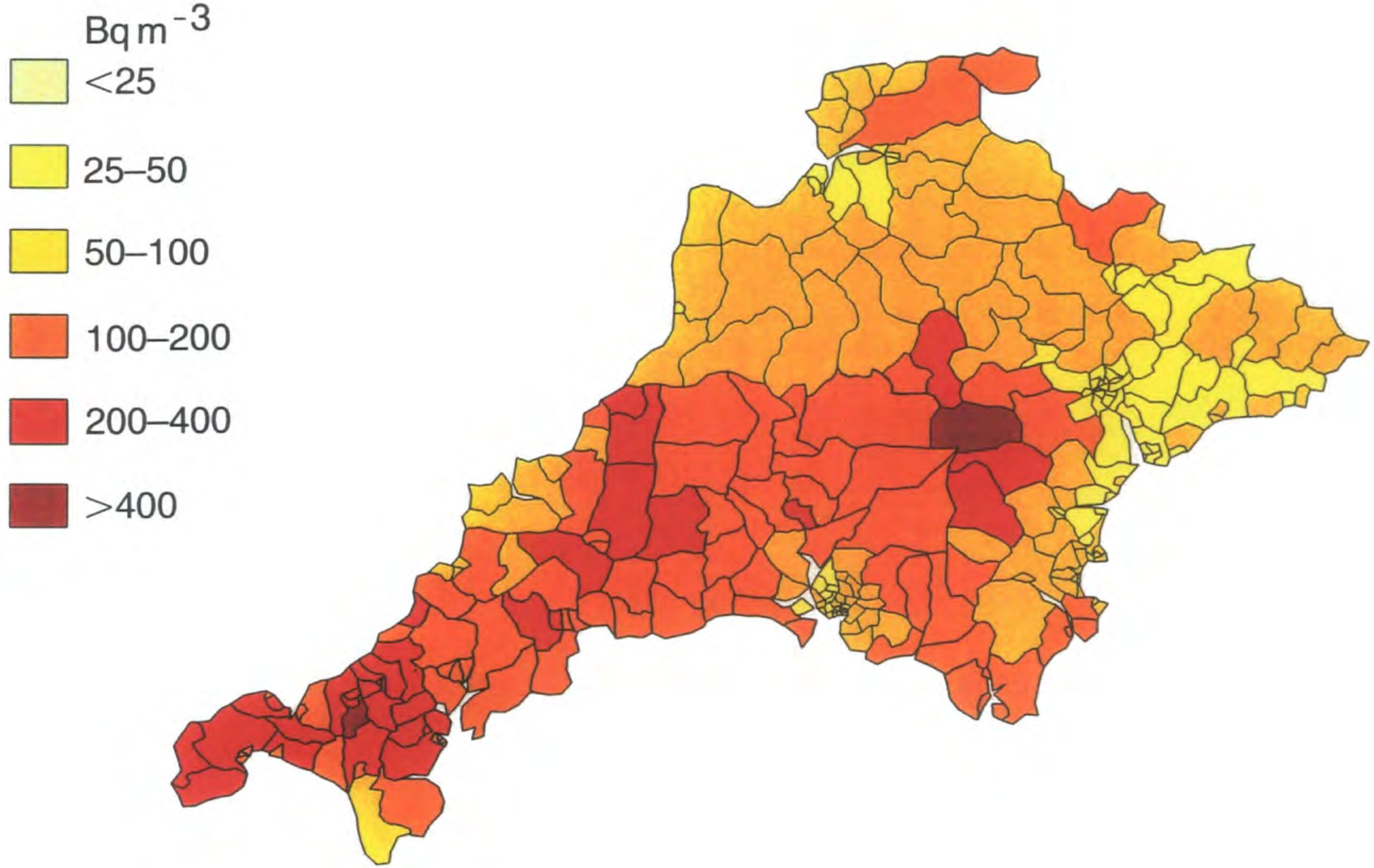


Figure 2.7 Radon concentrations in Cornwall and Devon by postcode sector

(Source: Green *et al.*, 1992)

information on the results of radon measurements in Scotland and Northern Ireland are due to be published shortly by the NRPB (Miles, correspondence, 1993; Green *et al.*, 1992; Wrixon *et al.*, 1988).

- Representative (using a stratified sample from the Post Office Address File) and infill surveys in all five Affected Areas have been carried out since January 1988 and these surveys, which are ongoing, have produced complete cover for Britain for five kilometre OS grid squares. Results from the continuous monitoring of these areas - Cornwall, Devon, Somerset, Northamptonshire and Derbyshire - are added to the five kilometre grid square database compiled by the NRPB (Green *et al.*, 1992)
- Radon levels in homes in Scotland (for postcode areas and 10 kilometre grid squares) (Green *et al.*, 1991).
- Island surveys, covering the Isle on Man, Scilly Isles and the Channel Islands (Green *et al.*, 1992).
- Number of homes *measured* for indoor radon in the United Kingdom, with details on the numbers exceeding the Action Level and the mean radon concentration - this survey is ongoing (in July 1993 the total number of homes measured in the United Kingdom stood at over 160,000) (Green *et al.*, 1992; NRPB, 1993).
- The distribution of homes that have been measured for radon gas at a five kilometre grid square resolution (Green *et al.*, 1992).
- The distribution of radon concentration in homes in the United Kingdom for grid squares, postcodes, local authority districts, counties and countries (Green *et al.*, 1992).
- Various surveys of radon in the workplace, including mine surveys dating back to the late 1960s (Green, 1991).
- The effects of remedial action on radon concentration in homes with radon levels in excess of 200 Bq m^{-3}.

(Sources: Wrixon *et al.*, 1988; Green, 1991; Green *et al.*, 1991; Green *et al.*, 1992; Miles, correspondence 1993; NRPB, 1993; Wrixon *et al.*, 1988)

2.4 Data Transformation

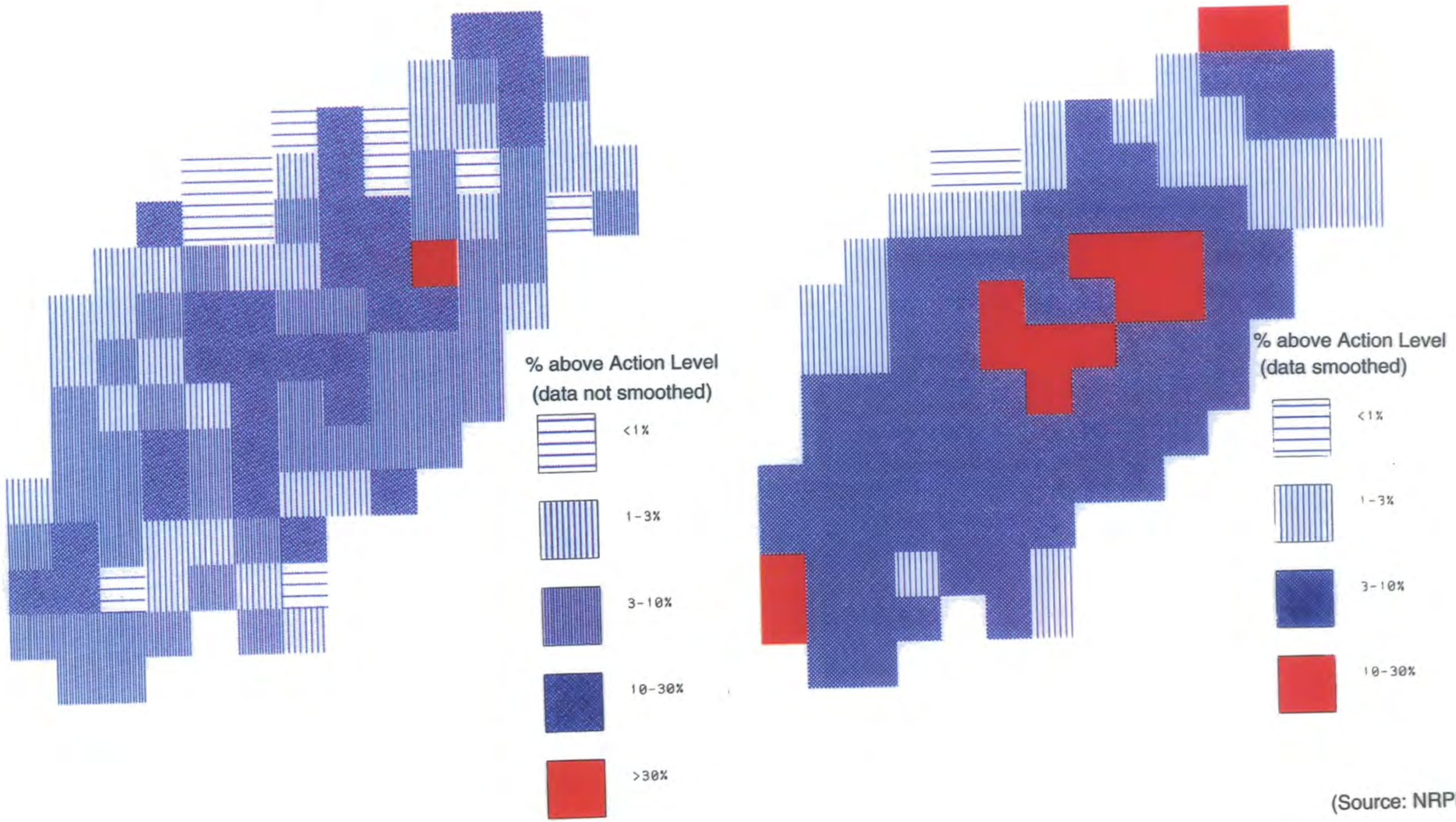
The choice of the type and format of radon data used in the project was limited by the data that the NRPB were in a position to provide for this study, but included all the up to date measurements of radon that had been assessed at the time of acquisition (early 1993) for the five counties. The data for Somerset, Derbyshire and Northamptonshire are more up-to-date than that for Cornwall and Devon due to the more recent date of the surveys in the former counties, however all the databases contain results that have been continuously updated. All of the data are in the form of averages for five kilometre grid squares as defined by the Ordnance Survey National Grid.

The radon data were smoothed by the NRPB using the method outlined below. In addition to pragmatism, there are a number of reasons for smoothing data. It can have a positive effect in terms of the visual impact that a map has on its viewer - reducing the noise in the data set enables the brain to make more sense of the pattern. There are also physical benefits, unless extreme care is taken over measurement, data are mixed with measurement error which can be reduced by using an appropriate smoothing method. Local or transient disturbances are of little interest when assessing the overall pattern. The method used by the NRPB enabled representative radon levels to be calculated for squares where no measurements had been taken, either because of their inaccessibility (significant radon levels can occur in upland areas such as Dartmoor) or due to low population densities as in parts of Cornwall, Devon, Somerset and Derbyshire. For the squares where there were no measurements, a value of D was estimated from the data for the eight adjacent squares. These squares, centred on the unknown square T form the base of a tetrahedron, with individual squares being the base of the fraction of the total volume. The value for each of the peripheral squares was then weighted by the corresponding fraction. Thus the value D for the unknown square T is given by:

$$D=\frac{3\sum X_i+2\sum Y_j}{20}$$

where X_i are the values for squares to the side of square T and Y_j are the values for squares on the corners of square T (Green *et al.*, 1989). If values for any of the adjacent squares are missing, this can be taken into account during the calculations by weighting the constants in the equation. The smoothing routine was used to give a more representative view of each of the squares, as a low number of readings can give a biased set of results for the squares (see Figure 2.8). The smoothing effect was carried out twice, as this gave the best indication of the radon levels in keeping with geological characteristics and NRPB radiological objectives (Green *et al.*, 1989). The final figure is one of five percentages which relates to the estimated proportion of homes above the Action Level. The categories are as follows: <1%, 1 - 3%, 3 - 10%, 10 - 30% and >30% of homes above the Action Level. The resulting maps, created by shading each grid square according to its radon category, can be seen in Chapter 12. Lines joining areas of approximate equal average radon levels can be added to the maps to show the approximate extent of the areas where more than 30% and more than 10% of homes are affected. The resulting isolines can not be accurate to more than a five kilometre resolution however because of the nature of the data and 'because even in areas of apparent geological homogeneity there is a wide spread in domestic radon

Figure 2.8 A comparison between smoothed and not smoothed radon data for Northamptonshire



(Source: NRPB)

concentrations' (Ball and Miles, 1993). Analysis of county-wide radon patterns in the Affected Areas is discussed in Chapters 12 and 13, together with the results of the overlays between radon and land capability and census data.

3. Land Use

The study of land use is a complex interdependent and interactive combination of classification, scale, survey, recording and interpretation, in which the selected scale largely determines the classification potential and the presentation of the observed data (Balchin, 1984). Initially patterns of land use were determined entirely by the natural environment, but since the Industrial Revolution other pressures have modified the basic pattern. Although many of the underlying correlations between land use and the physical environment exist, they have been highly modified by industrial, technological and socio-economic factors. There are 24 million hectares of land in the United Kingdom, three quarters of which is agricultural land. Even in the regions of less intensively used land, patterns of current land use are complicated by the large number of competing uses, producing a complex mosaic of different land uses and cover types across the United Kingdom as a whole.

It is useful at this stage, to define the term 'land use' and to differentiate between land use and land cover. Land use relates to the types of use to which the land is put, for example, agriculture, forestry, industry or residential. Land *cover* on the other hand, describes what the land looks like, for example roads and buildings, woodland, crops and areas of water. The following chapter presents a summary of some of the main sources of both land use and land cover data for England.

3.1 The History of Land-Use Mapping

The first attempts at land-use mapping occurred during the renaissance. William Cuningham, a physician from Norwich, was one of the first to put forward the idea of the map as 'a mirror of reality' in the *Cosmographical Glasse* in 1559. He was followed by the great county surveyors of the 16th and 17th centuries. The property survey however (rather than the county survey), is the forerunner of the modern day land-use map (Balchin, 1985).

There are essentially three main categories common to all true land-use maps. These may be termed 'super-categories' and are settlement, farmland and vegetation (Balchin, 1985). A classification system is necessarily a compromise between conflicting anticipated needs, as a result of the heterogeneity of the real world. The addition of many sub-categories may be deemed desirable or necessary depending on the scale and the area being surveyed. The extent of such a division reflects the required complexity and the skill of the surveyors or cartographers in the map production. Too many divisions and the map may become illegible, too few and only generalisation may be achieved.

The first recognisable land-use maps - maps specifically portraying the spatial distribution of the natural or man-made land use - appeared in England in the 1790s following the creation of the Board of Agriculture in 1793. In the first two decades of this century the Sociological

Society and later the Le Play Society (both of which were intimately associated with Geddes, as well as Mill, Fleure, Herbertson, Mackinder and many other social scientists) were responsible for numerous local surveys, planning surveys, exhibitions and novel presentations of material (Board, 1968). It was not until the early 1930s however, that the first systematic land-use survey of Britain was carried out and the results subsequently published fully in both map and text form. This mammoth undertaking is called the First Land Utilisation Survey and was a milestone in the development of nationwide, systematic and detailed mapping of the natural and man-made land use of the United Kingdom. Fox (1956) points out however that, "the Land Utilisation Survey of Britain began as an academic exercise, but the foresight and energy of L. Dudley Stamp and E. C. Willatts raised it to the status of a national inventory". Fox has also drawn attention to differences of meaning between *land use*, a formal concept and *land utilisation*, a functional concept. These distinctions are by no means universally accepted however and the terms are often used interchangeably. In this thesis, the term land use refers to both land use and land utilisation, except where specific reference is made to the two national Land Utilisation Surveys of Britain.

3.2 Land-Use Surveys of Britain

The First Land Utilisation Survey was formally launched in October 1930, with the object of making a complete cartographic record of the uses of the 'surface of the country' in 1931 - 1933 (Stamp, 1932). The survey is inextricably linked with the name L. Dudley Stamp and with the aid of 10,000 volunteer schools and 250,000 children he initiated pilot surveys to work out a national scheme of land-use categories. The great virtue of the system adopted by Stamp was its simplicity. Each land parcel (subdivided where surveyors deemed it necessary) was categorised under one of nine basic land-use types (Fuller *et al.*, 1994) (see Table 3.1 and Figure 3.1). Listed alongside the land use categories in Table 3.1 below, is the colour that was used to represent each category on the printed maps. The maps were surveyed in the field (each site was visited by at least one of the volunteers) at the scale of 1:10,560 and published at the scale of 1:63,360. A generalised map of the whole of Britain at 1:625,000 was published in two sheets in 1943. The resulting survey provides an immensely rich source of detailed information on the nature of the British landscape in the 1930s.

Figure 3.1 Extract from the First Land Utilisation Survey

Sheet 127 (River Torridge)

Scale 1:63,360

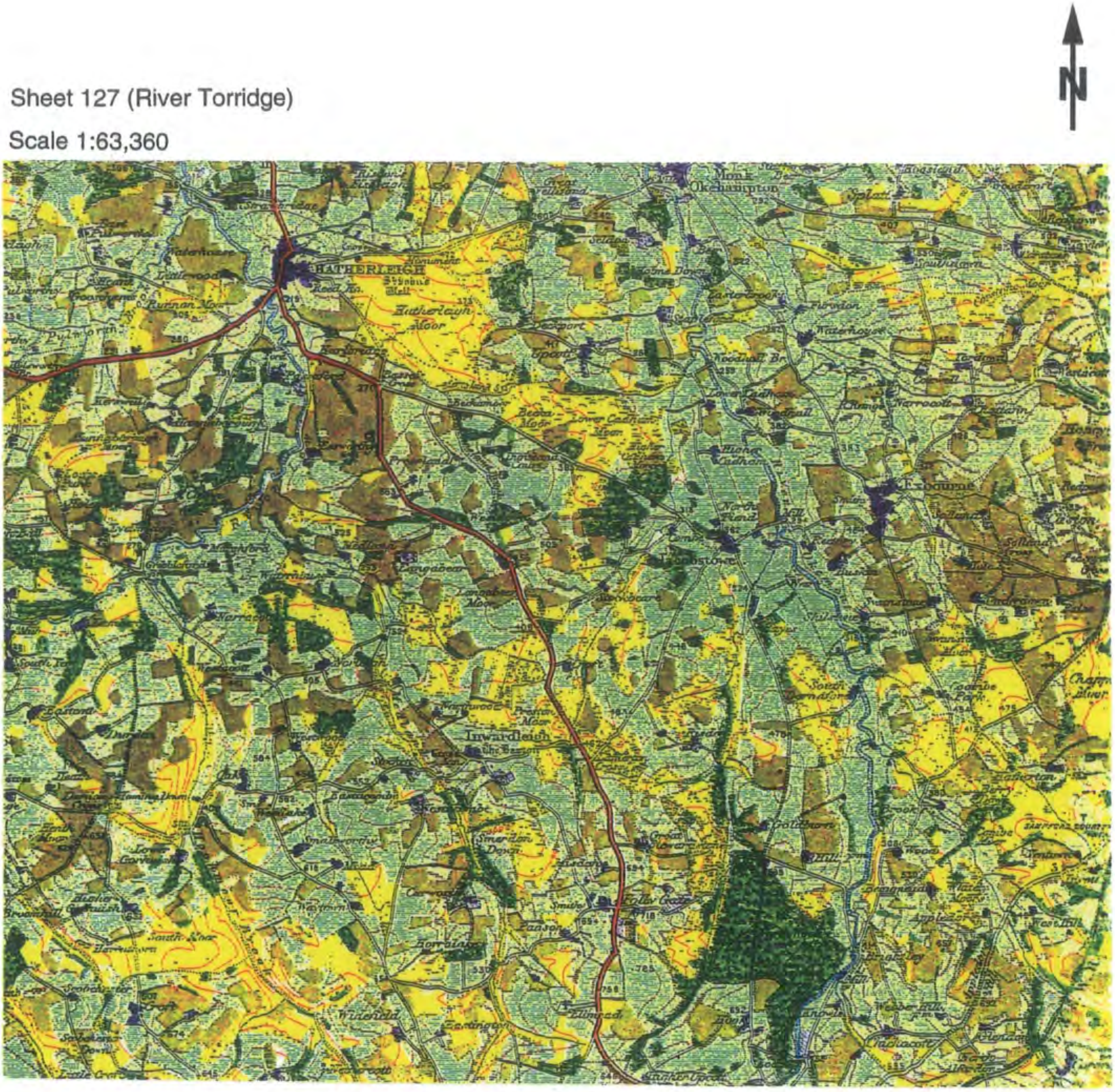


Table 3.1 Categories Used in the First Land Utilisation Survey of Britain

Printed Colour	**Major Land Use Category**	**Sub-Category**
Brown	**Arable and Ploughed Land (including fallow, short ley, rotation grass and market gardening)**	
Dark Green	**Forest and Woodland**	*DECIDUOUS*
		CONIFEROUS
		MIXED
		NEW PLANTATIONS
Light Green	**Meadowland and Permanent Grass**	*GRASSLAND IN PARKS*
Yellow	**Heath and Moorland**	*HEATH MOORLAND, COMMONS AND ROUGH PASTURE*
		ROUGH MARSH PASTURE
Purple	**Houses and Gardens**	*HOUSES WITH LARGE GARDENS*
		NEW HOUSING AREAS, NURSERIES AND ALLOTMENTS
Red	**Land Agriculturally Unproductive**	*CLOSE SETTLEMENT AND INDUSTRY*
		YARDS, CEMETERIES, PITS AND NEW INDUSTRIAL WORKS
		Main roads and stations
(Conventions)	**Special Types of Land**	*Fens, Cliffs, Golf courses, sports grounds etc.*
(Conventions)	**Orchards**	*Scrub and Orchard*
(Conventions)	**Nurseries**	*Tilled*

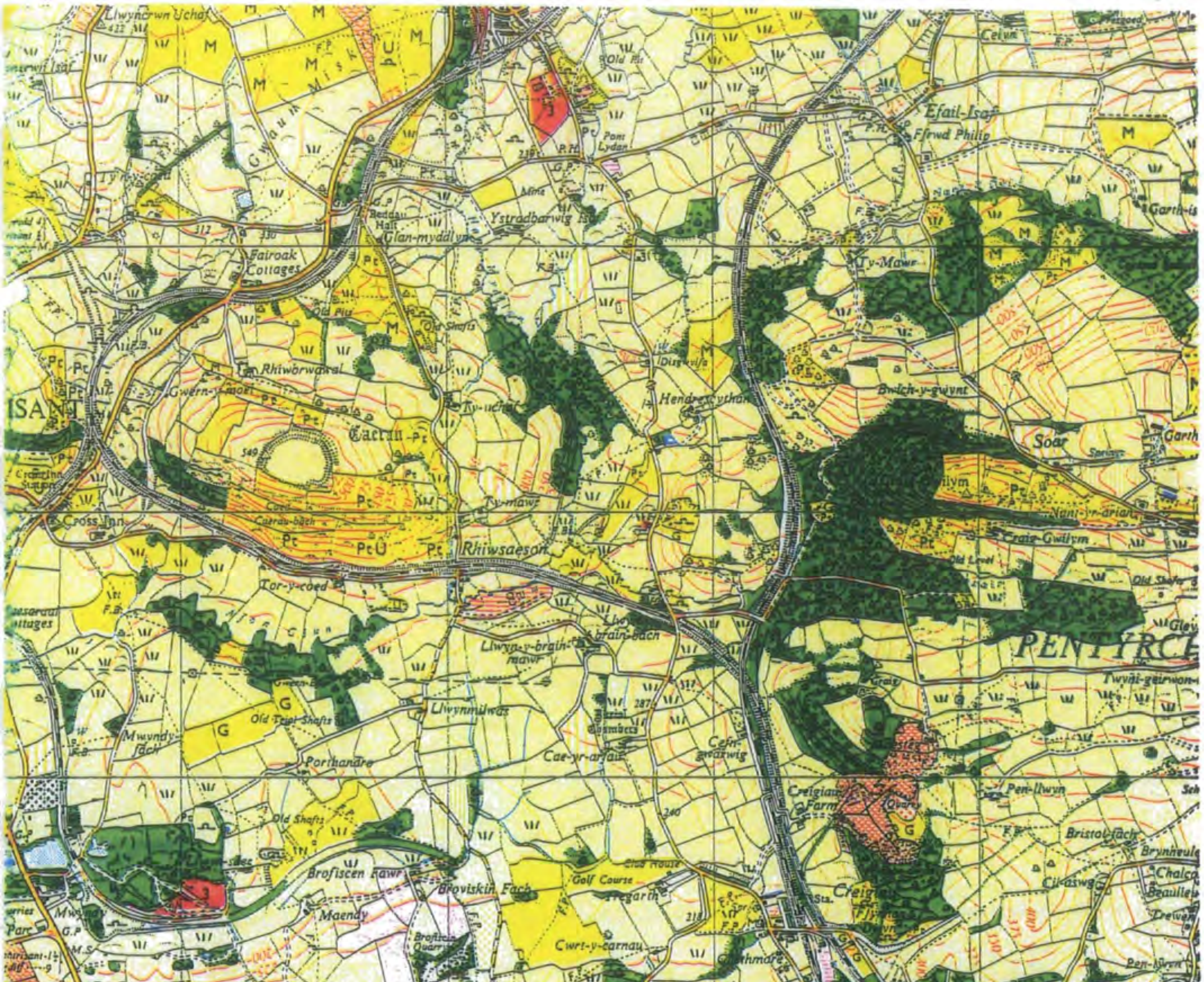
(Source: Balchin, 1985; Fuller *et al.*, 1994; First Land Utilisation Survey maps)

A further important contribution by Stamp was the production of a set of 92 county reports which provided more detail about the mapped land uses. In nearly all the reports, each category of land use is separately mapped and described, and each one has a map of land-use regions (Board, 1968). Although Stamp acted as the chief data collator and editor, he was helped in his task by a number of distinguished academics, such as E. C. Willatts, J. W. Birch, J. Clarke, J. T. Coppock. Well over half the county reports were in fact written by members or graduates of the Joint School of Geography (King's College, London and the London School of Economics) and many had been taught by Stamp himself (Board, 1968). The reports were followed up in 1948 by the publication of a national report entitled *The Land of Britain: Its Use and Misuse.* Stamp's methodical and analytical work demonstrated the practical importance of a land-use survey and encouraged a more thoughtful approach to the environment that set a precedent for ensuing land-use surveys. The reports provide an incomparable source of information on ideas and methods employed in land-use surveys and represent a considerable amount of work by some of the leading geographers of the decade. A complete set of the First Land Utilisation Survey maps is available for reference purposes in the seminar library of the Department of Geography at Durham University, but their use for research purposes is inevitably restricted because of their age (the maps are over 50 years old and are therefore considerably out-of-date), to historical analyses. Coppock (1978) examines the quality of the First Land Utilisation Survey in detail. The level of detail shown is considerable and the information is used to best advantage in studies of localised areas or regions, but not for inter-county comparisons. The maps and a full set of reports are also available from the Map Curator, London School of Economics, Houghton Street, London.

The extension of land-use surveys to many other countries after the second world war has been seen by many as a direct result of the ideas and enthusiasm of Stamp. At the International Geographical Union Conference in 1949 it was decided to establish a World Land-Use Survey Commission and a classification and programme of action were agreed upon (Board, 1968). By 1956 over 50 countries had began some form of land-use survey, though not all of these resulted in a complete national cover (Balchin, 1984). By 1960 a second national land-use survey of Britain was underway.

The Second Land Utilisation Survey of Britain, initiated by Alice Coleman, aimed to repeat the six-inch field coverage of the first survey but with a considerably more detailed classification producing roughly ten times as many categories. There are two levels of classification in the survey; major groups are represented by 13 classes, with various subdivisions which are shown by conventions. For example, the 16 industry classifications (based on the industrial classification of the 1951 Census) are overprinted on the red first order industry category using the numbers one to 16 to denote the type of industry; 5 denotes metal manufacture and 10 denotes textiles for example. The portrayal of 64 categories in total on the published maps was made possible by using the 1:25,000 scale (surveyed at 1:10,560) which did not exist in Britain at the time of Stamp's survey. It also had the immense advantage of showing field boundaries, but the increased cost prevented the publication of a complete cover for Britain and only 120 maps were printed. This has greatly restricted the use of this survey for research into changes over time and for comparisons between areas such as counties and is therefore of limited practical use in this project. All of the printed maps, as well as the manuscript maps for those areas which were not published, are available from the Department of Geography, King's College, Strand, London. Sinclair (1992) compares the data collected in the two national land-use surveys with other data collected for England and provides a quantitative summary of the findings of the early land-use surveys. Figures 3.1 and 3.2 show map extracts of the two Land Utilisation Surveys and the different level of detail (a function largely of the map scale) can clearly be seen. Figure 3.1 is a map extract from the First Land Utilisation Survey, from either Sheet 127 (River Torridge) or 128 (Tiverton). Figure 3.2 is an extract from the Second Land Utilisation Survey, Sheet 213 (Pontypridd). Table 3.2 below, lists the 13 first order categories used on the Second Land Utilisation Survey maps (shown in bold together with their representative colour) plus the major sub-categories.

Figure 3.2 Extract from the Second Land Utilisation Survey



Scale 1:25,000

Sheet 213 (Pontypridd)

Table 3.2 Categories Used in the Second Land Utilisation Survey of Britain

Printed Colour	Major Land Use Category	Sub-category
BROWN	ARABLE	*CEREALS* *LEY LEGUMES* *ROOTS* *GREEN FODDER* *INDUSTRIAL CROPS* *FALLOW*
DARK GREEN	WOODLAND	*DECIDUOUS* *CONIFEROUS* *MIXED* *COPPICE* *COPPICE WITH STANDARDS* *WOODLAND SCRUB*
LIGHT GREEN	GRASS LAND	
LIME GREEN	OPEN SPACES	*TENDED* *UNPRODUCTIVE LAND*
YELLOW	HEATH AND MOORLAND	*HEATH MOORLAND, COMMONS AND ROUGH PASTURE* *ROUGH MARSH PASTURE*
PURPLE	MARKET GARDENING	*FIELD VEGETABLES* *MIXED MARKET GARDENING* *NURSERIES* *ALLOTMENT GARDENS* *FLOWERS* *SOFT FRUIT* *HOPS*
PURPLE	ORCHARDS	*WITH GRASS* *WITH ARABLE LAND* *WITH MARKET GARDENING*
GREY	SETTLEMENT	*COMMERCIAL AND RESIDENTIAL* *CARAVAN SITES*
RED	INDUSTRY	*MANUFACTURING* *EXTRACTIVE* *TIPS* *PUBLIC UTILITIES*
ORANGE	TRANSPORT	*PORT AREAS, AIRFIELDS* *MAJOR ROADS* *OTHER METALLED ROADS*
BLUE	WATER AND MARSH	*WATER* *FRESHWATER MARSH* *SALTWATER MARSH*
WHITE	UNVEGETATED LAND	
BLACK	DERELICT LAND	

(Source: Balchin, 1985; Fuller *et al.*, 1994)

During the 1960s the Ministry of Agriculture decided that a national series of agricultural land classification maps for England and Wales was required in order that good quality agricultural land could be protected against urban encroachment. The initial plan was to have a three stage system of classification but only the first stage (a simple physical classification) was carried out. The aim was to devise a system of classifying land according to the degree to which physical characteristics impose long-term limitations on agricultural use. National coverage has been published at 1:63,360 with 113 sheets covering England and Wales. An updated version of this series was published in the early 1980s at 1:250,000 in seven sheets. A map at the scale of 1:625,000 was published in 1979 and covers the whole of England and Wales. The classification originally consisted of five grades of agricultural land, though a later modification introduced the two sub-sets 3a and 3b in order to reduce the dominance of Class 3 (good to moderate agricultural land) and increase the accuracy of the classification system. The set of revised guidelines were published by the Ministry of Agriculture, Fisheries and Foods (MAFF) in 1988 and they incorporate better assessments of climatic parameters and chemical status assessment. Further detailed information on this series is contained in section 3.4.

A Land Capability Classification has also been developed by the Soil Survey, modelled on the United States Department of Agriculture scheme. The prime aim of the scheme was to present the results of soil surveys in a suitable form for planners, agricultural advisers, farmers and other land users. The classification is primarily for agricultural purposes and distance to markets, types of road and farm structure are not taken into account in the assessment procedure. Like the British method, the US Capability Classification has seven classes, with class 1 being land with very minor or no physical limitations to its use and class 7 land with extremely severe limitations that can not be rectified (Bibby and Mackney, 1969).

The MAFF agricultural census is one of the major sources of information for those engaged in rural studies, in that it provides long runs of data about farming and land use. In the United Kingdom, the earliest experiments in collecting agricultural information in a recognisably modern form were conducted by Justices of the Peace or the clergy between 1795 and 1803 (Clark, 1982). The modern-day United Kingdom agricultural census is taken every year on 4 June (1 June in Northern Ireland) and the resulting statistics are published yearly as tables. In 1991 MAFF published the first edition of the 'Digest of Agricultural Statistics', which brings together the main results of the annual agricultural censuses of England, Wales, Scotland and Northern Ireland. The Digest includes data for five and ten year comparisons for the United Kingdom for land use, cattle, pigs, sheep, poultry and labour. It also includes annual summaries (for the above categories and others), summary data down to county level (regions in Scotland), maps of the main items and frequency distributions of the main holdings by size (MAFF *et al.*, 1991). The basic unit of enumeration is the holding, or more loosely the "farm". The guideline definition is pragmatic and operational and is subject to agreement with the individual farmer. The term "farming" encompasses horticultural activity. In 1991 there were 241,000 main agricultural holdings in the United Kingdom. This figure does not include holdings of less than six hectares (MAFF *et al.*, 1991). From analysing the census statistics, it is clear that whilst Cornwall, Devon, Somerset and Derbyshire are characterised by dairy farming and cattle and sheep rearing, Northamptonshire is characterised by cropping (arable farming). Indeed, Northamptonshire accounts for only 2% of the total area of England but is responsible for 3.4% of wheat production and 5.2% of oil-seed rape production. Other counties also specialise in certain crops. For example Cornwall, with 3% of the total area, accounts for 9.5% of the annual production of hardy nursery stock. Somerset, with 2.9% of the total area, is responsible for almost 4% of all orchards. Devon on the other hand, accounts for 5.5% of the area total but 8.8% of the grassland total for England. A summary of the main findings for the five counties is given in Table 3.3 below.

Table 3.3 Agricultural Statistics for all the Main Holdings in the Affected Areas (1991 Agricultural Census)

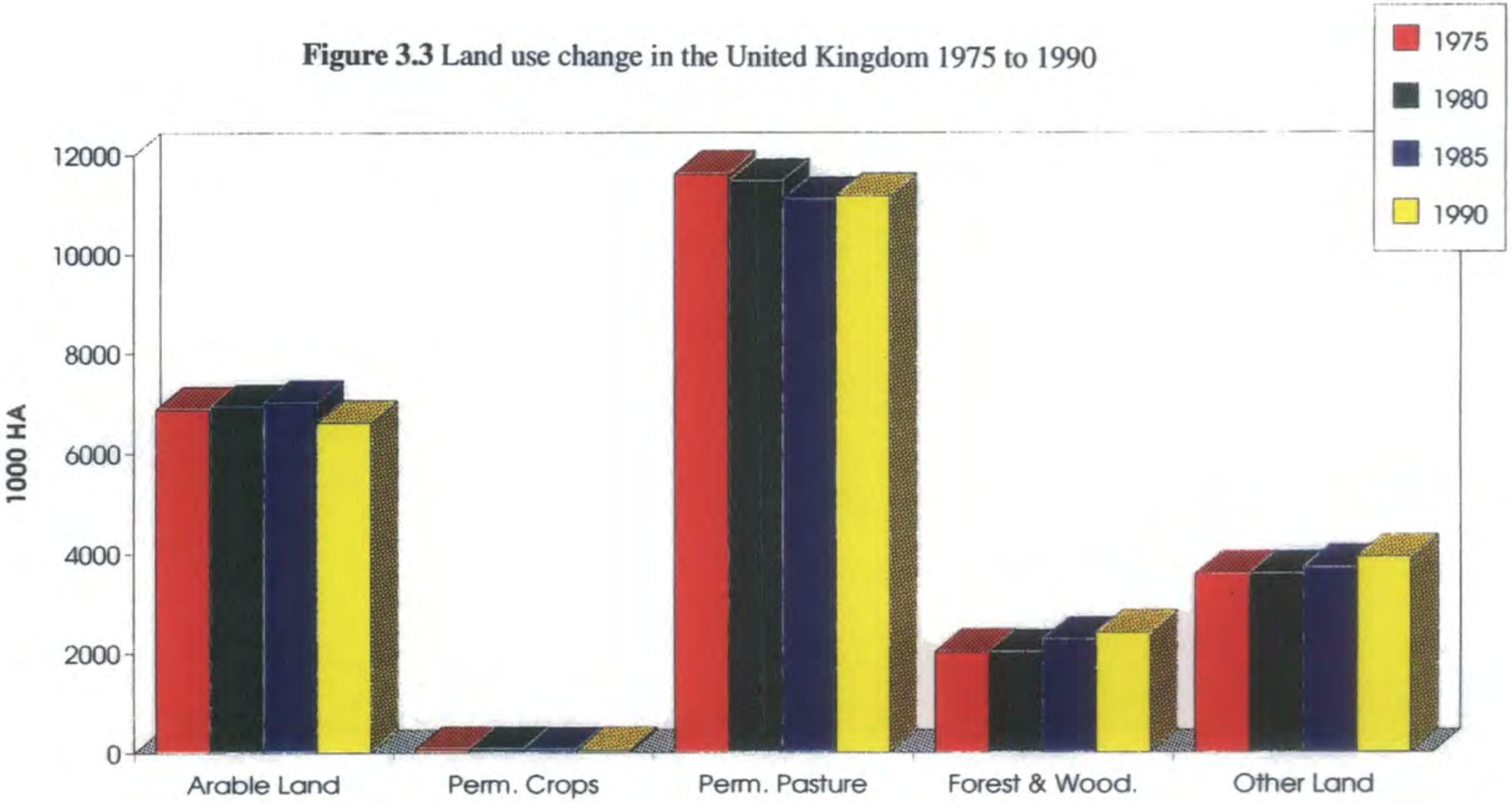
	Cornwall	**Devon**	**Somerset**	**Northants**	**Derbyshire**
LAND*					
Total agricultural area	3.0	5.5	2.9	2.0	2.0
Total cereals	1.3	2.0	1.3	2.8	1.0
Other arable crops	1.0	0.6	0.7	2.1	0.6
Horticultural crops	3.4	1.5	2.2	0.3	0.4
Grassland	5.0	8.8	4.7	1.4	2.5
All other	2.9	5.5	2.1	2.0	1.0
FARMS+					
Dairying	34.2	31.4	40.4	4.6	31.4
Cattle and sheep	41.5	48.9	31.7	16.0	39.6
Cropping	15.5	11.4	16.8	75.3	21.3

* % as a % of England (Source: MAFF *et al.*, 1991)

+ % as a % of the total number of farms in the county

The MAFF agricultural statistics are of limited use in this project because they lack precise geographical locations. They play an important secondary role however, in providing regional and national figures for land-use change over time. Figure 3.3 shows land-use change from 1975 to 1990 for the United Kingdom. The graph shows relatively little change in the area of each land-use category since 1980-1985; permanent pasture and arable land have decreased by 617,000 hectares since 1980 and forest and woodland and other land (*e.g.* urban) have increased by 625,000 hectares since 1980, in the United Kingdom as a whole. The capability of

Figure 3.3 Land use change in the United Kingdom 1975 to 1990



Source: FAO Production Yearbook 1991

land for agriculture however, is unlikely to alter very much over time, unless new agricultural methods, crops or chemicals are created. The MAFF Land Capability Classification (see section 3.4) takes into account possible future improvements to the land and is therefore not prone to drastic temporal change, as can be the case with some land-use classifications. The Council for the Protection of Rural England also publish regional and county level land-use change statistics for intervals of five years and the statistics extend back to 1945. Their latest publication by Sinclair (1992) is entitled 'Regional Lost Land'.

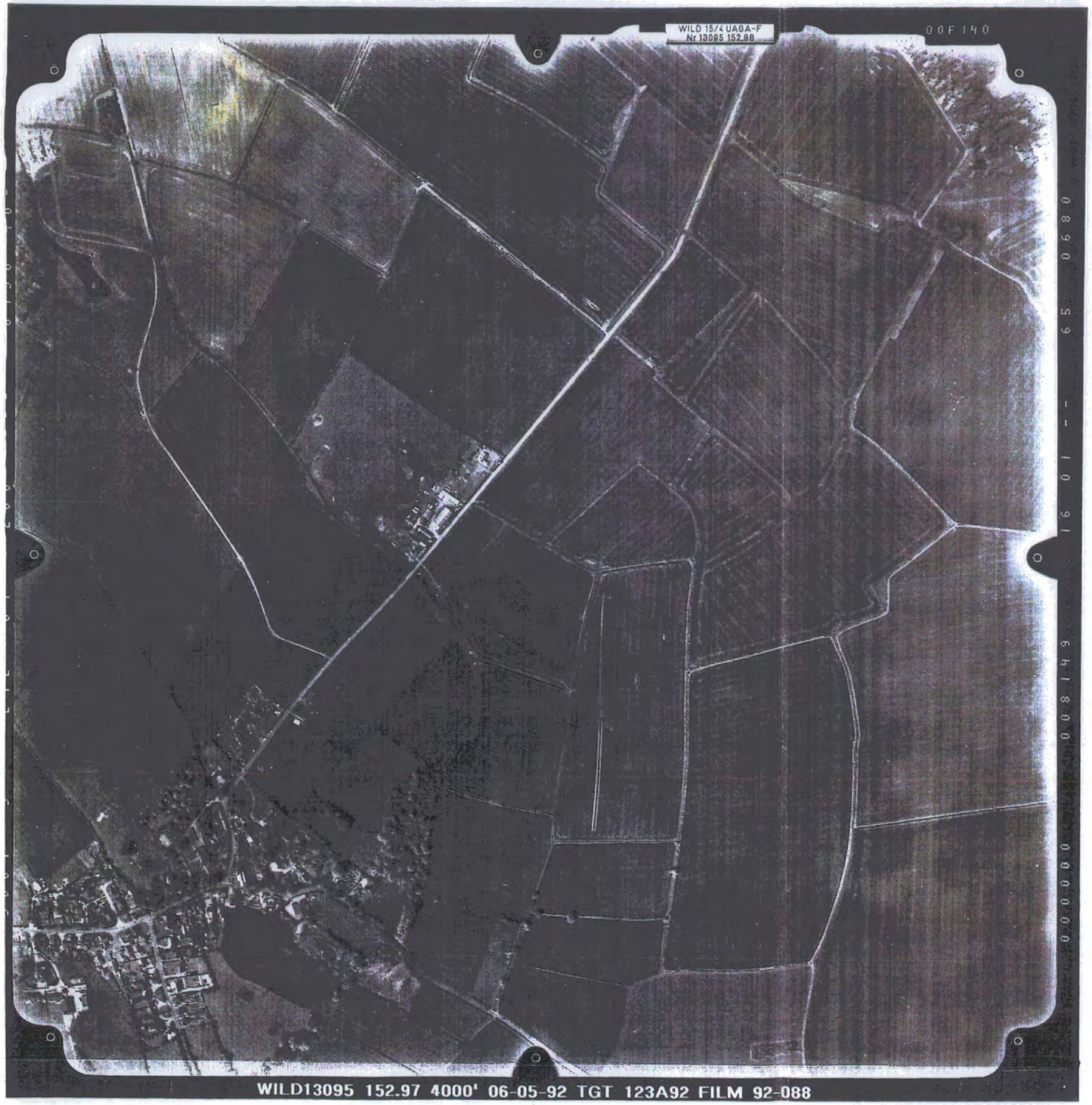
The Ordnance Survey, established in 1791, have published many maps over the last two centuries, at many different scales. Their maps have shown progressively more land-use information but they have not yet become comprehensive land-use records in their own right. The latest agricultural census, carried out in 1993 required each farmer or land owner to record the land use and the crops being grown on their land, on an Ordnance Survey map at 1:10,000. Results from this survey are in the process of being analysed. The completion rate by farmers however, is thought to have been low. Associated with this form of land-use survey, is another form of survey that provides land-use information as a spin-off from its main purpose. In an ambitious move to combat agricultural-support fraud (estimated to cost the European taxpayers between £2 billion and £6 billion per year), satellite-based monitoring programmes are currently being conducted in every European Community country. In the United Kingdom, MAFF has contracted the National Remote Sensing Centre Limited to monitor farmland in what is the largest such programme ever conducted in the United Kingdom (Burrill and Terres, 1994). The work focuses on the determination of both area and land use from Landsat and SPOT satellite imagery.

In the post-war period emphasis has been on the development of remotely-sensed methods of land-use data capture. Remote sensing is the ability to gather information about an area or object without being in direct contact with it (Hilton *et al.*, 1988). Aerial photographs are one example of a remote method of data capture and by the 1970s aerial photography was frequently used in local and regional ecological studies (Fuller *et. al.*, 1994). In addition to the original monochrome vertical and oblique photographs there is now stereoscopic, colour and infra-red photography. They provide a static but comprehensive view of land use which can be analysed to give detailed information. All aerial photographs contain certain distortions and cannot be used for taking measurements without specialist knowledge and equipment. Manual analysis of photographic data involves a great deal of work, namely in the conversion of the data into digital format. It is difficult to categorise and quantify land use from photographs and a methodology would have to be developed for the purposes of this study, followed up by ground surveys to check for errors and inconsistencies.

Aerial photographs of all types and ages are available from County Councils and the Ordnance Survey, as well as a multitude of private companies. The level of accuracy and detail is dependent on the quality and resolution of the photograph and the amount of cloud cover at the time of the flight. Most aerial photographs of the United Kingdom are affected to some degree by cloud cover which can obscure information. The photographs are available in a wide range of scales but need to be large scale in order to be legible. Figure 3.4 is a laser proof of a photograph that was purchased from the Ordnance Survey and shows Great Addington (grid reference SP9675), south-east of Kettering in Northamptonshire. A location map has been included to show the general area of the photograph. The photograph was taken in May 1992 at a height of 4,000 feet and is at the scale of approximately 1:7,500. This example highlights the fact that several hundred photographs would be required in order to cover the five counties that are being studied.

It was not until satellite images, with their digital output and potential for automated analysis, became available from the 1970s onwards that the real possibilities began to emerge of a national survey based on radically different methodologies from those used by Stamp in the 1930s (Fuller *et al.*, 1994). Satellite remote sensing involves the use of orbiting satellites as platforms and radiometers as onboard instruments. There are a variety of satellites currently orbiting the Earth (see Table 3.4 below) which are capable of relaying detailed and up-to-the-minute land cover information (as well as other data, such as hydrological and geomorphological data) to stations on the ground. Remotely-sensed images are continuously recorded by the on-board sensors all year round. The way in which an individual land use varies between wave lengths is its spectral signature and these spectral signatures differentiate uses that may seem similar on natural colour photographs. The American series of Landsat satellites has provided high quality multispectral data since Landsat 1 was launched in 1972, however it was not until the 1980s and the launch of the Thematic Mapper (TM) that Landsat provided an opportunity to map land cover at a field-by-field scale (Fuller and Pasell, 1990). A full scene of a Landsat image covers approximately 31,450 square kilometres (170km by 185km). There are five Landsat satellites in total with both Multispectral Scanner Systems (Landsats 1, 2 and 3) and TM equipment (Landsats 4 and 5) on board. Landsat 6 was launched in 1993 but the mission failed and Landsat 7 is at the planning stage. Landsats 1, 2 and 3 complete a full near polar orbit of the Earth every 99 minutes at an altitude of 913 and the resolution of the MSS data is 80 metres square. The MSS wavelengths for vegetation mapping range from 0.7 to 1.1 µm in the red to near infra-red spectrum. Landsat 4 and 5 orbit at 705 kilometres and the resolution of the TM data is 30 metres square. The TM wavelengths used for mapping various aspects of vegetation range from 0.45 to 0.69 µm in the visible band. The TM is built to record radiation in some parts of the electromagnetic spectrum only, and these are known as 'bands' and are labelled 1 to 6 in Figure 3.5. The TM records the reflectance or emittance for each of its bands in digital format, with zero for no radiation reflectance and 255 for maximum reflectance. These numbers when transmitted to the earth, are used to construct satellite images. (Fox, 1991; Hilton *et al.*, 1988; Strain and Engle, 1994)

Figure 3.4 Aerial photo showing Great Addington, Northamptonshire

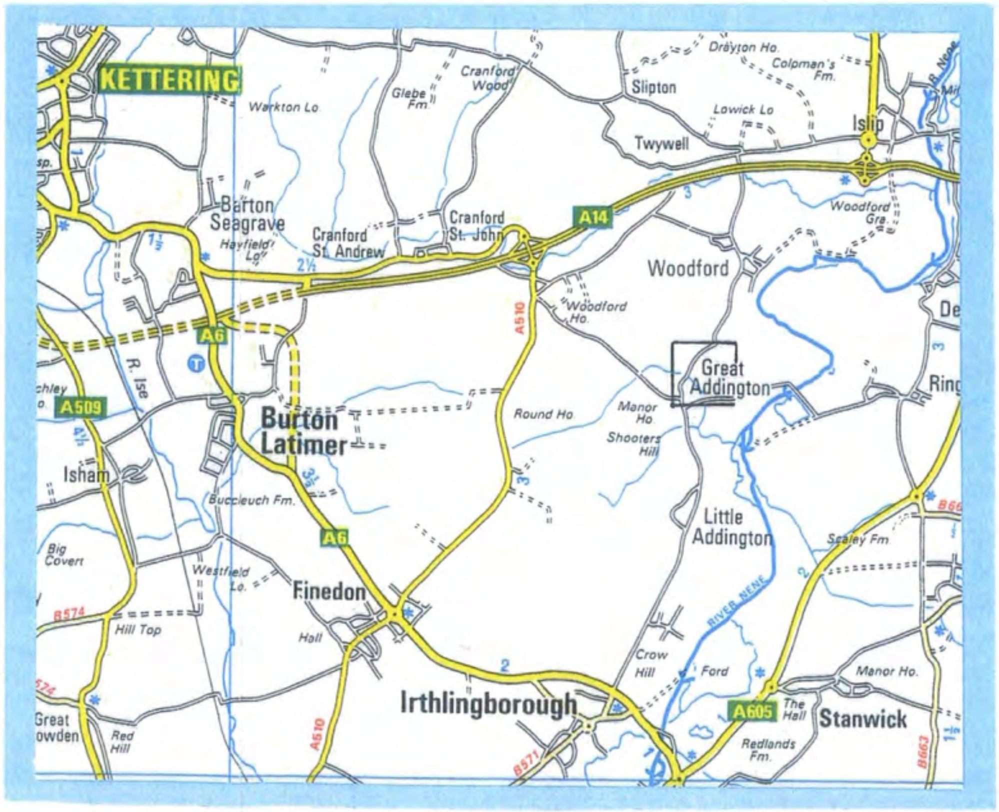


© Crown Copyright Reserved

Taken on 6 May 1992 at 4,000ft

Approximate scale 1:7,500

N



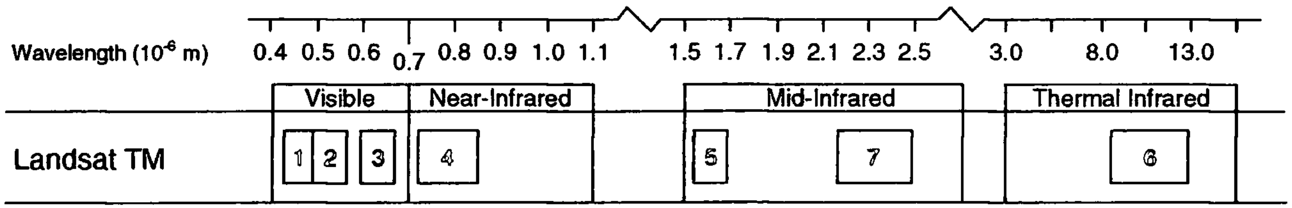
(Source: Philip's Motoring Atlas of the United Kingdom, 1994)

Figure 3.6 shows a Landsat TM image of the Peak District and Sheffield taken on 28 June 1986 at the scale of 1:117,000. The spatial resolution is 30 metres square. The highland on the left of the image is underlain by Millstone Grit with shale forming the valleys, and moorland vegetation is shown in dark tones (for example D 19). South of this region is the Vale of Edale and then the limestone plateau. On the right of the image Sheffield dominates the lowland farming region on the coal measures (Kay and May, 1992).

There are other satellites in orbit which also provide high quality land cover data. Amongst them are the French SPOT (Satellite Pour L'Observation de la Terre) satellite series, capable of a resolution down to ten metres square in the panchromatic mode and 20 metres square in the multi-spectral mode. SPOT-1 and SPOT-2 are currently orbiting the Earth and SPOT 3 (launched in autumn 1993) was formally declared operational on 29 November 1993. The new developments in SPOT-3 include a higher responsivity and wider spectral response of band B2 than SPOT 2. SPOT-4 is currently being planned and a possible launch date is 1996. The European Space Agency's research satellite ERS-1, launched in July 1991, has a polar orbit at an altitude of 780 kilometres above the surface of the Earth and has on board two sensors. The first is the Synthetic Aperture Radar (SAR) which has a resolution of 30 metres and the second is a scanning radiometer with a resolution of one kilometre (Strain and Engle, 1994). A second satellite ERS-2, is planned for 1994. The ERS satellite is a useful source of data because it can observe and measure the environment regardless of cloud and illumination conditions (Mather, 1992). The Japanese have developed an earth resources satellite called JERS-1, launched in February 1992, capable of producing high quality images with a resolution down to 20 metres square. JERS-1 is particularly useful for mapping geological structures on the Earth's surface and has on board a visible and near-infrared radiometer and a SAR. The Japanese also own and maintain the Marine Observation Satellites (MOS-1A and 1B), but the use of these data for land cover mapping is limited by their poor ground resolution (see Table 3.4 below). The Indian National Remote Sensing Agency have launched two satellites, known as IRS-1A and 1B, which fly a polar orbit at an altitude of 904 kilometres. On board are three linear imaging self scanned sensors, with a resolution of between 36 and 72.5 metres (Strain and Engle, 1994). The American NOAA (National Oceanographic and Atmospheric Administration) AVHRR (Advanced Very High Resolution Radiometer) series of satellites (NOAA-6 to -12) produce both high quality hydrological data and land cover data. Aspects of these and other established satellites capable of producing data for land cover mapping are given in Table 3.4 below.

Perhaps the most significant advance in recent years however, has been the release of military satellite data since the end of the Cold War. These data are often very detailed and can be used for many purposes and are especially useful for land cover mapping. The KFA Russian satellites have flown two main flights and although the data collected are photographic in origin they have been digitised to produce high quality digital data. On board the main flight were two camera systems, the MKF-6, made up of six cameras capturing data in the range of 0.46 and

Figure 3.5 Landsat bands



0.86 μm at a resolution of 20 metres. The second system, the KATE-140 operated in panchromatic mode with a resolution of 60 metres. A second flight by the Russian satellite carried a KFA-1000 (resolution five metres), MK-4 (resolution six metres, operating in four bands between 0.4 and 0.9 μm) and KATE-200 (resolution 15 to 20 metres, operating in three bands between 0.51 and 0.85 μm) series of cameras (Strain and Engle, 1994). Each scene of the KFA-1000 (panchromatic mode) covers an area 80 kilometres by 80 kilometres. The data are produced by Camera Systems based in Russia and marketed by Sojuzkarta in Moscow, but are also available through marketing outlets in the United Kingdom. Although the resolution of the data is high, the cost (one scene currently costs around £1,800 plus VAT) is likely to limit its use within the research community. American military satellites have also released some data hitherto unseen. Strain and Engle (1994) provide an up-to-date summary of satellites world-wide, together with an interesting range of examples of land cover maps that have been produced from satellite images.

Table 3.4 Characteristics of the principal imaging sensors operating in the solar (0.4 - 2.5 μm) and thermal (3 - 100 μm) regions of the spectrum

Platform	Sensor	Wavebands (μm)	Repeat Cycle	Overpass time	Resolution
NOAA	AVHRR	0.58 - 0.68 0.73 - 1.10 3.55 - 3.93 10.5 - 11.5 11.5 - 12.5	12 hours	1430 / 0230 1930 / 0730	1.1 km
Landsat	MSS	0.50 - 0.60 0.60 - 0.70 0.70 - 0.80 0.80 - 1.10	16 days	0945	80 m
Landsat	TM	0.45 - 0.52 0.52 - 0.60 0.63 - 0.69 0.76 - 0.90 1.55 - 1.75 2.08 - 2.35 10.4 - 12.5	16 days	0945	30 m
SPOT	HRV (MSS)	0.50 - 0.59 0.62 - 0.66 0.77 - 0.87	26 days	1030	20 m
SPOT	HRV (PAN)	0.51 - 0.73	26 days	1030	10 m
MOS-1	MESSR	0.51 - 0.59 0.61 - 0.69 0.72 - 0.80 0.80 - 1.10	17 days	n/a	50 m
MOS-1	VTIR	0.50 - 0.70 6.00 - 7.00 10.5 - 11.5 11.5 - 12.5	17 days	n/a	900 m 2.7 km

(Source: Mather, 1992; Strain and Engle, 1994)

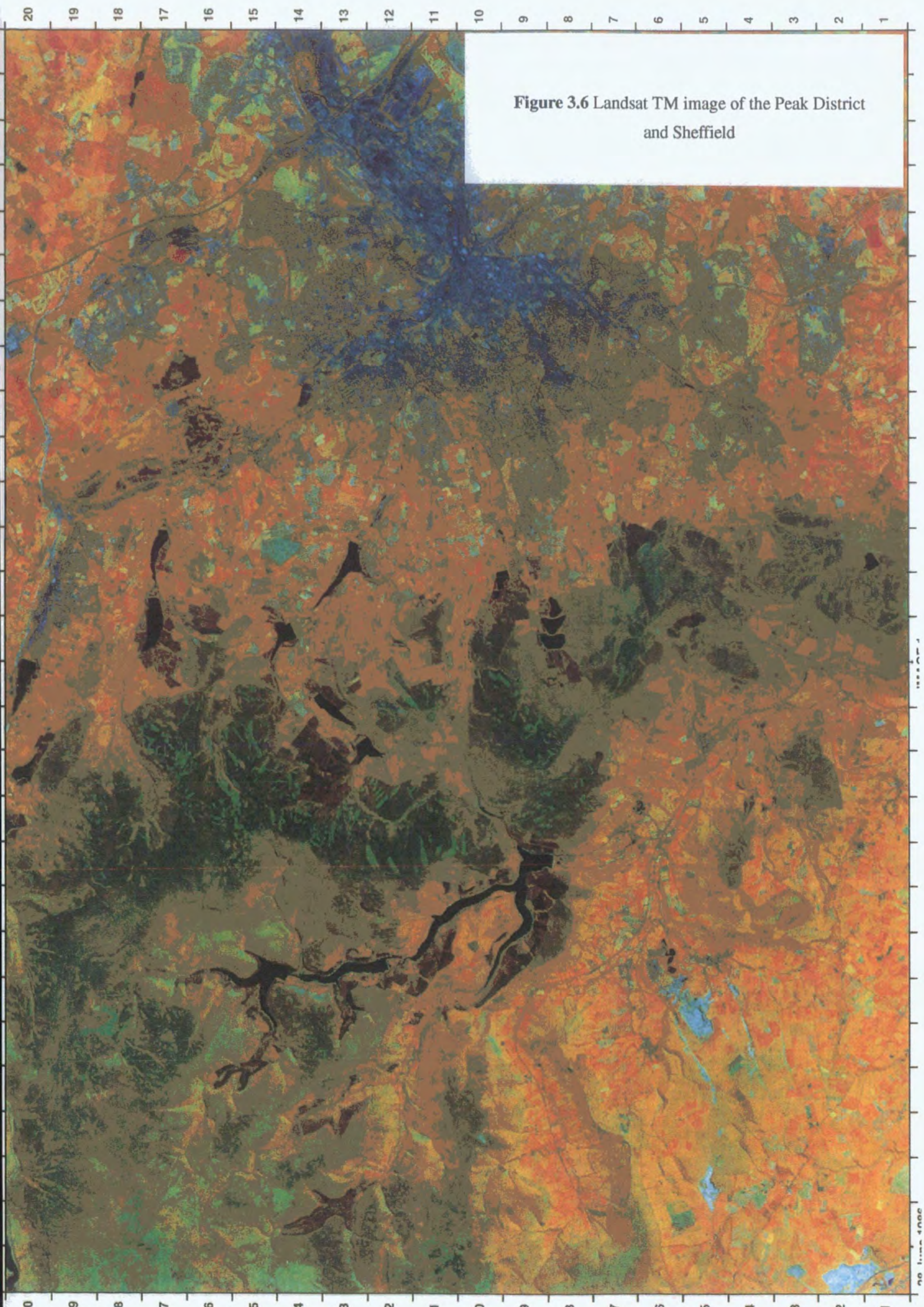


Figure 3.6 Landsat TM image of the Peak District and Sheffield

Until recently map compilation from remotely sensed images was seen more as an interesting experiment than as an accepted production technique. In April 1980 the National Remote Sensing Centre was formed to promote the use of satellite remote sensing in the United Kingdom. Within this general aim, five key objectives were identified (Fox, 1991):

1. the acquisition and supply of data to users;
2. the provision of support to United Kingdom industry and other remote sensing users;
3. the provision of research and development facilities;
4. the development of new techniques and applications for remotely sensed data;
5. the development of user awareness.

The centre has also developed as a major supplier of data to a broad spectrum of users, including academics and research establishments. CHEST (see Chapter 5) have purchased some Landsat data from the National Remote Sensing Centre Ltd., on behalf of the academic community and these data are available at a greatly reduced price. The impact of remote methods of assessing land use and land cover since the 1950s can not be stressed enough. Balchin (1984) estimates that over half the world's known land-use series have been based on remote-sensing (which includes satellite imagery and aerial photography). Over the last five years there have been a number of important number of developments in the use of satellite data for land cover mapping, in the United Kingdom.

The Institute of Terrestrial Ecology's (ITE) Survey of the Rural Environment, completed in 1985, provides summary land cover information for Great Britain. A total of 256 individual one kilometre grid squares (out of a total of more than 12,000 for the whole of the United Kingdom) were used to give an overview. The aim of this study was not to produce comprehensive land cover data for the whole of the country, but to act as a pilot survey to evaluate the use of satellite data for a land cover map of the United Kingdom. The classification that was used contains 473 classes, mainly for types of land cover and landscape features. The problems associated with this survey stem from the lack of detail (variations in land use of less than one square kilometre were not recorded), the poor ground resolution and the incomplete national coverage.

The Land Cover Map of Great Britain, the outcome of the work carried out on the pilot survey, was unveiled in Britain in July 1993. Between 1988 and 1991 cloud-free cover for all of Britain was collected by the ITE and the British Space Centre, using Landsat TM digital data for the whole of England, Scotland and Wales. The onboard sensors recorded reflected solar radiation in seven wavebands at a 30 metre resolution and the data were supplied as digital scenes, each 185 square kilometres. Cover maps from 46 scenes were joined to create a digital map, covering the whole of Britain. The survey is the most detailed computerised map of land use ever produced and constitutes a major improvement to the country's land-use data bank (ITE, 1993). Its role in the field of planning and ecology and in the assessment of environmental change is revolutionary and it is likely that the database will prove an invaluable source of information to academics, planners and businessmen alike. Currently at least 30 widely differing applications for the Land Cover Map are underway (Fuller and Groom, 1993). Figure 3.7 shows the Land Cover Map greatly reduced in scale and showing the 17-class summary data for more than 12,000 one kilometre grid squares. Dartmoor and Bodmin Moor are clearly visible as areas of moorland and a large percentage of Cornwall, Devon and Somerset are shown either as grassland or grass moorland and bracken. Derbyshire has a variety of land cover ranging from arable farmland mainly in the east of the county, but depending largely on the elevation, to grassland and grass moorland with bracken. Northamptonshire is mainly brown denoting an arable landscape.

For any one area, a summer scene (May to July) and a winter scene (October to March), gathered between 1988 and 1992, were registered geometrically to coincide with the Ordnance Survey national grid, producing 25-metre square output cells (ITE, 1993). In order to accommodate the diversity of land forms and land use, the Land Cover Map was divided into 25 classes (see Table 3.5 below). For this target list of 25 classes, typically 80 subclasses were required to encapsulate the range and variety of spectral classes to be found in each scene (Fuller and Groom, 1993). As a result of the combined summer and winter scenes and careful structuring of the large number of classes in the classification, comparisons with independent reference data indicate that the accuracy of the Land Cover Map is between 70% and 85%, but because perfect reference data do not exist, exact levels of accuracy are not known (Fuller *et al.*, 1994).

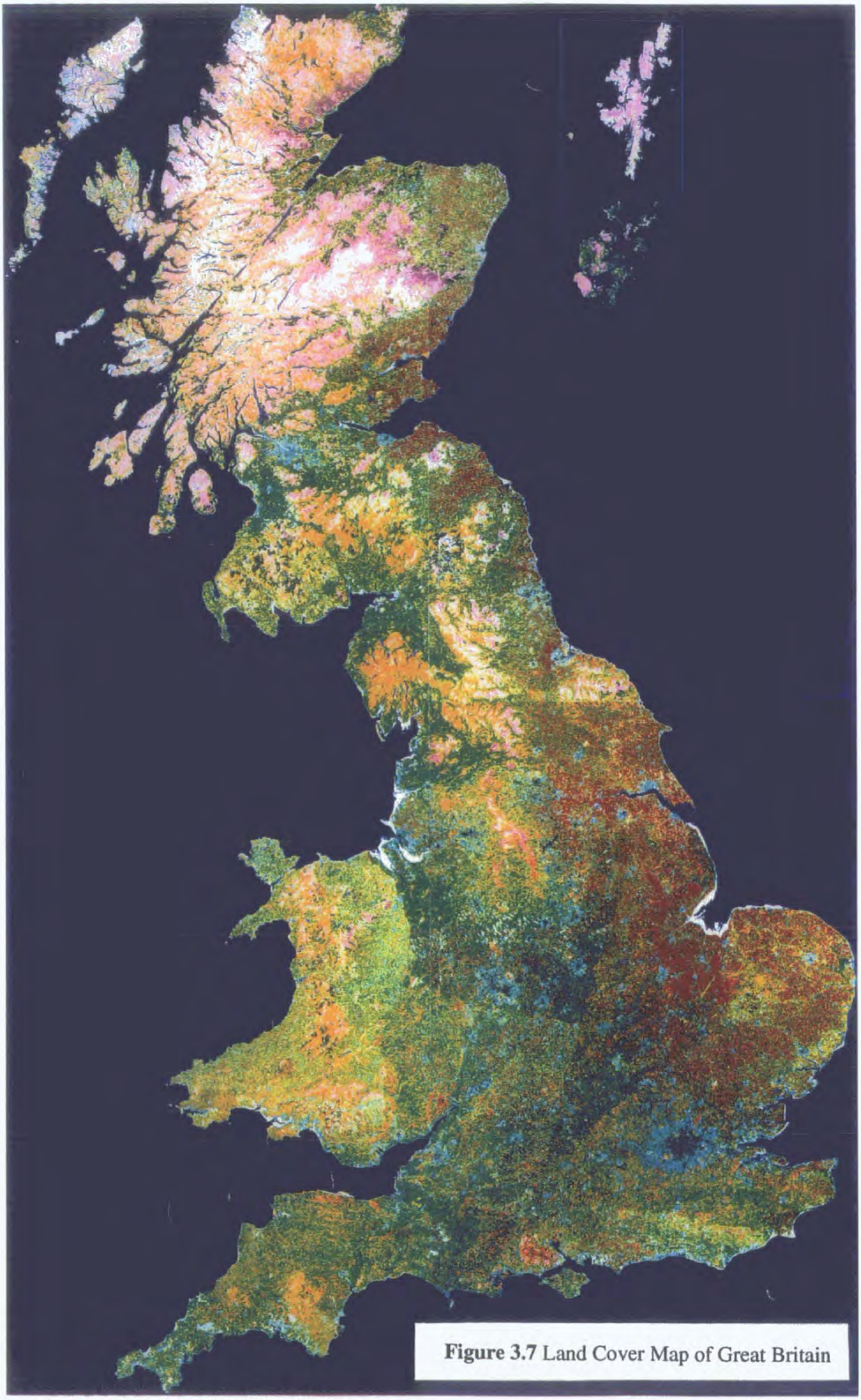


Figure 3.7 Land Cover Map of Great Britain

(Source: ITE)

Table 3.5 Categories Used in the Land Cover Map of Great Britain

Land Cover Category (17 class system)			Target Classes (25 class system)	
A[a]	1[b]	Sea / Estuary	1[c]	Sea / Estuary
B	2	Inland Water	2	Inland Water
C	3	Beach / Mudflat / Cliffs	3	Beach and Coastal Bare
D	4	Saltmarsh	4	Saltmarsh
E	5	Rough Pasture / Dune Grass / Grass Moor	5 9	Grass Heath Moorland Grass
F	6	Pasture / Meadow / Amenity Grass	6 7	Mown / Grazed Turf Meadow / Verge / Semi Natural
G	7	Marsh / Rough Grass	19 23 8	Ruderal Weed Felled Forest Rough / Marsh Grass
H	8	Grass Shrub Heath	25 10	Open Shrub Heath Open Shrub Moor
I	9	Shrub Heath	13 11	Dense Shrub Heath Dense Shrub Moor
J	10	Bracken	12	Bracken
K	11	Deciduous / Mixed Wood	14 15	Scrub / Orchard Deciduous Woodland
L	12	Coniferous / Evergreen Woodland	16	Coniferous Woodland
M	13	Bog (Herbaceous)	24 17	Lowland Bog Upland Bog
N	14	Tilled (Arable Crops)	18	Tilled Land
O	15	Suburban / Rural Development	20	Suburban / Rural Development
P	16	Urban Development	21	Continuous Urban
Q	17	Inland Bare Ground	22	Inland Bare Ground
			0	Unclassified

[a] class reference within the 17 'key' cover-type categorisation

[b] 'band' within the 17 'key' cover-type 1 x 1 km summary data

[c] label value within the 25 'target' cover-type 25 x 25 metre data

(Source: Fuller *et al.*, 1994; ITE, 1993)

Combining the Land Cover Map with other computerised databases will enable a wide variety of environmental questions to be answered. Incorporating past land-use surveys will enable changes over time to be monitored and will facilitate, for example, measurements of urban sprawl or the erosion of heathland. The data can be supplied to purchasers as hard copy, in IIS and ASCII format and in ARC/INFO export format. The paper maps will be available (they had not been produced at the time of writing in September 1993) at the scale of 1:625,000 and will probably be in two sheets covering the whole of the British Isles. The map data will be summarised as cover per one kilometre grid square for a simplified selection of 17 land cover classes (ITE, 1993). The cost of the whole map is £500,000 but areas of 100 kilometres square are available to academics for around £200, although the exact cost of the data for research institutions has yet to be finalised. The cost of the data is likely to be the limiting factor in their widespread use, although the survey was three times less expensive to produce in real terms than the First Land Utilisation Survey (see below). Probably the greatest potential of the data will be realised through its integration with other data in geographical information systems (GIS). A comparison between the First Land Utilisation Survey and the Land Cover Map of Great Britain is given below.

Table 3.6 A Comparison of the Logistics of the First Land Utilisation Survey of Britain and the Land Cover Map of Great Britain

	Land Utilisation Survey	**Land Cover Map**
Surveyors	L. D. Stamp & 20,000 coordinators & 250,000 schoolchildren	R. M. Fuller & 2 key personnel (5 person-years)
No. of Classes	9 field-surveyed land uses, divided into 22 using OS map context	25 target types
Targets	Land cover and land use	Land cover
Date of Survey	1931 - 1933 (limited additions 1935)	1988 - 1991 (2.5% imaged 1986 - 1987)
Resolution	Field-by-field (>1ha) with selective detail	Field-by-field (>1ha) with selective detail
Cost (1993 prices)	£1.5 million & 230 'free' person-years & 250,000 schoolchildren	£0.6 million (& field data, equivalent value £60,000 - £90,000)
Availability	Published maps at 1:63,360 and 1:625,000, summaries in book-form, surviving manuscript maps at 1:10,560 in the British Library of Political and Economic Science, London	Digital 25 metre data or digital / paper maps made to order; Countryside Information System on a micro-computer

(Source: Fuller *et. al.*, 1994)

A report, partly based on the ITE's Land Cover Map, was published on 17 November 1993 by the Department of the Environment (DoE) and was entitled the 'Countryside Survey of Great Britain in 1990': The report was heralded a 'world first' by John Gummer, the Environment Secretary (The Times, 18 November 1993). The so-called 'modern Domesday Book' survey covers the 12 years to 1990 and records a wide range of statistical and other information relating to environmental change. The survey was conducted for the DoE by the ITE and the Institute of Freshwater Ecology and will be repeated every ten years, the next appearing in the year 2000. Results show a 23% reduction in hedgerow length since 1984 and a 4% growth of built-up areas, although the proportion of urban and suburban land in Britain as a whole is still surprisingly low at 6.5% (Hornsby, 1993). Losses of plant habitat of 13.4% and 13.8% respectively were recorded in woodland and semi-improved grassland. To complement this survey the ITE has developed the Countryside Information System (CIS), a software package which incorporates data such as soils, flora and fauna (collected in the national Countryside

Surveys of 1978, 1984 and 1990) with administrative and National Park boundaries. The data are based on information from a stratified sample of the Land Cover Map, simplified for 17 land cover classes (Fuller and Groom 1993; Howard, 1993). It is also planned to include some socio-economic data to assess possible underlying causes for land-use change. The potential for modelling, both in the static and dynamic sense, is considerable (Bunce *et al.*, 1992). The system will be launched commercially in January 1994 and will provide unparalleled data on the natural environment of the United Kingdom for academics and planners alike.

3.3 Survey Methodologies and the Uses and Limitations and of Land-Use Data

3.3.1 Ground Survey

The different approaches to land-use mapping impose their own limitations and as yet there is no panacea for land-use surveys. The majority of current land-use monitoring projects use a combination of aerial photograph interpretation, satellite image analysis and ground survey work (in the main for checking levels of accuracy). The exact method or methods chosen for land-use surveys will depend primarily on the level of aggregation at which the information is needed, the urgency with which the results are required, the desired minimum levels of accuracy and the available resources. In practise, the last, which should be determined by the other requirements, will often be the deciding factor in the choice of method (Coppock, 1991). The result of the array of possible survey methods is a confusion of functional and formal characteristics in land-use description.

Land-use maps compiled from ground surveys have their advantages as well as their disadvantages. They are detailed, accurate and direct records but they are also cumbersome, expensive, difficult to analyse and present only a static picture. The main disadvantage of ground survey is the length of time it takes to complete. Field survey is necessarily time-consuming, as surveyors must first reach the survey area and then access all parts of it. A survey that is completed within a single season or year shows purely spatial characteristics, but if spread over a longer period of time there is an increasing risk that the spatial pattern will be affected by temporal change (Balchin, 1984). Seasonal changes in land use are often very marked; the variation in the water table is just one powerful variable. For these reasons the use of ground survey techniques are restricted today to small scale surveys and also for checking the accuracy of remotely-sensed land-use data. Existing national land-use surveys, such as the two Land Utilisation Surveys of Britain, are important sources of historical data for measuring land-use change over time. The use of sophisticated technology, which has developed since the World War II, has led to the demise of the ground-based survey and has encouraged a more 'remote' method of surveying land use and more commonly land cover (as distinct from land *utilisation*) of the United Kingdom. Land cover does not aim to assess the use of land in terms of agricultural productivity, specific crop patterns or recreational activities for example, but aims solely to represent the type of surface (*e.g.* water, heathland, urban) that covers the surface of the Earth. Land Capability is a different type of classification altogether and this is covered in section 3.4.

3.3.2 Remote Sensing

Aerial photography as a method of surveying land use is limited by the number of cloud free days, the availability of a plane and photographic equipment. Air photographs, which are available from a variety of sources, can be purchased for the year, season and area that is required as well as in colour or black and white. In general, they are an accurate representation of reality, although cloud cover can obscure some detail (see below). The cost of purchasing photographs is high. Ordnance Survey quote a price of approximately £25 per photograph and a large number of photographs are required to cover the five counties. Ordnance Survey aerial photographs vary in scale from 1:5,000 to 1:32,000. In addition, there is a necessarily time consuming processing stage in order to convert the data from the photographic print into a computer readable format for the entire five counties. For reasons of cost and lack of adequate experience, resources and time, aerial photographs were not considered to be a viable source of land-use information for this project.

New methods of gathering land-use information were introduced with the launching of the first Landsat satellite in 1972. The principal appeal of satellite data acquisition, compared with aerial methods, is the ability of satellite images to cover a large area on the ground with a single scene, thus reducing unit costs for the acquisition of source material. The synoptic view provided by a satellite image aids interpretation by providing a regional context and data can be acquired for all parts of the earth, however inaccessible. Furthermore, data collection is continuous and up-to-date. However, satellite imagery in practice, will obtain information on land cover rather than land use, as only ground surveys may accurately identify land use.

There are a number of disadvantages to using satellite imagery. The most obvious being the technological requirements and training that are required for data conversion and analysis. This, together with the current high cost of the satellite scenes, can limit the number of potential users of the data. At present there is also a lack of continuous stereoscopic cover and the data provided are generally of a relatively low resolution (on average 20 to 40 metres square), although some of the recently released Russian satellite data (little is known about the extent of the coverage) have a resolution of five metres square. Significant advances have recently been made in the resolution of satellite images, especially since the launch of the French SPOT satellite, but SPOT data have not yet been used for any published study of land use. SPOT 3 (launched in autumn 1993) was formally declared operational on 29 November 1993, and has a higher responsivity and wider spectral response of band B2 than SPOT 2. These refinements enable different types of vegetation to be discriminated more accurately. The images are also free from scale distortion and anisomophosis (AUSLIG, 1994).

The most important problem, that has yet to be solved, is one of accuracy. A much greater proportion of information on satellite images is filtered out and generalised compared with ground surveys and aerial photographs. Estimates of the accuracy of satellite images vary from 70% to 90% depending on the satellite, the type of sensor and the resolution, as well as weather conditions and the time of year that the image was captured. The amount of cloud cover can seriously affect the quality of a scene, in a similar way to aerial photography, although there are satellites (*e.g.* ERS - only the visible wavebands are affected by cloud cover) which are able to record land cover regardless of cloud cover. When purchasing satellite images the availability of cloud-free data is not necessarily guaranteed. The launch of several scheduled new radar and mapping sensors before the year 2000 should negate many of the current technical problems with satellite data.

At around £3,500 per scene (August, 1993) for Landsat TM digital data, the high cost of satellite data has acted as a barrier to their widespread use within the academic and research community. To overcome this hurdle the National Remote Sensing Centre (NRSC) and the CHEST recently arranged for NRSC to supply CHEST with up-to-date cloud free Landsat Thematic Mapper imagery available to those institutions subscribing to CHEST for a minimal charge (see above and Chapter 5). Deals such as this will encourage more widespread use of remotely sensed data within research strategies. The increased availability and decreased costs of satellite data mean that they will be a key resource in the future.

The ITE Land Cover survey is a very useful source of data - already stored in digital format (the data are available in ARC/INFO export format although it will not be made available until spring 1994) and up-to-date. However, the first comprehensive data from this survey were not available until August 1993 at which stage the project had progressed too far to incorporate it and prior to the publication of the final map the data were also limited in extent to an area 100 kilometres square (chosen by the user). The cost of the data also limited its use. The Land-Cover Map will be repeated periodically (every 10 years) and will provide an unprecedented source of up-to-date land-use information for the United Kingdom.

3.3.3 Rural Bias

The collection of land-use data has traditionally been concerned with rural areas. The rural bias in geography has historically been so pronounced that the actual term 'land use' is frequently misinterpreted as referring primarily or wholly to *agricultural* land use, with the urban sector almost completely excluded. Urban statistics were virtually non-existent in any comprehensive form until analyses were made of the development plans compiled under the 1947 Town and Country Planning Act (Best, 1968; Rhind and Hudson, 1980). Coppock's judgement that 'the collection of adequate data on urban land use and land-use change is always likely to present difficulties' (Coppock, 1978) to some extent remains true, although recent developments in remote sensing constitute a major step forward in the assessment of urban encroachment and the more detailed images enable different land-use types within a city to be classified - for example, Hyde Park stands out clearly in a Landsat image of London. The different reflectivities of surfaces in a city produce different colours on the satellite image, which can be interpreted as different forms of land use under the 'urban' umbrella. However, the different uses for multi-storey buildings can not be identified from satellite images and the one-dimensionality of the images is a major constraint to using satellite data for urban land-use surveys. Ordnance Survey are in the process of digitising their existing maps, from which some land use can be inferred. Priority is being given to urban areas (starting with London), coverage of which is expected to be complete by September 1995. Rural areas should be completed by December 1995 (Coppock, 1991; OS correspondence 1994).

3.3.4 Spatial and Temporal Compatibility

The problems of comparability and compatibility of land-use data, not least over conventions for spatial referencing, make spatial and temporal comparisons problematic. To this end, Kivell (1991) has stated that 'research workers using existing land-use data encounter the problem that the nature of the data is heavily influenced by the purpose for which it was originally collected and this inevitably limits its more general use'. There is a wide variety of uses and users for land-use data ranging from local planning, monitoring and environmental impact assessments to academic research and modelling. The data for these can come from the same data bank or, more commonly, from disparate and localised databases collected for specific purposes. It is difficult to know what weight to attach to the potential value of land-use data in the future, when unforeseen needs may arise. There are numerous instances in the literature when use has been made of land-use data originally collected for other purposes. For example, Coleman (1985) lists 12 national agencies, not all of which are planning agencies, that have made use of data from the Second Land Utilisation Survey. Coppock (1991) believes this to be true of virtually all national surveys and he cites it as one very good reason for the cheap and unrestricted use of data, to encourage maximum use and benefit.

No data are 100% accurate and where aggregation and simplification has occurred prior to the release of the data, information on the methods of data collection, aggregation and manipulation should be made available to the user as metadata (data about data). This allows the user of the land-use data to make critical assessments of the results of further analyses in the context of a research methodology. Research into the effects of aggregation on land-use surveys is ongoing; Bird (1991) for example, outlines a current area of research in the design of optimum classification schemes for land-use monitoring. The most popular type of areal unit for the collection of geographical data is the artificial unit, for example the one kilometre grid square, constant in size and shape and employed for ease of survey. Its validity has been argued by Evans (1979) who suggests that such regular spatial units are more than just a convenience of survey - 'they are to geography what collecting information by years (rather than by monarch's reigns) are to history'. They provide a means of sampling an area on an exhaustive and consistent basis. For example, a summary Land Cover map has been produced which summarises land-cover using 17 primary land-cover categories for each one kilometre grid square of Great Britain. The provision of 1971 Census data by one kilometre grid squares facilitated and simplified certain land-use and population matching operations (Rhind and Hudson, 1980) which have since been very difficult to carry out, owing to the lack of 1981 and 1991 population data available for one kilometre grid squares. 1981 Census data for one kilometre grid squares were only made available for areas where it was requested and paid for by customers (see Chapter 4.3).

There are disadvantages however, to using artificial units. The principal one being the generalisation of natural spatial variation by the imposition of such a rigid sampling structure. However useful or vital imposed boundaries may be to data collection organisations, the chosen spatial units, be they man-made or based on natural areas, impose boundaries which may give a false image of reality by artificially dividing two or more 'natural' zones and by smoothing internal variation. There is no set of 'natural' areal units however, which will meet all purposes and that are easy to survey (Rhind and Hudson, 1980). This situation is slowly changing as GIS become more widespread, allowing a much greater availability of data and encouraging geographers to become more aware of the effects of data aggregation on the results of spatial comparisons. There are currently several Geography Departments in the United Kingdom engaged in research on the effects of spatial aggregation of data - for example, Leeds University (Openshaw) and Southampton University (Wrigley). The problems associated with the comparisons between aggregated spatial data sets are discussed in more detail in Chapter 4.2.

The fundamental problem regarding land-use data (and other environmental surveys) in the United Kingdom at the present time, is that the surveys are not entrusted to a unified organisation but are split among a multiplicity of authorities. The way forward is to create a situation in which data, for all aspects of the natural environment, are pooled in a national data bank and available for research purposes. The ideal land-use database would describe 'atomic' areas which were homogenous in form and function. It would be classified to permit re-aggregation at whatever scale and grouping the user required, as well as allowing comparison with previous (and possible future) surveys of the same area or subject matter. In order to satisfy all of these requirements, the database would have to be very detailed and would have to meet the most stringent of accuracy specifications (Rhind and Hudson, 1980). Ideally, the land-use data would be collected by a single organisation which would be independent of bias towards any specific use of the data and would therefore optimise the data collection to suit not only themselves, but the user community at large.

3.3.5 Concluding remarks

Technology is advancing all the time at an ever-increasing rate so that some of the information in this section is already out-of-date. Several new satellites are launched every year and it is often the case that computer hardware and software have been surpassed by the time they are on general release to the public. Whilst many of the principles will remain the same, the methods of land-use data capture and the quality of the output will be much improved by the end of the decade. Although land-use surveys of the future will be based around remotely-sensed methods of data capture, ground-based methods will continue to be important for studies of urban areas and for checking the accuracy of satellite interpretation.

3.4 An Outline of the MAFF Land Capability Classification

The choice of data for this project was limited by several factors, namely the cost of the data, the availability, time constraints and data complexity. The data chosen as a source of land-use information were the Agricultural Land Classification maps published by MAFF (published between 1977 and 1983). The MAFF Land Capability series was by far the cheapest and most accessible land-use data available, although obvious sacrifices had to be made in terms of detail, complexity and age of the data. The maps were surveyed between 1977 and 1983 but since then there have been only minor changes to the proportions of land in each of the five main land-use categories (Figure 3.3, section 3.2). The over-riding advantage of paper maps is that they are inexpensive data sources and are readily available. It is inevitable that in the near future, data already stored on disc will become the main source of available information as the technology becomes cheaper and more widespread.

The Land Capability maps were purchased directly from MAFF at a minimal cost as flat paper sheets with England and Wales covered in four sheets. One main advantage of the data is that they have not been aggregated for any systematic areal unit. The observed 'natural' land capability areas transcend political and administrative boundaries and retain some degree of internal homogeneity. For the purposes of ease of data collection and display however, land capability has been divided into seven grades. The 'natural' zones produce a complex map of capability grades which had to be converted to computer-readable format before integration in ARC/INFO, as the information is not available in digital format. The maps are at the scale of 1:250,000 so that the level of complexity was not too great so as to render the digitising task impossible. The data held on the paper maps were captured manually using a digitising table and puck. Chapter 9 describes the integration of the Land Capability data within ARC/INFO in more detail.

Land capability classifications were originally developed in the United States Department of Agriculture in the 1930s and provide a framework for classifying land according to the extent to which its physical or chemical characteristics impose long-term limitations on agricultural use. The limitations can work in up to four principal ways: they may affect the crops that can be grown, the level of yield, the consistency of the yield and the cost of harvesting it (MAFF, 1988). The evaluation of land capability however, refers to broad agricultural systems and not specific crops or practices. The categories in the MAFF Land Capability maps are defined according to natural areal units, as opposed to imposed or rigid boundaries, the rich variation in the land has been arbitrarily divided into several grades for ease of classification and display, reducing the internal homogeneity of the land parcels. Such a categorisation detracts from the real world but is a necessary simplification of what is essentially an infinitely complex situation. The principal physical limiting factors are listed below and form the basis of the classification of all land in England and Wales into one of five agricultural grades, plus two non-agricultural categories. Scotland has its own system of classification. Land capability is related to land use, in that it determines the agricultural practises that can be supported by the land. For example, in areas of Grade 5 land (the poorest land quality) the land use is likely to be permanent pasture or rough grazing and in areas where the land capability is Grade 1 (the land most suited for agriculture) the land use is likely to be arable. The main limiting factors used in the MAFF classification scheme for England and Wales are as follows (after Davidson, 1992; MAFF, 1988):

1. Climate
    - average annual rainfall
    - accumulated temperature
    - local climatic factors
2. Site
    - gradient
    - micro relief
    - flood risk
3. Soil
    - texture and structure
    - depth
    - stoniness
    - chemical status
4. Interactions (between the above limitations)
    - wetness
    - draughtiness
    - soil erosion

To ensure a consistent approach to the classification a number of assumptions were made by MAFF prior to the survey. These are as follows (after MAFF, 1988):

1. Land is graded according to the degree to which physical or chemical properties impose long-term limitations on agricultural use. It is assessed on its capability at a good but not outstanding standard of management.

2. Where limitations can be reduced or removed by normal management operations or improvement, for example by the installation of an appropriate under drainage system, the land is graded according to the remaining limitations. Where an adequate supply of irrigation water is available this may be taken into account when grading the land. Chemical problems which can not be rectified, such as high levels of toxic elements or extreme subsoil acidity, are also taken into account.

3. Where long-term limitations outside the control of the farmer will be reduced or removed in the near future, the land is classified as if the improvements had already been carried out. Where no such scheme is proposed, or there is uncertainty about implementation, the limitations will be taken into account.

4. The grading does not necessarily reflect the current economic value of land, land use, range of crops, suitability for specific crops or level of yield. The grade cut-offs are not specified on the basis of crop-yields as these can be misleading, although in some cases crop growth may give an indication of the relative severity of a limitation.

5. The size, structure and location of farms, the standard of fixed equipment and the accessibility of land do not affect grading, although they may influence land-use decisions.

Special or local circumstances were therefore taken into account in the classification process. One problem with the subjective definition of grades according to possible future improvements lies in the fact that farmers and landowners do not always make optimum use of the land; it may not fit in to local farming practices or be economical to irrigate or drain a parcel of land. In such circumstances it would be useful to make two assessments; 'with' and 'without' irrigation, drainage and the like (McRae and Burnham, 1981).

A description of the land capability grades classified according to these assumptions is as follows:

Table 3.7 A Description of the Grades in the Agricultural Land Classification of England and Wales

Grade 1	Excellent quality agricultural land. Deep, well drained, easily cultivated soils on gentle slopes, with no major climatic or physical limitations. A wide range of crops can be grown and yields are high.
Grade 2	Very good quality agricultural land, with some minor limitations which may lead to lower yields or less flexible cropping.
Grade 3	Good to moderate quality agricultural land. There may be soil defects, altitude, slope or a limiting rainfall regime. Choice of crops may be restricted, as well as timing of cultivation or level of yield. (Later divided into subgrade 3a good quality land and subgrade 3b moderate quality land).
Grade 4	Poor quality agricultural land. Land with severe limitations due to adverse soil, relief and climate or a combination of these. A high proportion is under grass.
Grade 5	Very poor quality agricultural land. Land with severe limitations, including very steep slopes, excessive rainfall or exposure, shallow soil or extreme stoniness. This type of land is usually restricted to grass or rough grazing.
Urban	Built-up areas used for housing, industry, commerce, education, transport and the like. This category also includes all types of derelict land, including mineral workings, which are unlikely to return to agricultural use.
Other	Land that is neither in agricultural use nor urban.

(Source: Morgan 1974; Mather 1986; MAFF 1988)

Land may 'be assigned to a Grade for widely differing reasons. It may be placed in Grade 4 because it is a poorly drained clay in an area of high rainfall, or in the alternative, because it is an excessively drained sandy soil in an area of low rainfall' (MAFF, 1968). MAFF proposed a nominal productivity index of 20 for Grade 1 land , 18 for Grade 2 land, 10 for Grade 3, 3 for Grade 4 and 1 for Grade 5 (Jefferson, 1976). The classification scheme used for the capability maps produced a division of land into seven categories, which are divided as follows:- Grades 1 to 5 are agricultural classifications and are outlined above. Urban land is given its own category, as is land that is predominantly neither agricultural nor urban, for example golf-courses, airports, parkland and sports fields (Mather, 1986; MAFF, 1988). The latter was not included in the simplified classification used in this project, due to the very small areal extent of the category in the five counties that were studied and the fact that much of the land in this category behaves, in terms of the soil mechanics and characteristics, in much the same way as the surrounding agricultural land. Grades 1 and 2 were also amalgamated due to the close proximity of the two in terms of their spatial position and physical characteristics and the fact that there was very little Grade 1 land in the five counties. Grade 1 land has a strict classification regime and is confined to only 3% of the land area of the United Kingdom as a whole.

The most productive and flexible land falls into Grades 1, 2 and 3 and comprise almost half of the total agricultural land in England and Wales. A large area of the land in the counties under investigation was classified as Grade 3. This reflects the national trend, as approximately 50% of England and Wales is classified as Grade 3 land (Mather, 1986). The grade contains considerable internal variation as would be expected from a middle category in a relatively simple classification. MAFF have since revised the guidelines for the classification of land in England and have introduced a two-fold category to Grade 3. Two subgrades are now recognised Subgrade 3a and Subgrade 3b. The new division distinguishes between the wide ranging physical characteristics and spatial dominance of Grade 3 that was evident in the maps used for this project. These revisions were first listed in the MAFF Technical Report number 11/1 (1976), which also includes proposals for the development of an economic classification system linked to the physical one. It also identifies some significant problems in the collection of objective, accurate and up to date economic data which have prevented the subsequent incorporation of the economic classification. A complete revised set of guidelines and criteria for grading the quality of agricultural land were published by MAFF in 1988. These guidelines concentrate on the criteria used to assess climatic limitations and the limitations of a climate-soil interaction (MAFF, 1988). The revised methods were developed for MAFF by the Agricultural Development and Advisory Service (ADAS) in collaboration with the Soil Survey and Land Research Centre (SSLRC).

The advantages and disadvantages of the land capability classification compared to parametric systems are summarised in the list below (after McRae and Burnham 1981, pp. 83 - 86):

- The division of the data into a relatively small number of ranked categories is easily understood;
- It is versatile and can be modified to suit local conditions without affecting the basic structure of the system;
- It is a general purpose classification which shows a clear distinction between land capable of growing crops and that which is not and includes categories for non-agricultural forms of rural land use;
- The results can be clearly and simply displayed on maps, although it should be recognised that boundaries between classes are often transitional and lack the sharp dividing line depicted on the map;
- It is very widely used (land capability classifications are the most widely used categorical systems for evaluating agricultural land);

- It gives reasonable and acceptable results which usually match local opinion and it avoids the absurdly wrong evaluations which can be produced by following a set formula as in parametric systems.

The main disadvantages of the classification are as follows (after McRae and Burnham, 1981):

- The classification is subjective;
- The division into six categories is too coarse;
- Class five is anomalous and often difficult to apply.

In addition, the MAFF Land Capability Classification is static and does not allow for possible change or development over time; however this is a disadvantage of most land-use classifications. The classification does not give specific data about the farming practices or the type of crop that is being grown, neither does it specify why pockets of land are classified as they are. The most recent maps still fail to take into account the non-physical factors such as distance to market, types of road, farm structure and managerial efficiency (Rhind and Hudson, 1980).

The fact that the classification is qualitative is listed as an advantage but in this study the lack of quantitative data could be regarded as a negative aspect, especially when statistical analyses are being used in a GIS to interpret the results of correlations. The classification of the land was carried out at the end of the 1970s and early 1980s and therefore the main disadvantage of the MAFF Land Capability maps is that they are out of date. However, as discussed above, land capability data are less likely to be affected by change over time than land-use and land cover surveys, due to the fact that the classification takes in to account all probable future improvements to the land.

3.5 Related Environmental Research in the United Kingdom

With the development of increased computing power and the sophistication of mapping procedures, there has been a range of attempts to establish comprehensive land-use databases. The first of these was the Rural Land-Use Information System (RULIS) project based in Scotland in the 1970s. This study focused on the acquisition of comprehensive sets of information and was successful in showing that computing systems could enable such data to be stored and manipulated for specific requirements, at the regional level (Bunce *et al.*, 1992). The Scottish Office jointly sponsored the compilation of a land cover map of Scotland with the Nature Conservancy Council for Scotland and the Countryside Commission for Scotland. The project resulted in a digital database, based on interpretation of air photographs and was completed in 1992. More recently, the Northern Ireland Countryside Survey (funded by the DoE (N.I.)) completed in 1993, provides estimates for the main land cover types, wildlife habitats and field boundaries.

Thus, several research organisations in the United Kingdom have a current interest in land use, but three organisations have a particular interest in land-use data for integration in GIS:

- The Soil Survey and Land Research Centre based at Silsoe, has developed a computer database known as LandIS - the Land Information System. LandIS is one of the largest computerised land information systems of its kind in Europe (O'Carroll *et al.*, 1994). The database holds soil data, climatic data and altitude data for England and Wales. The Macaulay Land-Use Research Institute holds equivalent data for Scotland.
- The Macaulay Land-Use Research Institute, formed in 1987 and based in Scotland, holds data in several forms. One example is the National Peatland Database which provides data for peatland areas and these and other data are being integrated in the Macaulay Land-Use Information System, a GIS aimed at aiding the analysis of land use, land cover and habitats in order to provide informed input to conservation and planning issues. The Institute has produced a complete set of soil maps for Scotland at various scales, as well as Land Capability maps for agriculture and forestry. In addition, the Land Cover of Scotland (1988 Data Set) is the first detailed inventory of the existing land cover of Scotland. It is a digital data set at the scale of 1:25,000 and is referenced to the National Grid.
- The Environmental Information Centre at ITE in Cambridgeshire (discussed in more detail above), is concerned with, amongst other things, updating many different land-use data sets, for example the countrywide land-use map derived from satellite imagery and the Countryside Survey.

As well as these research centres, work is being carried out at Newcastle University on the NELUP project, a land-use programme run by the Centre for Land-Use and Water Resources Research. In Wales, WALTER or the Wales Terrestrial Database Project is jointly funded by all Welsh bodies with an interest in land-use studies. However, there is little co-ordination between these establishments and an essential requirement for the near future is a countrywide land-use database encompassing all of the above interests.

Further information on the method of land capability data capture and integration into ARC/INFO is given in Chapter 9 and a discussion on data error measurement and estimation can be found in Chapter 10.

4. Census Data

4.1 Aims

Census data from the 1991 Census of Population will be used in this study to assess the relationship between the social structure of the five counties and levels of radon gas. The chapter aims to assess the availability of population census data, concentrating on social class data at ward level, for the United Kingdom. Explanations for the choice of the data sets will be outlined and the main alternatives will be discussed. Research carried out by Wolff (1991) using 1981 Census data suggests that, at the county level, higher radon levels are significantly correlated with unemployment, car ownership, occupational class and population density. Lower domestic radon levels on the whole seem to reflect greater socio-economic deprivation. This chapter is concerned with assessing the importance of social class as a deterministic factor in the level of indoor radon gas, in the light of this and other research (see Chapter 12).

4.2 Introduction

The term 'census' is defined in The Dictionary of Human Geography (Johnston *et al.*, 1994) as 'the total process of collecting, compiling and publishing demographic, economic and social data pertaining to all persons in a defined territory at a specified time'. Despite a major initiative on the part of the European Union (EU) to harmonise census taking in member countries, conventional censuses are currently taken in only nine EU countries, and important differences still exist regarding the content of the censuses (Lievesley and Masser, 1994). There has been a Census of Population in the United Kingdom every ten years since 1801 (with the exception of 1941) and each one requires a separate order of parliament. The census is the largest social survey carried out in the British Isles and comes under the jurisdiction of the Office of Population, Censuses and Surveys (OPCS). By survey standards it is very cheap - the 1981 Census cost a little over £1 per head - but it is the most comprehensive and therefore, in total costs, the most expensive survey available to social scientists (Marsh *et al.*, 1988). People are enumerated by the census in their home (or someone else's home) on a specified day at a fixed time. The most recent census took place on 21 April 1991, when 54.9 million individuals (of all ages) in twenty-two million households filled in their census forms (see Table 4.2). The census form asked for information about each member of the household and anyone else present. It also sought information concerning the household as a unit and every census in Britain has essentially been a census of both population and housing (Rhind, 1983). Preparation for the 1991 Census began as soon as the 1981 Census was completed and constitutes a major investment of £135 million (for Great Britain, not including Northern Ireland) over ten years by Central Government.

Geography provides the base for carrying out a census. Data relating to dynamic human populations are different in their geographic properties to those relating to the physical world, as the location of any individual is always referenced via some other spatial unit. Martin (1991) stresses the need for care to ensure that the data collection and aggregation models that are used in social surveys are an accurate representation of reality. In England and Wales the geographic base (the smallest unit for which data are made available), for both the 1981 and 1991 Censuses, is the Enumeration District (ED). EDs may vary in size according to the population of the area being enumerated, but they are wholly contained within administrative area boundaries such as civil parishes, wards, local government districts and the like (Clark, 1992). They present the most 'accurate' picture (in terms of internal homogeneity) in as much as they relate to a small population in a relatively confined area, although there is a degree of spatial heterogeneity and also noticeable scale differences nation-wide (urban EDs tend to be much smaller in area than rural EDs). Wards are collections of enumeration districts and represent on average between 4,000 and 6,000 people. The main areal units represented in the 1991 Census and the number of each unit in England and Wales are listed in Table 4.1 below.

Table 4.1 The Hierarchy of 1991 Census Output Areas

Area Type	Number of Areas	Typical Population
England and Wales	1	49,890,000
County	54	923,889
District	402	124,104
Ward	9,135	5,461
Enumeration District	109,670	442

(Source: Martin, 1993)

The data available for EDs are grouped under the heading of Small Area Statistics (SAS) and are available in tables either at 100% enumeration or as a 10% sample where confidentiality or processing time creates restraints (see 4.4 below). The following is a list of the questions in the 1991 Census from which the two sets of variables (100% and 10%) are derived (Martin, 1993):

100% questions: sex and date of birth; marital status; usual address; term-time address of students; usual address one year ago; country of birth; ethnic group; type of accommodation; number of rooms; tenure of household; household amenities; availability of cars and vans *etc.*

10% questions: relationship within household; hours worked; occupation; name and business of employer; place of work; journey to work; higher qualifications.

The SAS were first produced for the 1971 Census when there were 1,571 tables. This number increased to 4,500 in the 1981 Census and further increased in the 1991 Census so that there are now over 9,000 counts in 86 tables (OPCS, 1991). The SAS have certain limits imposed on their contents and are abbreviated to around half the number of counts in the local base statistics (LBS) tables. The information contained in the individual census questionnaires remains confidential for 100 years and all published census data are therefore made available for different levels of aggregation. As a result of the strict confidentiality laws data are not released for areas containing less than 16 households or 50 persons and the LBS (providing more statistical detail, but only available down to ward level) have thresholds of 320 households and 1,000 persons. The downside of an increase in the number of tables and counts newly available in the 1991 Census, is that the minimum thresholds which apply to the release of census statistics for small areas, have been doubled from the 1981 thresholds of eight households and 25 people.

All OPCS count data are on ratio measurement scales *i.e.* they have a true zero base and strict multiplicability exists throughout the range of numbers so that, for example, 200 people are twice as many as 100 (Rhind, 1983). In order to preserve confidentiality in small areas (EDs in England and Wales and Output Areas in Scotland) where the detection of individual level data may be possible, the data can be modified or suppressed by OPCS prior to dissemination. In 1981 and 1991 modification of the SAS involved the addition of a quasi-random number in the range -1 to +1 to individual cells, and 'Totals' in the tables are the resulting sum of these modified cells (OPCS, 1977; Rhind, 1983; Martin, 1993). The adjustment is made to all variables for which 100% data are published. The processes of modification and suppression mean that the SAS are never wholly accurate or complete for every areal unit. Adjustments are likely to be more marked for small rather than large areas (such as counties), in areas of low population and in census variables with a low mean value. A full analysis of the implications of the data modification procedures and some approximate confidence limits are given in OPCS (1992) and elsewhere, *e.g.* Marsh, 1993; Martin 1993.

4.2.1 The 1991 Census of Population

The 1991 Census was the nineteenth in the history of the United Kingdom and took place on 21 April 1991. A huge investment of resources is necessary in order to carry out a decadal census, and this is well illustrated by the costs of the 1981 Census which was estimated to cost £45 million and involved the employment of more than 129,000 people. The census count of the population in England and Wales on census night in 1991 (excluding visitors but including residents who were recorded as absent on census night) was 54,888,844 million, before accounting for under-enumeration (estimated to be 2.2% of the population - see section 10.3 for a breakdown of the under-enumeration). To alleviate the discrepancies in calculations using the 1991 data, a set of mid-1991 population estimates for wards in England and Wales were published by the Registrar General, which take into account the under-enumeration. The provisional estimates for Great Britain were published in October 1992 and are shown in Table

4.2 below, together with a comparison of the various population bases for England and Wales in 1981 and 1991.

Table 4.2 The Population Bases for Great Britain in the 1981 and 1991 Censuses ('000)

	1981	**1991**
Residents counted in Census		
Resident population 1981 base	53,557	54,889
Coverage adjustments		
Net under-enumeration accounted for by CVS	215	299
Under-enumeration not accounted for by CVS	26	754
Other adjustments		
Definition of residents - *e.g.* students	8	69
Timing changes - census day to 30 June	9	44
Other changes	-5	0
Registrars General's mid-year estimates of residents (provisional)	**54,814**	**56,055**

CVS = Census Validation Survey (Source: Marsh, 1993)

A thorough review of the 1981 Census was carried out by OPCS and new initiatives were developed for the 1991 Census. These were tested in two major field tests in 1987 and 1989. The first test incorporated changes designed to improve or maintain the coverage standard set in 1981, to test problem areas such as the inner cities and to introduce a management information system for the control of fieldwork. Three new questions on individuals were added to the 1991 Census: ethnic group, limiting long-term illness and term-time address of students, in addition a question on weekly hours worked was reintroduced (Dale, 1993). The second test in 1989 was essentially a dress rehearsal for the full census and covered around 90,000 households in three areas of England and three areas of Scotland. Clark (1992) outlines a number of key objectives that were identified for the 1991 Census. These were as follows:

- to carry out the operation on schedule (the date of the census is fixed by law)
- to maintain a high level of coverage
- to develop procedures to cope with the expected problems in inner city areas
- to introduce a monitoring system to improve control and
- to introduce new checks on coverage, classification of household space and quality of data.

The results from the census of population are available from OPCS in a variety of formats, but are presented chiefly as tables of the number of people or households in specific categories or defined areas. The data are released by OPCS in the form of hard copy (printed) County

Monitors (the first results to be published at the 100% level) and County Reports (SAS at both 100% and 10% levels). In total the SAS comprise over 9,000 counts (or cells). The information is available for 15 different areal units - including counties (or Scottish regions), parliamentary constituencies, urban areas, parishes, electoral wards and enumeration districts. A summary of the geographical basis of the census statistics is given in Table 4.3 below. The 1971 Census is the only census for which data are available for grid squares for the whole country (see below).

Primary and secondary output from the census is available from a wide variety of sources and is outlined more fully elsewhere; for example, see Hakim (1982) and Martin (1993). Black and white maps of the census boundaries are available from OPCS or the ESRC Data Archive (at the University of Essex) in paper format at 1:50,000, 1:10,000 and 1:2,500 depending on the size of the census area required. The boundaries are drawn over Ordnance Survey maps and show the relationship between census boundaries and the local patterns of roads and settlement. Each map shows all the wards within a single local government district or London borough, but does not show any ward (or parish) boundaries outside the district (OPCS, 1993a). The cost of a map for a district is £2.95 (summer 1993) and more than one map may be required for some districts (OPCS, 1993a). Census boundary maps are also available on 35 millimetre microfilm at the above scales and in GIMMS digital format on-line from the CDU. The Ordnance Survey and various private companies also have digital ED and ward maps, available at a cost. A summary of the geographical basis of 1991 Census statistics available in machine readable form is given in Table 4.4 below. The machine-readable data are available to academics using SASPAC91 software, from the CDU at Manchester University.

The local geographic base of the census changes from one census year to another, to meet the operational needs of the census, to take account of changes in the distribution of populations and households and as a result of local government reorganisation. Many ward boundaries were re-drawn between the 1981 and 1991 Censuses as part of the local government reorganisations and it is therefore problematic to produce comparable figures for English and Welsh EDs, wards or parishes which have been subject to boundary changes. Data for county level and above are unaffected by boundary realignments since there have been no county boundary changes since the 1971 Census. Where changes in ED and ward boundaries have occurred, data for one census can not necessarily be accurately mapped using boundary data from another. In order to reduce the effects of boundary changes between the 1971 and 1981 Censuses, OPCS introduced the concept of 'census tracts', described more fully in *Census 1981: User Guide 79* (OPCS, 1982). A census tract is essentially an ED or group of EDs whose boundaries have remained the same over time. In rural districts, civil parishes provide a suitable areal base for comparisons as very few parishes were subject to change between 1971 and 1981 and similarly between 1981 and 1991. The data for these comparable areas are restricted to the 100% variables (which do not include any data for social class), rendering the

Table 4.3 The Geographical Basis of Census Statistics in England

Area Level	Published Reports	Abstracts
Enumeration District	-	SAS
Ward	County Leaflets County Monitor	SAS Local workplace and migration statistics
Civil Parish	County Leaflets	SAS
New Town	Preliminary reports New Town volume	SAS (user's aggregation)
Towns, conurbations, urban and rural areas	Preliminary Report SAS-based summary report	SAS (user's aggregation)
Local Government Districts	Preliminary Report County Monitors	SAS Extensions of published tables
Parliamentary Constituency	Special 'Monitors'	SAS
Health District	-	SAS Extensions of published tables
County	Preliminary reports County Monitors SAS-based summaries Certain topic volumes	Extensions of published tables
Standard Statistical Regions	Preliminary Report County Monitor (Regional) County Report (Regional) Certain topic volumes	Extensions of published tables
Nation	Preliminary reports County Monitor (National) County Report (National) Certain topic volumes	Extensions of published tables

(After Rhind, 1983)

census tracts unavailable for the temporal analysis of social data. In addition, some spatial detail is inevitably lost when the EDs are combined to form the tracts.

Table 4.4 Availability of 1991 SAS and LBS for England and Wales in Machine Readable Form

Area	**Small Area Statistics**	**Local Base Statistics**
Enumeration Districts	✓	-
Postcode sectors	✓	✓
Civil parishes	✓	-
Wards	✓	✓
Urban and rural areas	✓	-
Parliamentary and European constituencies	✓	✓
District and Regional Health Authorities	✓	✓
Standard regions, counties and local authority districts	✓	✓
Great Britain, England and Wales, England, Wales and Scotland	✓	✓

(Source: Martin, 1993)

4.3 Census Data: Its Uses and problems

Basic errors in the census data arise because some people are missed entirely (either by accident or by their own design) and some are double-counted during the enumeration process and also because details are wrongly recorded, wrongly coded or wrongly key punched during analysis of the census returns (Rhind, 1993). A number of techniques have been developed in order to measure the magnitude of these errors and these are discussed in Chapter 10.3.5 and in detail by Rhind (1993 pp. 38-39). Essentially the two methods used by OPCS to identify discrepancies in the census are the Census Validation Survey and the Registrar General's mid-year population estimates. Although under-enumeration in the 1991 Census is estimated to be much higher than in the 1981 Census (estimated at 0.5%), Martin (1993) suggests that 'for the majority of ... general studies, the estimated census coverage of 98% is unlikely to have serious effects.'

The usefulness of many forms of spatial study, quantitative or otherwise depends on the nature and intrinsic meaningfulness of the objects that are under study. Geographers have a long tradition of studying areal units; for example spatial objects such as wards, districts or regions. The problem lies in the fact that ever since the demise of 'the region' as the primary object of geographical study, very little concern has been expressed about the nature and definition of the spatial objects under study (Openshaw, 1984). A geographical framework may become intrinsically related to other data sets which are then plotted and analysed in relation to it; without such a data linkage, almost no other geographical data could be spatially analysed or displayed (Raper and Rhind, 1992). If the census is to be mapped and analysed with other spatial data it is usually the case that the most statistically significant results are produced by using the most detailed data that are available (namely EDs or postcode units), with both sets of data being analysed at the same level of aggregation.

Geographical anarchy has increased over the years however, and the advent of GIS has facilitated the integration of disparate data sets. Geographers, planners and others using spatial data often encounter severe problems in transferring data from one set of areal units to another to enable comparisons and statistical tests to be carried out. In effect, comparisons are only possible by making some assumptions and no doubt many errors. This problem encapsulates the difficulties that plague research projects concerned with integrating and comparing physical and human data sets. Indeed, Openshaw (1984) identifies the development of a geographical methodology for data aggregation and interpretation as crucial for future developments in the use and integration of spatial data and for the discipline of geography as a whole. Data collection is all too often carried out in an isolated manner without any liaison between different bodies, even between those that require similar or identical data sets, so that a variety of procedures and spatial aggregations have evolved. The large number of zoning systems in common use ensure that cross-area aggregation problems are widespread. In this project alone for example, the data that were used were aggregated (prior to acquisition) for wards, land use zones and grid squares, all of which are special purpose zonings used for different reasons. The fundamental problems faced by the users of census data, have been listed by Rhind (1983) and are as follows:

1. Data cost - now negated as the ESRC, UFC and Department of Education, Northern Ireland purchased a copy of the 1981 and 1991 Censuses for use by the academic community. The Common Higher Education Software Team (CHEST - see Chapter 5) has arranged a scheme through which higher education institutions can purchase site licences for SASPAC91 (Small Area Statistics Package for the 1991 Census) from the suppliers. Registration with the University of Manchester Computing Centre Census Dissemination Unit will provide online access to the census data via the Joint Academic Network (JANET). The Census Unit also provides digitised boundary data. Registration with NOMIS, at Durham University, enables the use of their data for research projects.

2. Data age - One of the main problems with using 1991 SAS is the delay in their release: although the census occurred in April 1991 the SAS were not available for academic use until

the spring/summer of 1993. In addition, the initial schedule for the release of the computer-readable data suffered considerable delay due to a processing difficulty at OPCS, concerning the incorrect classification of some economically active people as students (Martin, 1993). These delays are part of the problem exemplified by Rhind (1991) in which he states that for a large number of census users the data are out of date by the time they use it. A census is very much a cyclical activity, with a 'sharp peak' in the field work and a 'broader plateau' in the subsequent processing stage. Rhind (1983) estimates that census data halves in value about every two or three years but the data has a remarkably long half-life because there is no direct alternative. Blakemore and Rybaczuk (1993) question the justification for a census of population at 100% enumeration every 10 years, especially taking into considering the two-year delay in the release of the 1991 SAS. They volunteer that a move towards the 'lifestyles' approach may be more applicable for the future, which could involve a central data bank collating social information from a variety of routinely collected sources.

3. Fear of the ecological fallacy - This is the problem of inferring individual relationships from aggregated data. Figures for social class data at the ward scale are useful only when it is realised that a number of problems exist when inferring individual relationships between different data sets. The ever-present problem of the ecological fallacy undermines all current analyses of census data. Hannan (1971: cited in Rhind, 1983) noted that we cannot assume that a particular relationship exists between two phenomena simply because it assumes a particular form when observed at one level of aggregation, or that it will retain the same functional form at another. Individual relationships from aggregate data will not be inferred for this reason in the course of this study.

4. The effects of the chosen geography - The modifiable areal unit problem (MAUP) states that there is an almost infinite number of ways in which a data set can be aggregated, all of which produce different results. The uncertainty between social data and the highly irregular spatial units at which it is mapped is due to the aggregation of the individual data to imposed areal units (Unwin, 1981). The boundaries of these areal units (EDs or wards, for example) are not data-derived but are designed for the ease of enumeration. As a result, the data values for each zone may be as much a function of the zone boundary locations as of the underlying distribution and the resulting maps can be very misleading (Martin, 1989). In essence, different aggregations of data have been shown to yield different results but without any systematic trends emerging that could be used for prediction or correction purposes. The preferred solution is often to transform the data from enumeration districts to grid squares (for example Barke *et al.*, 1993 use 200 metre grid cells). Transformations such as these have been simplified by the rise of GIS as research aides. The resulting pattern contains no additional information but it serves to remove a spurious impression of accuracy conveyed by the use of enumeration districts (Barke *et al.*, 1993). The transformations create their own problems however, namely

in the assumption that the population characteristics, translated into the grid squares, are evenly distributed within the original enumeration districts.

Martin (1989) states that '*it will never be possible to fully reconstruct the detail of the spatial structure from the aggregate census data*', but he goes on to say that some spatial disaggregation should be possible either by the use of data initially collected for grid cells, as in the 1971 Census, or from the use of population-weighted centroids for enumeration districts. The use of SAS for wards partly negates the problems associated with the aggregation of census data at ED level, by effectively generalising the data, enabling a more accurate picture of the county-wide variation to be presented. However, using data aggregated to ward level increases the effect of the ecological fallacy, outlined in point three above. The choice of a geographical unit is therefore be a trade-off between the available data and the level of generalisation that is required.

5. Restricted thematic coverage (especially compared to some overseas censuses such as the Australian and American censuses). This problem remains, although the 1981 SAS were expanded in scope after the 1971 Census, and the 1991 SAS contain additional information that was not available to users of the 1981 SAS. The new questions in the 1991 Census were - the postcode of household (not coded in 1981), type of accommodation, ethnic group, limiting long-term illness, term-time address of students, and weekly hours worked (Martin, 1993).

The Economic and Social Research Council (ESRC) survey of 1987 (Marsh *et al.*, 1988) reported that 68% of respondents believed that postcode-based SAS would make a major improvement or be highly desirable for their research. The most common argument in favour of postcodes as areal units is the benefits that they offer for direct linkage of census data to the ever increasing amount of postcoded information available to academics from other sources (Wrigley, 1990). The ESRC survey presents seven arguments supporting the introduction of census data for postcode areas, which provide a summary of the main arguments put forward by the proponents of the postcode system in general. In essence these are as follows:

1. Linkage with other postcoded data sources
2. Linkage with personally collected social survey data
3. Linkage of academic social science to commercial research
4. Emerging standard of postcodes
5. Lowest common denominator argument - allow maximum flexibility in later aggregation
6. Comparability with Scottish census data
7. Increasing use of the Postcode Address File (PAF) for drawing samples in academic social survey research.

The benefits of a GIS (and the data integrated within it) depend on linking different datasets together. To enable linkage, locational references are required together with standard spatial units for holding and releasing data. The hierarchical nature of postcodes provides four levels of aggregation; the postcode areas (120 in the United Kingdom) which are divided into districts (c. 2,700), which in turn are split up into sectors (c. 9,000) and finally down to the unit postcode level of which there are approximately 1.6 million with an average of 14 dwellings per unit (Rhind, 1983; Green, Lomas and O'Riordan, 1992; Martin, 1993). The division of data into these hierarchical sets facilitates user-defined aggregation. A Central Statistical Office Task Force on 'The Use of Postcodes' made important recommendations regarding the provision of postcode data for the census of population as long ago as 1977. The task force recommended that the ward should be adopted as the smallest common statistical unit and that postcodes should be the mechanism used to aggregate data to ward level (Rhind, 1983). These recommendations were not adopted for the 1981 Census however, largely due to restraints of time and resources. The Chorley Committee (Committee of Enquiry into the Handling of Geographic Information)[1] more recently recommended that OPCS 'should ensure that the results of the 1991 Census of Population and any future censuses are available, subject to confidentiality, on a unit postcode basis' (DoE, 1987 pp. 92). A postcode-based census geography for England and Wales was proposed to enable postcodes to be combined to form EDs which would nest neatly within the higher level administrative boundaries. Despite the enormous benefits which might have been gained by such an approach, the cost of these proposals made them impossible to implement in the late 1980s and an alternative approach was adopted in which postcodes would be recorded on census questionnaires but ED planning would continue without reference to the postal geography (Martin, 1993). However, 1991 Census data will be available for pseudo-postcode *sectors* in 1994. In addition, OPCS have published a postcode to ED directory for the cross-referencing of spatial data (see below).

The NRPB collect and maintain an environmental database concerned with (amongst other things) indoor radon gas measurements for the whole of the United Kingdom. The results of their measurements have been published for postcode areas and districts in England and for postcode sectors in Cornwall and Devon (see NRPB-R254). The provision of census data from OPCS for postcode areas would enable the integration of the NRPB radon data in a GIS and would facilitate meaningful areal analysis. This can already be carried out in Scotland, where postcodes were adopted as the basis of the geography of the 1981 Census and have been in use since. Whilst there are problems with postcode areas in terms of their location and accuracy, their applicability to different social and environmental databases makes it likely that postcodes will play an increasingly important role in the future development of a more integrated approach to data collection and integration.

Some of the disadvantages of adopting postcodes as the geographical base for census data are listed by Marsh and Dale (1993, pp. 79). Postcodes reflect neither administrative

---

[1] The aim of the Chorley Committee was to produce a report for the Secretary of State for the Environment and advise the government on the future handling of geographic information in the United Kingdom, taking account of modern developments in information technology and market need.

boundaries nor social reality or the built environment. They are difficult to identify on the ground and their completeness and accuracy can vary from area to area. In addition many changes are made to unit postcodes (approximately 18,000 per year) to enable new residential and office addresses to be accommodated. However, their use by the Post Office as an established working system ensures that they are kept up-to-date, although local Post Office administrative variations sometimes result in inconsistencies in the definition of unit postcodes. In spite of the decision not to release 1991 data for postcode units and because of it, OPCS has produced a Postcode to ED directory as well as a Central Postcode Directory (CPD). The CPD contains a National Grid reference and ward code for the first address in each postcode at the nominal resolution of 100 metres for England and Wales. There are still problems however, with using the CPD for this purpose, such as the inconsistencies caused by duplicated and missing National Grid references, which are making the mapping of unit postcodes problematic in some areas. The Post Office maintains the PAF and together with the CPD they permit data to be cross-referenced to other spatial data; for example the conversion of data from wards to postcodes and *vice versa*. This goes some way towards meeting the demand for better linkage between different sources of data. On balance, it seems highly probable that the provision of data for postcode-based areas will be the most significant development over the next ten years with regard to both social and environmental data collection and integration.

It has been shown that census data are usually made available only for irregular spatial zones (the one exception being the 1971 Census), which are frequently subjected to boundary changes making comparisons over time problematic. The provision of data for one kilometre Ordnance Survey National Grid squares negates many of the problems inherent in using irregular and unstable areal units. One of the recommendations of the Chorley Report was that 'as far as practical, all geographic information ... should be referenced directly or indirectly to the National Grid...' (DoE, 1987). The use of grid-squares attracted considerable interest in the late 1960s as a solution to spatial unit incompatibilities and the need for locational referencing. Census data from the 1971 Census were collected using grid-references and the results were made available for 100 metre grid squares (although most of the data at this scale were suppressed for reasons of confidentiality) and one kilometre grid squares, in addition to EDs. The initial intention in the 1981 Census was to reference to one kilometre squares throughout England and Wales. However, in the end grid square referencing was carried out for a limited number of areas where it was requested and paid for by customers (approximately 20% of England and Wales was eventually covered). In Scotland, in 1981, population and household counts were made available for one kilometre squares. OPCS received no requests for grid square referencing for the 1991 Census and no data have been made available for grid squares for the 1991 Census (Denham, 1993). It is possible to interpolate data for grid squares from EDs using a form of population weighting (see below). This form of manipulation however, is prone to error and comparisons over time are problematic. The benefits of using grid squares include the ease with which they can be geo-referenced and mapped, their stability over time, their uniform distribution across the country and the possibilities they provide for customised data aggregation according to individual requirements. In addition, census data for grid squares would facilitate the integration of other data sources, for example indoor radon measurements (provided by the NRPB for five kilometre grid squares), within a GIS. The propagation of error due to incompatible boundaries using these data would be negligible and the results from an overlay operation within a GIS would be much more meaningful due to the limited extent of the effects of the MAUP. Regular grids are also of increasing relevance with the development of remotely sensed data (for example, for land cover mapping) where the smallest unit of collection is the pixel. However there are only a small number of statistics available on a grid square basis at present (DoE, 1987) and aggregations for grid squares will produce errors if the data do not nest entirely within the boundary of the squares (the phenomenon known as the edge effect).

The fact that OPCS does not provide digital boundaries of census areas (unlike the situation in the United States) has created many problems with the mapping and analysing of census data and there are current discussions as to which organisation should take on the work. The 1991 enumeration district boundaries were digitised on behalf of the academic community (funded by the Census Initiative, see section 4.6 below) and passed to the Department of Geography at the University of Manchester, who carried out checks on the integrity of the data before passing the digitised boundaries on to the Census Dissemination Unit, also at Manchester, for dissemination to universities and research institutions nation-wide. The demand for digitised boundaries is growing very rapidly among social scientists in the United Kingdom, in line with the spread and development of GIS and more advanced software for the processing and manipulation of data. There was widespread criticism of the quality of the 1981 ED maps (Marsh *et al.*, 1988) and care is being taken (at the time of writing, autumn 1993, digital map data for the 1991 Census had yet to be released from the CDU) to correct inconsistencies already identified in the 1991 maps before they are made available. Such checks are part of the reason for the delay in the release of the map boundaries, and it is unlikely that boundary data for the 1991 Census will be available for dissemination from the CDU before 1994. The delay in availability is a major hindrance to those wanting to map the 1991 SAS.

4.4 Social Class Data in the UK Census

Social data are typically derived from four sources: censuses, surveys, registers and administrative records. The 1981 and 1991 Censuses alone cover topics such as occupation and employment status, socio-economic group (SEG) of households, social class of households, social class and economic position of residents, SEG with social class and ethnic group and the standard occupational classification. Socio-economic groups and social class are essentially descriptive categories which are derived from analysing the answers to certain questions on the census form (see below and Appendix 4). Rhind (1991) and Lievesley and Masser (1994) provide an overview of social surveys in the United Kingdom as well as in Europe. The SEG categorisation was introduced in the 1951 Census and was extensively amended in 1961. It is a non-hierarchical classification which uses both occupation and employment status (Dale, 1993). The classification was designed by OPCS to group people with jobs of similar social and economic status and it sorts the economically active head of the household into one of 17 socio-economic groups, plus a group for those who were economically inactive on the day of the census (1991 Census) or never active (1981 Census). SEG data in the 1991 SAS are found in Table 86.

Stevenson introduced the Registrar General's social class schema in the 1920s as a categorisation of occupations that reflected 'the wealth or poverty and the culture associated with class' (OPCS, 1987 pp. 2). Over the years the social class categories have changed slightly - the division of skilled occupations into manual and non-manual occurred in 1971. In the 1991 Census, the title was extended to 'social class based upon occupation' to make clear that occupation provides the key to the classification (Dale, 1993). Social class data are analysed mainly from the answers to question 15 of the 1991 Census form (see Appendix 4). The derived 10% data are found in Table 90 of the 1991 SAS. Nichols (1979) puts forward the idea that the sociological concepts of class often stress the importance of different dimensions of social stratification and thus invite criticism from Marxists on the ground that they do not explain the generation of inequalities in the structural dynamics of society, only describe the results. This is a fundamental criticism of census data in general, in that it provides only counts and omits any possible meta-data that could be incorporated into the structure of the tables.

The social class categories are not composed of aggregations of socio-economic groups, but are related to much broader occupational groupings using a different system of classification (see Table 4.6). The social class classification (10% sample) is a categorisation of economically active, retired and permanently sick persons classified according to their present or former occupations. The economically active head of the household is classified according to occupation in to one of five classes (I to V), plus a group for those in the armed forces and inadequately described and retired. Table 90 (1991 Census) differs slightly from the corresponding 1981 table (Table 49) due to the inclusion of data relating to the number of

Table 4.5 The Socio-Economic Classification (1991 Census)

SEG	Description
1	Employers and managers in central and local government, industry, commerce, *etc.* (large establishments)
	1.1 Employers
	1.2 Managers
2	Employers and managers in industry, commerce, *etc.* (small establishments)
	2.1 Employers
	2.2 Managers
3	Professional workers: self-employed
4	Professional workers: employees
5	Intermediate non-manual workers
	5.1 Ancillary workers and artists
	5.2 Foremen and supervisors
6	Junior non-manual workers
7	Personal service workers
8	Foremen and supervisors: manual
9	Skilled manual workers
10	Semi-skilled manual workers
11	Unskilled manual workers
12	Own account workers (other than professionals)
13	Farmers: employers and managers
14	Farmers: own account
15	Agricultural workers
16	Members of the armed forces
17	Inadequately described or non-stated occupations
	Economically inactive head

people on a government training scheme and subdivisions to include a category for inadequately described and inactive. In addition, the 1981 social class classification categorises those retired on the day of the census into one of three retired groups, whereas the 1991 classification categorises all those who are retired into one single group. The classes are divided as follows:

Table 4.6 The Social Class Classification (1991 Census)

Social Class	Description
I	Professional *etc.* occupations
II	Managerial and technical
III N	Skilled occupations: non-manual
III M	Skilled occupations: manual
IV	Partly skilled occupations
V	Unskilled occupations
Armed forces	
On a government scheme	
Inadequately described	
Retired	
Other Inactive	

Part Two SAS data are available only as a 10% sample of the population and are taken from the records from which the 100% statistics are derived. The sample consists of one household selected at random[2] from each stratum of ten consecutively recorded households (and a similar sample of persons in establishments). Figures derived from the sample can only provide estimates of the population values, although the stratified nature of the sample tends to reduce the error margin. At ED level 10% SAS tables are subject to large errors (see Rhind, 1983 pp. 79) although the aggregation of the EDs to ward level helps to ensure that the level of variability between the cell values is reduced (Rhind, 1983). Social class data were chosen as the source of social data in Northamptonshire because the large number of classes in the socio-economic classification would make analysis time consuming and unnecessarily complex. There are 17 socio-economic groups in total and when this figure is multiplied by the seven districts in Northamptonshire the number of counts (119) becomes unmanageable. There are just six main social classes (plus a category for armed forces and inadequately described and

---

[2] The sample is not strictly random but is taken from blocks of 50 EDs, the strata running continuously from the first household in the first ED to the last household and from the first person in the first communal establishment to the last person in the last communal establishment (Rhind, 1993)

retired) which more than halves the number of counts and makes the number of entries more manageable. Thus social class data from the 1991 Census were chosen and were analysed at district level with radon data from the NRPB (see Chapters 9 and 12).

The main problems encountered with using social class or socio-economic group data are as follows:

- There is a lack of continuity in the layout and content of the social class and SEG data in different censuses rendering temporal analysis problematic.
- There is a lack of internal homogeneity in any class or group. The main disadvantage of the social class data is the broad headings into which the active head is classified. This allows for considerable room for heterogeneity within the groups - households at the top of a class are likely to differ significantly to those at the bottom, in terms of income and lifestyle. A compromise needs to be found between complexity (17 SEG groups is often regarded as too complex for spatial analysis) and accuracy.
- The integrity of the data is affected by omissions as a result of those unwilling to reply or those who answer unclearly. The random addition to some counts of 1, 0 or -1 to preserve confidentiality also affect the accuracy of the data. Accuracy is increasingly restored as the degree of aggregation increases, *i.e.* data for wards are more accurate than for EDs.
- The 10% sampling method may not be truly representative (see above).

However, in conclusion, social class data available from the 1991 SAS provides an adequate means of analysing the social structure of the population and constitutes a comprehensive database with which to work. A summary of the county-wide differences and similarities between the five Affected Areas in terms of social class characteristics is provided in Chapter 12.6.

4.5 Various Census Initiatives

The Census Initiative is jointly funded by the Economic and Social Research Council (ESRC) and the Information Systems Committee of the University Funding Council (UFC) and co-ordinated by Professor Philip Rees. It aims to encourage the development of new uses of census data, the production of new added-value products based on census data and the investigation of methodological problems common in the analysis of census data (Rees, 1991). The first round of projects funded by the initiative were carried out in the period 1992 to 1993 and the outcomes of the research should provide a better understanding of the various problems inherent in the integration of census data with other data sets. The initiative will, for example, address aggregation issues, aspects of population surface modelling and the evaluation of occupational definition coding and as a result will provide more information on the effects of the outstanding problems concerning the ecological fallacy and the MAUP. The Census Initiative also funded the purchase of a digitised 1991 ED boundary dataset for England and

Wales (and a postcode unit dataset for Scotland), on which the School of Geography and the Computing Centre at Manchester carried out a quality assurance check. The digital boundary data were not available until July 1994 however.

A new product for the 1991 Census is the Sample of Anonymised Records (SAR), specially commissioned by the ESRC and available from the Census Microdata Unit. The data files comprise two separate samples, a 2% sample of all individuals (1.21 million records) and a 1% hierarchical sample of households and individuals within those households (240,000 records). The SARs contain the full range of census data, *i.e.* both the 10% and the 100% statistics (Martin, 1993). The data allow analyses of the relationships between the characteristics of different household members, thus enabling statistically reliable individual analyses to be carried out for the first time (Marsh *et al.*, 1988). The SAR offers the potential to explore relationships in the data in ways which were not pre-planned, and this is perhaps its greatest strength (Martin, 1993). For a detailed explanation of microdata and their advantages, see Marsh and Teague (1993).

In addition to new products, the ESRC continuously assesses new methods for data dissemination. The ESRC Data Archive is currently pioneering the circulation of large datasets on compact disc (CD) and are keen to see the development of CDs which will allow the census data to be disseminated to the academic researcher's desk and to every higher education institution library.

4.6 Data Accession

The social data used for the analysis with radon levels in the county of Northamptonshire (discussed in Chapter 12.4), were taken from the 1991 Census of Population. Small Area Statistics on topics such as social class (and other 10% data) are available from OPCS in the form of County Reports (Part 2) and 100% data, such as population density, age structure and household characteristics, are available in the form of Ward and Civil Parish Monitors and as County Reports (Part 1), one of which is published for each county in England and Wales. Digital census data are also available on-line from the CDU at Manchester University and from the National Online Manpower Information System (NOMIS) at Durham University, provided the user is registered. There are a number of similarities between the SASPAC software and the NOMIS system - both procedures are command-driven, both with command languages written specifically for the purpose of accessing the data. The nature of the data that are held differs slightly in that NOMIS is concerned primarily with manpower and employment data (although all the SAS are also available) but both organisations rely on government-sponsored survey information. There is also some agreement between both organisations and the academic community. However, there are also some important differences between the two. NOMIS is a government sponsored institution and therefore plays an important role in the provision of employment information to the government. Another important role that it performs is to provide a source of census and employment information for private companies and non-academic institutions, for the cost of a annual registration fee. Census data from NOMIS, especially the digital boundaries, are subject to strict licensing laws and the principal method of using the digital boundaries is to process them through ARC/INFO on a NOMIS computer and to output the final map or graph for collection in their office. Census data held at the CDU at Manchester University however, are available using the software programme SASPAC91 and provide an unparalleled source of social information. The data are available to all registered academics at participating universities via the CHEST deal and the data are available for research use at the computer terminal of the users choice.

Census data are thus readily available in machine readable form. However, the delay in the processing of the data and in their dissemination, negated the apparent advantages of using digital data in this thesis. Because relatively few statistics were required for the county of Northamptonshire (data for the 148 wards and seven districts were used - see Chapter 12.4), the census data were manually entered into ARC/INFO in the TABLES option. Wards (and districts for social class) were chosen as the geographical basis for mapping the census within ARC/INFO. As mentioned above, the ESRC and CHEST have funded the purchase of a 1991 digital ED boundary map for England and Wales. But because of the sheer volume of work involved in the digitising process and the necessary quality checks that were carried out, the boundary data were not released until July 1994 (three years after the census was carried out), after the analysis for this thesis had been completed. The only way to avoid the time-lag in the release of the digital maps (other than purchasing the data from private companies), is to digitise the ward boundaries from paper maps, available at minimal cost from OPCS. The volume of work inherent in digitising the 148 wards of Northamptonshire, renders the task of digitising them impossible within the time frame and resources available for this study. On the other hand, digitised 1981 Census boundaries are available from the Bartholomew's digital data set which provides readily available digital administrative data for the whole of England and Wales (see Chapter 5 and Table 5.1). The data are available in ARC/INFO export format and can be readily imported into ARC/INFO. The number and scale of changes to ward boundaries since 1981 were not excessive, only three principal areas were affected in Northamptonshire. These were located in the towns of Northampton, Wellingborough and Kettering where administrative divisions tend to be much smaller in area. No district boundaries had been altered. A more detailed description of the differences between the 1981 and 1991 ward boundaries is given in Chapter 10. The methods of data accession and integration with the other data sets are outlined in more detail in section 9.4 (Chapter 9). The census resources available from the five county councils of the Affected Areas is listed in Appendix 2, together with a résumé of possible future developments in their available data resources.

5. Sources of Other Digital Data

It was originally planned to include a chapter on the effects of topography on radon gas levels. However despite continued efforts to find a suitable source of digital data, none were located and the possibility of carrying out this overlay was abandoned (see below). The situation of topographic mapping in England is unusual (compared to other European countries) in that there is only one agency responsible for mapping at both large and small scale - namely the Ordnance Survey (Lievesley and Masser, 1993). Map series maintained by the Ordnance Survey (OS) range from 1:1,250 to the 1:1,000,000 scales. Within this range of scales a wide variety of products is produced by the OS for retail to organisations and to the general public. Rhind (1992) reports that Britain is at the top of the European league in terms of cost recovery, with over 60% of Ordnance Survey costs recovered through the sale of map products and services and this percentage is due to increase over the next five years.

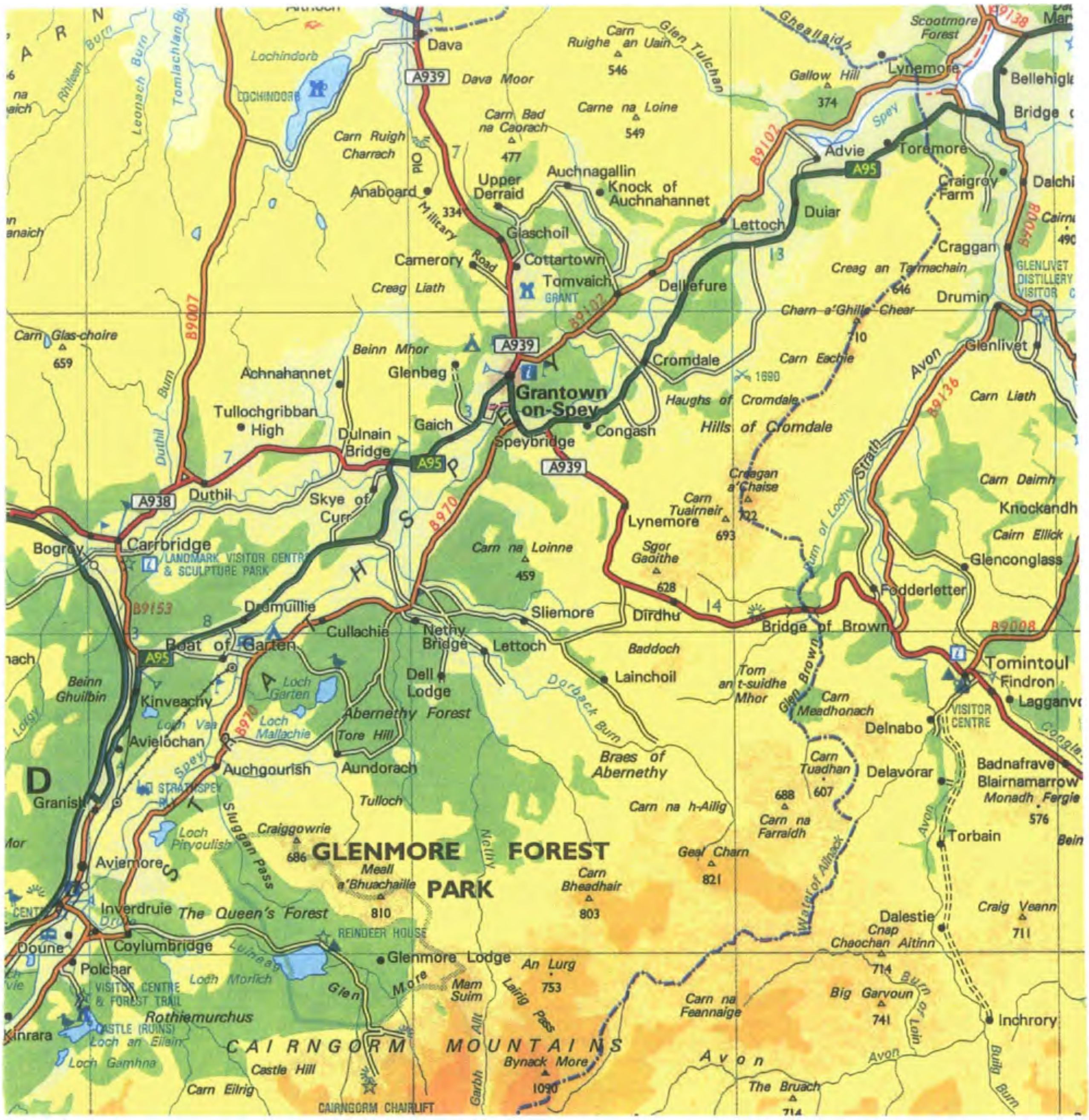
One readily available dataset is the Bartholomew digital data for Great Britain. The data are available on networked workstation computers via the academic network, as part of the country-wide CHEST (Combined Higher Education Software Team) agreement. The data are in digital format (scale 1:250,000) and have been digitised from Bartholomew's 1:253,440 scale motorists maps (see Figure 5.1). The data are supplied separately for each 100km square of the National Grid (all co-ordinates are metres on the National Grid) and are split into several layers (see below). The available coverages are as follows: administrative boundaries, drainage, danger zones, forest parks, place names, national park boundaries, national trust areas, other linear data (pipelines, long distance footpaths *etc.*), point data (spot heights, railway stations *etc.*), roads, railways and ferries, regional park boundaries, topography (beaches, rocky shorelines and woodland area boundaries), urban boundaries and areas of water. In addition to these, a coverage with the identifier CON contains all the contour information for each 100km square of Great Britain. The data are supplied in ARC/INFO EXPORT format which can be read straight into ARC/INFO. The Bartholomew administrative boundary coverage was imported into ARC/INFO and was used to produce the maps in Figures 12.4.7, 12.4.8, 12.4.10 and 12.4.15 (all Chapter 12) which show the wards and districts in Northamptonshire.

The contour coverage is very detailed (there are 16 different contour intervals ranging from 0 to 1,300 metres) and when mapped in conjunction with other data can obscure underlying information, giving rise to a cluttered map. It was decided not to use the contour data in an overlay with the radon grid for several reasons. Firstly, the coverage as it stands is too detailed for the requirements of this thesis. In addition, the method used to store the Bartholomew's data (as 100 kilometre squared grids) means that 11 grids would have to be imported into ARC/INFO to cover the Affected Areas. This would involve a great deal of time (both in accessing the data, incorporating it in ARC/INFO and then carrying out the overlays) plus additional training. The issue of time and the required training also negated the possibility of using a sample area. In addition, the use of a sample area would not further the assessment of the geographical differences and similarities of the Affected Areas. However, although no contour overlay was carried out, possible sources of topographic data have been discussed in this chapter (see Table 5.1 below) and a sample area of the Bartholomew data has been obtained. Figure 5.1 shows an extract from the Bartholomew's source road maps at the scale of 1:253,440 from which the digital data is derived. The relief of the land is depicted by hill shading with the boundary of each colour representing a contour line. The area depicted is in the Grampian region of Scotland - the grid reference of Grantown-on-Spey is NJ0327.

The OS produce a wide range of digital products. The OS Landranger 1:50,000 map series have been digitised in full colour raster format, for the whole of the United Kingdom. This dataset includes roads, contours, topography *etc.* (a similar range to the Bartholomew's 1:250,000 data described above). Scale and gazetteer data are also available and all the data are available to suit a variety of packages including GIF, BMP, SUN raster, PCX, PICT and TIFF. Data at the scale of 1:250,000 have also been released under the name of 'Strategi'. 'BaseData.GB' is the OS database at 1:625,000 scale. The 'Boundary Line' dataset contains administrative boundaries and is, in the words of the OS, "the definitive digital dataset of the administrative boundaries of England and Wales". There are plans to introduce a similar dataset for Scotland in 1994. The data are derived from the 1:10,000 Boundary Record cards held by the OS and contain information on county, district, ward, parish, parliamentary constituency and European constituency boundaries. ED-Line is a similar range of digital boundary data for 1991 Census enumeration districts for England and Wales. Urban areas have also been captured for both the 1981 and 1991 Censuses in co-operation with the Department of the Environment and the Welsh Office.

Digital data is an expensive way of acquiring spatially-referenced data. For example, the Bartholomew's 1:250,000 data is available (outside of the academic community) at the price of approximately £20,000. In addition, the majority of the public domain data (*e.g.* that supplied by the OS) are liable for annual copyright charges. As a result of these costs, the supply of datasets in higher education has traditionally been subject to the vagaries of available funding and rising prices (Kitmitto, 1994). The cost of data (especially digital data) and computer software was the main impetus for the spawning of the Combined Higher Education Software Team (CHEST) in 1988. CHEST is an instrument which acts as a central body negotiating, on behalf of the United Kingdom academic community, with suppliers of commercial software and datasets for the best prices (Kitmitto, 1994). A purchase is funded for the academic community as a whole and the data are then available to institutions on a site-licence basis for a modest annual fee. The OS has recently reached an agreement with CHEST under which samples of the full range of digital data are available to the academic community. Six areas of data are available and will be updated annually - these are Port Talbot, Gower, Peak District, Lake District, Bristol and Glasgow (none of which is of use in this thesis). As mentioned above, the

Figure 5.1 Extract from Bartholomew's Motoring Atlas (1:253,440)



Bartholomew's Great Britain database is also covered by the CHEST agreement. A CHEST Software Directory is published annually and is distributed to every department in all higher education institutions. The latest edition of the directory detailed nearly 1,100 products. CHEST's sister project, the National Information Services and Systems (NISS), maintains electronic versions of the CHEST directory in various online services. The following table is a list of the primary spatial sciences datasets that are available from CHEST in the United Kingdom. In addition to datasets, some spatial sciences software is also available via CHEST, namely ARC/INFO and ArcView from ESRI, and IMAGINE from ERDAS.

Table 5.1 CHEST Spatial Sciences Data Deals

Vendor	Product	Coverage	Scale	Format
Bartholomew/Times	Digital map data	United Kingdom	1:250,000	Arc/Info Export and Ungenerate
Bartholomew/Times	Digital map data	Europe	1:1,000,000	Arc/Info Export and Ungenerate
Bartholomew/Times	Digital map data	World	1:1,000,000	Arc/Info Export and Ungenerate
GEOPLAN UK Ltd.	Digitised postcode boundaries	United Kingdom	1:100,000 and smaller scales	Atlas
National Remote Sensing Centre Ltd.	Landsat data	United Kingdom	1 pixel = 30 m	Ground receiving station
Ordnance Survey	Sample digital data	sample areas (see text)	1:1,250 to 1:625,000	NTF and DXF

(Source: Kitmitto, 1994)

A variety of additional data sources for academics are available via the National Dataset Service at the University of Manchester. These include 1981 and 1991 digitised boundary data, together with a range of statistical data from the census and government survey datasets (*e.g.* General Household Survey, National Labour Force Survey, Family Expenditure Survey). A valuable directory of digital data sources in the United Kingdom has been assembled by O'Carroll *et al.*, (1994) and published by Oxford Brookes University.

6. Data Availability in the 1990s: A Résumé

The main factors influencing the choice of data and the method of analysis in this study were: time; the availability of money and necessary resources; the availability of hardware and software for integrating and analysing the data and for displaying the results together with associated training; knowledge about the resources and data that were available nation-wide for possible integration within GIS; data availability (free from copyright, relatively up-to-date and available for research use).

Information about the availability of information is not easy to obtain (Blakemore and Rybaczuk, 1993). Environmental and social data are collected by a diverse range of organisations, stored in a variety of formats and in a variety of locations throughout the country. However much of these data are application-specific and there is no umbrella organisation which keeps a directory of all available data for prospective users. A directory such as this would enable maximum use to be made of the data collected nation-wide. This problem is not confined to the United Kingdom; there is nothing in the literature to suggest that any country currently maintains a definitive directory of available spatial data. The Economic and Social Research Council Data Archive (based in Essex) goes some way towards creating a national repository for social and economic data and at the regional level - the Regional Research Laboratories maintain a relevant selection of data bases. The Environmental Information Centre at the Institute of Terrestrial Technology at Monks Wood (set up in 1989) is the only really comparable project for environmental data (Heywood, 1991), although there are agencies such as the Soil Survey Research Laboratory at Silsoe which maintain data banks for specific areas of study, in this example, for soil classifications. The Chorley Report (DoE, 1987) attempted to list the major databases of spatially referenced information available to GIS users in Appendix 3, but this is now out of date and it does not contain any information on the structure or quality of the datasets that are listed. There have been initiatives by government departments and national research institutes to compile lists of available data for specific areas of interest, but as yet these have not been collected and amalgamated by a central agency. The Department of the Environment (1991) compiled an inventory of spatial information held by central government departments as part of a wider Tradeable Information Initiative. Responsibility for the database was passed over the Ordnance Survey's Customer Information unit in January 1994 and information about the data is now freely available through the Spatial Information Enquiry Service (SINES). SINES currently holds details of over 450 data sets supplied by government departments throughout the United Kingdom (O'Carroll *et al.*, 1994). The Royal Institute of Chartered Surveyors has created a similar inventory for environmental datasets available in the United Kingdom (McLaren and Tulip, 1991).

The situation of data availability is constantly changing. One important development in the area of access to data and information on data quality is currently underway as part of the

United Nations Environment Programme (UNEP). UNEP has developed the Global Resource Information Database (GRID) network whose mission it is to prepare and make available reliable and up-to-date resource information. One of GRID's long term goals is to facilitate international access to all existing environmental data covering hundreds of countries and regions in dozens of subjects and themes. The GRID Meta-Database is currently being developed in the form of a computerised catalogue to help users locate and acquire environmental data. The Meta-Database contains complete descriptions of data resources held at GRID centres around the world in a variety of media, covering digital and paper maps, satellite images, tabular and statistical data. It aims to eventually contain descriptions of data held in other archives and information on the databanks of relevant institutions. There will be a variety of methods of accessing the data held on the database from personal contact to global computer networks. The result, once completed, should be a comprehensive and accessible directory of data resources known to exist, enabling fast and reliable searches on available data (UNEP, 1992). The Meta-Database is primarily targeted at national, international and global environmental organisations, but there is tremendous scope for a wide range of users who could benefit greatly from this information about information.

Problems of access, copyright and pricing limit the use of large amounts of data in the United Kingdom. There has been much lively debate regarding access to databases. The CHEST deal, outlined in Chapter 4, is one way in which Universities and data producers have co-operated to enable free access to national databases, that would otherwise be too expensive for educational institutions to use. A major step forward in the availability of spatial data in England and Wales was achieved in December 1990 with the partial 'opening' of the Land Registry information which brought it in line with the Land Registry in Scotland. McLaren and Tulip (1991) cite this as a step in the right direction towards greater openness and ease of access to national datasets.

Data are becoming the major factor in the success of GIS and a subject of much debate. The fact that the situation in the United States is so different from that in Europe serves only to fuel the debate. In the United States federal data are made available at the cost of dissemination (Rhind, 1993). In the United Kingdom the situation is increasingly becoming the opposite, as the government tries to save money by increasing the cost of data. The policy and practice of the British government has been made clear - if the customer is not prepared to pay for the information then it is not required and should not be produced. Recovery of costs is one of the important criteria of success within a data collection organisation (Rhind, 1993). The Ordnance Survey, for example, aims to recover 100% of their costs by the year 1997 by charging for the use of their data and from increased merchandising (Rhind, 1993). This situation could, in many cases, lead to the effective unavailability of appropriate, comprehensive, definitive and comparable datasets. This is a sore point with many GIS users.

At the Association of Geographic Information annual conference in 1992 several general issues were raised which caused widespread concern amongst GIS users. It is useful to cite them here as an overall view of some of the problems affecting geographical research at the present time. The issues include: the late availability of government (1991) Census information; networking and the better exchange of data; new technology for the ordinary user in the future; the management of databases; the quality of data and the likely propagation of error (Smith, 1993).

Geographical data are increasingly becoming part of the national and international information economy. Geographic or spatial data are the fuel of GIS and their availability, access and price are therefore of considerable importance. All of the data used in this study derive from government sources and whilst not all were available in computerised format, it is a function of our post-industrial society that data are increasingly being made available in digital format which can easily be exchanged, manipulated and traded in a multitude of ways by a wide variety of users (Rhind, 1992). Data collectors could greatly assist data users by providing information on the quality of the data so that more rational decisions can be made about which data to include in an analysis. Similarly, users would greatly benefit from discussions with data collection agencies about the nature, quantity and quality of the data that are required, and a two-way feedback discussion could be instigated at each level of data collection, organisation and manipulation. In this way all who use the data would benefit. The end users of the data would be provided with a higher quality and more relevant data set tailored in some respects to their needs and the data collection agencies would be rewarded with increased revenue (because the data would be applicable to a much wider range of user), better customer satisfaction and a longer shelf-life for the data, especially if the data are released shortly after collection.

Part Two: Methodology

PART TWO

7. Introduction

The purpose of Part Two is to outline the methodology employed to collate the three main data sets (identified in Part One) and overlay them in the geographical information system (GIS) named ARC/INFO. Chapter 8 provides a description of ARC/INFO, in addition to a general introduction to current GIS technology. The benefits and limitations of using GIS as a research tool are also discussed. The methods of data capture and integration in ARC/INFO are outlined in Chapter 9.

GIS have enjoyed a rapid rise in popularity and their development has greatly increasing the supply of all kinds of mixed-origin data without explicitly solving any of the outstanding data integration issues (Flowerdew and Openshaw, 1987). Most geographic data are generalised to some degree because of the methods used to collect and display spatial data. Transformations in the data, especially during overlay procedures, often result in error due to operational error and the problems inherent in the data structures themselves. This can lead to poor quality output which in turn will influence decision making. The issue of error propagation in GIS has only recently been highlighted as an important area of study. Chapter 10 assesses the main sources of error in the context of this study and looks at the wider implications of data error on GIS output in general. The range of new literature on the subject of error reflects the increasing user awareness and concern. Chapter 10 also examines some of the recent literature and looks at areas where research is currently being undertaken. Some questions on the subject of error remain unanswered. These include the possible effects of an areal partition or spatial aggregation on the results of spatial analyses and on the mapping of spatial data.

Chapter 11 serves as a résumé of Part Two.

8. Geographic Information Systems

8.1 Geographic Information Systems

Computers have been commercially available since the 1950s, but it was more than a decade before they exerted any influence on geography. During the early 1960s the discipline of geography experienced important philosophical and methodological changes which have subsequently been termed 'the quantitative revolution' (Johnston, 1987). Geography changed from an essentially qualitative and descriptive discipline into one which became increasingly concerned with the development of generalised laws and theories about spatial patterns using mathematical and spatial statistical methods (Maguire, 1989). The development of faster and cheaper computers heralded the start of a 'technological innovation'. Microcomputers were being used by a wide variety of academic departments by the late 1970s and by the 1980s geographical information systems (GIS) were in widespread use. By the end of 1993 the GIS market in the United Kingdom alone was worth £85 million and it is poised for growth of over 40% in 1994 (Pritchard, 1994). A consequence of the enormous commercial value of GIS for a wide range of applications, is that much of the research and development work on GIS has been undertaken by commercial organisations and is strongly application-oriented (Maguire, 1989).

There are many definitions of a GIS (for example see Goodchild, 1985; Parker 1987; Peuquet and Marble 1990) but the one offered by Burrough (1986) is concise and widely accepted. Burrough defines GIS as 'a powerful set of tools for collecting, storing and retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes.' A GIS is essentially a system of computer hardware and software combined with procedures designed to input, store, analyse and display spatial information. To be classified as a GIS the system must include all of the above functions and furthermore it must perform efficiently in all four areas. Appendix 4 of the Chorley Report (DoE, 1987; Rhind, 1981) lists the functions that a GIS are required to perform in more detail. There are five generic questions that a sophisticated GIS can answer (ESRI, 1987). These are as follows:

1. What is at...?
2. Where is it?
3. What has changed since...?
4. What spatial patterns exist?
5. What if...?

The first question seeks to discern what exists at a particular location; the second involves spatial analysis to discover what exists at a certain location; the third questions relates to change over time; the fourth requires more sophisticated methods to determine the location and interdependence of sections of the database and the fifth is based on modelling to determine responses to certain stimuli.

In recent years GIS technology has emerged from the realm of research and development to one of application and is rapidly becoming a powerful tool for integrating and analysing spatial data. Over the last 30 years the technology of GIS has changed significantly, as have styles of hardware and software, the diversity of analytic methods, the range and quality of products and the cost of purchasing the hardware and software. To emphasise the relatively recent rise in popularity of GIS as an important geographical tool, it was not until 1987 that an international academic journal concerned purely with GIS emerged. There are now several on the market. The importance of the role of GIS as research tools is underlined by the user base that has steadily grown since the development of the early GIS in the 1960s. GIS have now become a viable 'enabling' technology for addressing complex and multi-disciplinary environmental issues at local, national and global level. Indeed, GIS have been classed as the biggest step forward in the handling of geographic information since the invention of the map (Chorley Report, 1987). Since Tobler first outlined the potential use of computers for mapping in 1959, technology has developed at an ever increasing rate, so that what is fashionable and up-to-date at the present time will be surpassed by a more advanced, faster and more accurate computer system by the end of the decade, if not sooner. This is likely to be the case for many more years to come.

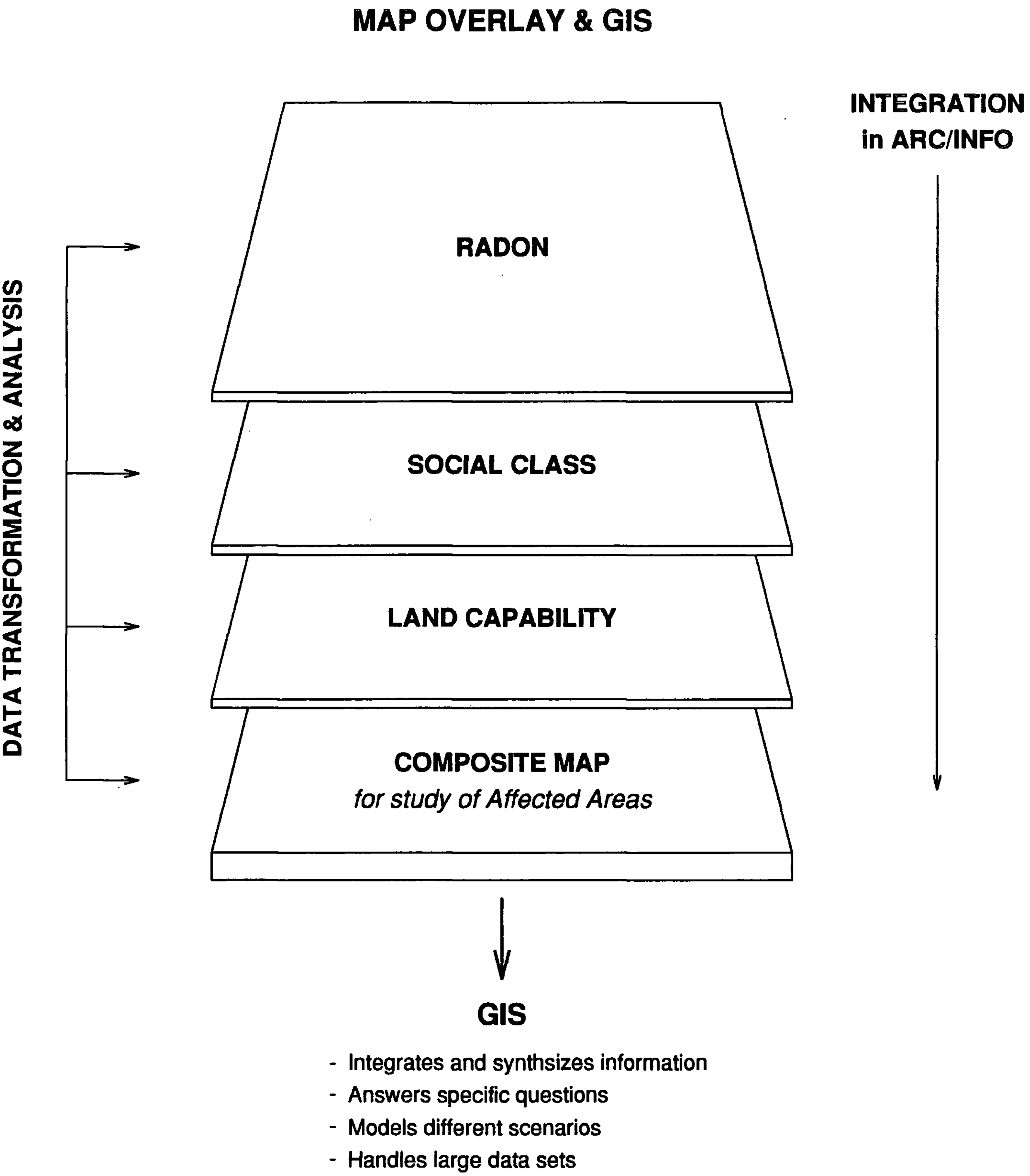
Two of the first large-scale GIS to be developed in the United States were the TIGER (Topologically Integrated Geographic Encoding and Referencing) project and DIME (Dual Independent Map Encoding) project both developed to aid analysis of the 1990 US Census results. The requirements of the TIGER system are defined as the ability to provide an integrated description of the census geography of the whole of the US, to provide an automated field mapping system, to provide a system for tabulating data by any of the standard administrative units and to provide a publication mapping system (Rhind, 1991). One of the more modern technological innovations which is already having a significant impact on GIS is the Global Positioning System (GPS). This promises to provide a revolution in position-fixing by enabling anyone with a GPS receiver to locate themselves to an accuracy of about 100 metres anywhere on the Earth's surface, using locational information from two GPS satellites in space. This technology can be incorporated into a GIS and enables 'pin-point' accuracy for the location of both natural and man-made features.

In the United Kingdom, GIS is beginning to penetrate significantly into local government (see Appendix 2) partly as a result of the Local Authority Service Level Agreement between the Ordnance Survey and the Local Government Management Board, which has made data available to local authorities at concessionary rates. County Councils are in a good position to act as regional databases for digital data and also to maintain meta-databases on other available data sets, both commercial and academic (see Chapter 6) (O'Carroll *et al.*, 1994). More generally, the technology provided by a GIS enables the integration and analysis of large amounts of spatial data that would not have been possible with traditional analogue techniques. GIS is especially relevant to geographical enquiry because of this spatial element inherent in its design, and its applications (which are not confined to the discipline of Geography) are many and varied. One of the most important features of a GIS is its ability to search for and display relationships among mapped attributes and it is this function that makes GIS technology so well suited for spatial analysis of the kind that is being undertaken in this project. Since the adoption of the early GIS in the 1960s, a wide range of applications have been developed. Perhaps the most fundamental advance was the development of the ability of the GIS to 'overlay' layers of data. Map overlay involves the superimposition of two or more input maps, or data layers, with the aim of producing a composite map showing the intersection of the mapping units on the individual data layers. Figure 8.1 shows the coverages that were used in the overlay procedure carried out in this project. The overlay procedure can only be carried out by using coordinates (*e.g.* National Grid coordinates) to register the data layers so that they correspond to exactly the same area. A variety of spatial analyses can then be carried out on the overlays, including buffering and nearest neighbour analysis. The results of the analyses can be produced as maps, statistical summaries or in tables.

It is increasingly recognised that the development and exploration of GIS technology is heavily dependent on the availability of geographic information; it remains true that the results from analyses using GIS will remain only as good as the raw data that are used. By far the largest suppliers of national spatial data sets are government agencies such as the Ordnance Survey, OPCS, the Department of Employment and MAFF, as well as research councils, health authorities and tourist boards. Consequently there is a growing debate about the extent to which the activities of these agencies could or should facilitate the use of GIS technology (see Blakemore and Singh, 1992; Rhind, 1992) The main problems with using data collected by organisations are the cost, copyright and licensing problems and scale (the resolution of the data is not always as detailed as the user would like). Furthermore, some data may be withheld or aggregated to preserve confidentiality, both personal and military. The ensuing difficulties in interpreting the results from analyses with these data have been covered in more detail in Chapter 4 and the sources and effects of induced error is covered in Chapter 10; these aspects of GIS usage will not be repeated here.

The main benefits of GIS over traditional analogue and early computer mapping techniques are based on the ability to link disparate datasets and the capability for quantitative analysis of spatial data. Traditionally geographers relied on manually produced maps to integrate and summarise data, for example using map overlays. Most maps are out-of-date by the time they are published, yet the information they portray is dynamic and needs to be up-to-date. This is the fundamental dichotomy that exists between spatial information and conventional maps. There are several problems inherent in storing information in the form of a paper map (not least because it is laborious and subject to errors), some of which have been alleviated by using GIS technology for spatial database management and mapping, whilst others have remained. The original data still has to be reduced in volume in order to be classified or grouped, (although

Figure 8.1 Map overlays used in the analysis of the Affected Areas



(Source: Freeman et al., 1993)

some data sets which do not require classification can be stored in their raw format, if raw data are available) and presentation, either as an image on the screen or as a hard copy, still has to be accurate and clear in order to be legible. However the problems of expensive updating and information retrieval have largely been negated by the modern GIS. Large or complex data sets are much more effectively handled if they have been integrated into a computer package which will carry out overlays, rescaling and the like in a matter of seconds. The dynamic nature of GIS enables data to be continuously updated as required. The final output can be presented in a number of formats including maps, tables, diagrams and charts which can be produced almost instantaneously.

8.2 Hardware and Software

A variety of computers were used in the data capture, storage, manipulation and writing up of the thesis. The word processing was carried out on an IBM-compatible 486 machine, running Microsoft *Word for Windows* (version 2.0). This machine was also used for simple data analysis and for drawing graphs using the spreadsheet, Microsoft *EXCEL* (version 4.0a).

The majority of the data capture, manipulation and storage was done on a networked Sun workstation at Durham University, Computing Department. Two different software packages were used to maximise ease of data capture and to capitalise on the resources available. The GIS package ARC/INFO was used for the integration and manipulation of the various data sources, for data analysis and for the production of the cartographic output. ARC/INFO, developed and maintained by Environmental Systems Research Institute Incorporated (ESRI) of Redlands in California, combines the functions of a mapping package with a database that can be used to assign attributes and carry out analytical functions. The ARC/INFO software was first released in 1982 and the user base has grown to over 17,000 licences and an estimated 35,000 regular users (ESRI, 1993). ARC/INFO was the first database-oriented GIS and was probably the most successful commercial GIS of the 1980s. It is designed as a vector-based GIS and is made up of two primary components. The first component called *ARC*, was written by ESRI to store coordinate data and to perform operations on the spatially referenced data. It is the main programme environment and has capabilities for the following (Peuquet and Marble, 1990):

- Data conversion, digitising and editing
- Error detection and data verification
- Workspace and file management operations to list, rename, copy, delete and describe map data files
- Coordinate projection and transformation functions for merging adjacent map sheets and for overlay functions
- Capabilities for the management and manipulation of feature attributes

- Analytical operations including feature buffering, map overlay, nearest neighbour analysis and reporting of summary statistics.

The second component *INFO*, is a relational database management system developed in the USA by Henco Corporation and is used to perform operations on attributes, which are the descriptive non-coordinate data (Peuquet and Marble, 1990). Relational databases (which preserve spatial relationships between the data), together with the mapping and graphics facilities within GIS, provide a potentially valuable structure for developing automated spatial statistical analysis (Haining, 1990). In total ARC/INFO has over 2,500 tools for capturing, managing, organising, analysing, displaying and plotting geographical data. The main advantages of ARC/INFO (other than its widespread availability within the academic community) are its ability to integrate many types of input data and its ability to operate in both UNIX and PC environments. A diagram of the ARC/INFO methodology is shown in Figure 9.1.

The second package that was used is Lites2. This package was developed by LaserScan in Cambridge and is a powerful interactive graphical digitising and editing program. It is particularly well adapted for digitising and mapping; ARC/INFO, on the other hand, is capable of simple mapping but is best used for the analysis of spatial data. The two thus complemented each other. To convert the conventionally mapped land use data into computer-readable format, a digitising table and puck were used in conjunction with the Lites2 software. The digitising process is outlined in Chapter 9. The data were then saved onto a floppy disc, converted into ARC/INFO export format and integrated into ARC/INFO.

9. Data Capture and Integration in ARC/INFO

9.1 Introduction

Like most other information, geographical data normally have little intrinsic value. Their value lies in their relevance to the applications to which they are applied and this value can be greatly increased by linking together multiple data sets (Raper and Rhind, 1992). One of the most important characteristics of GIS is their ability to integrate data from various sources, yet as Martin and Bracken (1993) cite, the divide between human and physical applications remains as noticeable as ever amongst GIS users. It remains problematic to integrate human and physical data due primarily, to the difficulty of geo-referencing and plotting the spatial distributions of socio-economic information with other types of data. These difficulties arise from a fundamental difference between data for 'natural' and 'imposed' areal units (Unwin, 1981) as the two types of unit rarely share any common boundaries.

The first stage in the spatial analysis of the radon, land capability and social data sets was some form of integration. The use of overlays within a GIS does not detract from the MAUP (outlined in Chapter 4), but enables datasets collected for various different types of spatial zoning to be studied in relation to one another. The initial obstacle in this project was the fact that the radon data were aggregated for grid squares, the census data for wards and districts and the land capability for naturally based capability areas identified solely for the purpose of the MAFF maps. These three types of data are difficult to integrate within a single spatial referencing system without advanced knowledge of computer programming. Figure 8.1 shows the three overlays that were used. Two approaches to handling a comparison between variables defined for incompatible sets of areal units are outlined in a research report by Flowerdew and Openshaw (1987). The first approach involves aggregating the units up to areas that are the same for both sets of data and the second approach involves breaking the units down into smaller ones. Neither of these methods can work within the context of this study, as the raw data are unavailable and aggregation or disaggregation of the data is not a viable option. The data were thus overlaid using their original areal units in spite of their spatial incompatibility. The conclusions drawn from the overlay results take into account the possible effects of the modifiable areal unit problem and the problem of the ecological fallacy.

9.2 Radon Gas Data

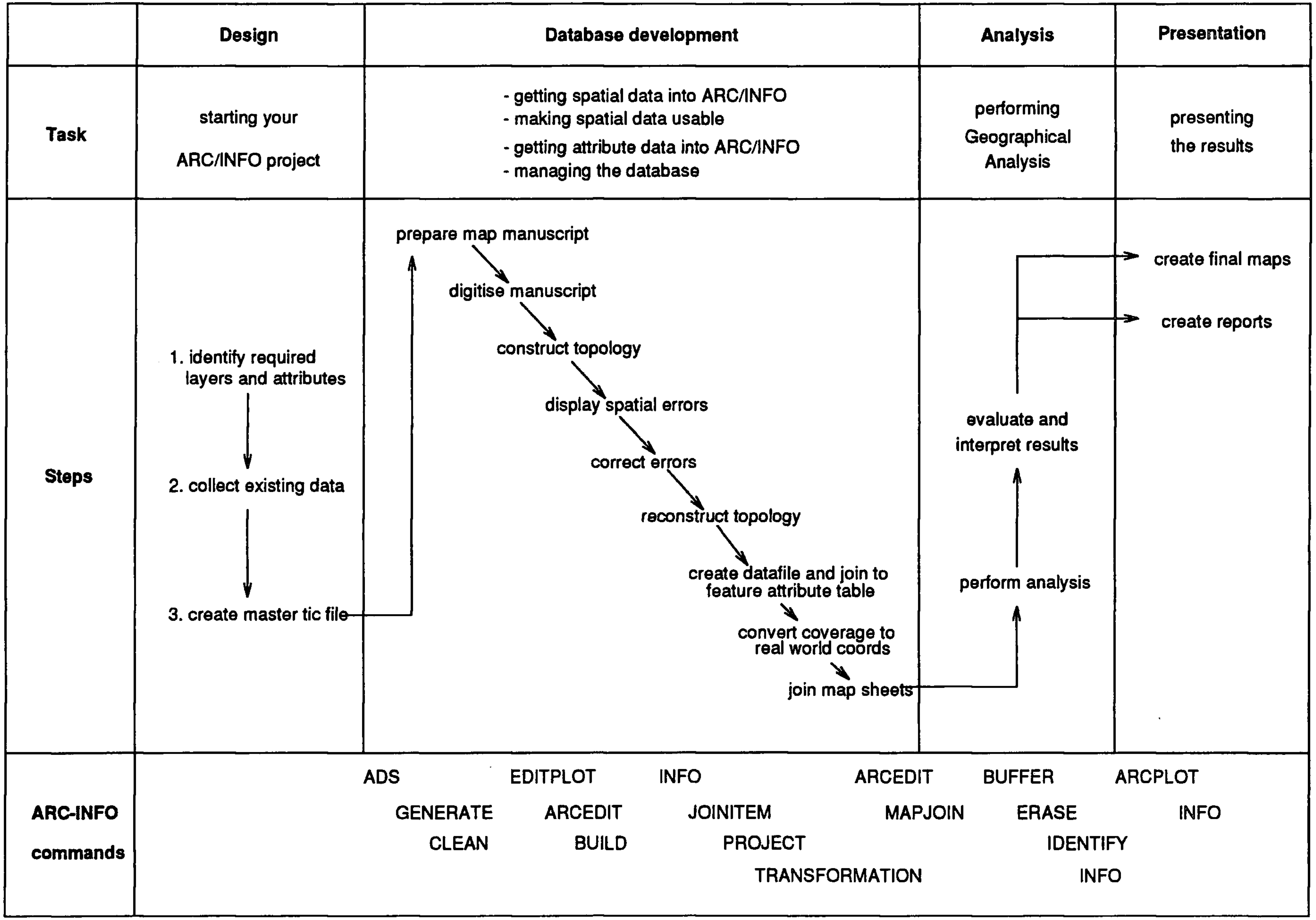
The data for indoor radon gas levels for the five counties were provided by the NRPB on floppy disc as the percentage of homes above the Action Level (200 Bq m^{-3}) for every five kilometre grid square in each of the counties. The arrangement of the data into regular grid squares facilitated the creation of a rectangular grid in ARC/INFO for each of the coverages (one coverage for the counties of Somerset, Derbyshire and Northamptonshire and one coverage for Devon and Cornwall combined). The grids were created using Ordnance Survey National Grid coordinates which also acted as a referencing system to enable the data layers to be overlain. Figure 9.1 shows the methods used within ARC/INFO to design, develop, analyse and present spatially referenced data. Each five kilometre grid square was then assigned an identification number within ARC/INFO and a polygon attribute table (PAT) was created. ARC/INFO stores the attribute (non-coordinate) information on each polygon by way of geo-referenced tables of statistical and thematic data within the INFO relational database system. A polygon attribute table has a *.PAT* extension and contains information that describes a specific coverage. The average indoor radon levels for all of the five kilometre grid squares were entered into the PAT of each of the five counties. The data were checked by creating a map of the radon coverage in *Arcplot* (the plotting facility within ARC/INFO) and comparing it to the printed maps in Miles *et al.* (1990 and 1992). Any mistakes were then corrected in the *Tables* option within ARC.

9.3 Land Capability Data

The first obstacle to the integration of the land capability data in ARC/INFO was the conversion of the information on the paper map into digital format. The three methodologies currently in use for data conversion (other than remote sensing) are manual digitising (using a digitising tablet and puck), semi-automatic line tracing and automatic scanning. The land capability maps are characterised by a clutter of cosmetic information typical of geographical categorical coverages. In addition to detailed land capability mapping, all main roads are also shown as well as towns and villages, rivers, administrative boundaries, historical monuments and the like. For this reason the data were captured manually in order to pick out the most important aspects of the coverage and leave out the excess clutter.

The data were captured with Lites2 software, using a cursor on a flat digitising table that has a two-axis digital recording device. The maps were digitised at the scale of 1:250,000 from conventional paper maps purchased from MAFF. Lines (*e.g.* the coastline of Cornwall, Devon and Somerset) and areas (land capability zones) were outlined by hand with the cursor and saved on disk as a file of coordinates. The data were then converted into ARC/INFO export format and loaded into ARC/INFO. In a vector GIS such as ARC/INFO, a line is defined by a series of pairs of x and y coordinates and an area consists of a series of pairs with the same start and end points (ARC/INFO is only able to store data as points). Information from the relational database can be attached to any line or area.

Variations in the land capability classes on the 1:250,000 maps occurred over localised areas down to approximately 250 metres and minor variations such as these were not digitised. The bulk of the digitising concentrated on areas that were homogenous over an area of at least



	Design	Database development	Analysis	Presentation
Task	starting your ARC/INFO project	- getting spatial data into ARC/INFO - making spatial data usable - getting attribute data into ARC/INFO - managing the database	performing Geographical Analysis	presenting the results
Steps	1. identify required layers and attributes 2. collect existing data 3. create master tic file	prepare map manuscript digitise manuscript construct topology display spatial errors correct errors reconstruct topology create datafile and join to feature attribute table convert coverage to real world coords join map sheets	perform analysis evaluate and interpret results	create final maps create reports
ARC-INFO commands		ADS GENERATE CLEAN EDITPLOT ARCEDIT BUILD INFO JOINITEM PROJECT TRANSFORMATION ARCEDIT MAPJOIN	BUFFER ERASE IDENTIFY INFO	ARCPLOT INFO

(Source: ESRI,1987)

Figure 9.1 ARC/INFO methodology

three square kilometres, as areas smaller than these were not deemed to play a significant role in the geographical distribution of radon Affected Areas at the local - ward level - or above. The largest homogenous zones in Devon and Cornwall were those areas classified as Grade 5 - areas of poor agricultural use. These corresponded to the upland regions of Dartmoor, Bodmin Moor, St. Austell and Carmenellis. In total, Grade 5 land covered approximately 30% of Cornwall and Devon. In Northamptonshire on the other hand, there was no significant area of Grade 4 or Grade 5 land and the majority of the county was classified as Grade 3 (good to moderate agricultural land). Land capability in Somerset is similarly dominated by Grade 3 land. Derbyshire is characterised by a more even distribution of land capability classes. Cornwall and Devon were treated as one coverage (and therefore as one area) to prevent the deceptive 'cut-off' of land use zones at the county boundary and to correspond to the radon coverage, which was similarly combined.

The first task that was performed on the data after the digitising process was complete, was *eliminate*. This command removes any sliver polygons (spurious polygons which are small in area and occur when two digitised arcs do not follow exactly the same route) that were created during the digitising process. *Transform*, *clean* and *build* tasks were then carried out and a PAT file was subsequently created for each of the coverages. All the polygons in a coverage are sequentially numbered from one by the command *createlabels*. This user-identification number acts as a link between the polygons in the coverage and their associated attribute data. The land classification grade for each polygon (numbered from one to five) was assigned to the relevant identification number in the *Tables* option. The accuracy of the data input was checked by drawing a map in *Arcplot* and comparing the result with the original MAFF land capability maps. The completed digital map was stored as a coverage.

9.4 Census Data

There are few references in the literature citing research findings on the subject of integration of social data with physical data. This is probably due primarily to the difficulty of adequately geo-referencing and representing the spatial distributions of census survey data (Martin and Bracken, 1993). Census data are a secondary data source and are collected and aggregated by OPCS prior to being disseminated. Spatial referencing of census information is invariably indirect, with locations given as (for example) enumeration districts, wards, districts and counties. The census data used in the thesis were taken from the 1991 Census Ward and Civil Parish Monitor for Northamptonshire (the only county that was analysed), published by OPCS in 1993. The three aspects of the social environment that were studied were: population density at ward and district level, the proportion of households containing only pensioners at ward level and social class at district level. The volume of data involved, especially social class data, necessitated that only one county was studied and Northamptonshire was chosen, in order to offer some insight into its unforeseen high radon levels.

The first stage of the data preparation was to isolate the Northamptonshire ward and district boundaries from the Bartholomew's 1:1,000,000 digital database, available via the CHEST agreement on the networked Sun workstations. The Bartholomew's data comprise 1981 ward boundaries for England and Wales - the 1991 CHEST digital data, to be disseminated from the Manchester University Computing Centre, were not available at the time when the analyses were being carried out. Thus 1981 ward boundaries were used with 1991 Census data and some inevitable errors arose from changes between the administrative boundaries used in the two censuses. Three areas in Northamptonshire were affected by changes to ward boundaries between 1981 and 1991 (see Chapter 10.3.5). The location of these wards (approximately 50 km^2 in total area) is shown in Figure 10.2. It was necessary to delete some 1981 boundaries to make them more compatible with the 1991 boundary and the census data for these areas was calculated by averaging the figure for the two adjoining wards. In total there are 148 wards and seven districts (none of which were affected by boundary changes between 1981 and 1991) in Northamptonshire.

Once the excess information had been deleted, the *clean* and *build* commands were carried out and a PAT was created. The census data were tapped into the Northamptonshire PAT via the *Tables* option. The raw data were first grouped into one of four or five categories, classified according to the range of values for each data set. The grouping was necessary to enable the data to be mapped and to enable comparisons with the radon data. Once the database was established, each dataset was mapped in *Arcplot* and checked for inaccuracies.

The overlay between the ward and district data and the five kilometre grid squares containing the radon information, was carried out using the *identity* option within ARC (see Figure 9.1). Statistics (number of polygons, total area and mean area) for the ensuing polygons created by the overlay procedure are recorded in Appendices 5 and 6 and the conclusions drawn from the results are discussed in Chapter 12.4.5.

10. Error Measurement and Estimation

10.1 Introduction

The widespread availability of GIS has enabled many users to integrate geographic information from a wide range of sources. An inevitable consequence of data integration has been the sober realisation that geographic data are not of equal quality and that they contain errors and uncertainty that need to be recognised and properly dealt with (Maffini *et al.*, 1989). The term *error* is defined in the Oxford English Dictionary as the "amount of deviation from a correct or accurate result". Errors can arise during data collection, manipulation and presentation and if undetected, can lead to misleading and even incorrect results. The body of literature devoted to describing the various aspects of spatial error is growing and Burrough (1986) and Goodchild and Gopal (1989) for example, provide detailed information on the important considerations in this field.

Environmental and social analyses are currently hampered by the necessarily multidisciplinary approach and the lack of pervasive quality standards within the field of GIS. It is in part the mix of problems surrounding spatially referenced data that gives the field of spatial data analysis its distinctive quality and as Haining (1990) points out, it should be noted that the treatments for many of these problems are still in their infancy. It is necessary to assess sources of intrinsic and propagated error within a GIS as computers are very precise machines and errors and uncertainties in data can easily be transformed leading to inaccurate results and poor decision making as a result. Although all conventional mapping agencies have published map accuracy standards they do not pervade computer cartography. This chapter aims to discuss some of the problems inherent in using GIS for geographical analysis and outlines some possible sources of error in the data used in this study.

10.2 Data Quality and Sources of Error

Accuracy standards for digital cartographic data have only recently begun to be introduced; in June 1992 a three-part British Standard was published which details the transfer of geographic data on magnetic media. The three parts are entitled '*Electronic transfer of geographic information*', '*Spatial datasets for geographic referencing*' and '*Specification for a Street Gazetteer*' (Rowley, 1993). The main catalyst in the development of accuracy standards has been the spread of GIS into academia, local government and industry over the last ten years. The rapid rise in the volume of data available to the GIS user has grown, along with the ability and desire to automate various analytical tasks. The merging of disparate databases has highlighted the need for carefully collected and maintained databases. Dahlberg (1986) states that 'a principle deficiency of land information systems [and GIS] in their present state ... is that they fail to include information ... about the quality of the data'. The lack of information about the quality of digitised outlines and geographical databases has been identified as a key problem in recent years (Maguire, 1989). Chrisman (1983) suggests that information about the quality of feature coding, digitiser reliability, positional accuracy and the age of source material are all important aspects of data quality. Trying to define 'quality' in the context of spatial data however, is not an easy task. Ralphs (1993) suggests that a working definition of data quality could be "the fitness of a dataset for use in a particular set of circumstances". There are problems with this definition however, namely that a range of standards could develop according to the requirements of the individuals or organisations collecting the data. Data collection organisations have often failed to liaise either with their users or with other data organisations, resulting in the wide variety of standards, collection units and classifications currently being used. This situation is gradually changing as awareness of error propagation increases. The Chorley Report (DoE, 1987) advocated greater dialogue between data users and data collectors and promoted the creation of bodies such as the Association for Geographic Information (AGI), which encourages greater participation between the different groups using spatial data and provides a forum for discussion. The report also called for the development of a statistical methodology relevant to spatial data and in particular for user friendly software for processing large spatially referenced data sets. A set of standardised 'rules' on spatial data collection and maintenance have yet to be established. These are essential if data are to have a longer 'half-life' and be more applicable to a wider section of the user community. To this end, Ralphs (1993) points out that what constitutes a 'high quality' data set for one user may be entirely inappropriate for another. Burrough (1990) suggests that database managers and experts could greatly assist GIS users to achieve quality results by formalising information on the following:

- Data collection, level of resolution and quality
- The use of basic analytical functions of the GIS
- Data requirements, sensitivity and error propagation

The quality of data however, can be low on the list of priorities of some data collectors, especially if resources such as time and money are major constraints. Therefore, although the definition given by Ralphs (1993) allows room for individual users requirements and changes in requirements over time, it also permits the more dangerous possibility of neglecting standards of data quality altogether.

In Europe, the European Statistical Office appears to be strengthening its efforts to develop an EU statistical system within the framework programme (1993 to 1997) for priority actions, in the field of statistical information (Lievesley and Masser, 1994). The Statistical Office aims to implement a system of standards, methods and organisational structures which will be capable of producing comparable, reliable and relevant statistics throughout the EU. In addition, it aims to increase data dissemination to aid those concerned with economic and social matters (Lievesley and Masser, 1994). It seems likely therefore, that European data in the future will be more accurate, comparable and hopefully, more freely available, than at present. The development of standards at the European level may also have a knock-on effect within member states of the EU, thus encouraging a more widespread adoption of data standards.

In *Reflections on Exile* Henry St. John said, "truth lies within a little and certain compass, but error is immense" (Webster and Oliver, 1990). Errors can arise in a number of ways, some of which may not be immediately obvious and some of which are more blatant. It is highly problematic to quantify the error in the input data used in this project without full access to the raw data and to information on the methods used to sample, collect and maintain the data sets. It is possible however, to identify the main routes through which error could be introduced and to discuss the possible effects of this error on data transformations at later stages of analysis. To this end, it is useful to divide the possible sources of error into two parts; data error and operational error. Carver (1991) goes one step further and sub-divides the two broad categories to give a more meaningful framework for error analysis. For this purpose a typology of errors is given in Table 10.1 below.

Table 10.1 A Typology of Errors

Obvious sources of error:	• age of data • areal coverage • map scale • density of observations
Errors resulting from natural variations or from ordinal measurements:	• positional accuracy • attribute uncertainty • generalisation
Errors arising through processing:	• numerical computing errors • faulty topological analyses • errors in interpolation

(Source: Carver, 1991 after Burrough, 1986)

10.3 Data Error

GIS are only as good as the raw materials used and the precision and quality of spatial data entered into a GIS will profoundly affect any analytical processes that are carried out. Newcomer and Szajgin (1984) suggest that the highest accuracy of a GIS product can only be that of the least accurate data layer used in the analysis. Data error can be introduced in a number of ways, either as a result of transformations carried out by the organisation or individual in charge of collecting the data, from natural variations, or simply because the data are out-of-date. A source map may contain mapping errors as well as those arising from expressing a curved surface on a flat piece of paper and shrinkage and distortion effects associated with the paper itself.

It is well known that accuracy, like (and partly because of) resolution, has to be paid for. For most surveys, the increase in costs rise sharply for any unit increase in accuracy above quite mediocre levels. Partial re-survey is often carried out to check accuracy, but in practice this only checks consistency. Office checks on the data, check only the feasibility (Rhind and Hudson, 1980). Exact levels of accuracy are often impossible to measure because of the lack of perfect reference data (Fuller *et al.*, 1994).

A restriction on the resources available, for example funding, can create very real restraints on the nature of the survey, the range of data that are collected and the speed at which they are gathered and released. The scope of the 1991 Census was partially determined by the available funding and the decision not to make data available on a postcode basis was taken for purely economic reasons (Martin, 1993; Wrigley, 1990).

10.3.1 Age of Data

The date of a survey is important in that the data collected relate to a specific moment in time, which cannot be recreated. A dichotomy exists in that both social and natural environments are constantly changing but most national surveys (at least in the United Kingdom) are carried out periodically, often at ten-yearly intervals. For the majority of the time that it is available therefore, a large proportion of official data are out-of-date. The example of the 1991 Census is a case in point. The Census was carried out in April 1991 but the first 100% data were not published until June 1992 (national coverage of these data was complete by the end of 1992) and the 10% data dissemination began in the spring of 1993. Data were still being released, for the first time, in early 1994. In addition to the inconvenience of the late availability of data, the time lag between each full census in the United Kingdom is ten years (the next census will be carried out in 2001). Such a gap in the release of important social data can have far-reaching implications for all those using census data, especially those using it in the late 1990s, when many of the statistics will no longer be accurate. In some other European countries (Denmark and Sweden, for example) the periodic census has been replaced by continual registration systems and this effectively reduces the error caused by antiquated data. According to Martin (1993), OPCS assessed the possibility of a 1996 census, but in the light of users' responses and the need to limit public expenditure, the government decided to abandon any plans for a five-year census.

The time-lag in the release of the 1991 digital ward boundary data for England and Wales is of particular interest - it was not released by the Census Dissemination Unit at Manchester University until July 1994 (see Chapter 4). The redesign of statutory boundaries during the

1980s resulted in extensive changes to lower level census areas. Thus all census users using data from the 1991 Census between 1991 and mid-1994 would have had three options open to them: digitise the ward and / or enumeration district boundaries themselves (a lengthy process and one prone to errors); buy the digital data from a private GIS company (a very expensive solution); or use 1981 digital boundaries, with the errors that this entails (see Chapter 4). Often the only practical solution is to use 1981 boundaries and to suffer the problems encountered when using 1991 data with ward boundaries that have been superseded. The errors introduced by this action are covered in section 10.3.5 below. In some areas of the country, statutory boundary changes will already have made the areas represented in the census output obsolete by the time the data were published, and imminent reorganisation of local government is likely to cause further difficulties for the application of census statistics to statutory areas.

The land capability maps are the least accurate coverage used in this thesis, in terms of data age. The data were collected by MAFF in the late 1970s and the early 1980s and since the maps were researched there have been changes in land use (*e.g.* urban encroachment onto farming land), farming techniques (the 1980s saw an increase in the amount and type of fertilisers used on the land) and planning regulations (such as set-aside which allows the land to revert to a semi-natural state). These mechanisms of change will have had an impact on the classification of the land in the five counties, although in some parts of the country it will be more noticeable and wide-ranging than in others. The most likely change to the data is an increase in the area covered by the 'urban' class. According to the MAFF Farm Land Use Survey (1945 to 1990) the total area in England under farm management declined from 78.0% in 1970 to 77.4% in 1980 and to 76.5% in 1990 (Sinclair, 1992). Table 10.2 below shows the transfer of agricultural land to other uses between 1976 and 1986 (latest available figures). The intrinsic error inherited from the MAFF land capability survey will therefore have been translocated through the various stages of data preparation, integration and analysis. Whilst it must be remembered when interpreting the results of this study that the data are now somewhat out-of-date, the nature of the land capability classification ensures that the data survives changes over time better than analyses using tradition land use or land cover classifications. This is due to the fact that the classification relates to the *capability* of the land for agriculture (in terms of the location's climate, site and soil) and these factors are not susceptible to temporary or short-term change.

Table 10.2 Net Transfer of Land to and from Agriculture in England: 1976 to 1986 - Five Year Moving Average ('000 hectares)

Five years ending	Urban, Industrial & Recreational	Government	Forestry	Unreported	TOTAL
1976	-13.8	-0.1	-0.7	-12.3	-26.9
1977	-12.9	-0.1	-0.5	-11.6	-25.0
1978	-10.5	0.0	0.0	-12.2	-22.8
1979	-9.7	0.2	-0.1	-11.6	-21.2
1980	-8.4	0.1	0.1	-10.3	-18.5
1981	-7.1	0.1	0.0	-9.2	-16.3
1982	-5.9	0.2	0.1	-10.0	-15.6
1983	-5.4	0.2	0.4	-9.3	-14.1
1984	-4.4	0.3	0.6	-9.7	-13.2
1985	-4.4	0.3	1.0	-10.5	-13.5
1986	-4.3	0.4	1.3	-9.4	-12.1

(Source: Sinclair, 1992)

Errors in the radon data related to age have not been quantified in any published work. However, the measurements used to arrive at the estimated proportion of homes above the Action Level for each five kilometre grid square, are continuously being undertaken and the estimates are continuously revised. Thus changes to the grid square estimates are likely to occur in time as the data are refined. These changes are unlikely to have a significant impact on the results of the work for this thesis. In general radon levels in the home tend to remain constant over time (indoor radon measurements are estimated for an average annual reading), except where changes to the nature of the building occur. For example, results from the NRPB National Survey (Wrixon *et al.*, 1988) found that the presence of either double glazing or draught-proofing around doors and windows typically increases the radon concentration by about 30%. Once again, however, the likely impact of these changes on the overall results is slight.

The effects of old age on the quality of data in general, is difficult to quantify, although Rhind (1983) estimates that census data halves in value every two to three years. If all the data sets were collected over a short time span, for instance in the early 1980s, the effect of age would be minimised as the data would all relate to a short period of time. However, the census data and the radon data were collected approximately ten years later than the land capability data, with all the social and environmental implications that this entails.

10.3.2 Areal Coverage

A major barrier to linking different data is the use of definitions of area which do not match the various data sources. The choice of an areal framework is a function of the spatial scale of the problem, the objectives of the study and the specific attributes of the study (Visvalingham, 1988). As Burrough (1986) states: "Many ... geographers know from field experience that carefully drawn boundaries ... on maps are elegant misrepresentations of changes that are often gradual, vague or fuzzy". Mapped boundaries usually relate to human perception or to administrative objectives rather than to natural changes, although abrupt boundaries do sometimes occur in nature. Methods are needed that can detect and deal with the problems associated with the spatial organisation of the data, such as spatial outliers, uneven spatial coverage and the intrusive effects of an areal partition (Haining, 1990). For a discussion on the current areas of research see Chapter 4.

All of the data sets used in this study contain some measure of error inherited from the original method of data collection and aggregation. The four areal frameworks (ward, district, five kilometre grid square and land capability zone) that have been used to map the data, present problems during data analysis and transformation. The overlay procedure presumes that the data are evenly distributed throughout the areal unit. However, this is often not the case; population for example, is discrete. As areas become larger the level of generalisation increases and, since the areal units are normally irregularly shaped, the effect of the areal partition is directionally variable and very difficult to quantify (Langford and Unwin, 1994). Openshaw and Taylor (1982) effectively demonstrate the sensitivity of spatial statistics to zonal definition by calculating the variation in the correlation coefficients for the distribution of voters in Iowa, USA for differing numbers of zones. The different aggregations of the data for zones ranging in number from 99 to six, yielded correlation coefficients as diverse as -0.99 to +0.99. A comparison of the results of overlays in ARC/INFO between radon levels and population density at both ward and district level, is given in Chapter 12.4.4.1.

The presence of the MAUP and the effect it may have on the results of spatial analyses are often ignored because it is deemed either insoluble (the lack of any relevant spatial analysis techniques to study the effects of the MAUP means that at present error from the MAUP is very difficult to quantity) or of trivial importance, or because to acknowledge its existence would cast doubts on the applicability of nearly all applications of quantitative techniques to zonal data. Openshaw and Taylor (1981) agree with the latter but find it unacceptable to ignore or trivialise the importance of the effects of the MAUP. Where it is necessary that the data be stored as aggregate data, for example for reasons of confidentiality as in census data, the propagation of error can be kept to a minimum by ensuring that individual relationships are not inferred from aggregated data. The presence of the MAUP and the fact that the data are aggregated (and modified) prior to dissemination is one of the most damaging features of census data (and other data sources) in terms of their use and applicability for research - a view that is echoed by the academic community at large. The Chorley Report (DoE, 1987) looked into many aspects of geo-referenced information but highlighted the importance of spatially referenced information as an area for GIS research in the future. Ideally, areas should be small and uniform (with respect to attributes) to enable flexible aggregation of areal units and to facilitate the merging of data from different sources. However, in practice individual census returns are aggregated for unequal size units. In England and Wales, administrative areas were not designed for presenting census material, they are just the handy framework of local government. Parishes alone for example, in England (1981 Census) range in size from nine hectares to 25,500 hectares (Dewdney, 1985). Evans (1981 pp. 55) suggests that "the aggregation effect ... is a much less important source of variance in correlations when the zones are grid squares than when they are irregular". With respect to the spatial organisation of the radon data therefore, it would appear that the imposition of regular grid squares introduces less error than would be expected if the data were aggregated for irregular units. However, the process of aggregating data will, in itself, obscure individual relationships and generalise the data.

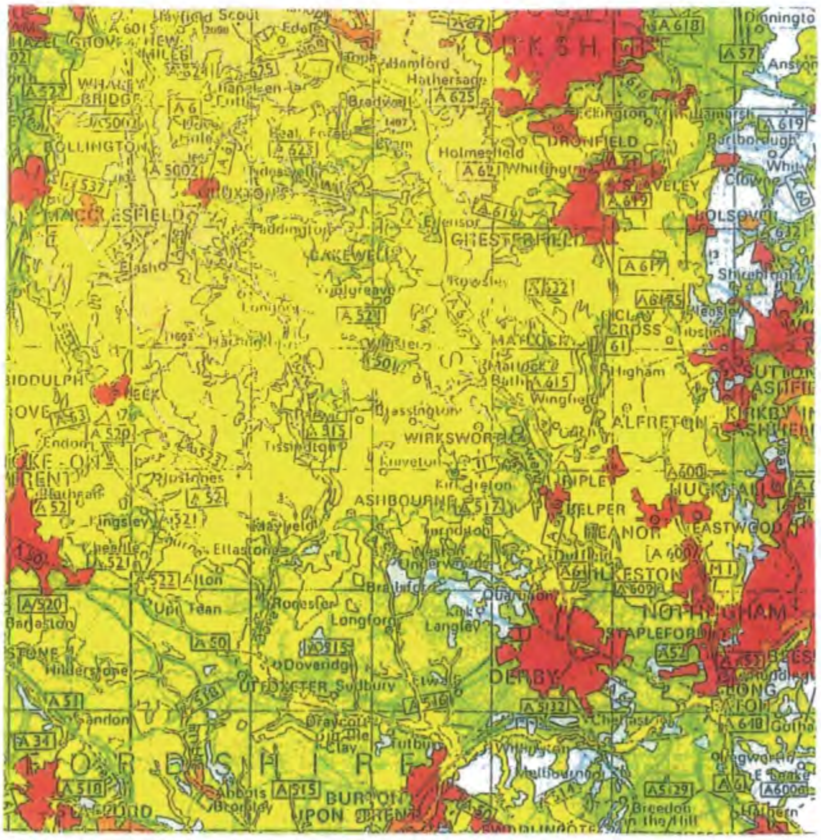
A study based on aggregate data is also liable to the ecological fallacy problem (see Chapter 4), which occurs because areal studies cannot distinguish between spatial associations created by the aggregated data and real associations possessed by individual data (Openshaw, 1984). In this study, inferences will not be made about individuals from the results based on aggregated data, for the reasons given above and any statistical correlations between the three main data sets will be analysed in terms of the areal unit at which they were observed. In this way it is hoped that the propagation of error will be kept to a minimum.

Questions about the applicability of the results are also raised as a consequence of the areal extent of the Affected counties and the scale at which analysis has been carried out. Only the five counties identified as Affected Areas by the NRPB were studied, in terms of the association between radon and land capability, and the results identified in Chapter 12 relate solely to these areas. It is problematic to extrapolate the results for the Affected Areas to the rest of the United Kingdom. Local geographical differences between areas may influence relationships elsewhere. Similarly, the impact of social factors on radon levels in Northamptonshire only provides results for this county. Relationships identified to occur in Northamptonshire, will not necessarily occur in other parts of the country. For these reasons, no extrapolations for other areas have been made.

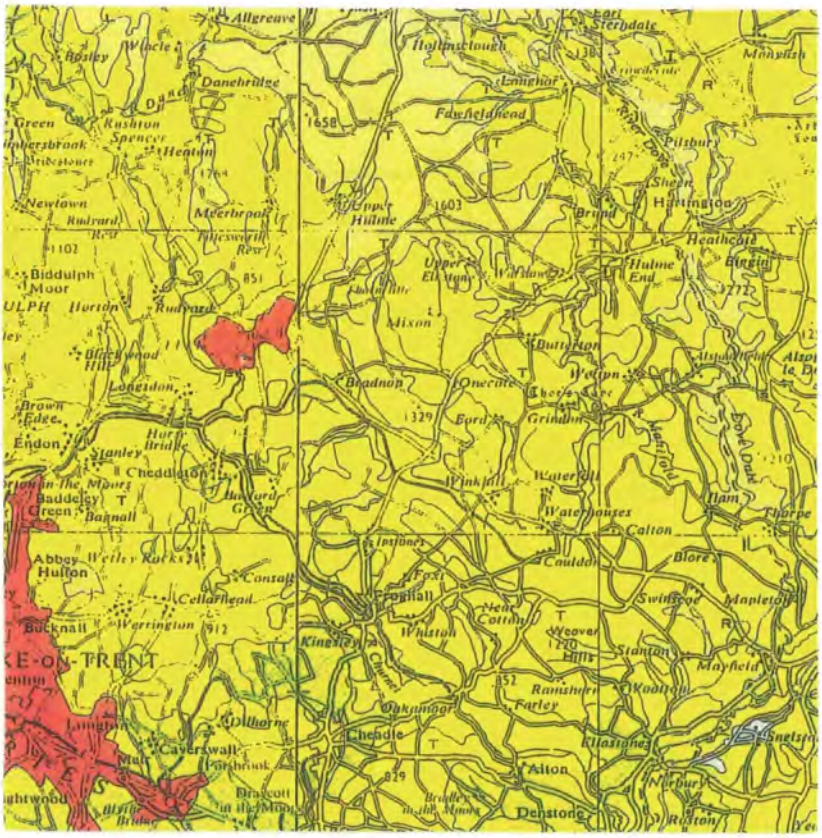
10.3.3 Map Scale

The locations of lines and areal boundaries on a map are often inaccurate because of scale effects and in some cases because of the map projections used to display the initial data. Different scales of study will produce a different spatial picture of the data which will vary not only in detail but also in the extent to which they hint at relationships between variables. For

Figure 10.1 A comparison of two map scales: Land capability in Derbyshire



Land Capability Classification, England and Wales 1:625,000 (MAFF, 1979)



Land Capability Classification, Midlands and Western Region 1:250,000 (MAFF, 1983) example, the classification and delimitation of land capability classes is a process heavily dependant for its final outcome on the scale of the data. Figure 10.1 shows two extracts of the MAFF land capability maps, at the scales of 1:625,000 and 1:250,000 for an area in Derbyshire.

Areal boundaries and line features are stored in GIS in terms of points which are joined by straight lines. This introduces inaccuracy which is a function of the complexity of the line and the number of points used to represent it. No current GIS carries the scale of the source document as an attribute of the dataset and few even adjust tolerances when the scales change. Most vector systems perform operations such as line intersection or overlay at the full precision of the coordinates without attention to their accuracy, consequently, the resulting inaccuracy (when checked) is often a surprise (Goodchild and Gopal, 1989).

The scale of cartographic output should be related to the nature of the data to be mapped; data at one scale may not be visible at another. The resolution of a data set defines the smallest object or feature which is included or is discernible in the data. In theory, at a scale of 1:250,000 the effective resolution is 125 metres on the ground, but in practice spatial variations will have been smoothed for ease of cartographical representation to give a more generalised view, resulting in a lower ground resolution. The land capability survey is reliable for land units of 80 hectares or more (Mather, 1986) and the capability maps were published at 1:250,000. In order to give a clear representation of variations in land capability at this scale it was necessary for MAFF to smooth any small scale variation. Thus the source maps were already simplified before the digitisation of the maps was carried out. The final maps produced in ARC/INFO are at varying scales between 1:450,000 and 1:1,000,000 (see Chapter 12).

The spatial units that were used to analyse the social data (the ward and district) are not the smallest available but they provide a compromise between the excessive detail of enumeration-level data and the generalisation of county-level data. The degree to which the level of aggregation affects the results obtained within a GIS project has yet to be quantified, but for the purposes of this project and in relation to the detail provided in the radon and land capability databases, the ward and district provide an adequate spatial scale at which to study the social variation within Northamptonshire.

The choice of five kilometre grid squares for the mapping of radon is a function of the need to encompass a significant number of measurements at the same time maintaining a level of detail. To paraphrase Evans (1981 pp. 55) who looked at the relationship between the size of grid square and 100% population data, it appears that while areal scale is an important consideration in geographic studies concerned with correlated data, differences between similar scales (for example, one and two kilometre grid squares or between 16 and 32 kilometres) are fairly small.

The representation of data at scales between 1:450,000 and 1:1,000,000 will be restricted to polygons above a certain area. Thus any small polygons formed in the overlay operations and their associated data, are not visible in the cartographic representation of the counties. The maps are generalisations of reality and are therefore unable to show all the intricacies of the natural environment. The small polygons were included in the analysis and interpretation stage of the project however, and the inclusion of error from this source was therefore kept to a minimum. The use of more detailed (or indeed individual level data) decreases the level of error but at the same time increases complexity, costs and computing time, so that a compromise has to be found to suit the particular needs of the study. In this case, in order to asses the geographical differences and similarities of five English counties, sacrifices were made on the scale and detail of the data portrayed on the maps, and emphasis was placed on the differences and similarities *between* the counties.

10.3.4 Density of Observations

The number of homes that were measured for radon in each of the five kilometre grid squares varies; in general the more densely populated the square the higher the number of measurements and the total may vary from only a few to several tens of measurements (Green *et al.*, 1989). Inevitably, average radon levels and figures for the estimated proportions of homes above the Action Level have less significance in areas where fewer radon measurements were taken, as sampling error is reduced by increasing the size of the sample. Error will be present in the data received from the NRPB (in early 1993) due to the small total number of homes that had been measured at that time for radon, especially in Derbyshire and Northamptonshire. New measurements are continuously being added to the NRPB radon database and analyses using these data at a later date are likely to produce slightly differing results, as a result of revisions to the database. However, as discussed above, the impact of these data refinements on the conclusions drawn in this thesis, are likely to be insignificant.

The density of observations is not relevant when discussing census data as it is the aim of the census to count every person in the country (although in practice this is often not the case - see Chapter 4). However, some questions in the census are analysed for a sample of the total population. The method used for collating data on social-class is to take a systematic sample of 10% of the total population and key in the results with the view that they are representative of the population as a whole. The problems inherent in using data which have been sampled in this way are discussed in Chapter 4.

An additional and possibly serious data problem, is that of incompleteness, as in the suppression of some census data for reasons of confidentiality. Census data at enumeration district level are suppressed when the number of homes or individuals does not reach the required minimum threshold. The effect of such error on this project is minimal however, as only ward level data or above were used. Another reason for incompleteness in the census data are incorrect (or illegible) answers on the census questionnaire forms. The vagaries of human nature dictate that some questions will be misinterpreted or misunderstood by the general public and will be incorrectly answered. In addition, people may leave some answers blank out of principle, in spite of the fact that it is necessary by law (under the Census Act of 1920) to complete the form correctly. Knowledge of the census is very weak in Britain. In two polls undertaken in January and March 1991 by Gallup, less than half of those questioned knew that the census takes place every ten years and only 14% were aware that the census was one month away. However, three quarters of people in Britain knew that participation in the census was compulsory (Marsh, 1993).

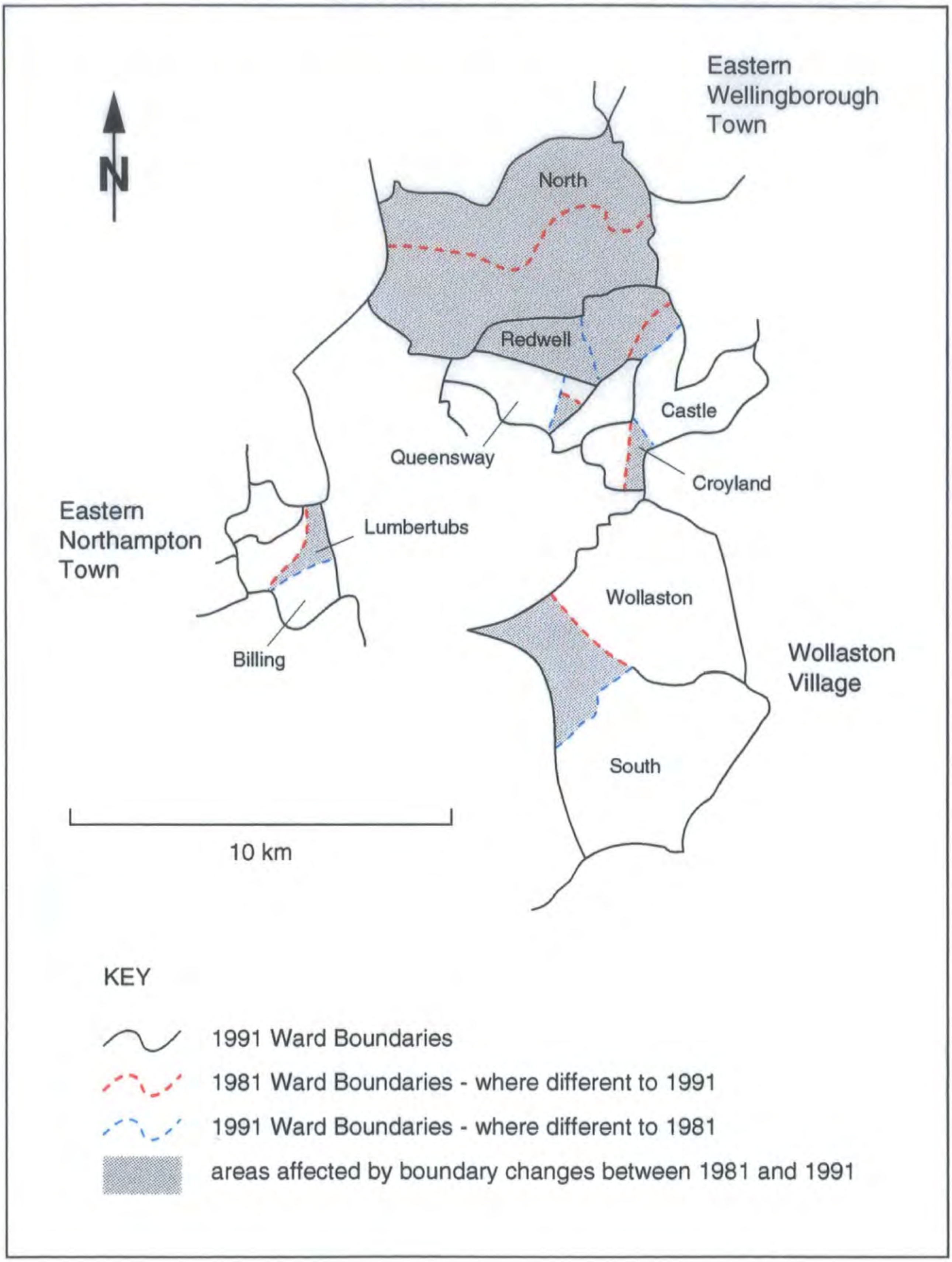
10.3.5 Positional Accuracy, Attribute Uncertainty and Generalisation

Methods of data collection and organisation are fundamental to quantitative geography - all data have a limited accuracy, arising from human error and instrument error. There are few studies that attempt to document the seriousness of this problem but Hampel *et al.* (1986), who briefly review the problem of measurement error, suggest that as a matter of routine between 1% and 10% of values will contain gross errors; that is, occasional but powerful errors that in addition to the slight distortions (due to rounding, scale *etc.*) that naturally occur (Haining, 1990).

An important source of error in the analysis of ward level social data and radon levels in Northamptonshire (see Chapter 12.4.4) is that introduced by the use of 1981 Census boundaries with 1991 Census data. The administrative boundary changes affected three regions of the county, namely the towns of Northampton and Wellingborough, where administrative divisions tend to be much smaller in area, and Wollaston, a village in the east of the county (see Figure 10.2). No district boundaries were altered between 1981 and 1991. Figure 10.2 shows the location of the 1981 wards as used in the social overlay in this thesis, together with the 1991 ward boundaries which define the areal extent of the 1991 Census data that were used. The total affected area is approximately 50 km^2. Where changes have occurred to the ward boundaries, they have either been ignored where the difference is small (as in Northampton, Wollaston and in part of Wellingborough) or figures for the two 1991 wards have been averaged to give a value for the old 1981 ward (as in the ward of Redwell in Wellingborough). The problem of boundary changes is not restricted to the 1981 Census. Indeed since the 1991 Census was carried out (*i.e.* in just three years) more than 12,900 people have been affected by local authority areas involved in boundary changes in 24 separate areas (OPCS, 1994).

A further source of error stems from the nature of measurement in geography (Maffini *et al.*, 1989). Any measurements that are acquired with instruments inevitably introduce error. The capability of the person using the measuring device can adversely affect the error margin, as can the scale at which measurements are made and the frequency of sampling. The decision by the NRPB to leave radon detectors in the home for three months and in two rooms (the living room and the bedroom), reduce the possible effect of short-term atmospheric conditions which may influence the amount of radon in the indoor air. Thus the annual radon dose in a

Figure 10.2 Changes to ward boundaries between 1981 and 1991 in Northamptonshire



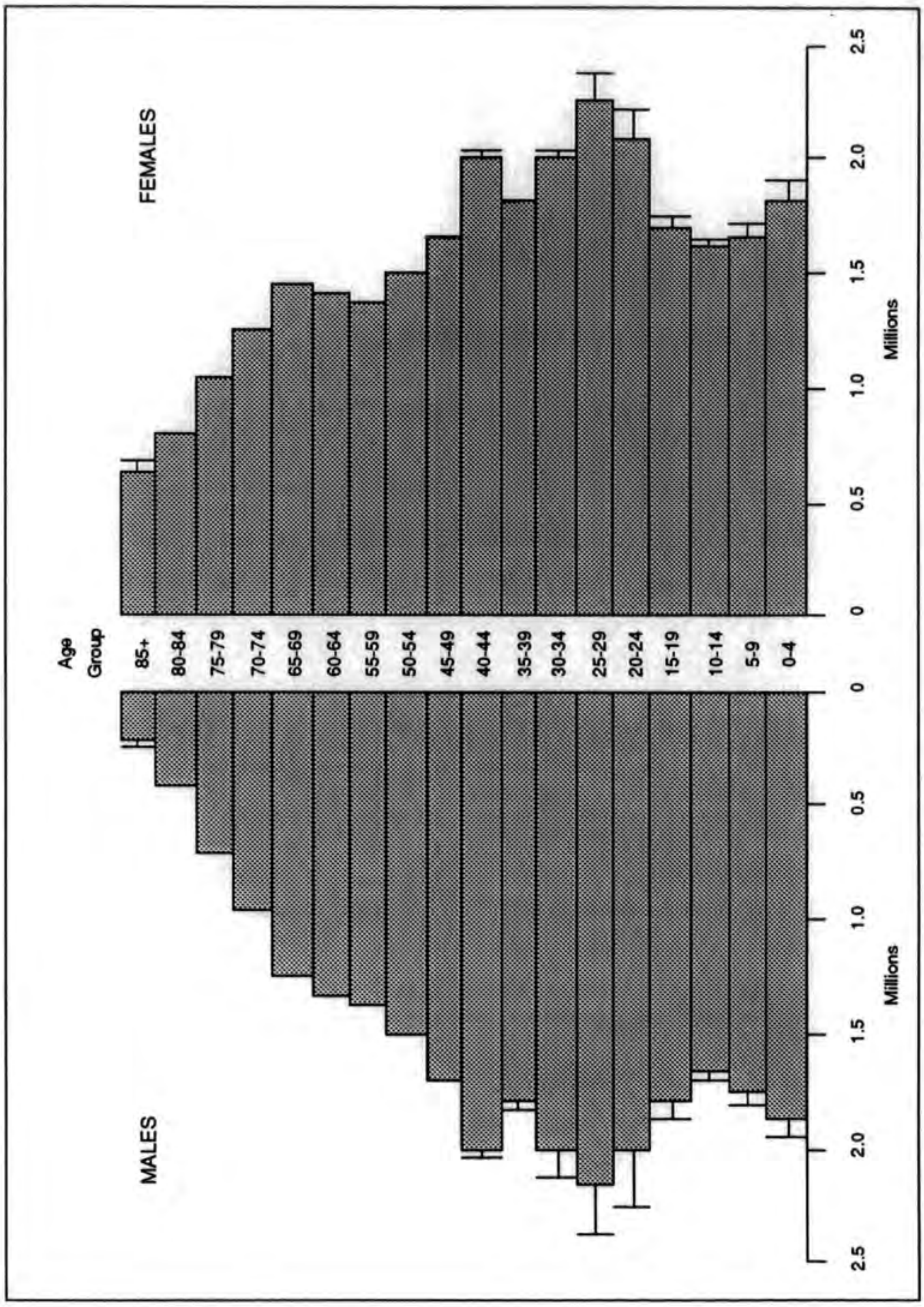
Source: OPCS 1993(f)

home or building can be accurately estimated from the three-month reading (the pits left by radon gas on the detector surface are counted automatically).

The initial data collection and land classification process carried out by MAFF has been criticised as being too subjective and open to personal interpretation. Details on the methods of mapping are not available but some error is likely to have occurred in the initial categorisation of the land parcels into the five grades of agricultural land shown on the maps, despite the production of guidelines outlining the classification and the education of the researchers on the methods to be used. It seems likely that some measure of subjectivity will have been introduced by the individuals in charge of assessing the capabilities, and the nature of the terrain (and accessibility) of the area, the weather conditions at the time of mapping and the time of day (light conditions and the resulting tiredness of the surveyor) may all affect the results of a classification. Where there is operator-dependent bias in the recording process, dependency effects between the areas could be generated, with (mild) discontinuity at the block boundaries. However, the quantification of such measurement error is highly problematic and no estimates are available from MAFF.

Two main methods are used by OPCS to identify possible errors in the census - the Census Validation Survey (CVS) and the Registrar General's mid-year population estimates. The CVS for the 1991 Census (see Table 4.2) was conducted in June and July of 1991 on a stratified sample of 20,000 houses and results suggest that the census under-enumerated the total population of Great Britain by 340,000 (0.73%). This inaccuracy can be accounted for by a number of different error types, namely the failure to identify all residential accommodation, the failure to identify all households and household spaces within accommodation, and errors in the estimates of the number of persons in wholly absent households or where no contact was made between the enumerator and the residents. The greatest percentage errors were found in English metropolitan counties outside London (0.98%) (Martin, 1993). From the mid-year estimates (compiled from routinely collected information) the under-enumeration in the 1991 Census was found to consist almost entirely of people aged up to 45 years or over 85 years (OPCS, 1993b), with men aged between 20 and 29 years being most affected (Martin, 1993). Thus the effect on social analyses will not be evenly spread between the social classes but will have a disproportionate impact on the lower classes (especially unskilled or semi-skilled as young heads of household are less likely to be professionals or managers compared to the over-45s) and on analyses involving a male/female split. Figure 10.3 shows the age distribution of under-enumerated persons, based on the Registrar General's provisional mid-1991 population estimates for Great Britain. Under-enumeration in the 1981 Census is thought to have been much lower than in 1991 (see Chapter 4). The raw census data are unavailable by law for 100 years (for reasons of confidentiality) and although it may be the case that there is a correlation between people that are missed from the census or that have provided inadequate information, and the lower classes, it is impossible to reach any conclusions from the aggregated data. The

Figure 10.3 Under-enumeration in the 1991 Census



Extensions indicate mid-1991 population estimates produced by the Registrars General

Source: OPCS (1993b)

Small Areas Statistics (SAS) data on social-class, obtained from OPCS, constitute a 10% sample of responses to the 1991 Census which have been aggregated to form enumeration districts and wards. In spite of efforts by OPCS to make the sampling method as systematic and comprehensive as possible, the 10% sample of social-class data for wards is not necessarily an accurate representation of the socio-economic make-up of the area; there will always be the possibility of measurement error or bias towards one or more groups in the community (unless the data are made available for the full 100% of the population and even then under enumeration could affect certain social classes more than others). The precise extent of sampling error on the reliability of census data is impossible to ascertain, but Martin (1993 pp. 27) working with 1981 data, concludes that although there was no evidence of systematic bias, the data for enumeration districts are subject to large sampling errors. He suggests that 10% data should be analysed at ward level or above, as the degree of error decreases with increasing sample size.

The division of data into classes in order to display the data in map form introduces error related to the necessary simplification involved in the categorisation process. For both the radon and the census data, the data were allocated to a specific class using a quantile scheme and shaded appropriately on the choropleth maps (see Chapter 12). The categorisation of the percentage of homes above the Action Level in the radon data set smoothes the internal variation in the grid squares but introduces error related to the generalisation of the data. In addition, some of the classes are not mutually exclusive, giving rise to 'fuzzy' areas between the categories; for example, 1 to 3% and 3 to 10%. For the purposes of this work the three middle categories relate to the following ranges: 1 - 2.9%, 3.0 - 9.9% and 10.0 - 29.9% (see Chapter 12.1). The use of grid squares, which are a regular size and shape, reduces the impact of the MAUP and provides a regular surface with which to map spatial data (see Evans, 1981 discussed above). The classification used by NRPB is as follows:

Estimated proportion of homes above the Action Level
- less than 1% of homes above the Action Level
- 1 to 3%
- 3 to 10%
- 10 to 30%
- more than 30%

Only the first four categories are used in Somerset and Northamptonshire. The categorisations used in the census maps for Northamptonshire (Chapter 12.4.4) are dependent on the nature of the data, but the number of classes does not exceed the upper limit of seven suggested by Langford and Unwin (1994 pp. 23). Generalisation of the data inevitably occurs but the disadvantages of this are weighed against the impracticality and complexity of analysing and mapping the uncategorised census data for each of the wards or districts. The categories used for data analysis are listed in Chapter 12.4.4 and will not be repeated here. Two examples are provided however, namely population density and the proportion of households containing only pensioners. The classification systems used do not necessarily begin at zero, nor do the classes necessarily increase at regular intervals. The intervals were chosen in order to allocate an approximately equal number of counts to each class, although for some classes there are no counts (*e.g.* 11.0 to 12.9% and 13.0% to 14.9% in social class V in Northamptonshire) because of the nature of the data.

Population density
- 0.0 to 0.9 persons per hectare
- 1.0 to 2.9
- 3.0 to 19.9
- 20.0 to 49.9
- >50

The proportion of households containing only pensioners
- <15%
- 15 to 19.9%
- 20 to 24.9%
- 25 to 29.9%
- >=30%

Inevitable generalisations have therefore occurred in the radon and census data as a result of the classifications that have been used. Inconsistencies are found when attempting to repeat the results found in categorised data using other methods of classification, and the extent of the validity of the results obtained in this thesis is therefore drawn into question. However, as Openshaw and Taylor (1981) maintain, the existence of outstanding methodological problems inherent in spatial analysis cast doubts on the applicability of nearly all applications of quantitative techniques to zonal (and categorical) data.

10.4 Operational Error

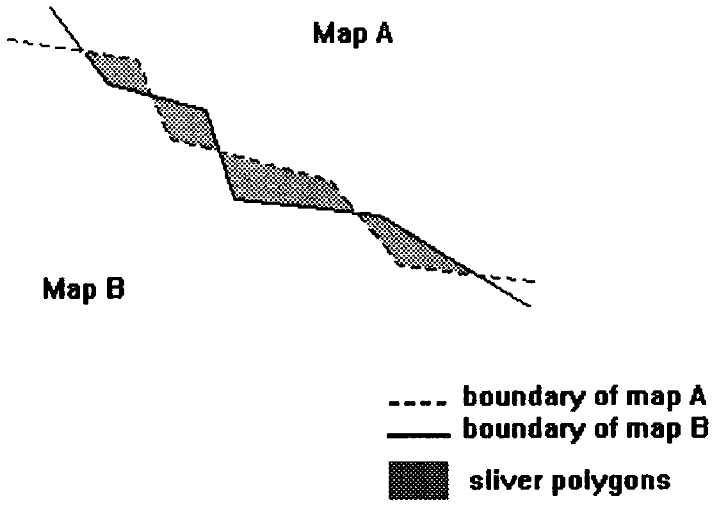
Errors in data manipulation and in the raw data itself can not always be measured or quantified and several very small errors, when processed within a computer package, can result in a significantly large error that can remain undetected. There are many changes in the state of information when using a spatial information system database. There are changes in media, *e.g.* from paper map to digital encoding in order to display the data on a monitor; there are changes in form, *e.g.* from tables of numbers to geographical representation; there are changes in representation of entities, *e.g.* from vector geometry to grid cells. Some alterations occur in conceptualisation, others in digital encoding and others in product creation (Laurini and Thompson, 1992). Data processing errors are probably the most difficult errors to assess in the main because GIS carry out large numbers of transformations and calculations on the data during processing and analysis. Unless the exact nature of these operations is understood, the effect on the data and possible errors contained within it are ignored or misunderstood. As Openshaw (1989) states, 'the ease with which data from different scales ... with different levels of innate accuracy can be mixed, integrated and manipulated totally disguises the likely reality of the situation'. Hampel *et al.* (1986) warn that where data recording methods are automated (*e.g.* scanning) the risk of undetected error increases.

10.4.1 Numerical Computing Errors

The accuracy of a digitised map can be determined to a certain degree by deductive estimating or by testing. The digitising process, especially when carried out manually, often results in numerous small errors including closure errors (where line segments do not meet but either overshoot or undershoot), labelling errors, the omission and duplication of edges or nodes and poor line shapes. The complexity of the problem escalates when regions abut each other in a mutually exhaustive partitioning, such as wards. In this case every boundary line is common to two polygons (except for the outer line) and the digitised lines may not all follow the same path. Errors that are visible on the computer screen can be checked by comparing the original map with the digitised version. GIS software packages often facilitate data preparation by automatically checking for and correcting digitising errors. Within Lites2 an error detection capability enabled, for example, polygons which had not been fully closed or which had more than one identification number to be located automatically and then changed as required. Errors that were identified in this thesis at an early stage of data input, such as closure errors, label errors and duplicated edges or nodes, were corrected before any transformations were carried out. However, as Goodchild and Gopal (1989) state, "... digitising errors are often relatively minor compared to the error inherent in the process of cartographic abstraction itself".

Errors in the coverages are not confined to digitising errors but can also be introduced during map overlay. Spurious polygons, characterised by their small areal extent, long, thin shape and usually only two arcs, are artefacts of the map overlay operation that do not represent features in the real world. The polygons are called sliver polygons and arise from positional discrepancies between representations of the same line on alternate data layers. Veregin (1989) outlines some error reduction strategies, with particular reference to map overlay. For example, a potential error reduction strategy at this juncture might be to delete or close those polygons below a certain threshold size. Errors from this source did not arise in this thesis because the boundaries of the different data sets did not overlap, as the boundary of the radon data was defined by the edge of the five kilometre grid squares and that of the other data sets, by the county border.

Figure 10.4 The Generation of Sliver Polygons



(Source: Goodchild, 1991)

The two most conspicuous errors in the land capability data used in this project relate to the omission of Chard (in Somerset) as an urban area, roughly six square kilometres in size, and the erroneous classification of a 60 square kilometre area in the east of Somerset as Grade 3 land instead of Grade 4. Whilst the areal extent and significance of these errors is small (see Appendix 3), their omission illustrates the point that the digitising process is prone to operational error. The impact of human error, especially regarding the ignorance which surrounds the basic functions and shortcomings of GIS, is perhaps the least understood source of all the sources of error. Carver (1991) however, regards it as a fault in the design of GIS packages that the user lacks information on how to achieve optimum results with a specific piece of GIS software.

The magnitude of error incurred during the digitising process is not covered in any depth in the literature on GIS. The primary sources of error in the maps and data produced in ARC/INFO result from the raw data that are integrated in the GIS. Although numerical operations undoubtedly play a part in the propagation and possibly also the magnification of error, the lack of information on the impact of numerical processes has resulted in a question that can not be fully answered without detailed mathematical analyses (for example see Veregin, 1989).

10.4.2 Faulty Topological Analyses and Errors in Interpolation

Good description of spatial variation is important for summarising data and for interpolation. Interpolation procedures can not be expected to provide good estimates if the sampling method has missed important elements of surface variation (Haining, 1990). The interpolation of radon values for squares where no measurements were taken is outlined in Chapter 2 and involved estimating the radon level from the eight surrounding grid squares to give a representative figure. This process was also carried out in grids where very few measurements were taken to reduce the effects of error. Smoothing was also carried out on all the data to further reduce the anomalous influence of outliers.

A major difficulty arises when data are collected over different (and incompatible) areal partitions. As we have seen above, in the case of vector GIS, the representation of a line or an edge implies a level of certainty or precision that may not be discernible in the real world (Maffini *et al.*, 1989). The analyst may be interested in analysing relationships between variables that have been collected by different agencies that have used different areal frameworks. Solutions to this problem range from ignoring the incompatibility to attempting some form of adjustment such as weighting or aggregation. The accuracy of such adjustments rests on the assumption of intra-unit uniformity (Haining, 1990). Because of the danger of the ecological fallacy (see 10.3.2 above) and the aggregation problem (see 10.3.2) to infer individual relationships between aggregated data sets risks the introduction of error. No individual relationships have therefore been extrapolated from the findings of this thesis.

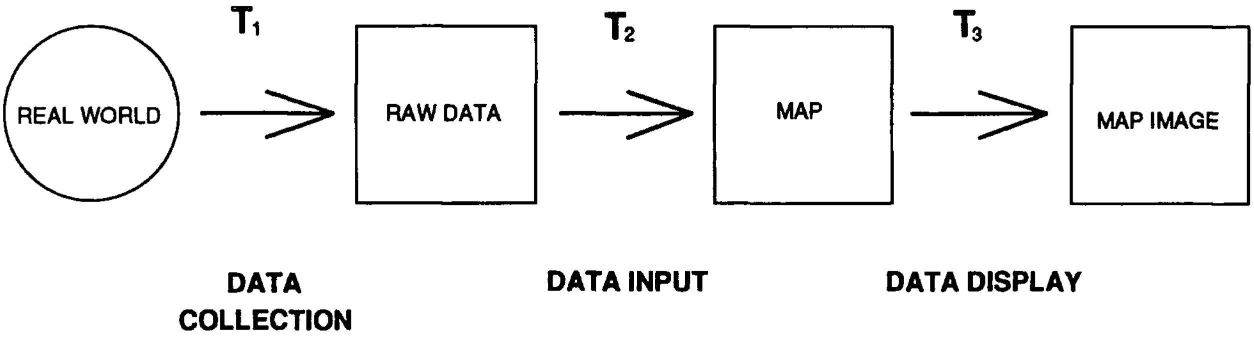
Errors that originate from the data sets themselves can be compounded in different types of GIS operations (especially in overlay). The more times a set of measurements are transformed through one process or another, the more likely new errors or uncertainty will be introduced into the derivative products (Maffini *et al.*, 1989). When the end result of analysis within a GIS includes some form of cartographic output it is difficult but important to separate the form (cartographic representation) from the content (the data collected as a representation of reality). There are a number of aspects which affect the way that spatial data are mapped, all of which can affect the complexity of the data that is shown. Rhind (1983, working with census data) has identified these as follows:

- the type of areal units
- the geographical resolution of the data
- the uncertainty introduced by data suppression and by adjustment or 'blurring'
- available symbolism
- technical capabilities available *e.g.* computer mapping programmes
- reproduction technology and cash resources available

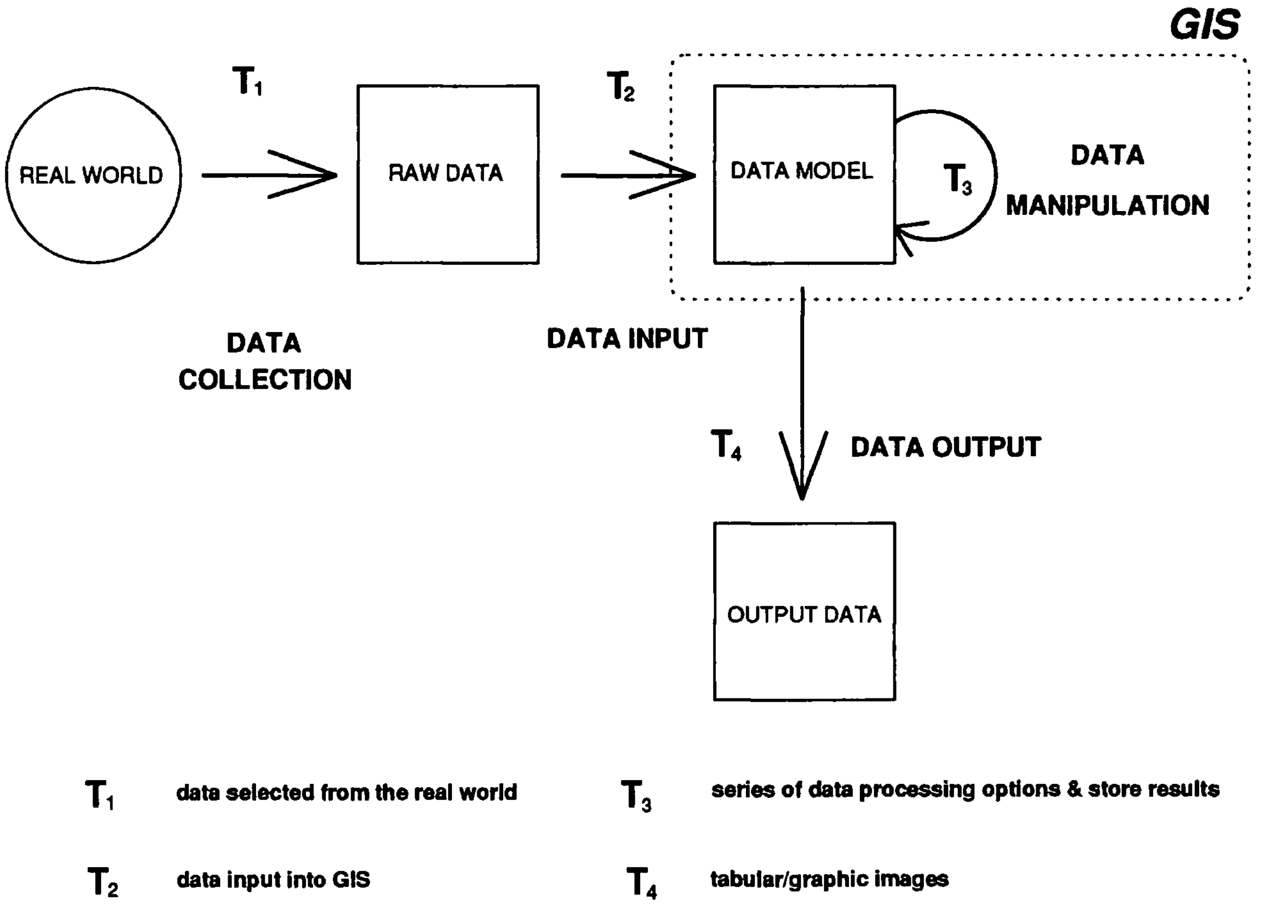
Figure 10.5 shows the stages of data transformation in the production of a traditional map and compares it to the transformations involved in the production of output from a GIS. As a result of the data transformation and manipulation that can occur during the production of a

Figure 10.5 Data transformation

Transformation stages in the traditional cartographic process



Transformation stages in GIS



(Source: Martin, 1991)

map, the form and content of the cartographic output should not be judged on equal merits. The data that is being represented and not the representation of the data (although this has a major part to play in the communication of ideas and results) is what is most important. As Carver (1991) illustrates, it is often the case that results are judged in terms of the quality of the graphics used for presentation rather than on their intrinsic value. This is a source of error which could be reduced by educating GIS users not to accept printed output at face value, but to question the relevance of the results that are portrayed. Those people in influential or managerial positions within the GIS community have an important educational role to play in preventing the perpetuation of false ideas and complacency regarding error in the final output from GIS.

10.5 Conclusion: The Detection and Reduction of Error

A greater number of companies and institutions are now taking a more serious approach to error detection and data quality control. As Maguire (1989) states, 'standards for data quality are important prerequisites for a more united approach to data integration in GIS'. The Ordnance Survey is currently assessing the use of metadata (information about data) within their computerised mapping system, and an ongoing National Land Information System (NLIS) research project is developing and testing a metadata approach to data quality assessment (Ralphs, 1993). The Data Archive of the ESRC at the University of Essex, is contributing towards the development of more accurate data banks by asking academics depositing data there to fill in an exhaustive questionnaire on the techniques and methodology used.

Some crude error estimation processes have been developed at Newcastle and elsewhere (see for example, Veregin, 1989; Chrisman, 1989). Carver (1991) suggests that at the very least a GIS should offer a measure of result robustness, along the lines of confidence intervals. He outlines two approaches to error estimation that have been developed at Newcastle; the Monte Carlo simulation and a probabilistic epsilon error approach (PEE) for use in ARC/INFO. However neither have been readily incorporated into a GIS package and both require a certain amount of knowledge of the theory behind GIS. The PEE approach requires programming using the ARC/INFO macro-language AML and the Monte Carlo simulation has a number of problem areas associated with its use, not least that it requires a two-fold increase in the processing time.

The main sources of error in the data used in this project relate to the spatial units for which the data are collected (which are all different and incompatible) and the aggregate nature of the data sets themselves. In addition, the age of the land capability data and two land capability data input errors, reduce the reliability of this data set, although as discussed in Chapter 3, the age of the data does not represent a large source of error because of the nature of the data. The main source of error due to the age of the land capability data results from the spread of urban land into the surrounding countryside and not from changes in the capability of agricultural land, which is unlikely to alter significantly with time. The accuracy of the land capability data was further reduced by simplification during the digitising of the paper source maps and the resulting generalisation. Inconsistencies between the 1981 Census boundaries and 1991 Census data introduce a further source of error to the results. In addition, inevitable generalisations have occurred in the radon and census data as a result of classifications that have been used to group the data prior to analysis. As discussed above, there are many problems inherent in quantifying sources of error and in assessing its impact. The identification of possible sources of error is the first and most important step in identifying inaccuracies in the final results and maps. Some degree of error is unavoidable and is inherited from the methods used to measure and aggregate the data, thus whilst these methods continue to be used, data error can never be eliminated. However, the incorporation of metadata in future will facilitate error measurement.

The challenge to the GIS community is therefore to recognise that data errors exist in virtually all applications and to consider the limitations that they create rather than ignoring them (Ralphs, 1993). It is important that users are at the very least provided with practical error handling and analysis technology to enable estimates to be made of the effects of error on GIS output (Carver, 1991). Much of the literature on the accuracy of GIS begins the summary chapter with the word 'ideally'. But it is less ideological and more a necessity that GIS are produced with a means of testing the confidence limits of the GIS output. However, the problem is compounded by the complex set of rules which are used during overlay in many applications. As Goodchild and Gopal (1989) state, in severe cases the concept of confidence limits may have to be replaced by a system of warnings, to help unwary users to identify meaningless operations or ones where the results would be hopelessly erroneous.

11. Résumé

The three main data sets that were identified in Part One (namely radon gas, land capability and social data from the 1991 Census) were digitally captured and integrated using modern GIS technology, in order to analyse the geographical distribution of the data and the spatial relationships between them. GIS is the enabling technology of the 1990s and has implications not only for the development of new methodologies within geography, but also within other disciplines concerned with spatial data.

Three groups of support structures for the utilisation of GIS have been identified by De Man (1990), namely professionals and users, system maintenance and organisations. These support structures do not emerge or continue automatically however - they all require commitment. Closer collaboration and increased discussion between software developers, academic users and industry has developed following the creation of bodies such as the Association of Geographic Information and the recently formed United Kingdom Academy for Information Systems, and the publication of new journals concerned with spatial information systems. As GIS technology has become more user friendly and the potential for information synthesis has increased, awareness of data quality has become more important. The development of standards is one way to ensure that the progression of GIS software is beneficial to the user as well as to the software companies. Chapter 10 identified the need for standards and outlined the recent advances that have been made in this area. De Man (1990) offers a detailed summary of the strategic choices for the development of a GIS and the analytical requirements for spatial information.

The development of automated GIS have led to notable increases in the application of map overlay techniques. The results of the overlay procedures carried out in this study are discussed in Chapter 12, but it is useful to stress the importance of the initial data capture and integration on the resulting accuracy of the output. The choice is often not whether to perform map overlay but how to select maps that will produce a meaningful result (Chrisman, 1990). The process of overlaying the data sets to establish spatial relationships was complicated in this thesis by the variety of classification units that were used. It is a geographical fact of life that the results of spatial study will *always* depend on the areal units that are being studied. Indeed, the MAUP remains one of the most important unresolved problems left in spatial analysis. There has been very little research compared with that afforded to many far less significant problems and as Openshaw (1986) states, whilst it appears to be primarily a technical problem, it is also a major conceptual problem that is central to many aspects of geographical study. He suggests that quantitative geographers should devise a body of relevant spatial analysis techniques that can cope with geographical data. Indeed the MAUP is a potentially powerful geographical tool once the aggregational uncertainty inherent in it can be usefully exploited for geographical purposes.

Problems of comparing and relating data from different sources are not new but have assumed greater importance since the Information Revolution of the 1970s and early 1980s. Maps have always answered the question 'where', but in the information era they must also answer a variety of other questions such as, 'why', 'when', 'by whom' and 'for what purpose'. The inter-relationships between the data in this thesis were analysed for five counties whose boundaries have been artificially imposed for administrative purposes. Physical and human characteristics do not stop at these boundaries but change gradually and there are often significant similarities between adjacent areas, counties and even regions. Spatial dependence is a common characteristic of spatial data that is of intrinsic interest but also creates problems for the application of statistical procedures. For example, it can cause the distortion of statistical results using data that are measured and tested for imposed spatial units. New methods are needed for analysing spatial data that can detect and deal with the problems associated with the spatial organisation of a particular data set.

To a certain degree map error is unavoidable thus no stage in the transformation of the raw data through to the output of the finished results should be ignored in terms of assessing the impact of error propagation on the end results. However, despite the current activity in system development and standardisation, the propagation of errors from source materials through to the finished output is still rather poorly understood (Chrisman, 1990). MacDougall (1975 pp. 30) presents an analysis of error with the following conclusion: "... that some overlay maps may indeed differ little from random maps and that most overlay maps contain more error than the compilers and users probably realise". However, overlay maps are usually sufficiently reliable for the needs of small-scale analysis of the geographical differences between areas. Indeed, Mohan and Maguire (1985; cited in Martin, 1991) suggest that 'the *merit* of GIS is that, by linking data sets at small spatial scales, they offer a reasonable compromise between the accuracy of individual data and the generality of aggregate data' (not original italics).

A spatial information system will only be successful if it enhances the ability to make decisions or to reach conclusions in research, planning or management (Smith *et al.*, 1987; Heywood and Cornelius, 1988). Consequently, an information system should facilitate a heuristic dialogue (Peuquet and Marble, 1990). Burrough (1990) believes that metadata (information about data) could be incorporated into a knowledge base, with procedures for the detection and measurement of errors alongside the GIS, so that a user could be advised on the best way to achieve a desired aim. In this manner intelligent GIS would present a range of alternative strategies such as better methods, more data, different data, better models or better spatial resolution, as well as basic error measurement algorithms. These should be the requirements of the GIS of the future in order to halt the growing problem of error propagation and the poor interpretation of results through ignorance of the capabilities and limitations of GIS. At present no GIS can present the user with information about the confidence limits that should be associated with modelling and analysis. In the words of Chrisman (1989), one impediment to the widespread adoption of error models is that no single error model and no single definition of 'accuracy' is applicable in all instances. Areas where future research is needed include the following (after Dale, 1986):

- Ensuring that the raw data is kept up-to-date
- Linking point-sampled data with those sampled on the basis of area
- Establishing standards for the classification, quality, evaluation and coding of spatial data
- Reconciling of raw data which is often in point form, with processed data which is more generalised
- Remedying the quality of both two- and three-dimensional generalisation
- Determining user needs
- Remedying deficiencies in past studies of map perception which tended to exclude context from the assessment
- Developing artificial intelligent systems which provide direct answers to questions

In short, the rapid rise in the development (and the lowering of costs) of GIS over the last ten years has resulted in the integration of a wide variety of data from many different sources, by an ever-growing user base. The need for better data standards and accuracy checks is becoming more obvious as the GIS community become increasingly aware of the problems with integrating aggregated data. The publication of the Chorley Report in 1987 brought the issue of data accuracy to the forefront and encouraged greater debate among academic users and industry. However, there is still a great deal of research to be carried out in the field of GIS (as exemplified by Dale, 1986), but it is the ever-developing technology that is a characteristic of GIS, which gives the field its dynamism. Despite the problems that remain with integrating data collected for different areas and at different scales, GIS continue to be appealing because of their ability to carry out spatial analysis and produce high quality maps in a short space of time.

Part Three: Results

PART THREE

12. Results

12.1 Introduction

The purpose of Part Three is to display the results of the overlays carried out on the three main data sets and to discuss the findings. It is clear from NRPB research throughout the United Kingdom that elevated radon levels occur in areas that are not necessarily underlain by granite (and other uranium enriched rocks) and that have no clear geological explanation. In these areas other explanations are being sought for the unusually high radon readings in homes. The spatial analysis of radon with land capability and social class data (in Northamptonshire only) is used to identify possible relationships between radon and factors other than geology and to highlight the geographical differences and similarities between the five counties under investigation.

The relationship between radon and physical characteristics (namely geology) have been well documented elsewhere (*e.g.* Ball *et al.*, 1991; Ball and Miles, 1993). The implications of a different kind of physical attribute, that of land capability for agriculture have been assessed in Part Three for the five Affected Areas. The general trend to be studied in this chapter is that as land capability decreases, the proportion of homes above the Action Level increases. The analysis of land capability and radon levels in Devon and Cornwall, serves as a kind of 'control' to assess whether the techniques used to ascertain the presence of spatial relationships can be applied to other areas of the country, with a degree of confidence. The results show that there is a clear relationship in the control counties between land capability and the proportion of homes estimated to be above the Action Level and the methods used to achieve these results can therefore be used in the analysis of the three other Affected Areas. By using the proportions of each land capability class (for example see Table 12.2.3) comparisons between the Affected Areas are still possible. The results of the analyses between land capability and radon levels are outlined below for each county (Devon and Cornwall have been combined) and the main findings for each of the Affected Areas are compared and contrasted in Chapter 13.

In this thesis, the focus of enquiry in to the geographical differences and similarities of radon Affected Areas and the spatial relationships between the factors influencing the proportion of homes above the Action Level, lends itself more to the ideas of Cox and Jones (Wrigley and Bennett, 1981 pp. 135-142) than to the statistical inference methods expounded by Cliff, Haggett and Ord (*e.g.* Cliff and Ord, 1973; Haggett *et al.*, 1977). The methodology employed, places more emphasis on exploratory (rather than confirmatory) modes of statistical analysis, coupled with graphical displays of the findings. To this end, statistical analyses have been kept to a minimum. It is hoped that the abundance of maps and tables serves to illustrate the main geographical differences and similarities between the five counties without the need for lengthy description or regression analyses. Indeed Cox and Jones (1981) have argued that geographers should engage in more uninhibited exploration of their data and that they should often be content with simple but robust numerical measurements and imaginative graphical displays. The increasing use of exploratory data analysis (as opposed to confirmatory or inferential approaches) is part of a wider movement within geography towards a more holistic approach to the subject.

The overlay procedure used to achieve the results outlined in this chapter, was carried out in ARC/INFO using radon data from the NRPB, digitised land capability data and 1991 Census data at ward and district level. The results from the overlay processes therefore, relate specifically to these data. There will inevitably be differences between the data and the real world, as imposed boundaries, generalisation and the effects of scaling in the data sets can distort results. Consequently, any extrapolations from the overlay findings can only be an approximation to the real-world situation.

The display of geographic information in the form of a map is a distinctive feature of spatial interpolation and with the new technology that tools such as GIS now offer, maps can be designed for specific applications that can meet user requirements more effectively. A map can never be truly objective because one of its distinguishing features is that it focuses attention *selectively* on regions, objects and themes. The results of the overlay operations have been printed out in the form of maps in order to show the spatial variation in land capability and radon levels for the five counties, as well as social class, population density and the proportion of households containing only pensioners in Northamptonshire. To paraphrase Buttenfield and McMaster (1991), a map is an abstraction of reality and its effectiveness as a communication medium is strongly influenced by the nature of the spatial data, the form and structure of the representation, the intended purpose and the experience of the viewer. To ensure that the data portrayed in this thesis are as objective as possible, a section on methodology (Chapter 9) and possible sources of error (Chapter 10) is included to present the reader with an accurate picture of the processes involved in the creation of the maps and tabular results. The full data sets for each county are included either in the following pages or in the Appendices. Dissatisfaction with choropleth mapping techniques and artificial areal units in general has led to the development of some new methodologies, such as boundary-free representation of the SAS (*e.g.* Martin, 1989). However, the concept of a map allows the relationships between a wide variety of both qualitative and quantitative data to be organised, analysed, presented, communicated and used in such a way which no other product can match (Taylor, 1991).

In addition to printed maps, ARC/INFO can produce a variety of statistics relating to the original coverages and to the resulting overlays. It is possible to pose questions using the *if, and, or, not, equal to* conditions, known as relational operators, which allow complex searches on sections of the database. An example of a query is as follows:

Display all areas where the *radon level is greater than 1* (< 1% of homes above the Action Level) ***and** land capability is Grade 1 **or** Grade 2*

The resulting statistics can take various forms:

- The number of polygons created in the overlay process. For example, in Derbyshire the number of polygons with a radon level (the proportion of homes above the Action Level) of <1% and a land capability Grade 2 is 25.
- The total area of the overlay polygons. For example, in Derbyshire the total area of Grade 2 land which has a radon level of <1% is 159,503,311 square metres.
- The mean area of the polygons created in the overlay process.
- The frequency, maximum, minimum and standard deviation of the polygons created in the overlay process (*e.g.* see Table 12.2.2).

For the purposes of this work, the number of polygons created in the overlay procedure is meaningless, as they can vary in size from less than 500 square metres to 25 square kilometres. To use the total area of the polygons is problematic as the figures are very large and bear no relation to the other polygons in the coverage, making comparisons difficult. For these reasons, percentages have been used as a means of analysing spatial variation. To calculate the proportion of each overlay variable (*e.g.* Grade 1 land and a radon level of <1%, or Grade 2 land and a radon level of 1 - 3% *etc.*) the total area for that variable was divided by the total area for the land capability column (other minor transformations that were employed are explained in the following sub-sections). This method was deemed preferable (to using either the total of the radon column or the total area of the county) because the resulting percentage cancels out the distortion due to the disproportionate area totals of the different land capability grades. The use of the radon totals or the total county area as the denominator, results in a large difference between the smallest and the largest category for each variable due to the dominance of Grade 3 land in four of the counties (Derbyshire being the exception). Thus, unless otherwise stated, the figures in the graphs and tables below relate to these percentages and not to areas or counts. In addition, the term 'radon level' is used to mean the proportion of homes above the Action Level; thus if a polygon has a radon level of >10%, more than 10% of homes in the polygon are estimated to have domestic radon levels above the Action Level (set at 200 Bq m^{-3}). The radon data obtained from the NRPB were aggregated into five categories, namely <1%, 1 - 3%, 3 - 10%, 10 - 30% and >30% of homes above the Action Level. In this thesis, the three middle categories relate to the following ranges: 1 - 2.9%, 3.0 - 9.9% and 10.0 - 29.9%, to ensure that the classes are mutually exclusive.

12.2 Devon and Cornwall

12.2.1 What spatial patterns exist in the radon data set?

The log normal distribution of radon in buildings and the definition of Affected Areas as being areas (more specifically counties) where 1% or more of homes are above the Action Level, led to the declaration by the NRPB in 1990 that the whole of Cornwall and Devon should be classified an Affected Area. This amounts to an area of over 10,000 km^2 and a population of 1,512,800 (1991 Census). Devon and Cornwall were the first counties in the United Kingdom to be designated Affected Areas.

Figure 12.2.1 shows the spatial distribution of the estimated proportion of homes above the Action Level. This map corresponds to the graph in Figure 12.2.2 which shows the estimated area of each of the radon categories for Devon and Cornwall. In Table 12.2.1 below, it is evident that more than 50% of the area of the two counties is characterised by a radon level of 10% or more (*i.e.* more than 10% of homes exceed the Action Level) and more than 20% of the two counties are affected by radon levels in excess of 30%. These are by far the highest levels in the country; in comparison only 0.3% of Derbyshire (the next highest Affected Area) has radon levels where more than 30% of homes exceed the Action Level.

Table 12.2.1 Summary of the Number of Homes Above the Action Level in Devon and Cornwall

Radon Categories	Area (sq. km)	%
<1%	265	3
1 - 3%	2,727	27
3 - 10%	2,071	20
10 - 30%	3,135	30
>30%	2,056	20
TOTAL	**10,254**	**100**

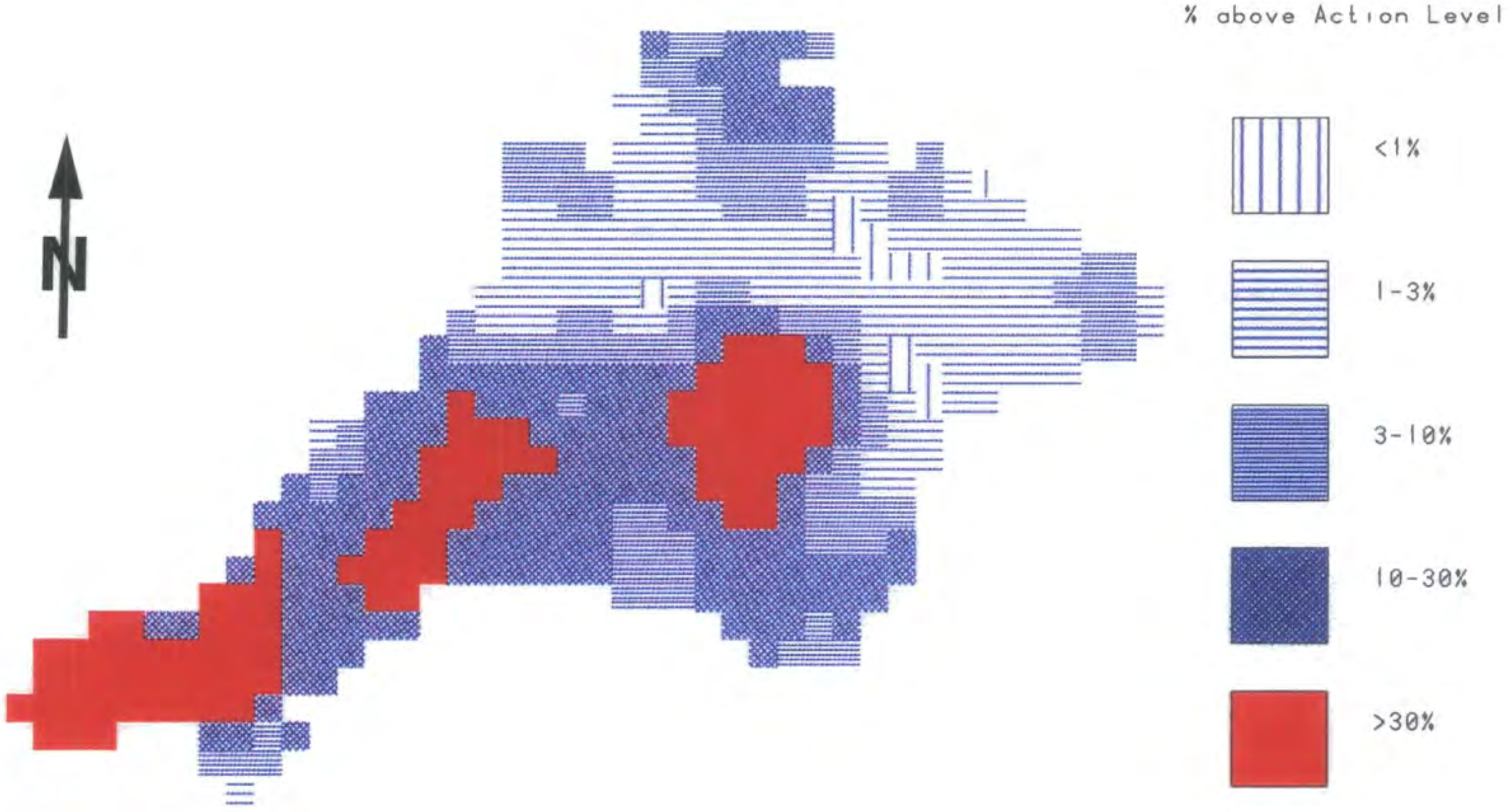
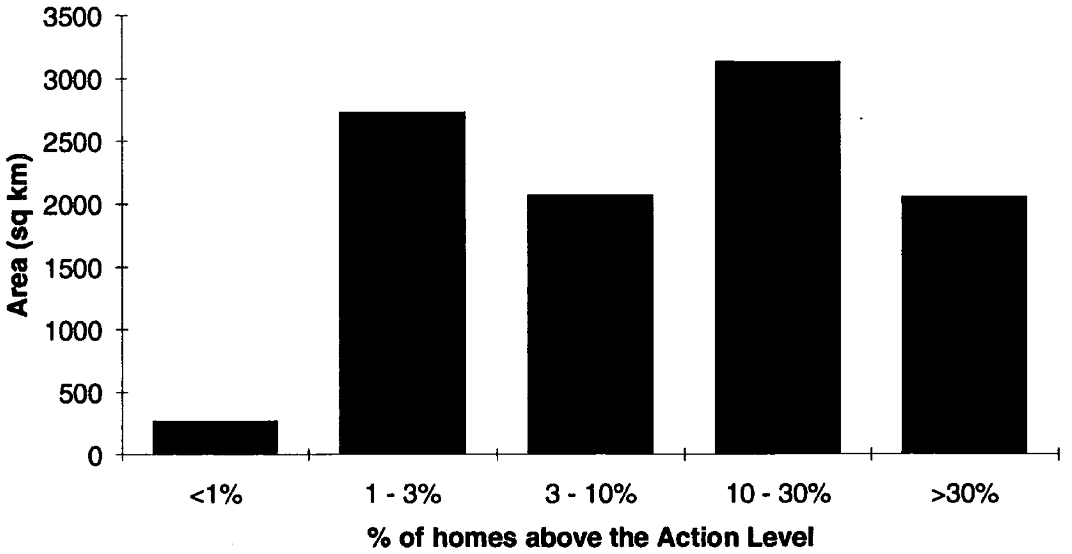


Figure 12.2.1 The spatial distribution of the estimated proportion of homes above the Action Level in Devon and Cornwall

---

Please note that the addition of an extra category (>30% above the Action Level) in this county means that the colours and shading used to represent the categories differ from those used in Figures 12.3.1 and 12.4.1.

Figure 12.2.2 Total Area (in square kilometres) of each Radon Category in Devon and Cornwall



There is a clear division between north and east Devon and the rest of Devon and Cornwall. Dartmoor, from Bodmin Moor to the south coast at St. Austell and the area west of Truro, clearly stand out as the areas of highest radon where the NRPB estimates that more than 30% of homes exceed the Action Level. These areas correspond to the granite masses of Dartmoor, Bodmin Moor, Blackmoor, Wendron Moors and west Penwith (Miles *et al.*, 1990). The underlying granite batholith connecting these outcrops, comes close to the surface at various points, resulting in increased radon levels in areas where granite rock is not visible at the surface, for example around Marazion in south-west Cornwall. Most of the area immediately west and south of Dartmoor is characterised by radon levels in excess of 10%. Radon levels are also elevated (between 10 and 30% of homes above the Action Level) in parts of north Devon, which corresponds to the western part of Exmoor, underlain by Devonian rocks. Generally, however, the north and east of Devon are characterised by much lower radon levels (<1 - 3%).

12.2.2 What spatial patterns exist in the Land Capability data set?

Figure 12.2.3 is a choropleth map showing the spatial pattern of land capability in Devon and Cornwall (digitised from the 1977 MAFF Land Capability map). The pattern of land capability is dominated in the two counties by Grade 3 land (good to moderate quality agricultural land). Approximately 72% (this figure is deduced from the digitised maps not the MAFF source maps) of the two counties is classified as Grade 3 (see Figure 12.2.4 below). This reflects the situation in England as a whole, as MAFF estimate that half of England and Wales is classified as Grade 3 land (Mather, 1986; see Chapter 3.4). The dominance of Grade

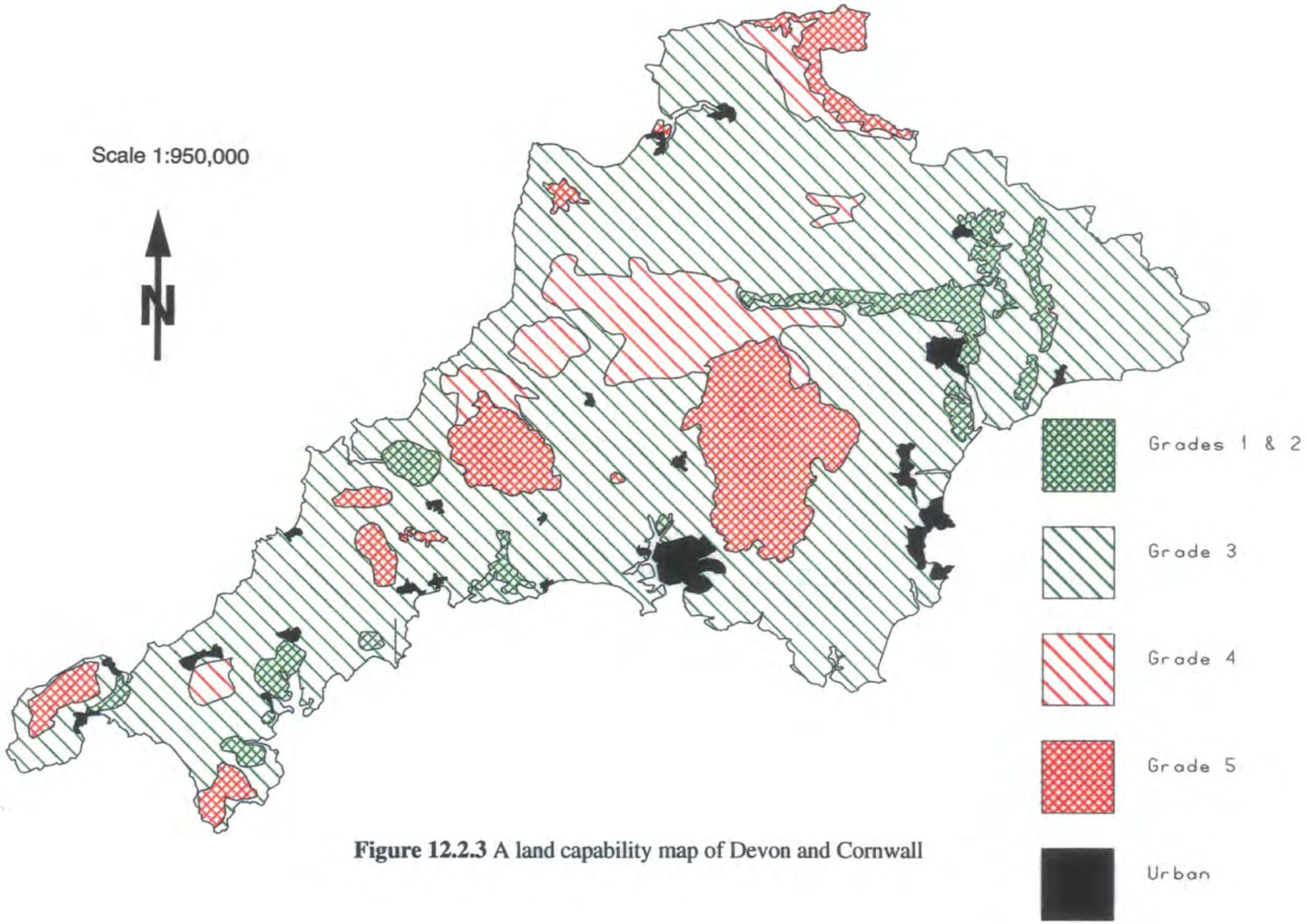
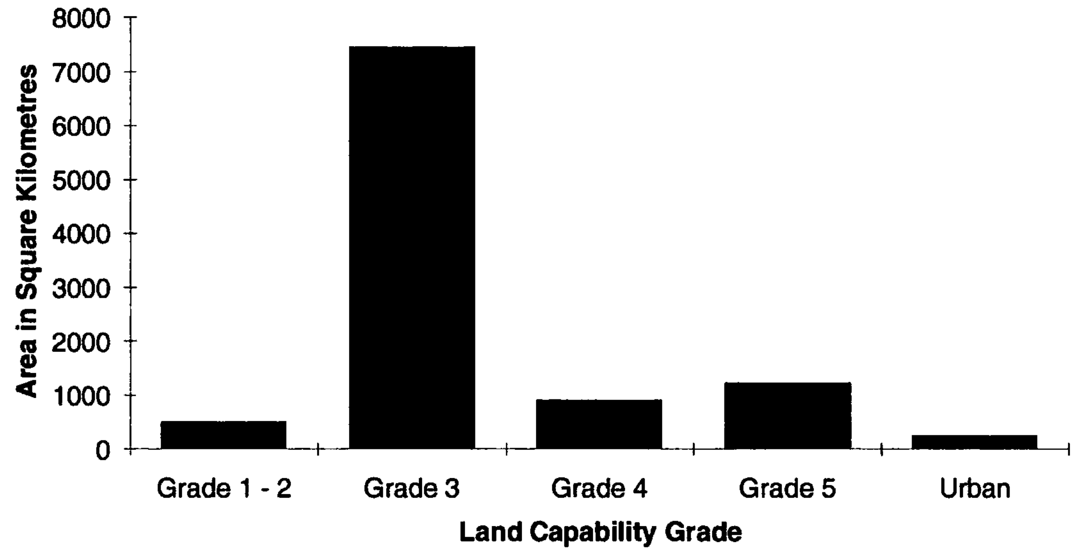


Figure 12.2.3 A land capability map of Devon and Cornwall

3 land in Devon and Cornwall may be exaggerated in this study by the fact that not all the minutiae of detail in the land classification maps were digitised. Small enclaves of Grade 1 - 2 land and Grade 4 land occur throughout the county, but were not digitised if they covered an area of less than approximately three square kilometres. The spatial distribution of land capability is influenced by the nature of the physical environment, and the vertical and horizontal bands of very good quality agricultural land that run north-south and east-west in east Devon are closely associated with the New Red Sandstones which underlie these areas. The best quality land Grade 1 - 2, accounts for only 5% of the total area of Devon and Cornwall and is found predominantly near the south coast, in the main north-east of Penzance, along the Fal estuary and just north of Exeter.

Figure 12.2.4. Total Area (in square kilometres) of each Land Capability Grade in Devon and Cornwall



The second most abundant grade of land is Grade 5. This is classified as very poor quality agricultural land and accounts for approximately 12% of the two counties. It is largely restricted in spatial extent to the moorland areas of Dartmoor, Bodmin Moor and west Penwith. Grade 4 land accounts for 9% of the total area and urban land, just 2%. The majority of the urban areas (23 in total) are small in size, with the exception of Plymouth, Torbay and Exeter.

Table 12.2.2 Summary of Land Capability Data for Devon and Cornwall[1]

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
frequency*	14	1	6	12	23	56
sum (km^2)	499.9	7,441.8	906.3	1,231.0	245.3	10,324.4
% of sum	5	72	9	12	2	100
mean	35.7	-	151.0	102.6	10.7	-
maximum	152.4	-	535.9	646.7	75.2	-
minimum	3.3	-	29.8	3.0	1.5	-

* frequency of polygons in the digital coverage

12.2.3 What spatial patterns exist between radon gas and land capability?

Figure 12.2.5 shows aspects of the relationship between the radon gas and land capability overlays produced in ARC/INFO and is a useful illustration of the spatial relationships between the two data sets. The map shows all the polygons (created during the overlay process) with a land capability Grade 5 and with >30% of homes above the Action Level. There is a strong correlation between these two categories because of the close association in Devon and Cornwall between poor land capability and granite at or near the surface. The uranium content in the granite bedrock often leads to high indoor radon levels. The relationships between the different radon and land capability classes however is difficult to assess merely from cartographic output because of the complex nature of the data. There are five radon categories and five land capability grades, resulting in 25 different possible associations. Analysis is best undertaken using the summary overlay statistics produced in ARC/INFO shown in Table 12.2.3 below.

If the most important category in terms of area is analysed for each grade of land capability, a pattern emerges which suggests that there is a spatial relationship between the two variables. The highest category in each grade of land capability is printed in bold in Table 12.2.3 below. The general trend suggests that as the proportion of homes above the Action Level increase, land capability for agriculture decreases. However, whilst the overall trend appears to show some relationship, there is a great deal of variation between the radon levels in each land capability grade. Urban land does not fit in to the overall trend (see below and Chapter 13.3).

---

[1] The totals in this table may not correspond exactly to those used in the land capability and radon overlay (see Table 12.2.3) because of slight differences in area between the two coverages.

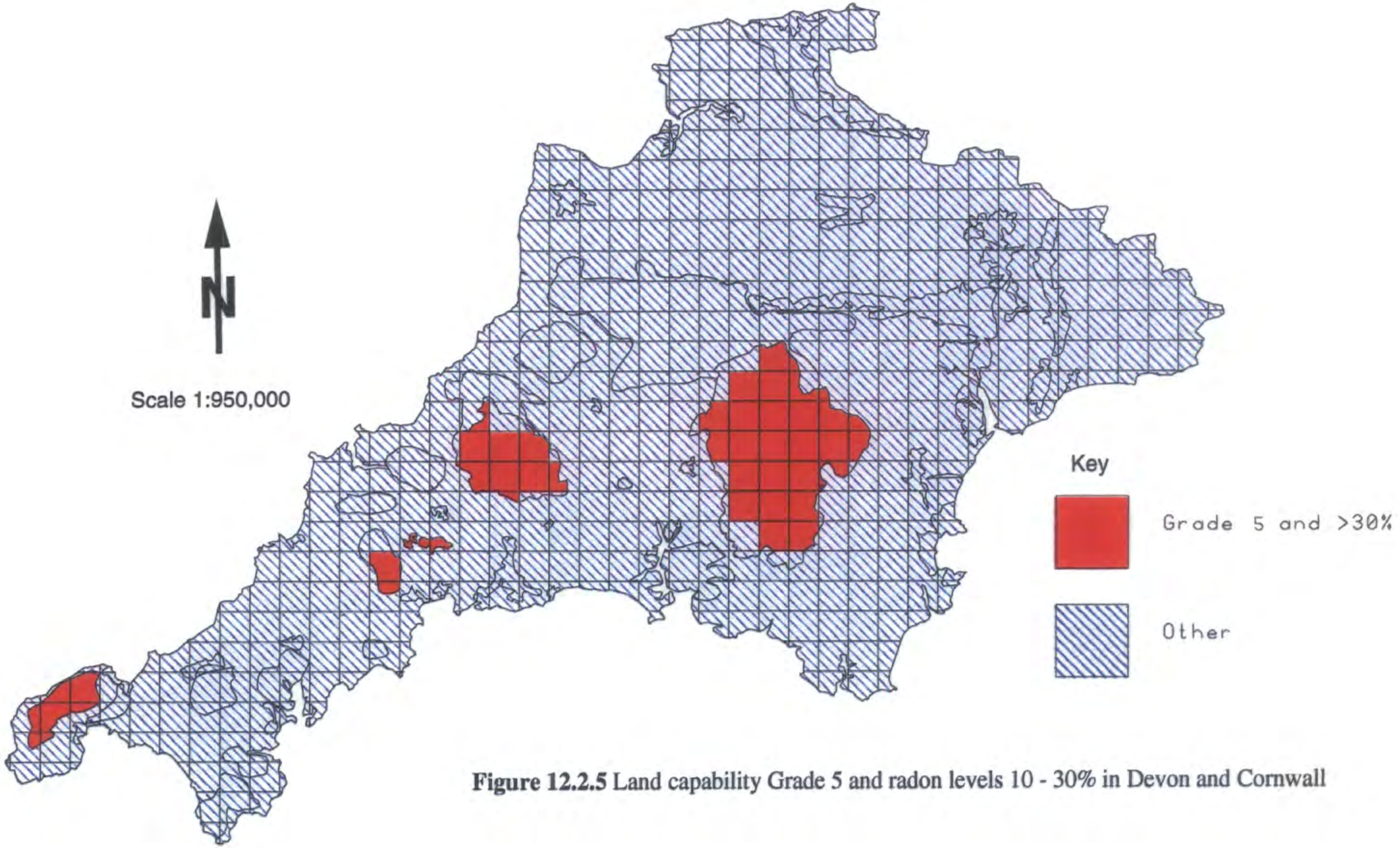


Figure 12.2.5 Land capability Grade 5 and radon levels 10 - 30% in Devon and Cornwall

Table 12.2.3 Summary of the Overlay Between the Estimated Proportion of Homes Above the Action Level and Land Capability, in Devon and Cornwall*

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
<1%	5	3	3	0	10
1 - 3%	**53**	29	28	2	17
3 - 10%	6	22	27	6	**46**
10 - 30%	24	**33**	**33**	23	12
>30%	12	14	9	**69**	14

* These percentages are calculated using the area of each radon category as a proportion of the total area of each land capability grade.

The general relationship that has been identified between radon and land capability needs to be analysed by further dissection of the summary statistics. Areas of low radon (<1% of homes above the Action Level) make up only 3% of the total area of Devon and Cornwall. As would be expected, this radon level is more closely associated with the better quality agricultural land (Grade 1 - 2) than with Grade 5 land. Indeed there is no Grade 5 land associated with areas having less than 1% of homes above the Action Level. The relationship between Grade 1 - 2 and radon is bimodal but the strongest correlation is associated with radon levels of between 3 and 10%. There is little Grade 1 - 2 land in Devon and Cornwall and the mean area is just 35 square kilometres due to the fragmented nature of the distribution. The radon data are at the resolution of 25 square kilometres and it may be that the relationship between the two data sets is distroted by the effects of aggregation and the MAUP.

Grade 3 and Grade 4 land have virtually identical proportions in each radon class, which suggests that the land in these categories may have a similar impact on radon levels in the home. As Grades 3 and 4 are in the middle of the land capability range, it is likely that their physical characteristics are also similar, thus producing the minimal difference that is apparent in Table 12.2.3 above. However, if this result is not echoed in the other counties (and it is not) it may be a result of the particularly marginal nature of Grade 3 land in Devon and Cornwall. Alternatively, the results may be due to the nature of the underlying bedrock (or other local factors) which have produced results peculiar to Devon and Cornwall. After the publication of the land capability maps used in this thesis, MAFF introduced a division of Grade 3, creating the two sub-divisions 3a and 3b. The inclusion of these sub-divisions would be useful for further analysis.

Radon levels of >30% make up 20% of the total area of Devon and Cornwall. Areas of high radon are closely associated with upland moorland areas (in particular, areas with granite bedrock) and there is clearly a strong association with Grade 5 land (see Table 12.2.3 and Figure 12.2.5 above). Indeed, 92% of this grade is associated with areas where more than 10%

of homes exceed the Action Level. The strong correlation between high radon levels and very poor quality agricultural land is associated with the observed relationship between high radon and upland areas.

The relationship between radon and urban land does not fit in to the general trend and is more difficult to account for. There appears to be a parabolic relationship between the two variables with the apex at radon levels of 3 - 10% (see Figure 13.4). Once again however, there is notable variation either side of the mean. The total area of urban land in Devon and Cornwall is small, accounting for just 2% of the total, with the smallest mean area of all the land capability grades at just 10.7 square kilometres. The results of the overlay operation with the radon coverage may have been affected by the small areal nature of the urban polygons and, as the resolution of the radon data is 25 square kilometres, the accuracy of the results must be questioned. However, the wide spread of radon levels in urban areas may be the result of other factors, such as underground disturbance of the bedrock and soil during construction (which creates channels through which radon can travel, sometimes causing two adjacent houses to have widely varying readings), local variations in radium concentrations and ground permeability, or social factors (discussed further in 12.4.4).

Figure 12.2.6 The General Trend Between Radon and Land Capability Grade in Devon and Cornwall

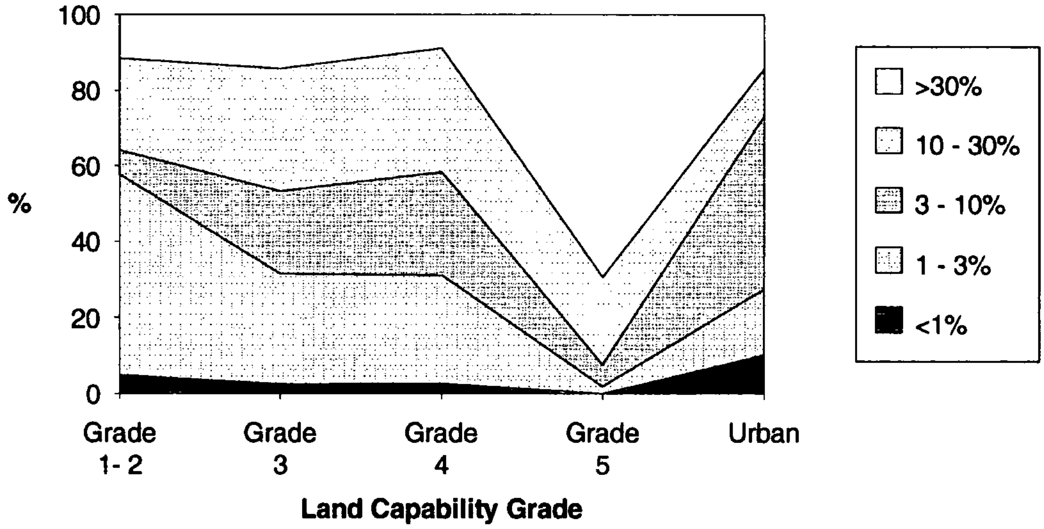


Figure 12.2.6 shows the relationship between radon and the different land capability classes. The dominance of the radon level 1 - 3% in land classified as Grade 1 - 2 and of >30% in Grade 5 land is clearly visible. General trends between these and the other radon and land capability classes, calculated in EXCEL using the *trend* function, indicate the following relationships (the calculations do not take urban land in to consideration). These findings emphasise the relationships already identified in the data:

Table 12.2.4 General Trends Shown in the Radon Overlay Data in Devon and Cornwall

Proportion of homes with radon levels in excess of the Action Level	General trend
<1%	As land capability grade increases, the area of radon at this level decreases slightly
1 - 3%	As land capability grade increases, the area of radon at this level decreases considerably
3 - 10%	As land capability grade increases, the areas of radon at this level increases slightly
10 - 30%	As land capability grade increases, the area of radon at this level decreases slightly
>30%	As land capability grade increases, the area of radon at this level increases considerably

To summarise, the general trend - which proposes that as land capability for agriculture decreases the proportion of homes above the Action Level increases - appears to be true in Devon and Cornwall. Grade 1 - 2 land is closely correlated with areas where fewer than 3% of homes are above the Action Level. On the other hand Grade 5 land is very closely correlated with areas where more than 30% of homes exceed the Action Level. However the identification of these statistics does not necessarily prove that a correlation exists between land capability and radon levels. Moreover, the correlations can not prove that one causes the other. In order to demonstrate a causal relationship, more detailed tests would need to be carried out. For example, one test could involve altering the land capability and measuring the impact on known radon levels. The overlay statistics are a useful indication however that there is some association between the land capability for agriculture and estimated average radon levels in the home. This finding may not necessarily hold true in the other Affected Areas.

12.3 Somerset

12.3.1 What spatial patterns exist in the radon data set?

The county of Somerset was designated an Affected Area in 1990, following measurements by the NRPB which identified parts of the county as having more than 1% of homes above the Action Level. Certain parts of the county are perceived as being more affected than others and the spatial distribution of radon for the whole of Somerset is depicted on Figure 12.3.1. Somerset is not characterised by extremes. The vast majority of the radon (and land capability) measurements fall in to the middle categories with little divergence. This situation is exemplified in Tables 12.3.1 and 12.3.2.

There are two bands of relatively elevated radon levels (3 - 10% of homes above the Action Level) which are clearly visible on Figure 12.3.1. One band runs roughly north - south from Shepton Mallet and is associated with Carboniferous and Mesozoic Limestone bedrock. The elevated area in the far west of the county corresponds to the eastern part of Exmoor and the Brendon Hills. These areas are underlain by Devonian slates and shales. The low-lying peaty Somerset Levels are characterised by low radon levels, where less than 1% of homes are estimated to be above the Action Level.

Table 12.3.1 Summary of the Number of Homes Above the Action Level in Somerset

Radon Categories	Area (sq. km)	%
<1%	319	9
1 - 3%	1,698	50
3 - 10%	1,381	40
10 - 30%	42	1
>30%	0	0
TOTAL	**3,441**	**100**

Almost 99% of the five kilometre grid squares in the county, contain fewer than 10% of homes above Action Level. This is in stark contrast to Devon and Cornwall, where 51% of the county has radon levels in excess of 10%. No five kilometre squared area in Somerset has been identified as having more than 30% of homes above the Action Level and there are only two grid squares where the proportion of homes above the Action Level is estimated to exceed 10%. The area covered by each of the radon categories is shown in Figure 12.3.2 below. The skewness of the data is 0.615 indicating an asymmetric distribution, with the longer tail extending towards the higher radon levels.

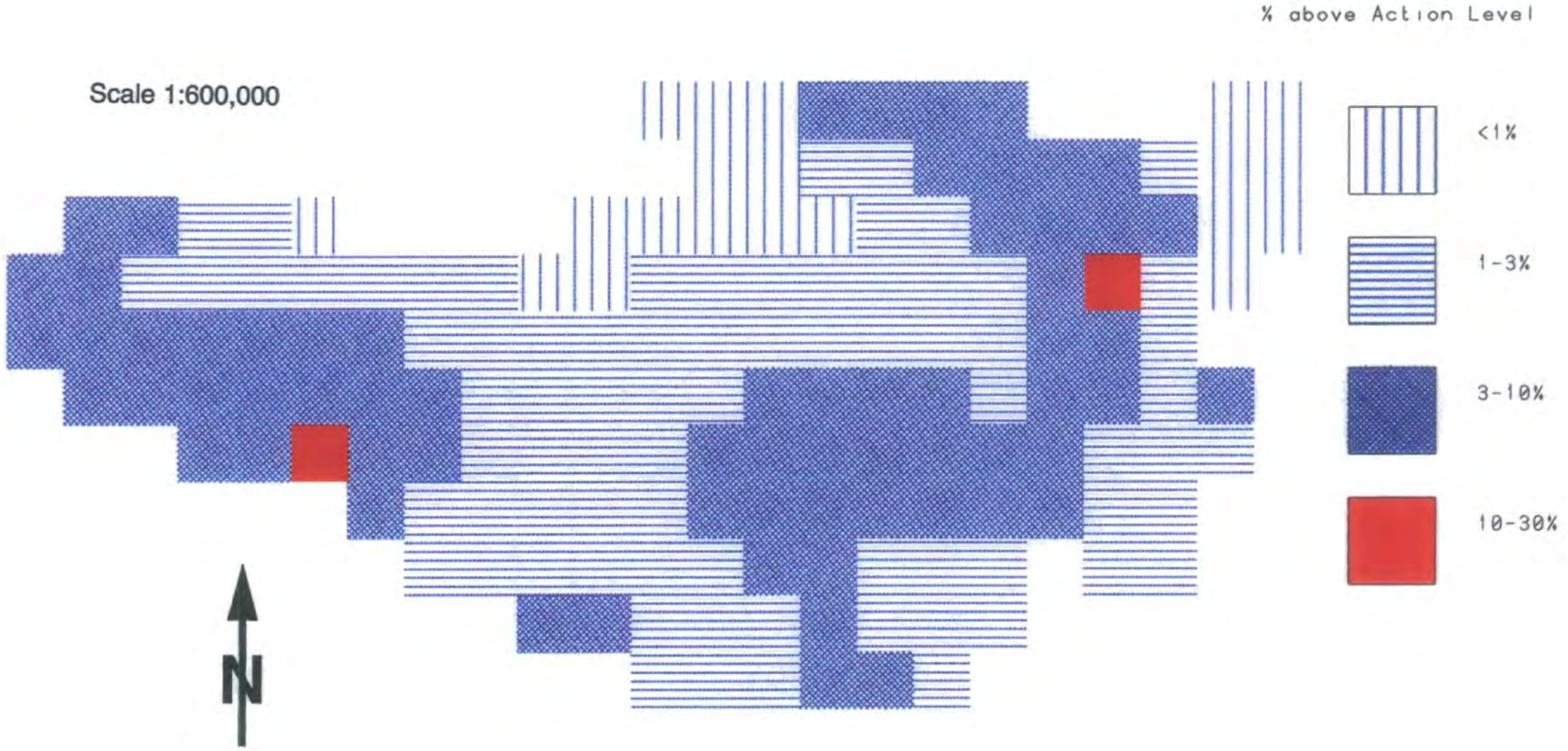
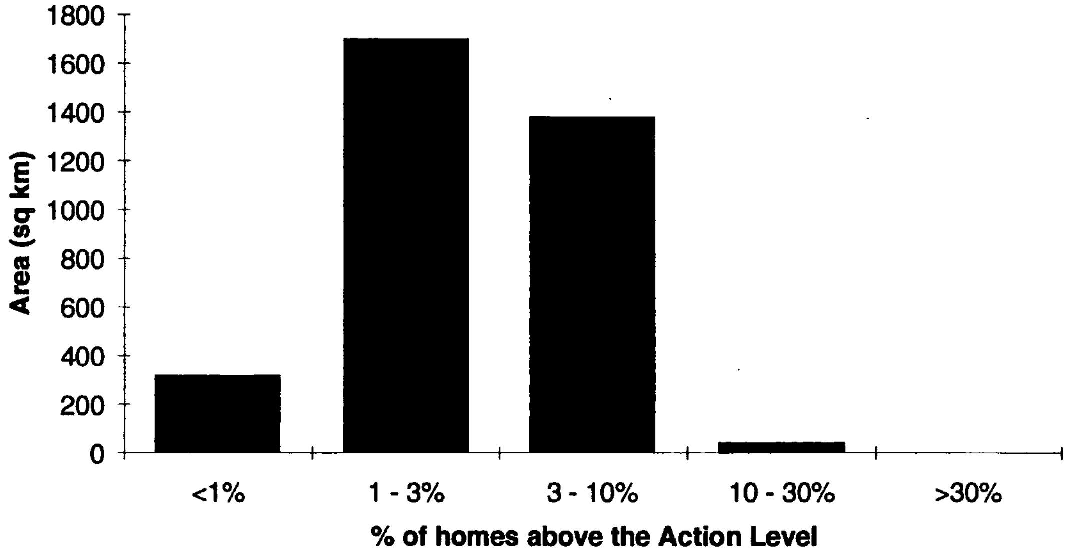


Figure 12.3.1 The spatial distribution of the estimated proportion of homes above the Action Level in Somerset

(Source: NRPB)

Figure 12.3.2 Total Area (in square kilometres) of each Radon Category in Somerset



12.3.2 What spatial patterns exist in the Land Capability data set?

Figure 12.3.3 shows the Somerset land capability coverage used in this thesis, which was digitised from the 1977 MAFF land capability map. Land capability in Somerset is dominated by Grade 3 land which accounts for 67% of the total area. This reflects the situation in England as a whole, although the dominance of Grade 3 land may be exaggerated by the fact that not all the minutiae of detail in the land classification maps were digitised. Small parcels of land were not digitised if they covered an area of less than three square kilometres. Land capability in Somerset is more homogenous than in Devon and Cornwall, so the effect of this simplification will be less marked than in the latter coverage. Errors during the digitising process, resulted in an area in the east of Somerset being erroneously classified as Grade 3 land instead of Grade 4 and the town of Chard (six square kilometres in area) was omitted altogether. Subsequent calculations however, using the revised figures (see Appendix 3), show an insignificant effect on the overall totals used in the analyses.

Figure 12.3.3 A land capability map of Somerset

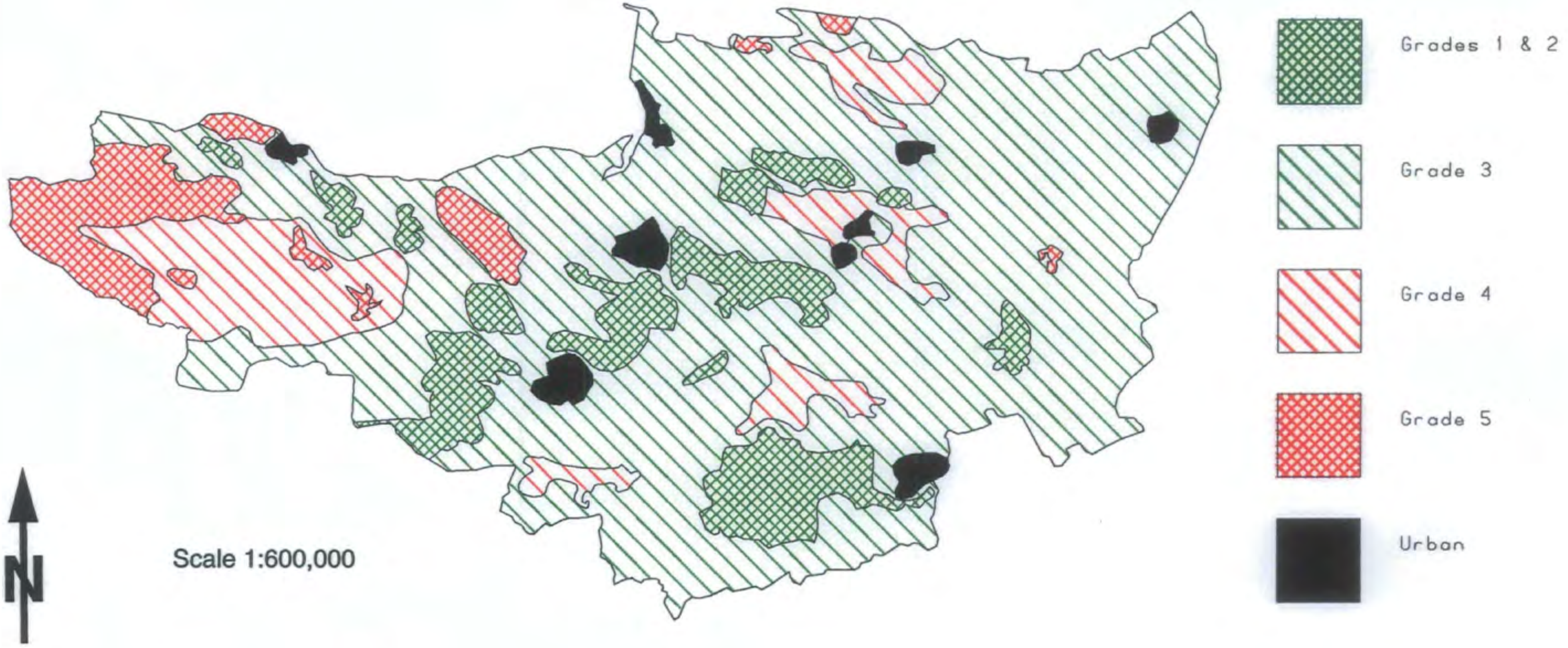
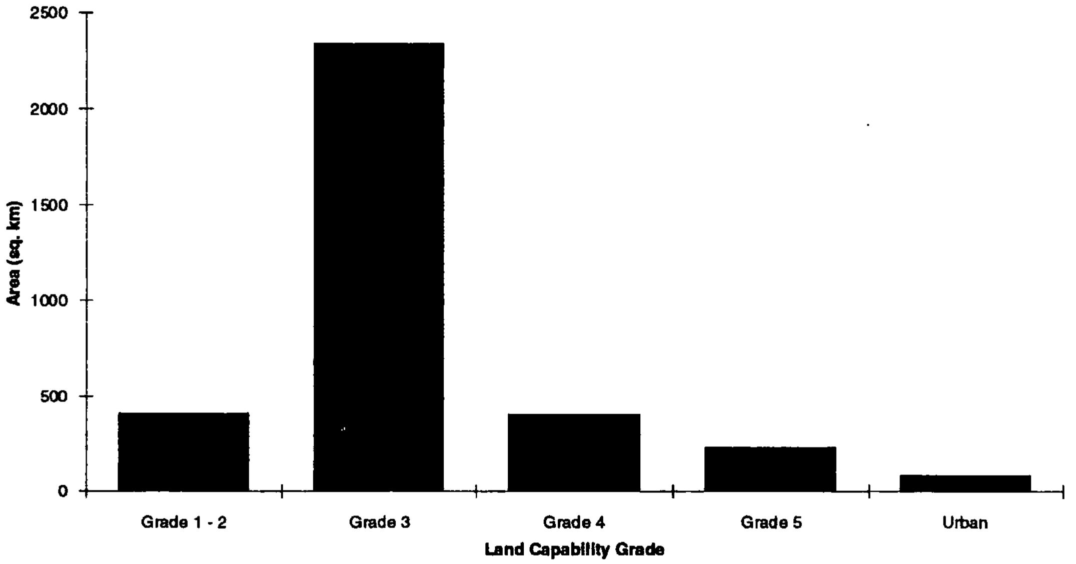


Figure 12.3.4. Total Area (in square kilometres) of each Land Capability Grade in Somerset



The positive skewness of two indicates an asymmetric distribution skewed towards the higher land capability categories (Grades 4 and 5). The general quality of land in Somerset is good and twelve percent of the county is classified as Grade 1 - 2 land (with a slightly higher percentage of Grade 2 than Grade 1 land). Grade 1 land has a strict classification regime and is confined to only 3% of the land area of the United Kingdom as a whole, the majority of which is located in the south east of England, more specifically The Fens.

Grade 4 and Grade 5 land are largely restricted to the hilly areas of the county namely, Exmoor, the Brendon Hills, the Quantocks and the Mendips. Urban land accounts for only 2% of the total area of Somerset. The mean size of the urban polygons in the land capability coverage is just over nine square kilometres; Taunton, Yeovil and Bridgwater being the only towns of any significant size.

Table 12.3.2 Summary of Land Capability Data for Somerset[2]

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
frequency*	13	1	5	9	9	37
sum (km²)	407.7	2,338.9	404.1	230.7	84.5	3,465.9
% of sum	12	67	12	7	2	100
mean	31.4	-	80.8	25.6	9.4	-
maximum	117.6	-	224.5	155.9	18.5	-
minimum	4.6	-	22.5	2.9	4.1	-

* frequency of polygons in the digital coverage

12.3.3 What spatial patterns exist between radon gas and land capability?

Figure 12.3.5 is a map of Somerset, produced in ARC/INFO, showing the result of the overlay between the land capability coverage and the radon grid. The map shows all urban polygons produced in the overlay which have a radon level between 1 and 3%, as well as polygons with Grade 4 land and a radon level of between 3 and 30%. Table 12.3.3 is a statistical summary of the overlay between radon and land capability.

Table 12.3.3 Summary of the Overlay Between the Estimated Proportion of Homes Above the Action Level and Land Capability in Somerset*

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
<1%	1	13	0	2	24
1 - 3%	**69**	**50**	26	44	**73.5**
3 - 10%	30	36	**72**	**54**	2.5
10 - 30%	0	1	2	0	0
>30%	0	0	0	0	0

* These percentages were calculated using the area of each radon category as a proportion of the total area of each land capability grade.

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
Skewness	1.09	0.09	1.35	0.10	1.48

The lowest radon level (<1% of homes above the Action Level) is associated almost entirely with Grade 3 (good to moderate quality agricultural land), which is the most abundant

---

[2] The totals in this table may not correspond to those used in the land capability and radon overlay, because of slight differences in area between the two coverages.

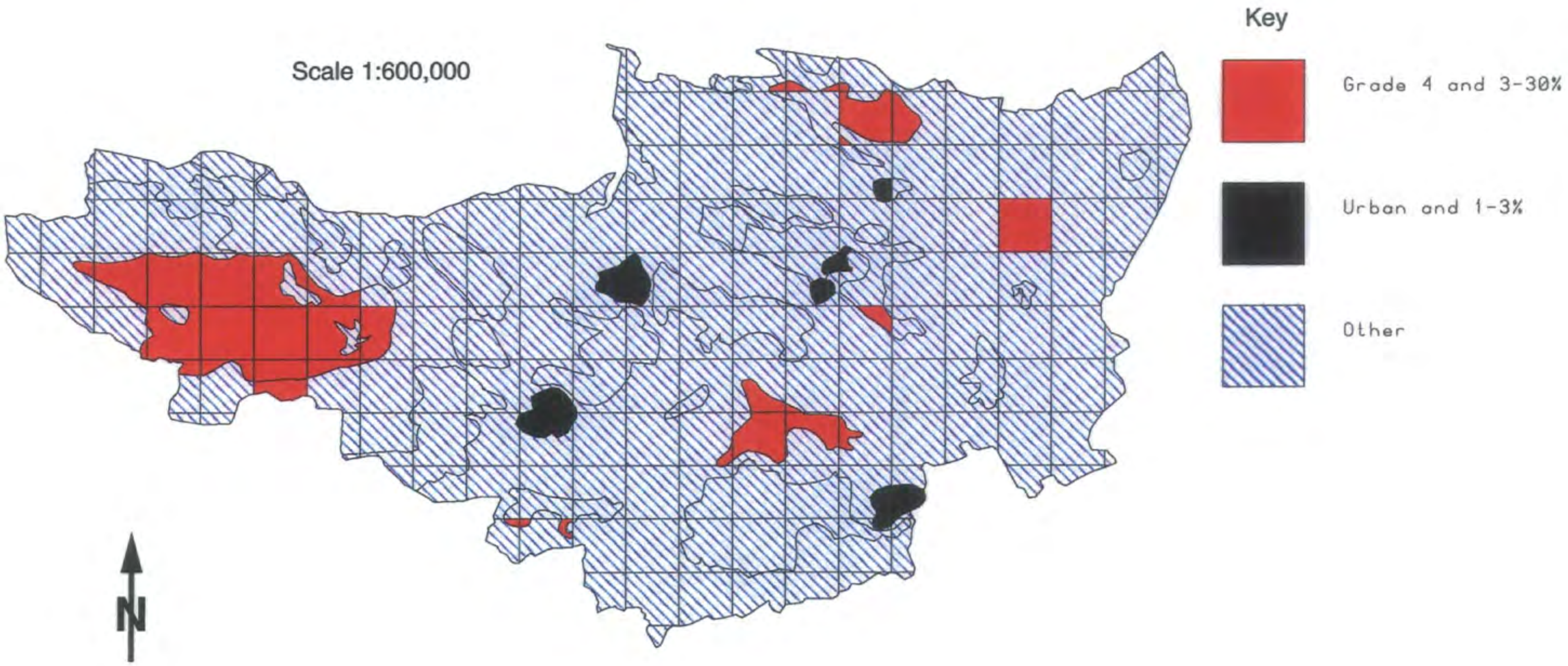
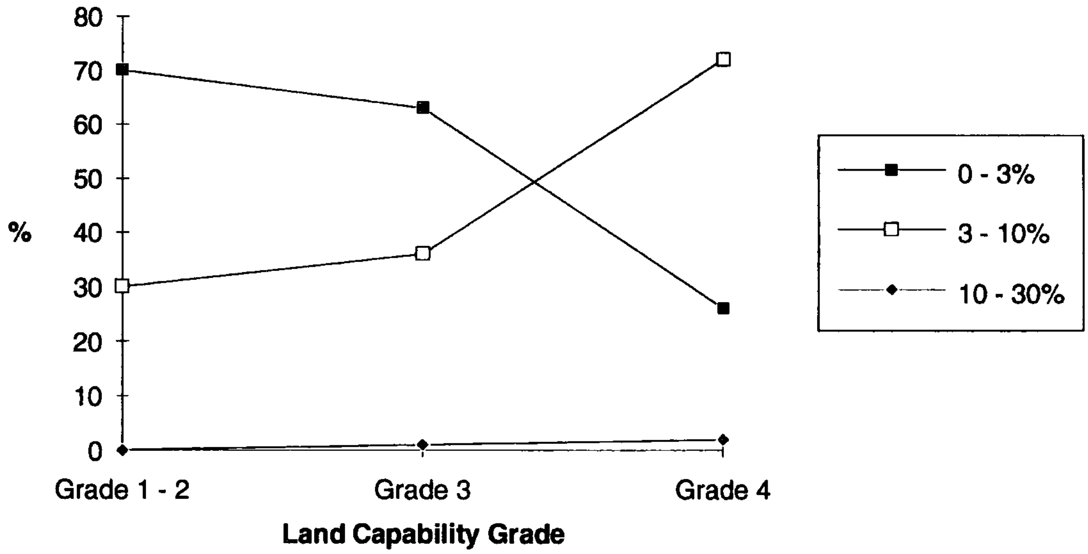


Figure 12.3.5 Radon gas and land capability in Somerset

land capability category in Somerset. This association may be due to a number of factors (*e.g.* soil type, ground permeability, geology), but is influenced by the fact that a large proportion of the low radon area corresponds to the low-lying coastal zone of the Parrett Estuary, where low relief and the resulting poor drainage conditions (a result of the intercalated peats) prevent intensive agricultural practices. The dominance of Grade 3 land, which accounts for 67% of the county area, will also affect the results of the overlay.

There is a strong correlation between Grade 1 - 2 land and radon level 1 - 3% but only a very weak correlation between <1% and Grade 1 - 2% a similar situation to that found in Devon and Cornwall and Somerset. The skewness of land capability Grades 1 - 2 indicates an asymmetrical distribution skewed towards the higher radon levels, although 70% of the area of this grade is associated with radon levels where less than 3% of homes are estimated to be above the Action Level. Grade 3 land however, is more symmetrical in distribution and is only slightly skewed towards the higher radon levels. This is a result of greater spread between the radon categories, so that although 63% of the area of this grade is associated with radon levels below 3%, 37% is associated with radon levels of 3 - 30%. Figure 12.3.6 graphically displays the relationship between radon and land capability.

Figure 12.3.6 The Relationship Between Land Capability and Radon Levels Between 0 and 30% in Somerset



The low radon levels show an inverse correlation, so that as land capability decreases the proportion in the 0 - 3% radon levels also decreases. However, for the higher radon levels (3 - 30%) the correlation is a positive one, so that as land capability decreases the proportion in these radon categories increases. This fits the general trend being investigated (see Chapter 12.1). Over half of the area of Grade 5 is associated with the higher radon levels of 3 - 10%, a situation that it has in common with Grade 4 land. However, a high proportion of the grade is also associated with the lower radon level of 1 - 3%. Therefore, whilst the proclivity towards the higher radon levels is maintained, the difference between the two radon categories is greatly reduced. Reasons for the lower radon levels may be varied, but it is useful to point out that the Quantock and the Mendip hills (both include areas of Grade 5 land) are characterised by relatively low radon levels. The bedrock which forms these hills - in the Quantocks it is Devonian Grits and in the Mendips it is Carboniferous Limestone - is not believed to contain very high levels of uranium or radium, the precursor to radon.

Two percent of Grade 4 land and 1% of Grade 3 land is associated with radon levels of 10 - 30%, the highest bracket of radon found in Somerset. In total there are just two five kilometre grid squares, one located in the Haddon Hill area, in the generally poor quality land, of western Somerset and the other in the east of Somerset near Shepton Mallet. These two areas have little in common other than their location at the edge of upland areas, approximately 200 metres above sea level. Haddon Hill, at the south-eastern tip of Exmoor and the Brendon Hills, is on Devonian rock and is Grade 4 land. The area around Shepton Mallet is at the south-eastern end of the Mendips and has a land capability Grade of 3.

Urban areas are strongly correlated with the areas of 1 - 3% of homes above the Action Level. Urban land accounts for just 2% of Somerset however and the radon level 1 - 3% accounts for 50% so the result may be due to the imbalance between these two proportions. In addition. the relationship may be partially attributed to the fact that all of the significant urban areas (in terms of size) are located in low relief areas, surrounded by land that is either Grade 1 - 2 or 3.

12.4 Northamptonshire

12.4.1 What spatial patterns exist in the radon data set?

In 1990, as a result of measurements carried out by the NRPB, the county of Northamptonshire was designated an Affected Area. Virtually the whole county has in excess of 1% of homes above the Action Level and an area in the centre of the county (accounting for 12.5% of the total area) has an estimated 10 - 30% of homes above the Action Level. This central area contains the towns of Northampton and Kettering. The spatial distribution of radon for five kilometre grid squares is shown in Figure 12.4.1.

The radon measurements show a distinct pattern, with the highest values along a northeast-southwest trend, which is attributable to the regional strike of the rocks. The bedrock types associated with much of the highest radon area, are phosphatic ironstones, sandstones and limestones of the Lower Estuarine Series (Miles *et al.*, 1992b). Table 12.4.1 below is a summary of the data contained in Figure 12.4.1 and serves to highlight the variation between the proportions in each radon category.

Table 12.4.1 Summary of the Number of Homes Above the Action Level in Northamptonshire

Radon Categories	Area (square km)	%
<1%	10	0.5
1 - 3%	447	19
3 - 10%	1,586	68
10 - 30%	292	12.5
>30%	0	0
TOTAL	2,335	100

The distribution of the proportions in each radon category is roughly symmetrical about the mode (3 - 10%). Only a very small proportion of Northamptonshire (two five kilometre grid squares in total) has radon levels of less than 1% and there are no grid squares with more than 30% of homes exceeding the Action Level. Therefore, 99.5% of the area of Northamptonshire has between 1% and 30% of homes above the Action Level. Nineteen percent of the county has between 1 and 3% of homes above the Action Level and the grid squares which make up this category are found mainly in the north and along the western side of the county on the border with Leicestershire and Warwickshire.

The most notable category in Northamptonshire is the 3 - 10% grouping which accounts for 68% of the total county area (see Figure 12.4.2 below). This category makes up a large part of the central area of the county and its spatial distribution is relatively uniform and compact.

Figure 12.4.2 Total Area (in square kilometres) of each Radon Category in Northamptonshire

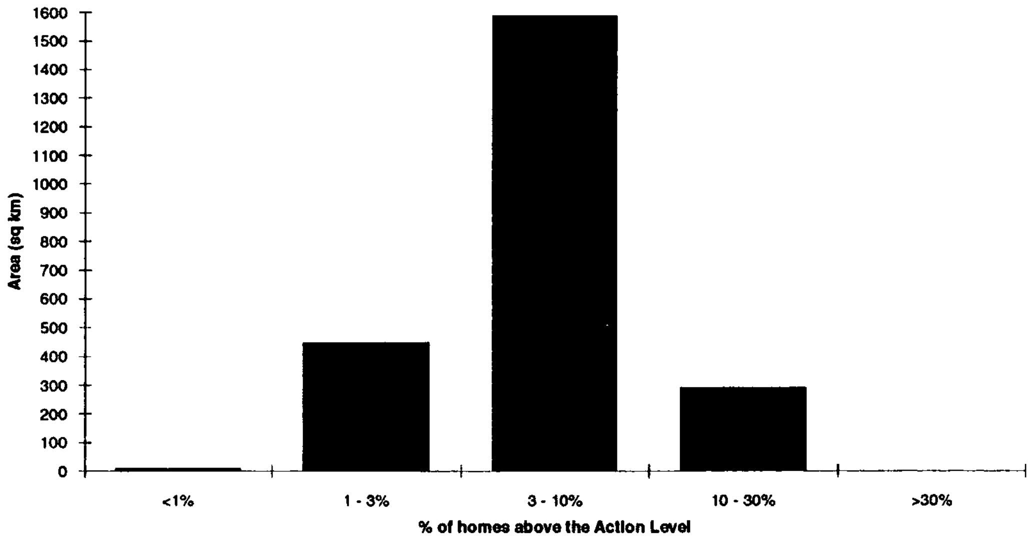
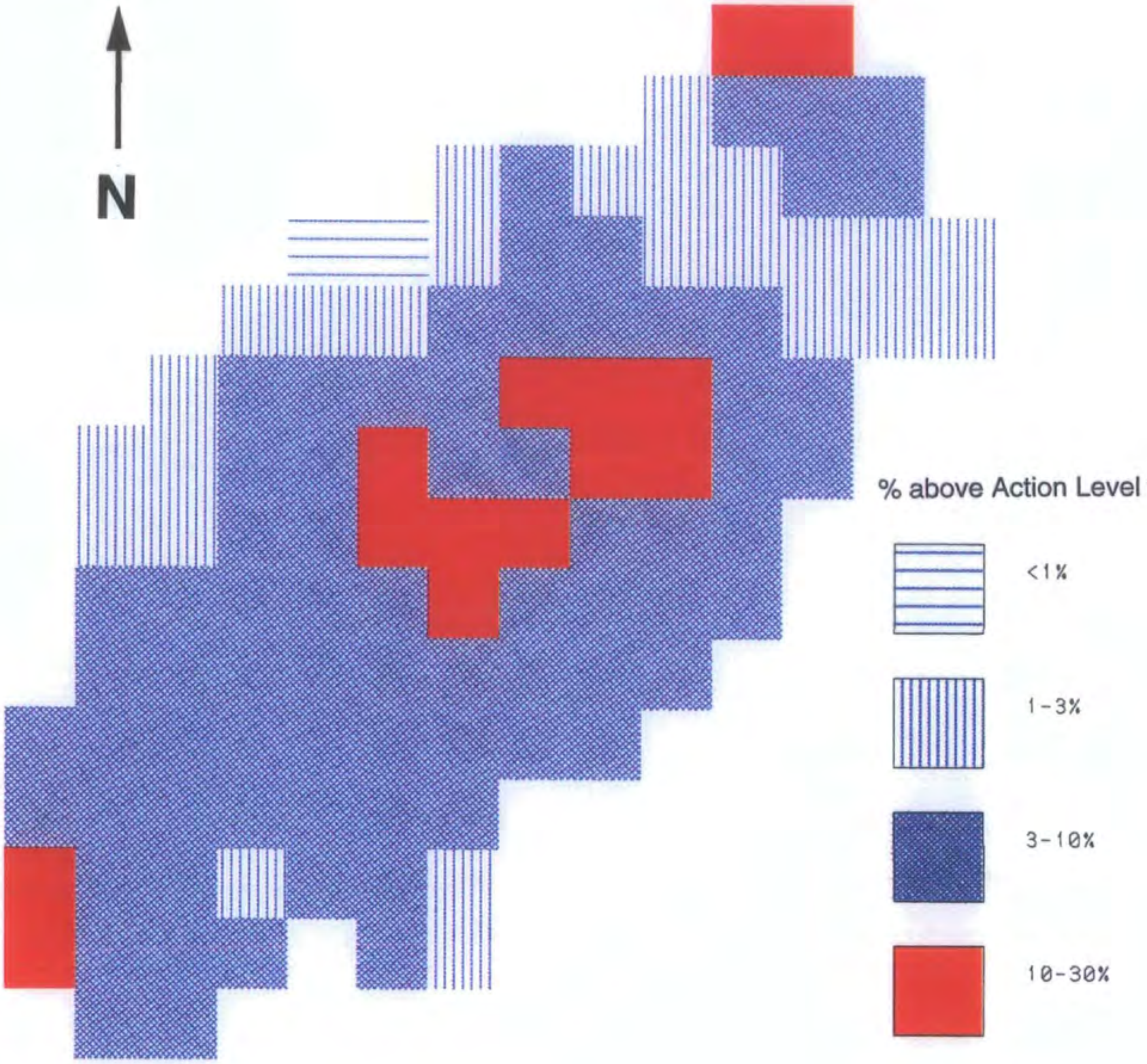


Figure 12.4.1 The spatial distribution of the estimated proportion of homes above the Action Level in Northamptonshire

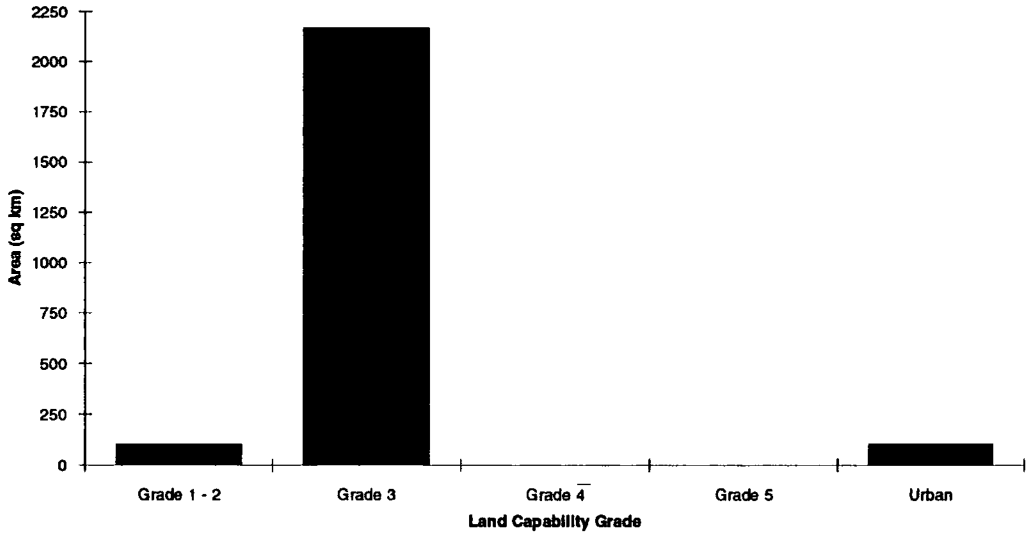


(Source: NRPB)

12.4.2 What spatial patterns exist in the Land Capability data set?

Figure 12.4.3 shows the spatial pattern of land capability for Northamptonshire (digitised from the 1983 MAFF Eastern Land Capability map). The pattern of land capability is dominated by Grade 3 land which makes up 92% of the county area (see Figure 12.4.4). Although Grade 3 land accounts for 50% of the total area of England and Wales (Mather, 1986), this figures relates to the upland regions of the country as well as the lowlands. The proportion of good quality agricultural land is much higher in the south and east than it is in the north and west, consequently this seemingly exaggerated figure for Northamptonshire is partly a reflection of the above average quality of agricultural land in the south-east of England. An example of this can be found in the neighbouring county of Cambridgeshire where over half the total area is classified as Grade 1 or 2 land. The absence of any poor quality agricultural land (Grade 4 and 5) in Northamptonshire is also a function of the location of the county in the low-lying, fertile land of the south-east. The aerial photograph in Figure 3.4 (Chapter 3.2) typifies the largely agricultural nature of the county.

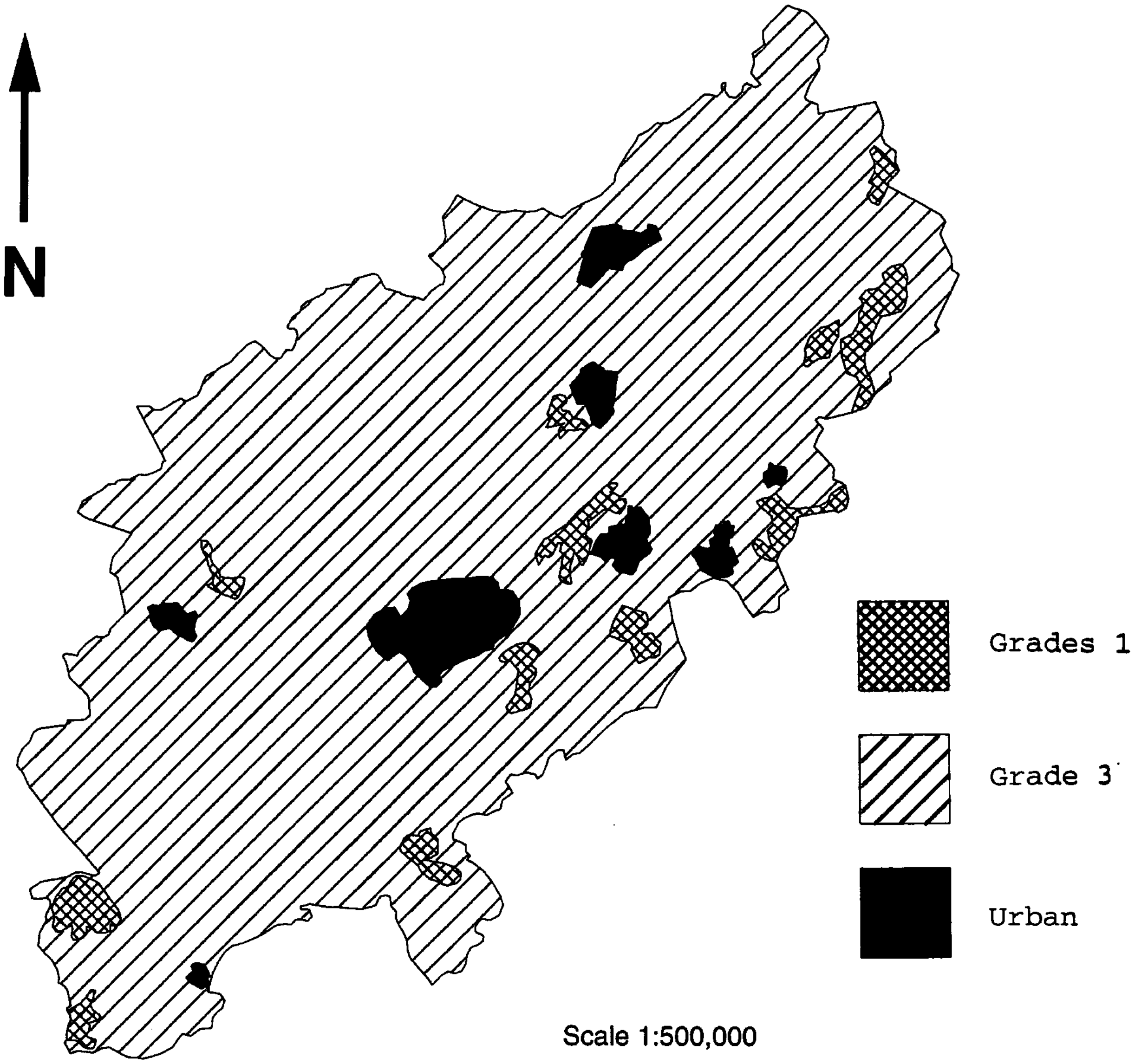
Figure 12.4.4 Total Area (in square kilometres) of Each Land Capability Grade in Northamptonshire



Three percent of Northamptonshire is classed as very good quality agricultural land. This land is dispersed, but is more frequently found along the eastern edge of the county and in the lower lying land of the Nene valley. These areas are intensively used for agriculture and are primarily arable land.

The remaining 5% of the area of Northamptonshire is classified as urban land - Northamptonshire, Wellingborough, Corby and Kettering being the main towns. The mean size of urban areas in Northamptonshire is 13 square kilometres which is a similar size to urban areas in Derbyshire but at least three square kilometres larger than the mean in Somerset and

Figure 12.4.3 A land capability map of Northamptonshire



Devon and Cornwall. Possible explanations for this include the counties' proximity to the south-east and the good communication lines (*e.g.* the M1 and A6 primary roads and Network South-East rail lines) which link it to London. The propinquity to the national capital has encouraged the growth of Northamptonshire's market towns by providing a large outlet for agricultural and manufactured products.

Table 12.4.2 Summary of Land Capability Data for Northamptonshire[3]

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
frequency*	12	1	0	0	8	21
sum (km^2)	102.6	2,168.4	-	-	104.7	2,335.3
% of sum	3	92	-	-	5	100
mean	8.5	-	-	-	13.1	-
maximum	17.5	-	-	-	51.2	-
minimum	4.0	-	-	-	1.9	-

* frequency of polygons in the digital coverage

12.4.3 What spatial patterns exist between radon gas and land capability?

The map in Figure 12.4.5 shows polygons with Grade 2 land capability and radon levels of 1 - 3%, as well as those urban polygons with high radon levels (3 - 30%). Table 12.4.3 below provides a statistical summary of Figure 12.4.5, Grades 4 and 5 have been omitted because they do not feature in the land capability maps. If the most significant category (in terms of area) is highlighted for each of the three grades, there appears to be no significant relationship between poor land capability and high radon levels or between good land capability and low radon levels. The largest radon category associated with all three land capability grades is 3 - 10%. This result is not surprising as the radon level 3 - 10% accounts for over two thirds of the total area of the county.

---

[3] The totals in this table may not correspond to those used in the land capability and radon overlay, because of slight differences in area between the two coverages.

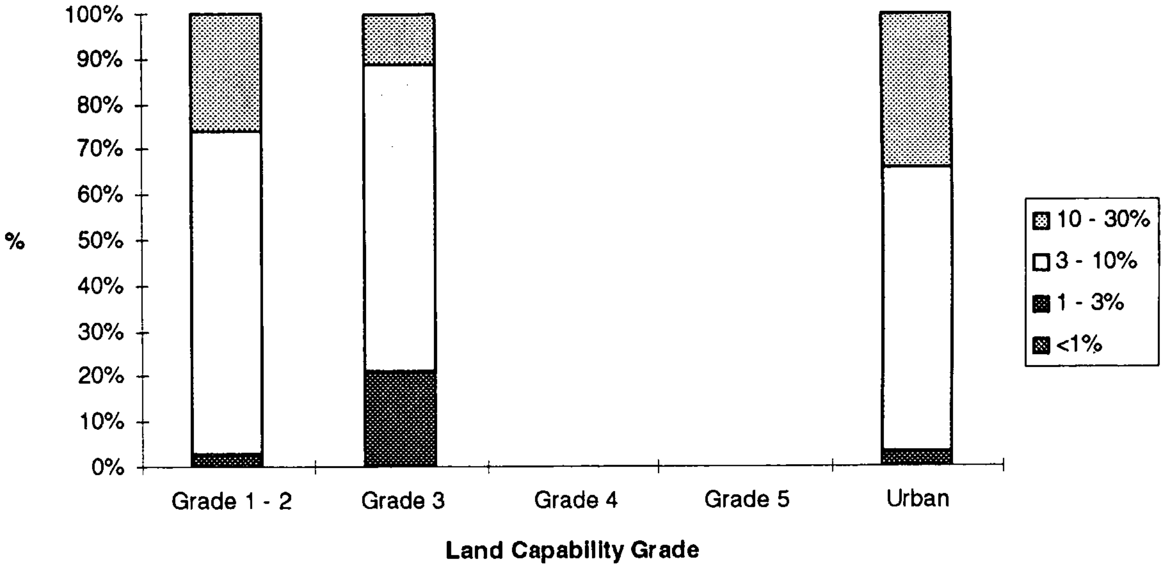
Table 12.4.3 Summary of the Overlay Between the Estimated Proportion of Homes Above the Action Level and Land Capability in Northamptonshire*

	Grade 1 - 2	Grade 3	Urban
<1%	0	0.5	0
1 - 3%	3	20.5	3
3 - 10%	71	68	63
10 - 30%	26	11	34
>30%	0	0	0

* These % are calculated using the area of each radon category as a proportion of the total area of each land capability grade

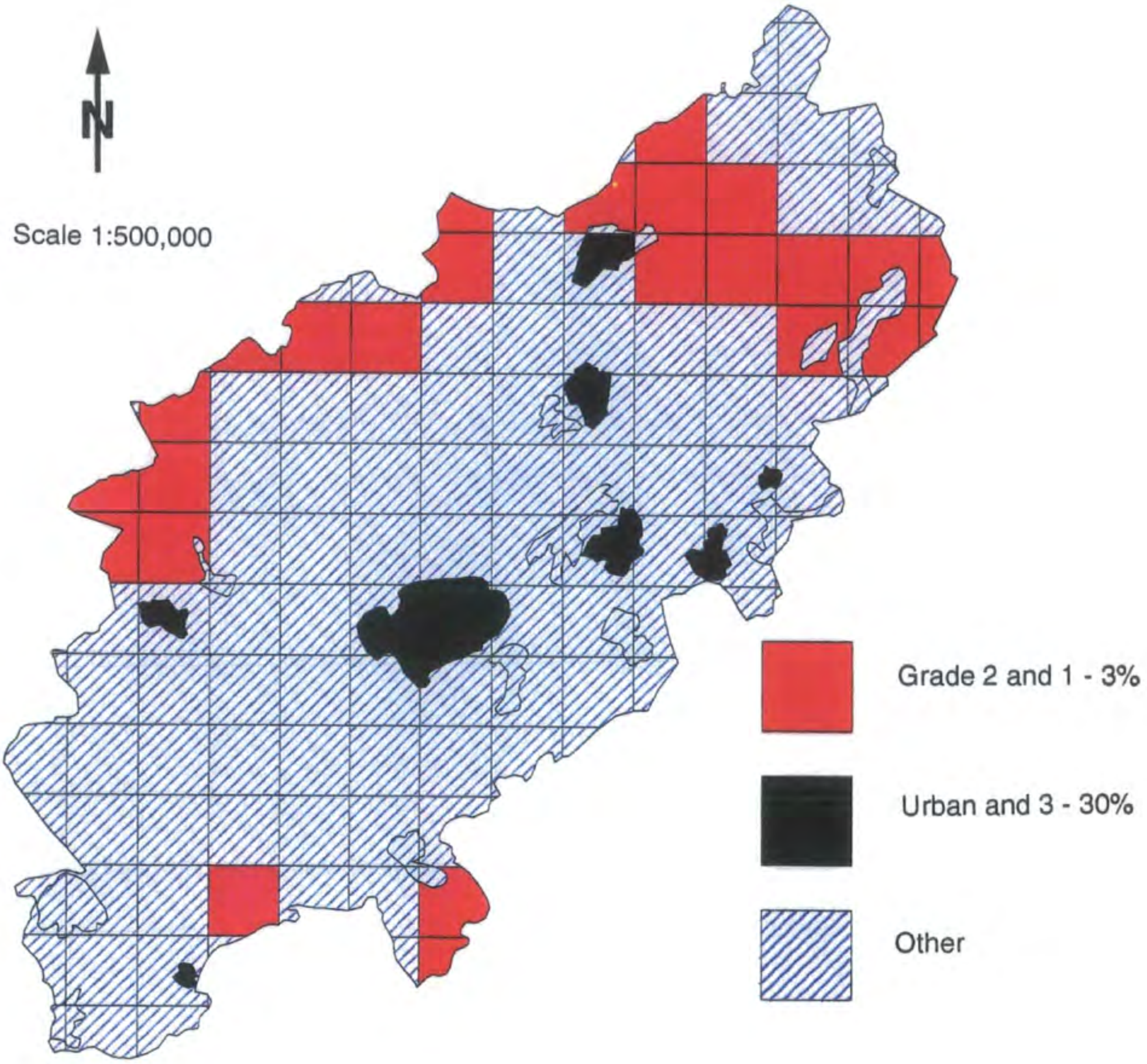
The relationship between the three grades can be seen on Figure 12.4.6. The graph extends only as far as the 10 - 30% bracket because no grid squares in Northamptonshire have yet been identified as having more than 30% of homes above the Action Level and to extend the graph would give a false picture of the distribution of radon levels.

Figure 12.4.6 The Relationship Between Land Capability and the % of Homes Above the Action Level in Northamptonshire



The majority of Grade 1 - 2 land (all Grade 1 - 2 land in Northamptonshire is in fact Grade 2 land) is unusually associated with the higher radon levels (3 - 30%), whilst Grade 3 land is associated more with the lower levels (<1 - 10%). This is a reversal of the usual situation and the reasons for the lower radon levels in the poorer agricultural areas are not immediately apparent. In the absence of any Grade 4 or Grade 5 land, it could be anticipated that Grade 3

Figure 12.4.5 Radon gas and land capability in Northamptonshire



land (as the poorest quality land found in Northamptonshire) would act as a substitute and associate with the higher radon levels. This shift in the pattern of association has not happened.

Radon variability depends partly on local variations in radium concentrations and ground permeability which can be affected by soil type and farming practices. Alternatively, it may be that the slight differences in the definition of land capability between Grade 2 land and Grade 3 land are not distinct enough to affect the radon levels. In addition, there is not a significant area of Grade 2 land in the north or the west of the county where radon levels are generally lower. More significant however is the fact that the majority (92%) of the county is categorised as Grade 3 land and it is therefore significantly more probable that the low radon areas will be associated with this grade. Other theories (for example, the influence of relief) cannot be examined without further investigation.

Urban areas are characterised by relatively high radon levels in Northamptonshire, in comparison to the other Affected Areas (see Chapter 13). Ninety-seven percent of urban areas have more than 3% of homes above the Action Level and 34% of urban land is characterised by radon levels of 10 - 30%. The corresponding totals for Somerset are 3% and 0% and for Derbyshire 19% and 3%. Devon and Cornwall have altogether higher radon levels (over 30% in some squares), but the overall urban totals for these two counties remain lower than in Northamptonshire, at 73% and 13%. Possible causes for the elevated levels in urban areas may stem from the location and site of the towns in relation to the underlying geology and the natural environment, but other factors may also influence domestic radon levels. The greater mean size of urban areas in Northamptonshire compared to Somerset and Devon and Cornwall, may be a contributory factor - by covering a larger area, the towns are more likely to be located on a number of different rock and soil types. The higher radon levels may be the result of increased sub-surface disruption due to the construction of buildings and the presence of underground utility pipes which enable a more rapid diffusion rate of radon gas to the surface. On the other hand, the cause may be socio-economic and may be influenced by population density, building type or occupancy factors, for example. Northamptonshire is a more affluent county than any of the other Affected Areas and the resulting social profile of the county may be one reason for the higher radon levels in urban areas; for example, in research carried out by Wolff (1991) he found that higher radon levels reflected higher socio-economic status in England as a whole. Urban radon levels in the five Affected counties do not conform to a pattern, but (with the notable exception of Derbyshire) are more closely associated with radon levels in excess of 1% of homes above the Action Level (see Chapter 13). High radon levels are not associated with any of the major urban areas in the United Kingdom.

In summary, the relationship between radon and land capability in Northamptonshire is affected by the dominance of Grade 3 land. The magnitude of this dominance is problematic to assess. The results of the overlay do not fit the general trend identified in Devon Cornwall and Somerset - indeed the opposite appears to be the case. Further studies analysing variations in land use in more detail need to be carried out for the county, for example using the Land Cover Map of Great Britain data for one kilometre grid squares.

12.4.4 Social Structure and Radon in Northamptonshire

So far, the impact of purely physical factors (namely land capability) on radon levels has been discussed. However, there are also a number of human factors which can influence domestic radon levels, both indirectly and directly. Building type, construction and construction material are believed to directly influence radon levels, whereas occupancy habits and various social factors may be indirect influences. In Northamptonshire, the impact of social aspects of the environment on radon levels has been discussed. Three questions have been posed:

1. What is the relationship between population density and radon levels?
2. What is the relationship between social class and radon levels?
3. What is the relationship between the proportion of households containing only pensioners and radon levels?

The data for this chapter have been taken from the 1991 Census of Population County Monitor (Part Two) for Northamptonshire. At present to infer relationships between census data and other aggregate data is subject to the ecological fallacy problem. For this reason no conclusions about individual households will be drawn from the results of the overlays between radon and the social data. Any correlations that exist can only be shown to occur at ward level or higher (*e.g.* district level), depending on the level of aggregation. There remain formidable methodological and interpretative problems with using aggregate data, and much research into the problem is still required (see Chapter 10). However, at the present time analysing the relationships shown to occur at ward or district level permits the identification of general patterns of association. More detailed localised studies can then be carried out in the areas identified as being of interest, using data aggregated for the smallest possible units (*i.e.* enumeration districts or postcode units).

There is scant reference in the literature to research carried out between radon and social factors. An exception is provided by Wolff (1991), based at the Department of Clinical Pharmacology, University College London. Wolff used 1981 Census data and found that detached houses have higher radon levels than semidetached or terraced houses, which in turn have higher levels than flats or maisonettes. Lower domestic radon levels seem to reflect greater socio-economic deprivation at the county level, lower population change and higher population density, whereas higher radon levels reflect the converse. He concluded therefore, that it is possible that higher domestic radon levels merely reflect higher socio-economic status. His correlation coefficients for socio-economic variables from the 1981 Census and domestic radon levels for twenty-two counties in England and Wales are as follows (Wolff, 1991):

Table 12.4.4 Correlation Coefficients For Socio-Economic Variables and Domestic Radon Levels (After Wolff, 1991)

Variable	Correlation Coefficient*
Unemployment	-0.44 (<0.025)
Car ownership	0.54 (=0.005)
Overcrowding	-0.29 (NS)
Home ownership	0.11 (NS)
Social Class I and II	0.42 (<0.05)
Social Class IV and V	-0.41 (<0.05)
Population change	0.42 (=0.025)
Population density	-0.75 (<0.0005)

*Spearman's rank with arithmetic mean radon level

Wolff's research was carried out at county level, and the result is that the resolution of the data is very coarse. His conclusions are interesting however, because they relate to the whole of England and Wales and are not restricted to radon Affected Areas. His findings for population density and social class differ from the findings in this thesis for the county of Northamptonshire, although comparisons between the results are problematic because of the different nature and scale of the analyses and the discrepancies between the ages of the census data that were used.

The variables that have been analysed in this thesis for Northamptonshire using 1991 Census data, are population density (at ward and district level), social class (at district level) and the percentage of households containing only pensioners (at ward level). Each overlay has been analysed separately and the results are discussed below. Figure 12.4.7 shows the districts of Northamptonshire in relation to the main towns and the surrounding counties.

12.4.4.1 What is the relationship between population density and radon levels?

Population density (residents per hectare) for each of the 148 wards in Northamptonshire was mapped and an overlay was carried out in ARC/INFO using the radon data for five kilometre grid squares (see Table 12.4.5 below). Population density for wards varies greatly in Northamptonshire, ranging from 0.2 residents per hectare (in the wards of Barnwell and Everdon) to 66.4 in St. Andrews ward in Kettering (see Figure 12.4.8). The density categories that were used in the analysis were chosen in order to provide an even number of counts in each category, thus the range of values in each category varies. The reasons for grouping the data are to provide a manageable data set and to enable correlations between radon levels and population density to be identified. However, the process of grouping the data inevitably obscures the variety inherent in the range of population density figures (see Chapter 10).

Figure 12.4.7 Northamptonshire Districts: 1991 Census

(Source: OPCS, 1993f)

In Northamptonshire, the dominance of the radon category 3 - 10% (69% of the total area) distorts the distribution of the data (see Table 12.4.5). Table 12.4.5 shows the proportion in each category as a proportion of the total area of the county. It is clear that the radon categories <1% and >10% account for considerably less area in the county than the two other categories. On the other hand, the lower densities account for a much greater proportion of the total area, as is shown in the last column (row totals) of the table below. Such a skewed distribution is problematic when assessing the relationship between population density and radon. Relationships that are identified using these raw data are likely to be a function of the dominance of the category under consideration, rather than a true association.

Table 12.4.5 Summary Totals of the Overlay Between Population Density (wards) and Radon Levels in Northamptonshire (before transformation)

	<1%	1 - 3%	3 - 10%	>10%	ROW TOTALS
0.0 - 0.9p/ha	0.43	16.67	45.29	0.80	63.19
1.0 - 2.9	0	3.85	16.83	2.47	23.15
3.0 - 19.9	0	0.79	8.61	1.43	10.83
20.0 - 49.9	0	0.01	1.34	1.10	2.45
>50	0	0.01	0.12	0.27	0.4
COLUMN TOTALS	0.43	21.33	72.19	6.07	-

In order to reduce the effects of the dominant radon and population density categories, a simple transformation was carried out. The summary figures in Table 12.4.6 show the percentage of each category as a proportion of the total area of each radon level (column total). This method produces a practical set of summary statistics, free from the effects of dominance, which have been used to assess the relationships between population density and the proportion of homes above the Action Level. The column totals in Table 12.4.6 may not add up to 100 due to the effects of rounding. Appendix 6 gives the raw data and summary statistics for the overlay between radon and population density for wards.

Figure 12.4.8 Population density for wards in Northamptonshire (1991 Census)

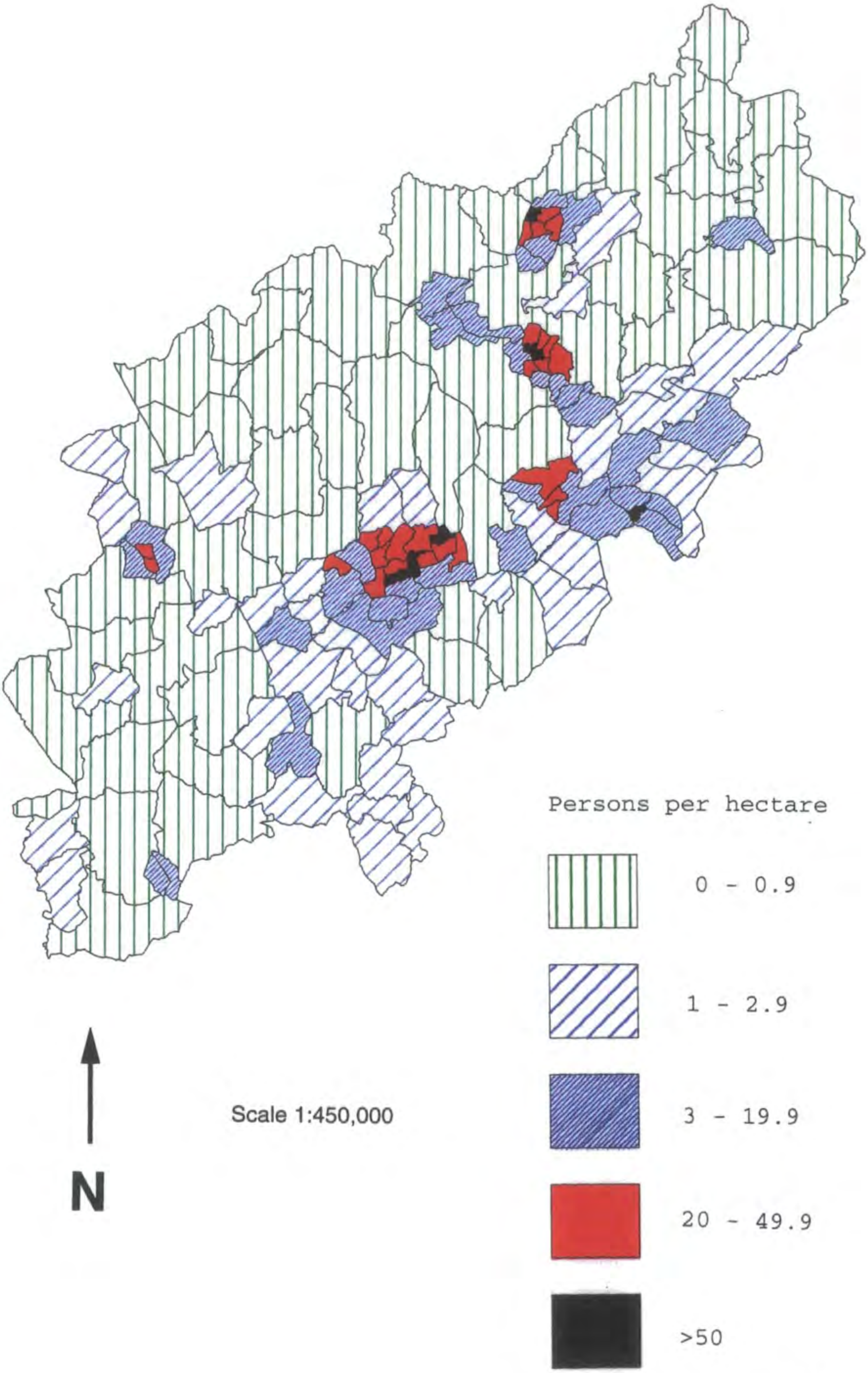


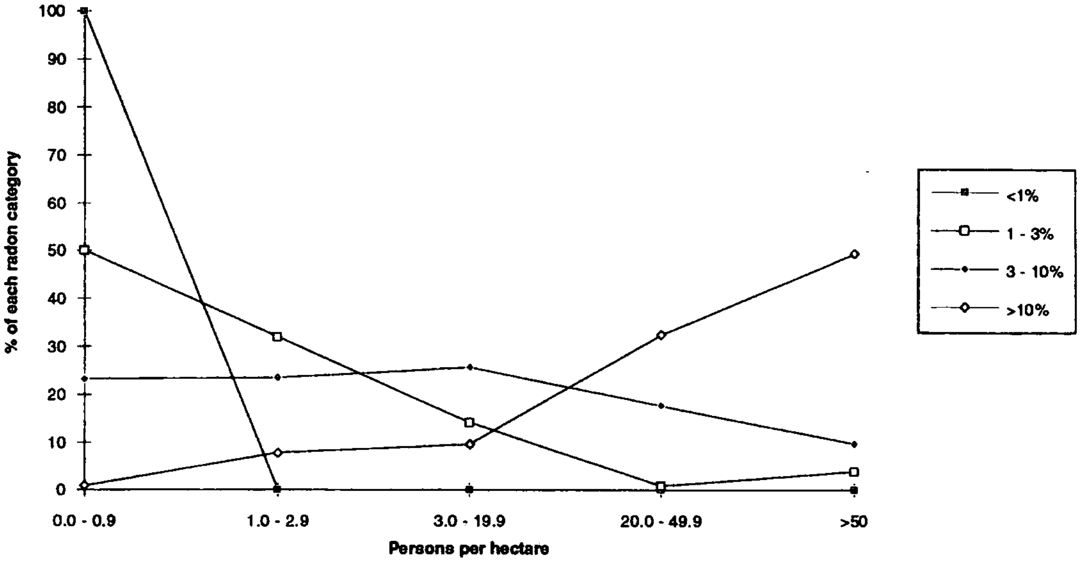
Table 12.4.6 Summary of the Overlay Between Population Density (wards) and Radon Levels for Wards in Northamptonshire (after transformation)

	<1%	1 - 3%	3 - 10%	>10%
0.0 - 0.9	**100**	**50**	23	1
1.0 - 2.9	0	32	24	8
3.0 - 19.9	0	14	**26**	10
20.0 - 49.9	0	1	18	33
>50	0	4	10	**49**

When analysing the results of the overlay (as shown in Table 12.4.6) it is important to note that the range in each population density category is uneven. An approximately equal number of counts were divided into each of the categories. The choice of the number of categories and the system chosen to apportion the data, inevitably affects the results given in Table 12.4.6. The division of the data was undertaken for purely pragmatic reasons and whilst it does raise some important methodological issues, these will not be discussed further here (but see Chapters 4 and 10.3.5).

The data shown in the table above (the largest category in each radon group is highlighted in bold), indicate that a positive relationship exists between high radon levels and high population density, at ward level in Northamptonshire. The highest radon category (>10% of homes above the Action Level) is correlated with polygons having a population density of more than 50 persons per hectare. This result agrees with that found in section 12.4.3, where a high proportion of urban land was found to correspond to areas with between 10 and 30% of homes above the Action Level. The lack of detail portrayed in the land capability coverage and the fact that the land capability data are now out of date confound the results identified in 12.4.3, but the overall conclusions are borne out by the findings of this section; namely that greater population density at ward level is correlated with areas having in excess of 10% of homes above the Action Level. The converse is also true, so that low population densities are, on the whole, more likely to be associated with the lower radon levels. These results are displayed on Figure 12.4.9 below.

Figure 12.4.9 Population Density (wards) and Radon Levels in Northamptonshire



The conclusion that higher radon levels are correlated with higher population densities contradict Wolff's (1991) findings. He concludes that lower domestic radon levels reflect higher population density. Wolff's conclusions relate to the whole of England and Wales however, and the findings in Chapter 13 (a comparison of the Affected Areas) show that the relationship between urban land and radon level is liable to differ widely. Northamptonshire appears to show a similar pattern to that found in urban land in Devon and Cornwall and it may be that other factors (either human or physical or both) are causing high domestic radon levels in the towns of these counties.

Population density for the seven districts of Northamptonshire was also mapped to provide a comparison with the ward data (see Figure 12.4.10) and to determine the effects of scale on the analysis. A summary of the results of the overlay (which have been calculated and transformed using the same method as above) are shown in Table 12.4.7. Once again, the strongest association in each radon category has been highlighted. The increased generalisation of the data has reduced the range of population densities from 0.2 - 66.4 (ward data) to 0.9 - 22.4 (district data). The district of Northampton has the highest population density (at district level) and Daventry the lowest.

Essentially the differences found in the analyses between the two referencing systems (ward and district) stem from the different spatial zones and the increase in the generalisation of the data (seven districts as opposed to 148 wards). Because the location of many boundaries is arbitrary, changes to boundaries produce different results. This is the well-known modifiable areal unit problem that pervades many aspects of the statistical analysis of zone data (see Chapter 4). This source of error is virtually impossible to eliminate if the data are available in

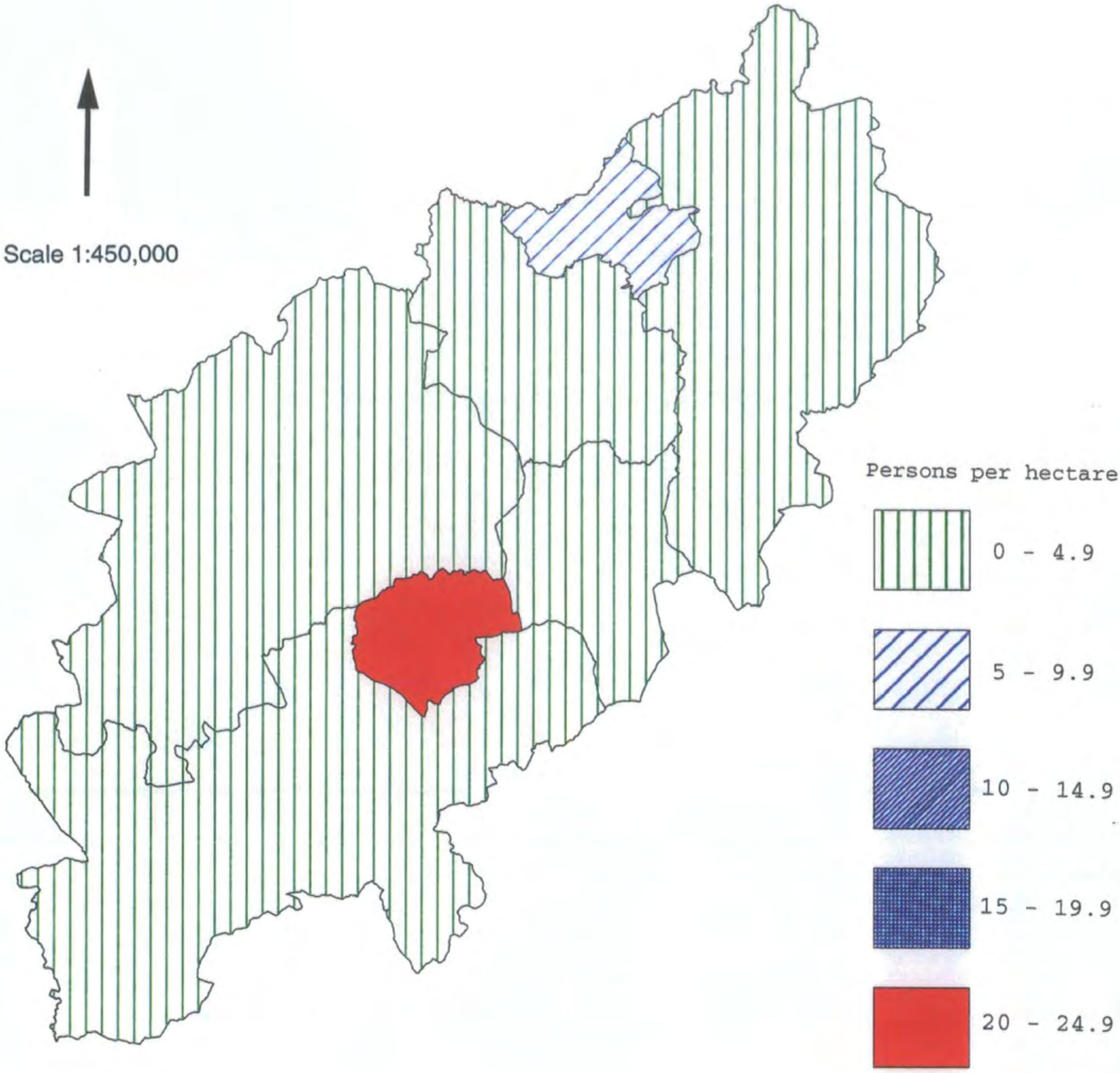


Figure 12.4.10 Population density for districts in Northamptonshire (1991 Census)

aggregate form only. Openshaw (1984) has termed the sensitivity of data to the areal partition that is used, the 'aggregation problem'.

Table 12.4.7 Summary of the Overlay Between Population Density (districts) and Radon Levels in Northamptonshire (after transformation)

	<1%	1 - 3%	3 - 10%	>10%
0 - 4 p/ha	100	23	40	27
5 - 9 p/ha	0	77	21	0
10 - 19 p/ha	0	0	0	0
20 - 24 p/ha	0	0	39	73

The general trend identified from the ward data in Table 12.4.6 above, exists also in the results of the overlay using district data, shown in Table 12.4.7 (the raw data, including the number of polygons and sum area of the overlay polygons, are given in Appendix 6). The categories of population density for the ward data are different from those used with the district data, to reflect the different range and distribution of density in the two data sets. There is an important difference between the two sets of data within the radon category 3 - 10%. The correlation between this radon category and population densities of less than 20 people per hectare is quite marked in Table 12.4.6. In Table 12.4.7 on the other hand, there exists an equally strong association with both the lowest population densities (0 - 4 people per hectare) and the highest (20 - 24 people per hectare). This is likely to be a result of the dominance of the 3 - 10% radon category in the county as a whole, coupled with the poor resolution of the population density data, which are less accurate at the district level. Daventry for example, is 66,561 square hectares and contained within this are a variety of landscapes, including both agricultural land (with low population densities) and numerous villages (with population densities rising to 30.5 residents per hectare in the ward of Abbey South).

12.4.4.2 What is the relationship between social class and radon levels?

Social class for each of the seven districts in Northamptonshire was mapped and an overlay was carried out in ARC/INFO with the radon data for five kilometre grid squares. Differences and similarities in social class for the five Affected Areas are discussed in Chapter 13. The social class categories from the 1991 Census are as follows (see also Chapter 4.4):

I	Professional etc.
II	Managerial and technical
IIIN	Skilled occupations - non-manual
IIIM	Skilled occupations - manual
IV	Partly skilled occupations
V	Unskilled occupations

In addition, there are categories for 'retired' and 'other inactive' (*e.g.* handicapped and the long-term ill) which, for the purposes of this study, have been grouped to form category **VI** (retired). The categories for 'armed forces', 'on a government scheme' and 'occupation inadequately described or not stated', account for only 1.2% of the total population in Northamptonshire and were not included in the analysis for this reason. The source for the data on social class is the 1993 OPCS 'Ward and Civil Parish Monitor for Northamptonshire' and the 'County Report for Northamptonshire (Part two)'. The data are available only as a 10% sample of the population (see Chapter 4). The distribution of social class for the county as a whole is shown below. The most abundant category is retired and other inactive (VI) which makes up 31.3% of the population of the county. A list of the corresponding figures for England are shown in Table 13.5.

Table 12.4.8 Social Class Data in Northamptonshire

Social class	Northamptonshire %
I	3.7
II	20.5
IIIN	8.6
IIIM	20.8
IV	9.8
V	2.9
VI	31.3
Other	1.2

In order to assess the relationships between radon and social class, the percentages in each social class category were grouped (except for class VI, which is discussed in more detail in 12.4.4.3 below). Table 12.4.9 is a summary of social class data for the seven districts and shows the percentage of heads of household in each of the districts, together with the corresponding group (in bold) to which each was assigned.

Table 12.4.9 Social Class Data in Northamptonshire Districts (%)

	Corby	**Daventry**	**E. Northants**	**Kettering**	**Northampton**	**S. Northants**	**Wellingboro'**
I	1.6 **1**	5.3 **5**	4.2 **4**	3.0 **3**	3.4 **3**	5.5 **5**	2.8 **2**
II	11.0 **1**	25.7 **4**	21.0 **3**	19.7 **2**	19.2 **2**	27.5 **4**	19.8 **2**
IIIN	8.2 **3**	8.1 **3**	6.4 **1**	8.3 **3**	10.0 **4**	7.6 **2**	8.8 **3**
IIIM	24.8 **4**	17.8 **1**	23.6 **3**	21.0 **2**	20.1 **2**	18.3 **1**	21.6 **2**
IV	15.9 **5**	10.4 **2**	9.0 **2**	8.9 **1**	10.0 **2**	7.8 **1**	9.2 **2**
V	3.6 **4**	2.3 **1**	2.3 **1**	2.8 **2**	3.2 **3**	2.7 **2**	3.4 **3**
VI	32.5	29.1	32.0	35.0	32.4	27.5	33.1

(Source: OPCS, 1993g)

The figures shown in bold in Table 12.4.9 correspond to the following groups (figures are percentage of households):

Social Class I

Group 1	1.0 - 1.9
Group 2	2.0 - 2.9
Group 3	3.0 - 3.9
Group 4	4.0 - 4.9
Group 5	5.0 - 5.9

Social Class II

Group 1	10.0 - 14.9
Group 2	15.0 - 19.9
Group 3	20.0 - 24.9
Group 4	25.0 - 29.9

Social Class IIIN

Group 1	6.1 - 7.0
Group 2	7.1 - 8.0
Group 3	8.1 - 9.0
Group 4	9.1 - 10.0

Social Class IIIM

Group 1	18.0 - 19.9
Group 2	20.0 - 21.9
Group 3	22.0 - 23.9
Group 4	24.0 - 25.9

Social Class IV

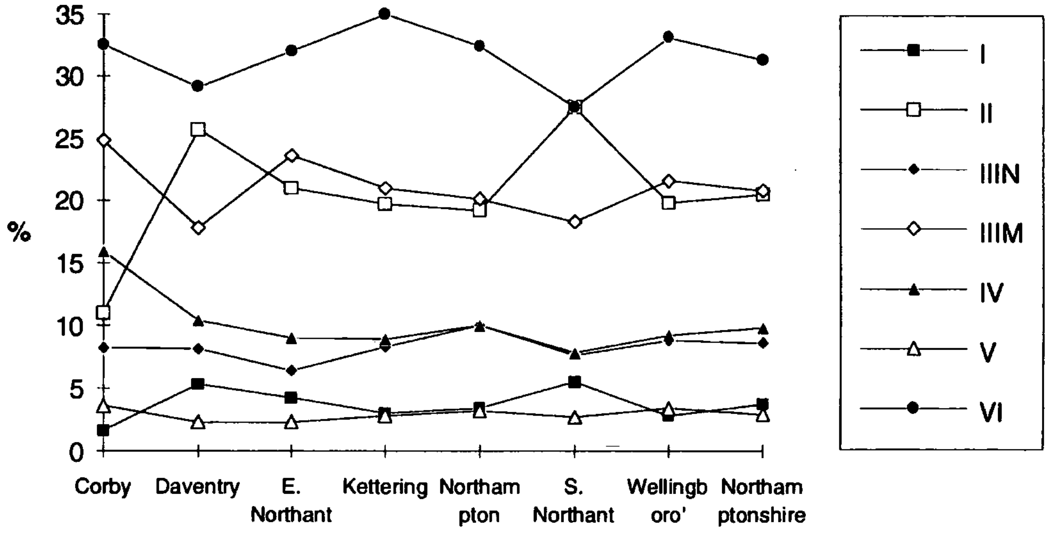
Group 1	7.0 - 8.9
Group 2	9.0 - 10.9
Group 3	11.0 - 12.9
Group 4	13.0 - 14.9
Group 5	15.0 - 16.9

Social Class V

Group 1	2.0 - 2.4
Group 2	2.5 - 2.9
Group 3	3.0 - 3.4
Group 4	3.5 - 3.9

The general distribution of social class is similar for all seven districts. However, the proportions of households in social classes I and II (professionals and managers) are significantly lower in Corby than in any of the other districts and the proportions in classes IV and V (partly skilled and unskilled) are higher, reflecting a generally lower class region of the county. Daventry and South Northamptonshire on the other hand, are the only districts in which class II accounts for a greater proportion of households than class IIIM, and the relatively high proportion of households in class I in these two south-eastern districts reflects a trend towards the higher classes. The different distributions in the seven districts and the comparison with Northamptonshire as a whole is depicted in Figure 12.4.11 below.

Figure 12.4.11 Social Class in Northamptonshire Districts



The overlay operation between the coverages for social class and radon levels produced 208 polygons, with a mean area of 11,384,792 square metres. As with the population density data above, the largest number of polygons were associated with the radon level 3 - 10% and the smallest number with <1%. Analysis of the results was problematic because of the large number of categories inherent in the social class data and the different range of values for the groups in each social class. Thus it is necessary to analyse each social class separately. As with the population density data, it was useful to analyse the relationships between radon and social class using the proportion of the total area of each radon category. Thus 99% of the area with radon levels of <1% has between 5.0 and 5.9% of heads of household in social class I. Succinct conclusions are drawn from the data for each social class. A summary section at the end of 12.4.4.2 draws together the principal findings and identifies the main points.

Social Class I

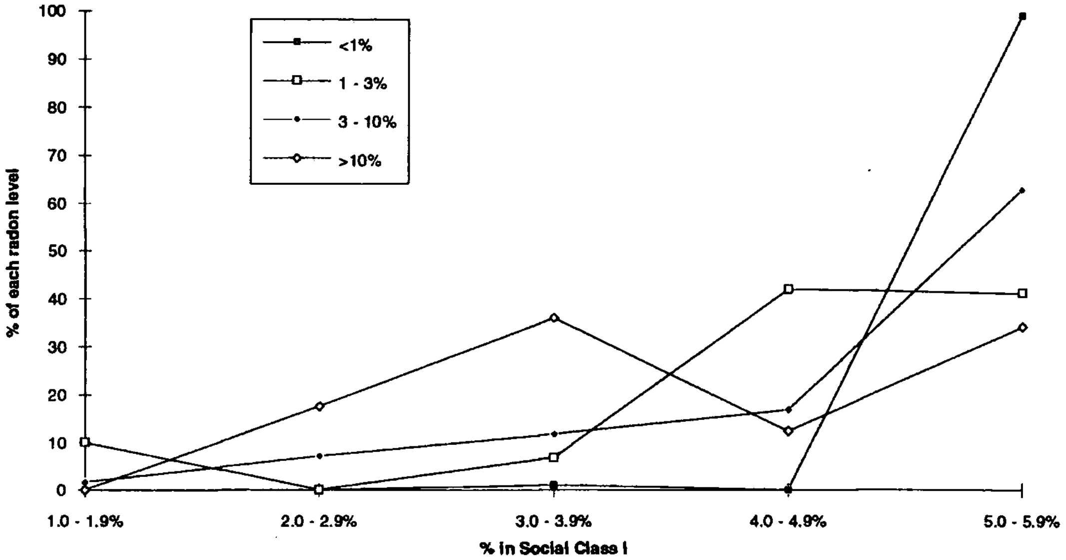
This class accounts for just 3.7% of the total households in Northamptonshire, but is important because it is confined only to people with professional occupations and often therefore, to wealthier households. Areas with less than 3% of homes above the Action Level are more likely to be associated with areas having more than 4% of households in social class I (the county average is 3.7%). The higher radon levels (>10%) are also associated with more than 5% of households in class I, but show a slightly stronger correlation with values between 3.0 and 3.9% (approximately equal to the county mean). There is a significantly greater range of values in the higher radon categories than in the lower ones.

Table 12.4.10 Social Class I and Radon Levels in Northamptonshire

	<1%	**1 - 3%**	**3 - 10%**	**>10%**
1.0 - 1.9%	0	10	2	0
2.0 - 2.9%	0	0	7	17
3.0 - 3.9%	1	7	12	**36**
4.0 - 4.9%	0	**42**	17	12
5.0 - 5.9%	**99**	41	**63**	34

The relationship between the four radon categories and the proportions in social class I, is shown on the graph below. The average proportion in this class is affected by the low percentage in Corby (1.6%) and as a result only 9% of the results in Table 12.4.10 above are concerned with less than 3.0% in class I. This situation has manifested itself in the graph below, as a positive overall relationship between the two variables.

Figure 12.4.12 Social Class I and Radon Levels in Northamptonshire



Social Class II

Social class II accounts for 20.5% of households in Northamptonshire, with South Northamptonshire having the highest proportion (27.5%) and Corby the lowest (11%). The distribution within the radon categories, in the managerial and technical class, is similar to that of social class I, but the categories account for a much larger range of percentages. By analysing Table 12.4.11 below, it is evident that the areas with <3% of homes above the Action Level, are associated with districts having more than 20% of households in class II. On the other hand, the areas with higher radon levels (>10%) are more likely to have between 15 and 20% of households in social class II (*i.e.* below the county average of 20.5%).

Table 12.4.11 Social Class II and Radon Levels in Northamptonshire

	<1%	1 - 3%	3 - 10%	>10%
10.0 - 14.9%	0	10	2	0
15.0 - 19.9%	1	7	19	**53**
20.0 - 24.9%	0	**42**	17	12
25.0 - 29.9%	**99**	41	**63**	34

Social Class IIIN

Social class IIIN (skilled non-manual) accounts for 8.6% of the households in Northamptonshire and the difference between the highest and lowest district is just 3.6%. The distribution of this class with respect to radon levels is dominated by the 8.1 - 9.0% category (four out of the seven districts fall within this category). The overall relationship between the proportions in social class IIIN and radon is weak, but a pattern appears to exist in that the majority of the area with <10% of homes above the Action Level is associated with <9% of households in class IIIN. On the other hand, areas with more than 10% of homes above the Action Level are much more likely (81%) to have more than 8.1% of households in this class.

Table 12.4.12 Social Class IIIN and Radon Levels in Northamptonshire

	<1%	**1 - 3%**	**3 - 10%**	**>10%**
6.1 - 7.0%	0	42	17	12
7.1 - 8.0%	0	11	36	6
8.1 - 9.0%	**100**	**47**	**44**	**72**
9.1 - 10.0%	0	0	3	9

Social Class IIIM

Social class IIIM (skilled manual) accounts for 20.8% of the households in Northamptonshire as a whole and is highest in Corby (24.8%) and lowest in Daventry (17.8%). The general trend of the data (as shown in Table 12.4.13 below), has shifted away from the higher percentages towards the lower percentages. This is most noticeable in the <1% radon level. However, there is no clear relationship between radon and the proportion in social class IIIM, and all four radon categories are associated mainly with less than 23.9% in class IIIM. There is a weak association between areas with radon levels of 1 - 3% and areas with an above average proportion of households in class IIIM. It is probable that the resolution of the social class data is too coarse to identify any further associations.

Table 12.4.13 Social Class IIIM and Radon Levels in Northamptonshire

	<1%	**1 - 3%**	**3 - 10%**	**>10%**
18.0 - 19.9%	**99**	41	**63**	34
20.0 - 21.9%	1	7	19	**53**
22.0 - 23.9%	0	**42**	17	12
24.0 - 25.9%	0	10	2	0

Social Class IV

Social class IV (partly skilled) accounts for 9.8% of the households in Northamptonshire and is highest in Corby (15.9%) and lowest in South Northamptonshire (7.8%). The vast majority of the county has less than 10.9% in this class. Corby on the other hand, has over five and a half percent more people in class IV than any other district and this extreme value has increased the county average by more than half a percent. Only the 1 - 3% radon level (the most common level in Northamptonshire) displays a significant association with more than 11% in class IV. Due to the fact that the data were aggregated for districts, the spatial distribution of these areas (radon levels 1 - 3% and >11% in class IV) will be restricted to the district of Corby only. However, this relationship may not hold true if the data are analysed for ward or enumeration district level.

Table 12.4.14 Social Class IV and Radon Levels in Northamptonshire

	<1%	1 - 3%	3 - 10%	>10%
7.0 - 8.9%	1	18	44	33
9.0 - 10.9%	**99**	**72**	**54**	**67**
11.0 - 12.9%	0	0	0	0
13.0 - 14.9%	0	0	0	0
15.0 - 16.9%	0	10	2	0

Social Class V

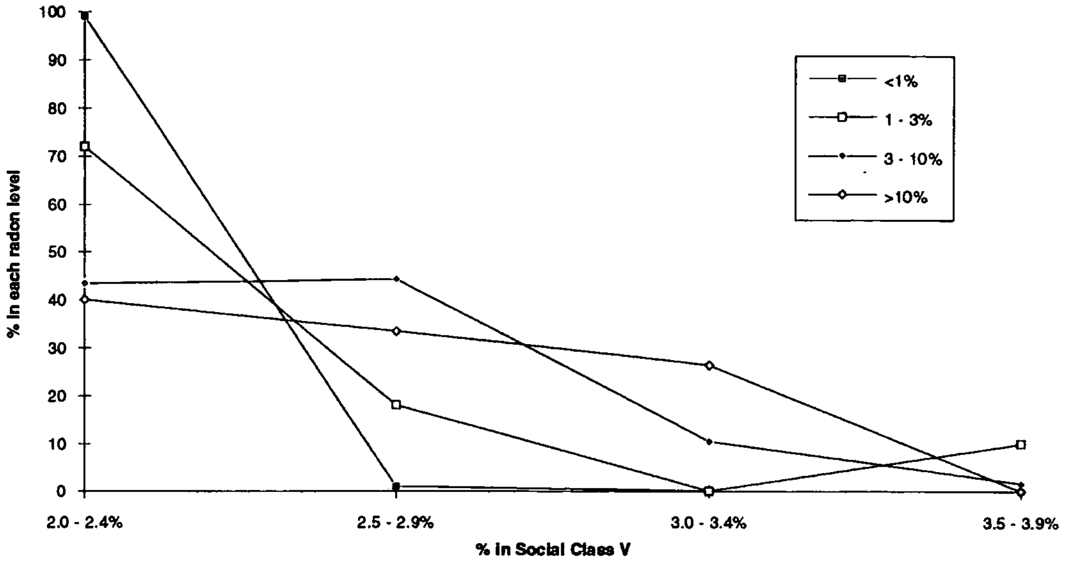
Social class V (unskilled) accounts for only 2.9% of households in the county of Northamptonshire. Areas with less than 3% of homes above the Action Level are closely associated with less than the average proportion of households in class V (the inverse of the situation in Table 12.4.10 - social class I). On the other hand, the higher radon levels (>3% of homes above the Action Level) are correlated with the lower percentages, but there is also a significant proportion of these categories associated with more than 2.9% (*i.e.* above average proportions).

Table 12.4.15 Social Class V and Radon Levels in Northamptonshire

	<1%	1 - 3%	3 - 10%	>10%
2.0 - 2.4%	**99**	**72**	43	**40**
2.5 - 2.9%	1	18	**44**	33
3.0 - 3.4%	0	0	11	26
3.5 - 3.9%	0	10	2	0

The relationship between radon levels and the proportions in social class V are depicted on the graph below. The graph provides a useful comparison to Figure 12.4.12 above. The *overall* trend is a negative one, whereby as the proportions in social class V increase, the association with all four radon categories decreases.

Figure 12.4.13 Social Class V and Radon Levels in Northamptonshire



Conclusion

The general trends that have been identified between social class and radon levels are most clearly depicted in social classes I and V (the two 'extremes' of social class groupings). Areas with above average proportions in social class I are more likely to have a low percentage of homes above the Action Level. On the other hand, areas with above average proportions in class V are slightly more likely to be associated with areas having more than 3% of homes above the Action Level. In summary, areas with a low proportion of homes above the Action Level are more likely to be associated with higher proportions in social classes I and II and lower proportions in IV and V.

The results of the overlay operation depend to a degree on the classification methods that were used to group the data. Thus the correlations that have been identified may simply be the result of the categorisation that was used. However, a methodological dilemma exists because it is necessary to divide the data into classes in order to carry out the overlays. The presence of the MAUP further undermines results achieved using data aggregated for areal units. In order to ascertain whether the correlations identified are a true reflection of reality, the relationships require further investigation at the smallest possible areal unit (enumeration districts or postcode units), although the problem of data categorisation will remain.

12.4.4.3 What is the relationship between the proportion of households containing only pensioners and radon levels?

This section follows on from 12.4.4.2 and analyses the impact of pensioners (approximating to social class VI) in detail. The data for the proportion of households containing only pensioners were taken from the 1991 Census and were mapped for the 148 wards of

Northamptonshire (see Figure 12.4.15). The average proportion of households containing only pensioners for the county is 22.9% and values range from 11.0% in Hill (Daventry) to 38.5% in Swanspool (Wellingborough). Summary statistics for the seven districts are shown below.

Table 12.4.16 The Percentage of Households Consisting Only of Pensioners for Northamptonshire Districts

District	%
Corby	**20.1**
Daventry	**21.5**
East Northamptonshire	**24.8**
Kettering	**25.5**
Northampton	**22.5**
South Northamptonshire	**21.8**
Wellingborough	**23.8**

An overlay was carried out in ARC/INFO with the radon data and the results are listed in the table below. The figures represent the total area of each row (proportion of households containing only pensioners) as a proportion of the column (radon level) total. The largest category in each column has been highlighted.

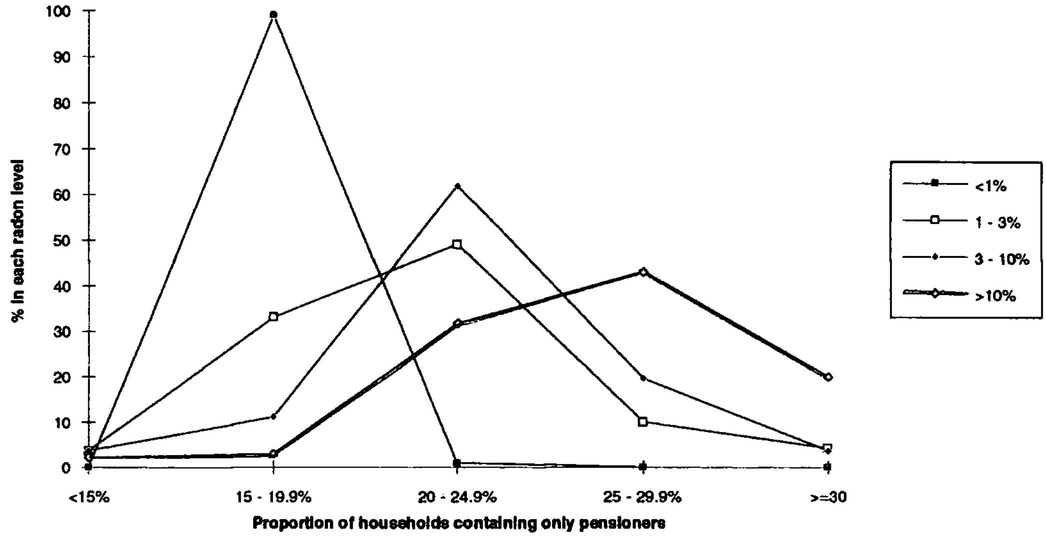
Table 12.4.17 The Proportion of Households Containing Only Pensioners and Radon Levels in Northamptonshire

	<1%	**1 - 3%**	**3 - 10%**	**>10%**
<15%	0	4	4	2
15 - 19.9%	**99**	33	11	3
20 - 24.9%	1	**49**	**62**	32
25 - 29.9%	0	10	20	**43**
>=30%	0	4	4	20

The overall trend implies a positive relationship between the proportion of homes above the Action Level and the percentage of homes containing only pensioners. From the results in Table 12.4.17, it appears that areas with less than 10% of homes above the Action Level are more likely to be associated with areas where less than 25% of households contain only pensioners. This is most marked in radon areas <3%.

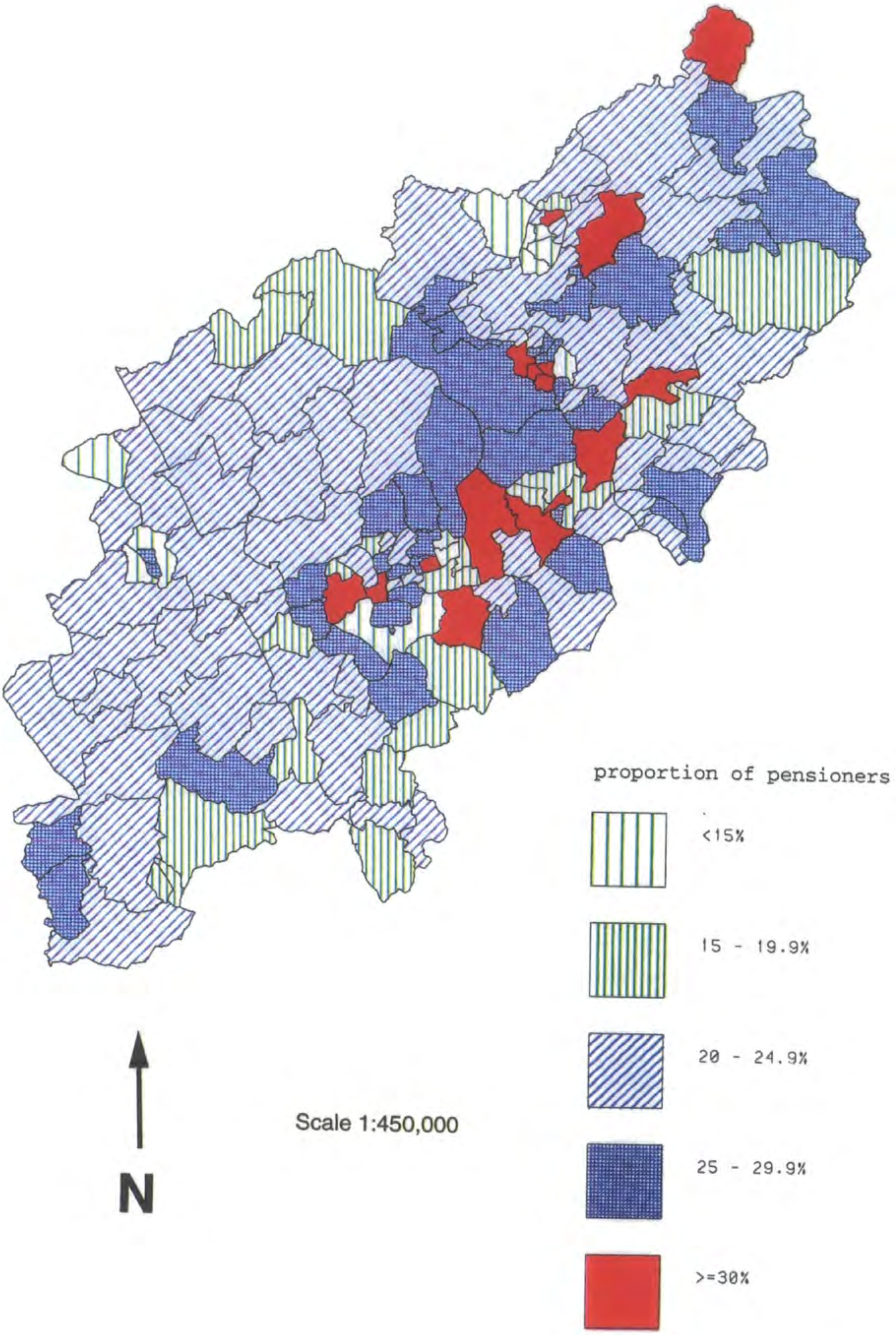
In areas where more than 10% of homes exceed the Action Level however, the situation is reversed - more than 63% of the area is associated with >=25% of households consisting only of pensioners. This relationship is shown clearly on the graph below (depicted in red). Whilst most of the area with 1 - 10% above the Action Level is centred around the average percentage of pensioners (22.9%), the distributions are skewed in favour of either the lower percentages (1 - 3%) or the higher ones (3 - 10%). Figure 12.4.14 presents a graphic display of the findings.

Figure 12.4.14 The Association Between the Proportion of Households Containing Only Pensioners, and Radon Levels in Northamptonshire



The association between higher radon levels and a greater proportion of pensioners may be a result of the aggregations used to map the data or the groups chosen to categorise the data. The MAUP and the problem of the ecological fallacy may also have influenced the results (there is no evidence to suggest that the homes with high radon levels also contain pensioners), but the results do suggest that the correlation deserves further research attention. Pensioners homes may be more susceptible to higher radon levels because of their occupancy habits - pensioners tend to remain indoors for a larger part of the day and thus the exchange of air (between outside and inside the home) will be reduced. In addition, the type of house often inhabited by pensioners (bungalows for example) may also play a part in determining the radon level. However domestic radon levels are a result of a combination of several different factors (levels of radium in the adjacent/subjacent soil and bedrock; soil moisture and permeability; building type and construction *etc.*) and thus any simple correlation will be prone to errors and will be confounded by the implicit complexity of the radon hazard. A summary of the findings between radon levels and social factors in Northamptonshire can be found in Chapter 14.

Figure 12.4.15 The proportion of households containing only pensioners for wards in Northamptonshire (1991 Census)



12.5 Derbyshire

12.5.1 What spatial patterns exist in the radon data set?

The county of Derbyshire was designated an Affected Area in 1990, following measurements by the NRPB which identified parts of the county as having more than 1% of homes above the Action Level. The spatial distribution of radon for the whole of Derbyshire, as measured by the NRPB, is depicted on Figure 12.5.1. Table 12.5.1 below shows the estimated area and percentage of each of the radon categories in Derbyshire. The figures do not add up to 100% due to rounding.

There is a clear relationship between radon levels and geology in Derbyshire (Miles *et al.*, 1992b). The area to the west of the county (see Figure 12.5.1), with more than 10% of homes above the Action Level, is associated with the underlying Carboniferous Limestone. The area in the east of the county, characterised by radon levels between 3 - 10%, is underlain by Permian Limestones. The southern part of the county where radon levels are generally much lower, is underlain by Triassic mudstones and sandstones (Miles *et al.*, 1992b).

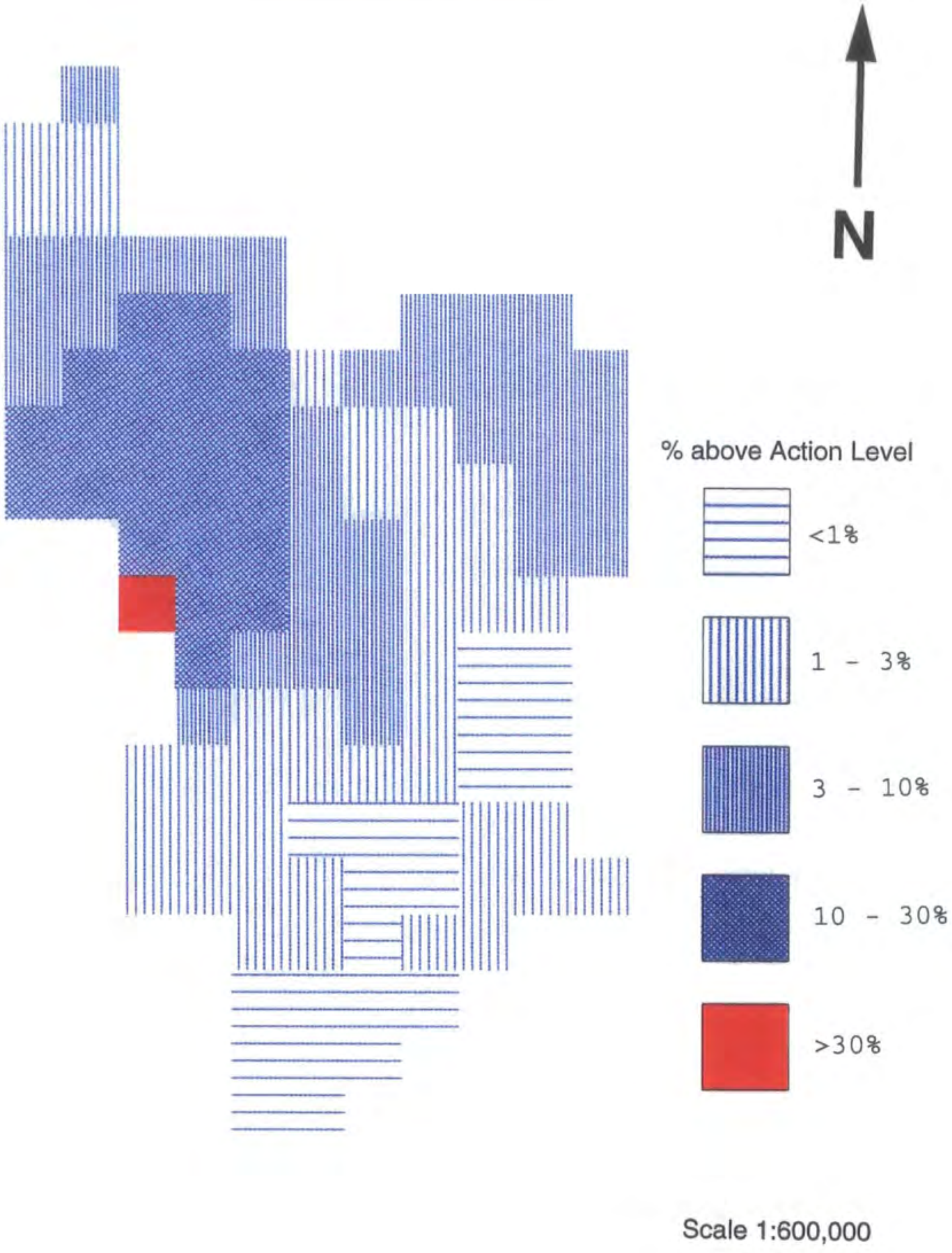
Table 12.5.1 Summary of the Number of Homes Above the Action Level in Derbyshire

Radon Categories	**Area (sq. km)**	**%**
<1%	372	15
1 - 3%	890	35
3 - 10%	754	30
10 - 30%	511	20
>30%	7	0.3
TOTAL	**2,533**	**100**

The distribution of radon levels is asymmetrical, with a mode of 1 - 10% (see Figure 12.5.2 below). Areas of low radon (<1% of homes above the Action Level) make up 15% of the total area of Derbyshire and are found in the south of the county, in a band running roughly south-west to north-east, encompassing the towns of Swadlincote, Derby, Ilkeston and Ripley. There is only a very small proportion of the county (just one five kilometre grid square) with radon levels in excess of 30%. The area shown in Table 12.5.1 for this square does not total 25 square kilometres because it is located at the edge of the county, so the county boundary cuts across the edge of the square. Thus only seven square kilometres fall within the county boundary as digitised from the land capability maps.

Of the five Affected Areas, only Cornwall and Devon have higher radon levels then Derbyshire. The proclivity towards the higher radon levels is depicted in Figure 12.5.2 below.

Figure 12.5.1 The spatial distribution of the estimated proportion of homes above the Action Level in Derbyshire

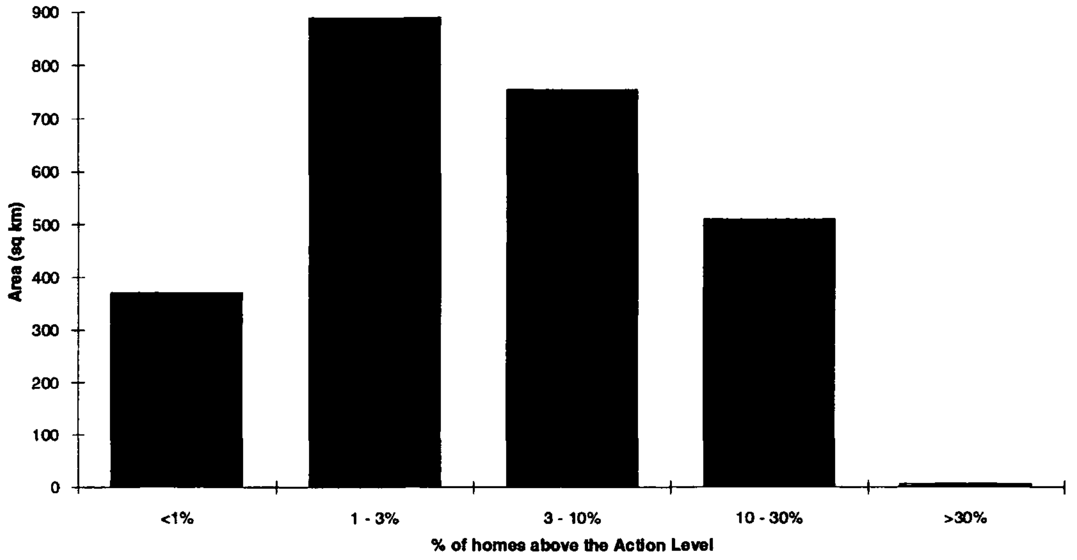


Please note that the addition of an extra category (>30% above the Action Level) in this county means that the colours and shading used to represent the categories differ from those used in Figures 12.3.1 and 12.4.1.

(Source: NRPB)

However, the graph also shows that an area in excess of 370 square kilometres has less than 1% of homes above the Action Level. Indeed, Miles *et al.* (1992b, pp. 26) suggest that 'parts of the south of the county with less than 1% of homes above the Action Level need not be designated as Affected Areas'. For the ease of classification however, the county boundary is used to delimit Affected Areas and thus the whole county is categorised as Affected.

Figure 12.5.2 Total Area (in square kilometres) of each Radon Category in Derbyshire

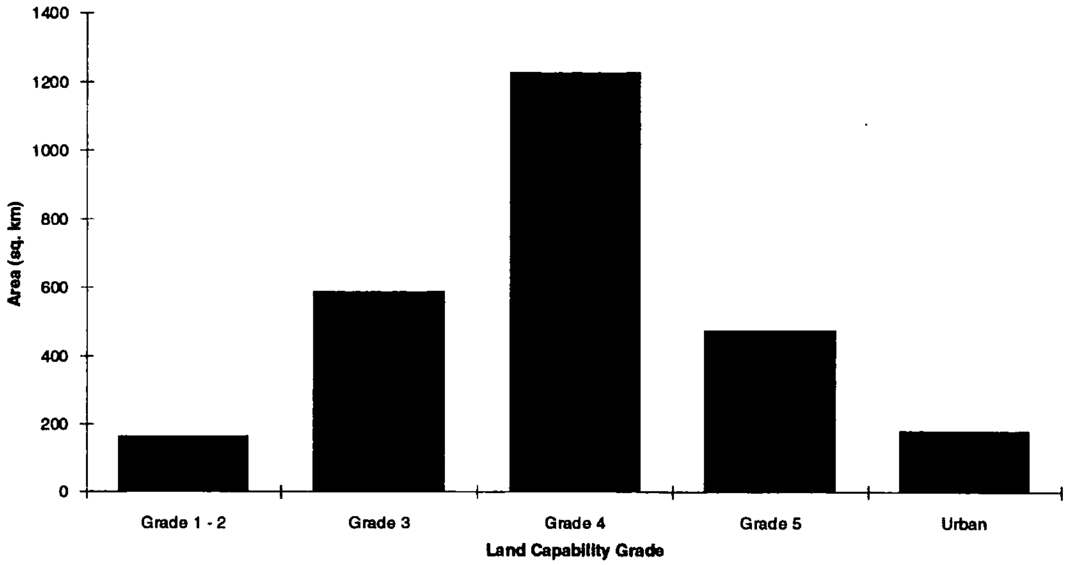


12.5.2 What spatial patterns exist in the Land Capability data set?

Figure 12.5.3 shows the spatial pattern of land capability for Derbyshire digitised from the 1983 MAFF Land Capability 'Midlands and Western Region' map. There is a clear distinction between the poorer land in the north and the better quality land in the south, a result of the influence of the physical environment. Grade 5 land dominates the north-west of the county and is associated with the Peak District at the southern end of the Pennines. The majority of the area south of this but north of Derby is Grade 4 land which correlates closely to the Carboniferous and Permian Limestones found in this part of the county. South of Derby, where relief is generally lower, the land is mainly Grade 3 with pockets of Grade 1 - 2 and some isolated urban areas.

The distribution of land capability grades is shown in Figure 12.5.4 below. The distribution is roughly symmetrical and has a kurtosis of +1.9. For the purpose of studying land capability, the county can be divided roughly into two. The northern half is characterised by poor quality agricultural land and the south by good quality agricultural land and urban areas.

Figure 12.5.4 Total Area (in square kilometres) of each Land Capability Grade for Derbyshire



Grade 1 - 2 land (very good quality agricultural land) accounts for 6% of the area of the county, which is a similar proportion to Devon and Cornwall, but much less than that found in Somerset. The best quality land is located entirely in the south of the county adjacent to the urban areas of Derby and Burton-on-Trent. In keeping with the general north/south divide, Grade 3 land is found predominantly in the south of the county around the pockets of very good quality agricultural land and the larger urban areas. Derbyshire is unusual in that Grade 3 land does not account for the greatest proportion of the county in terms of area, as it does in the four other Affected Areas: approximately 22% of the total area of Derbyshire is Grade 3 land, compared to 67% in Somerset, 72% in Devon and Cornwall and 92% in Northamptonshire. This may be due to the location of Derbyshire at the southern end of the Pennines, where relief and rainfall are generally higher than in Somerset and Northamptonshire. Grades 4 and 5 predominantly in the northern half of the county are underlain by Limestone rocks. Grade 5 land is closely correlated with land over 200 metres above sea level (especially with land over 400 metres above sea level), and is found mainly in the north-west of the county.

There are thirteen urban areas identified on the land capability map, making up 7% of the total area (see Table 12.5.2 below). The largest urban areas are Derby, Chesterfield, Burton-on-Trent, Long Eaton and Ilkeston. The mean size of the urban areas is 13.7 square kilometres. Chapter 13 provides a comparison of statistics for the other Affected Areas.

Figure 12.5.3 A land capability map of Derbyshire

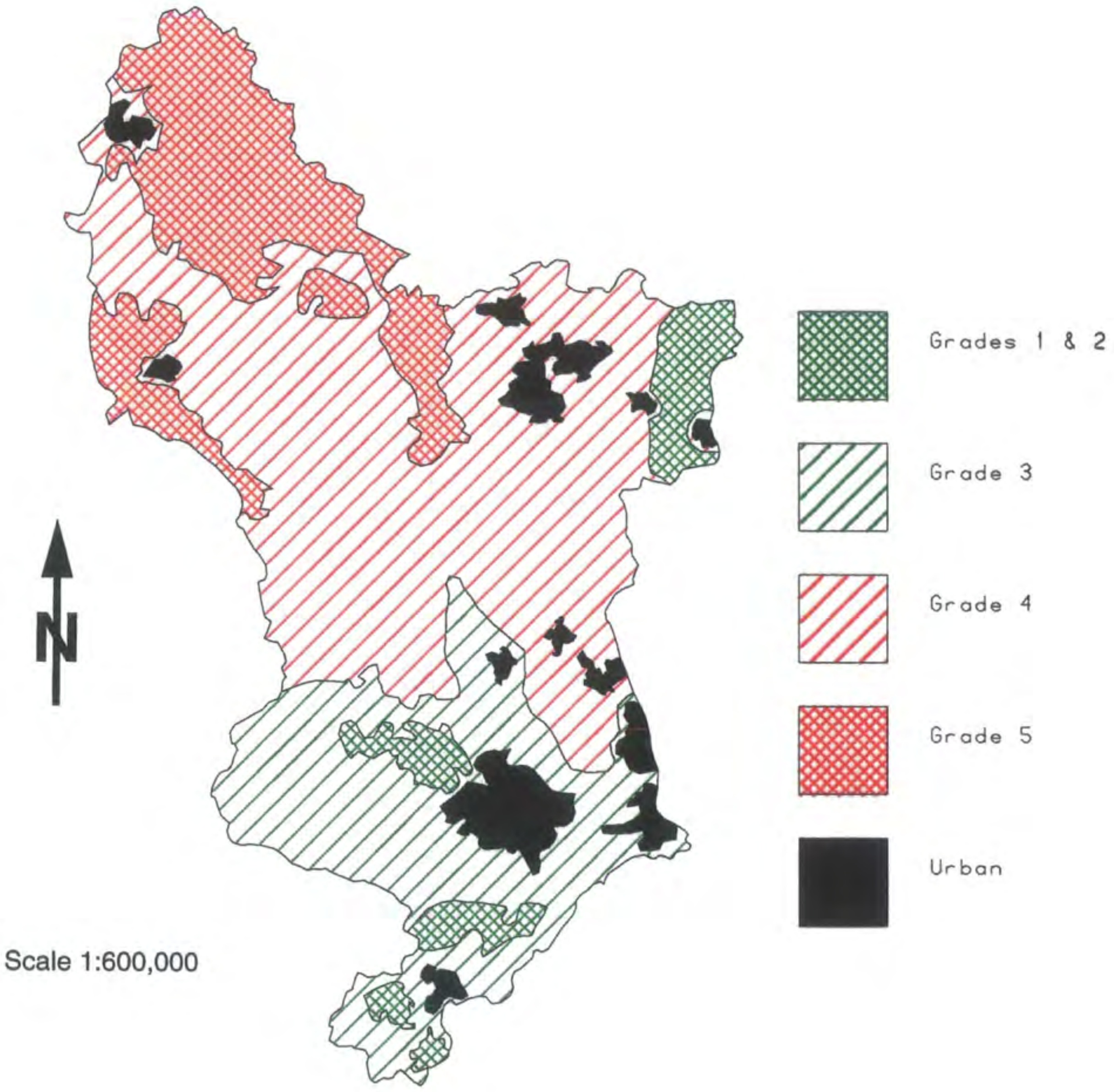


Table 12.5.2 Summary of Land Capability Data for Derbyshire[4]

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
frequency*	5	1	2	3	13	**24**
sum (km^2)	163.5	587.1	1,228.4	474.4	178.3	2,631.7
% of sum	6	22	47	18	7	100
mean	32.7	-	614.2	158.1	13.7	-
maximum	75.9	-	1,224.9	370.4	64.2	-
minimum	6.8	-	3.5	18.6	3.0	-

* frequency of polygons in the digital coverage

12.5.3 What spatial patterns exist between radon gas and land capability?

Figure 12.5.5 shows aspects of the radon and land capability overlay produced in ARC/INFO and provides a means of scrutinising the spatial relationships between the two data sets. The map shows all polygons in the county which have a land capability Grade 4 and radon levels of between 3 and 30% (a total area of 822 square kilometres). In addition, all urban areas which overlay polygons having fewer than 1% of homes above the Action Level are also shown. Analysis is best undertaken however, using the overlay summary statistics shown in Table 12.5.3 below. The figures shown in bold highlight the largest categories in each land capability grade.

Table 12.5.3 Summary of the Overlay Between the Estimated Proportion of Homes Above the Action Level and Land Capability, in Derbyshire*

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
<1%	36	27	6	0	**49**
1 - 3%	18	**66**	26	25	33
3 - 10%	**46**	7	**35**	**47**	16
10 - 30%	0	0	33	28	3
>30%	0	0	1	0	0

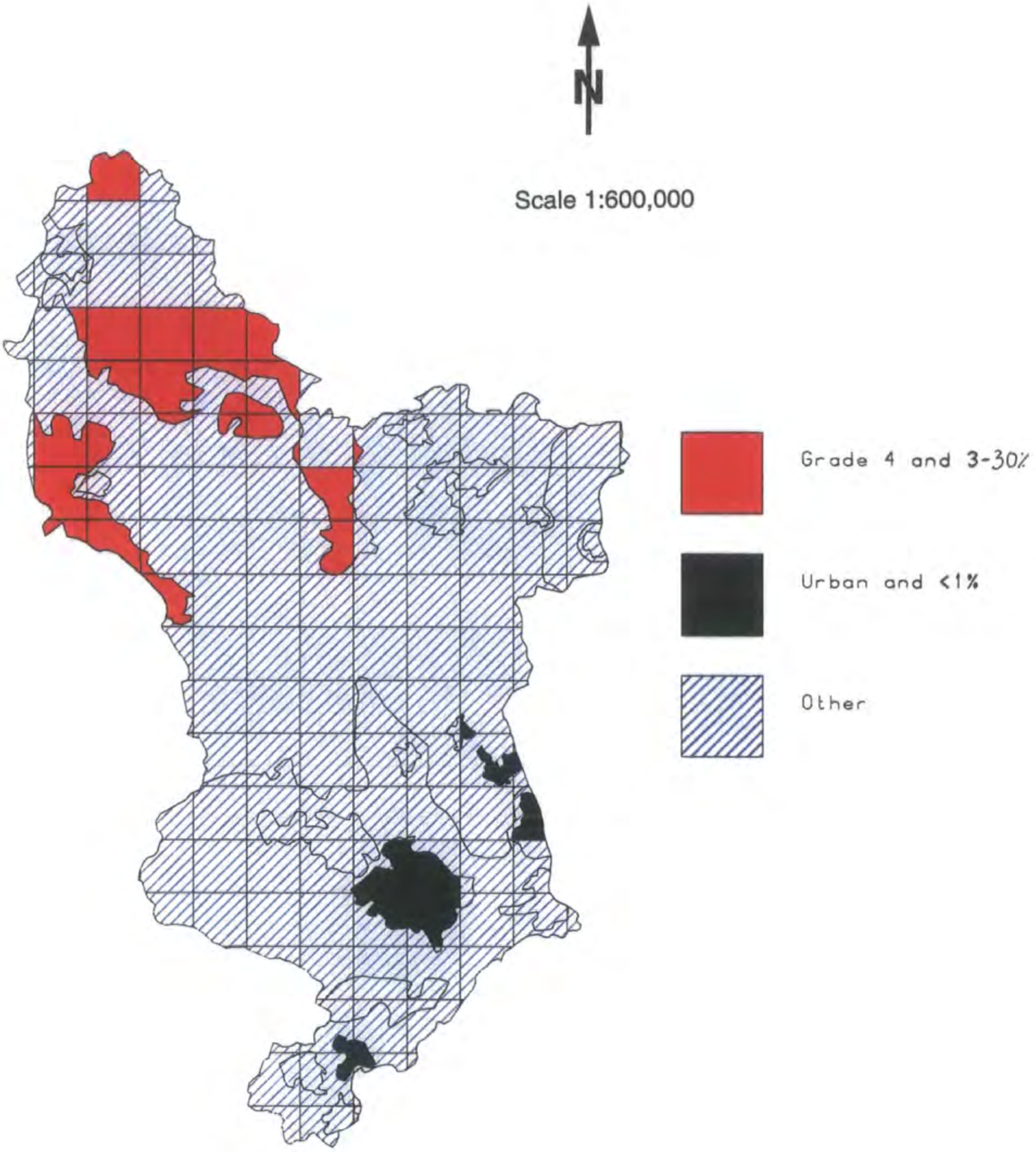
* These percentages are calculated using the area of each radon category as a proportion of the total area of each land capability grade.

No pattern is immediately apparent in the data shown in the above table. In order to display the relationships between the two variables more clearly, the distribution of radon in each land

---

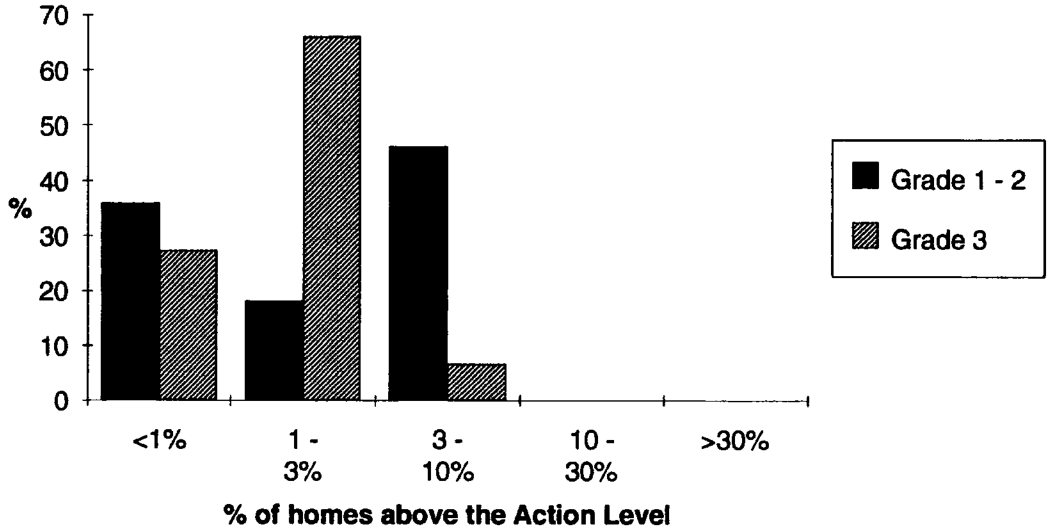
[4] The totals in this table may not correspond to those used in the land capability and radon overlay, because of slight differences in area between the two coverages.

Figure 12.5.5 Radon gas and land capability in Derbyshire



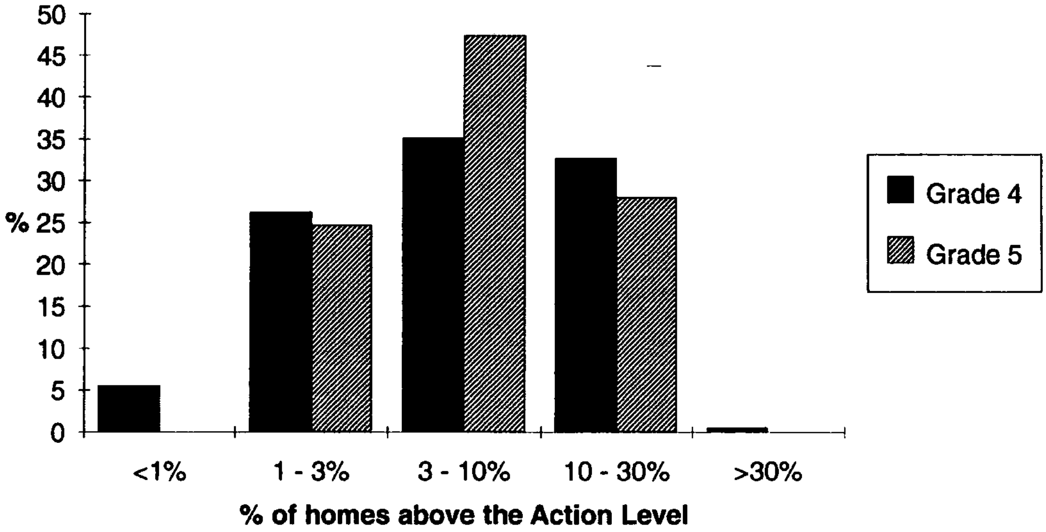
capability grade is depicted in the graphs below. Each graph shows the radon levels for each of the land capability grades, as measured from the overlays in ARC/INFO.

Figure 12.5.6 Derbyshire Radon Levels in Areas of Land Capability Grade 1 - 2 and Grade 3



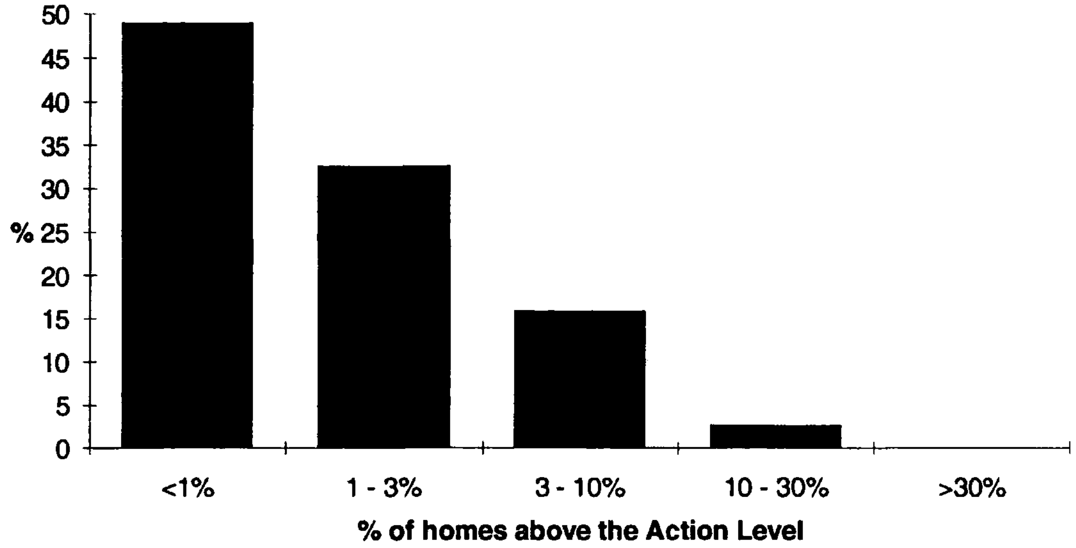
The graph above (Figure 12.5.6) shows a general association between the better quality agricultural land (Grades 1 to 3) and lower radon levels. Indeed the land capability Grades 1 - 2 and 3 are only associated with areas which have less than 10% of homes above the Action Level. This seems to fit the general trend (identified in section 12.1) which suggests that good quality land is associated with lower radon levels. However, there is a great deal of variation between the radon levels in each of the grades, and Grade 1 - 2 land is more closely associated with radon levels of 3 - 10% than is Grade 3 land.

Figure 12.5.7 Derbyshire Radon Levels in Areas of Land Capability Grade 4 and Grade 5



The graph relates to the poorer quality agricultural land (Grades 4 and 5), and shows a quite different radon distribution than that associated with Grades 1 to 3. The distribution has shifted away from the lower radon levels towards the higher levels (only 6% of Grade 4 and 5 land is characterised by radon levels of less than 1% above the Action Level, compared to 63% of Grade 1 to 3 land). This association between poor quality land and the higher radon levels is also characteristic of Devon and Cornwall and Somerset, both areas having large areas of poor quality land.

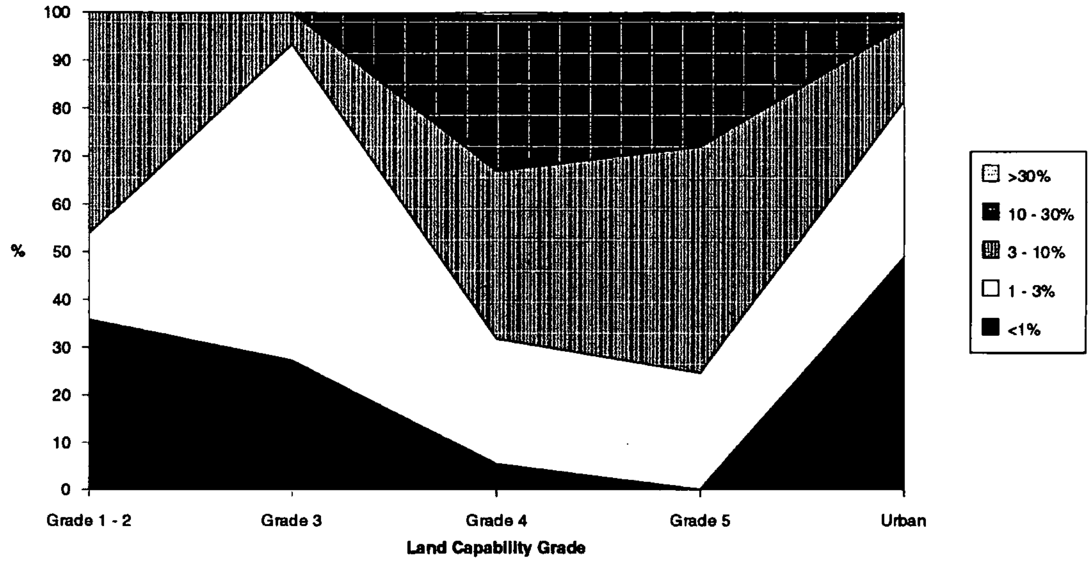
Figure 12.5.8 Derbyshire Radon Levels in Urban Areas



Urban land does not fit in to the overall trend but shows a strong inverse correlation. In Derbyshire, urban land is closely correlated with the lower radon levels (<1% - 3% of homes above the Action Level) and this is depicted in Figure 12.5.8 above. Approximately 82% of urban land in the county is associated with grid squares where fewer than 3% of homes exceed the Action Level. This may be a result of the location and site of the urban areas in the county, which are mainly situated in valleys and on the mudstones and sandstones of the southern part of Derbyshire. The mudstones and sandstones contain proportionally low levels of uranium. Alternatively, details of house construction, habits of occupants, local variations in radium concentrations and ground permeability, can also affect the level of radon in the home.

The area chart in Figure 12.5.9 below, shows the association between radon and land capability. Grade 1 - 2 land, Grade 3 land and urban land are clearly more closely correlated with radon levels of <1% than are Grades 4 and 5. This relationship is less clear when analysing radon levels of 1 - 3% and 3 - 10%. Radon levels of 10 - 30% however are almost entirely associated with land capability Grades 4 and 5. The radon category of >30% is not very clear on the graph as it represents less than half a percent of the total area of the county. It is associated with Grade 4 land only.

Figure 12.5.9 The General Trend Between Radon and Land Capability Grade in Derbyshire



The results suggest an association between land capability and radon levels, which implies that as land capability decreases the proportion of homes above the Action Level increases. Urban land in Derbyshire is closely associated with low radon levels. However, a more detailed study of the relationships between land capability and radon levels would need to be carried out in order to clarify these findings. In addition, the effects of other aspects of the natural environment, such as soil type and ground permeability would need to be analysed to asses their part in the relationship between radon and land capability. Chapter 14 provides a summary of the results.

13. A Comparison Between the Affected Areas

The land capability and radon data and the relationships between the two sets of data, have already been discussed in detail for each county. The aim of this section is to highlight the main differences and similarities between the counties and to analyse any relationships that have been found. Although the five counties have different physical and social characteristics, they share the fact that they are all classified by the NRPB as Affected Areas and have 1% or more of homes above the Action Level. This situation poses questions about the underlying cause of the elevated radon levels. Insight in to possible causes may be furthered by comparing different aspects of the counties. This comparison is achieved by discussing each data set in turn.

13.1 Radon Gas

The proportion of homes above the Action Level in each of the four areas is given in Table 13.1 below. Small percentages have been included in order to register their existence, no matter how small. The largest category in each county is highlighted in bold.

The proportions in each category vary between 0.3% and 68% and the distribution is different in all four areas. As would be expected, the majority of the area in all counties is dominated by the middle three radon categories (1% - 30%), but within this framework there are significant differences and similarities.

Table 13.1 The Proportion of Homes in each Radon Category as a Percentage of the Total Area of Each Affected Area

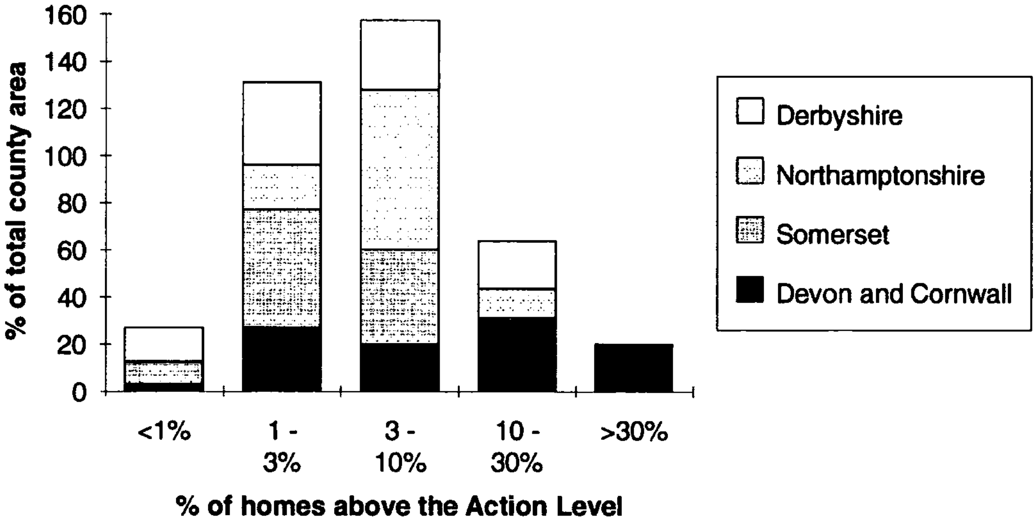
	<1%	**1 - 3%**	**3 - 10%**	**10 - 30%**	**>30%**
Devon and Cornwall	3	27	20	**30**	20
Somerset	9	**50**	40	1	0
Northamptonshire	0.5	19	**68**	12.5	0
Derbyshire	15	**35**	30	20	0.3

Radon levels of 3 - 10% account for almost a third of the total area of the five counties, although this category is more widespread in Northamptonshire (68%) than in any other county. The highest radon levels (>30%) account for almost as much of the total area as the lowest levels (<1%), although 98.5% of the high radon levels in the five Affected Areas is found in Devon and Cornwall.

Derbyshire dominates the lowest radon category and almost a sixth of the total county area has fewer than 1% of homes above the Action Level. Indeed, the NRPB has identified parts of the south-east of the county as not being 'Affected' by radon in the strictest sense (but the whole county is considered an Affected Area for administrative purposes) - but further radon tests may identify more homes above the Action Level in Derbyshire. Approximately 85% of the county however, has more than 1% of homes above the Action Level and in 0.3% of the county more than 30% of homes exceed the Action Level. A similar situation exists in Somerset, where approximately 9% of the county contains fewer than 1% of homes above the Action Level. The county is dominated by the radon level 1 - 3%, which accounts for 50% of the area.

Northamptonshire is characterised by generally higher radon levels then either Somerset or Derbyshire and virtually all of the county has between 1% and 30% of homes above the Action Level. No significant area in Northamptonshire has fewer than 1% of homes above the Action Level (measurements are collated for five kilometre grid squares so the resolution is only accurate to 25 square kilometres). Only a very small proportion (3%) of Devon and Cornwall has fewer than 1% of homes above the Action Level, on the other hand around half of the area of the two counties has more than 10% of homes above the Action Level.

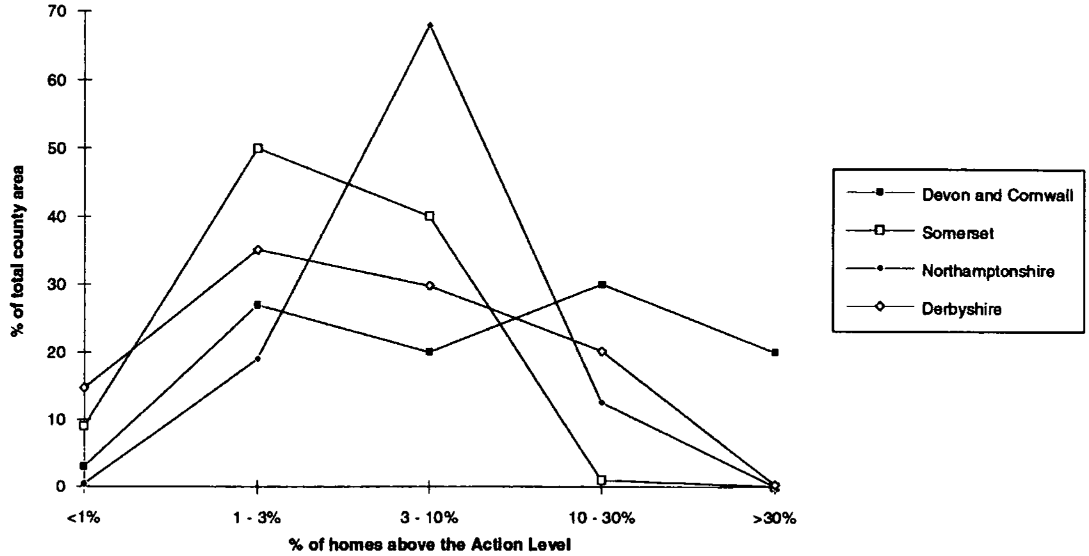
Figure 13.1 The Percentage of Total Area for Each Radon Category in the Affected Areas



The importance of each radon level in terms of the proportion of the total county area is shown in Figure 13.1. The dominant radon-levels in each county can be seen clearly on the graph. Devon and Cornwall, Somerset and Derbyshire have similar overall patterns, with between 3% and 9% of the area of the counties having radon levels of <1%, followed by a rise to between 27% and 50% associated with radon levels of 1 - 3%. The proportion of each county with radon levels of 3 - 10% then falls and proportions in Somerset and Derbyshire continue to decline. In Devon and Cornwall however, the percentage area with radon levels 10 - 30% rises, before falling again in the >30% radon level bracket

Northamptonshire, on the other hand, has a quite different pattern of radon distribution. The total area of <1 - 3% is small (19.4%) compared to Devon and Cornwall (30%), Derbyshire (49.7%) and Somerset (59%). However, over two-thirds of the county is characterised by radon levels of 3 - 10% and the graph shows this rapid rise in the proportion of area within this radon level, whereas the four other counties experience a decline. This is followed by a decline to 12.5% associated with radon levels between 10 and 30%. The difference between Northamptonshire and the other counties is depicted on the graph below.

Figure 13.2 The Distribution of Radon Levels as a Proportion of Total County Area



13.2 Land Capability

Land capability in the four areas is dominated by Grade 3 land, although within this broad generalisation there is a great deal of variation. In Derbyshire for example, nearly half of the total area of the county is Grade 4, and Grade 3 land accounts for just 23% (well below the national average which is approximately 50%). Table 13.2 provides a summary of the differences and similarities between the four areas in terms of land capability. The largest category for each Affected Area is highlighted in bold.

Table 13.2 The Proportion of Homes in each Land Capability Grade as a Percentage of the Total Area of each Affected Area

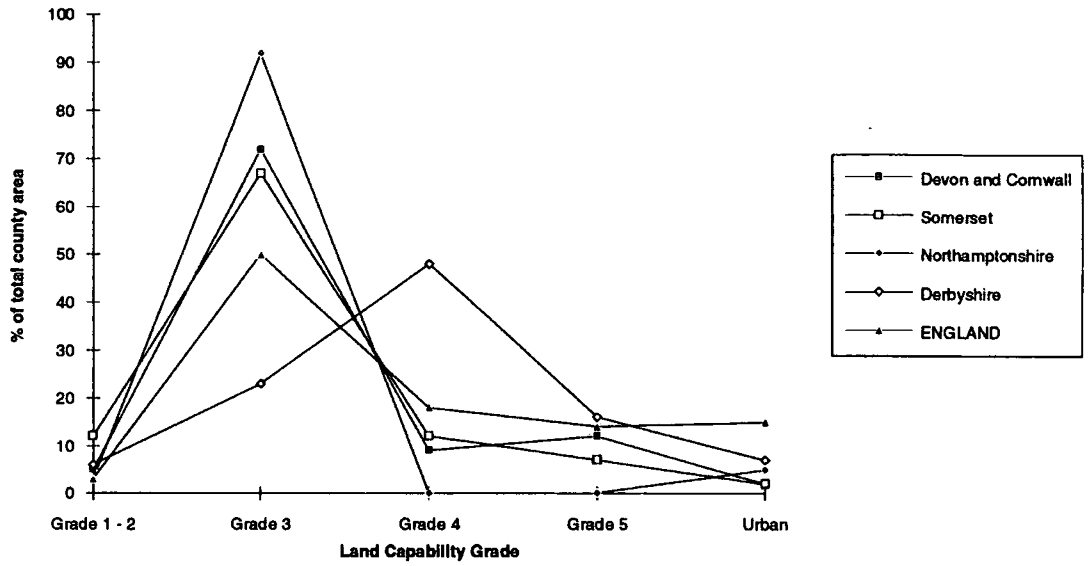
	Grade 1 - 2	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	5	**72**	9	12	2
Somerset	12	**67**	12	7	2
Northamptonshire	3	**92**	0	0	5
Derbyshire	6	23	**48**	16	7

A comparison between the land capability classes is shown on Figure 13.3 below. Whilst Devon and Cornwall, Somerset and Northamptonshire enjoy a similar distribution, Derbyshire displays a quite different pattern. This may be due to Derbyshire's location at the southern end of the Pennines, with its resulting, often marginal, physical characteristics. The upland nature of the county and the generally poor soils combine to give Derbyshire its overall poor land capability for agriculture. Derbyshire also has the highest proportion of very poor quality agricultural land (Grade 5), which accounts for almost one sixth of the county. By comparison, Northamptonshire has no poor quality agricultural land (according to the MAFF land capability land classification maps at the scale of 1:250,000).

The dominance of Grade 3 is most apparent in Northamptonshire, where it accounts for 92% of the total area of the county. Devon, Cornwall and Somerset, perhaps due to their close areal proximity in the south-west of the country, have similar proportions of land in each grade, with Grade 3 land accounting for between 67% and 72% of total area.

The proportions of urban land in the four areas varies between 2% and 7%. It is lowest in the south-western counties and highest in Derbyshire. In England, urban land accounts for 1,920,000 hectares or 14.9% of the total area (Sinclair, 1992), a figure which serves to highlight the predominantly rural nature of the Affected Areas. As discussed in Chapter 10, the figures used in this thesis are taken from the MAFF land capability maps and the boundaries of many of the urban areas will have spread into the surrounding farmland since the land capability classification was carried out in the late 1970s and early 1980s. However, whilst the proportion of land in the United Kingdom classified as urban increased slightly between 1981 (88.8%) and 1991 (89.1%) the change is unlikely to have had any significant impact on the results of the land capability overlay in this thesis.

Figure 13.3 The Percentage of the Total Area for each Land Capability Grade in the Affected Areas



A comparison with approximate figures for England have been included to show the overall national pattern. The England curve shows less variation between the classes and a greater proportion of urban land, and in this way it is similar to Derbyshire. However, the overall shape of the curve approximates to Devon and Cornwall, Somerset and Northamptonshire.

13.3 The Relationship between Land Capability and Radon Gas

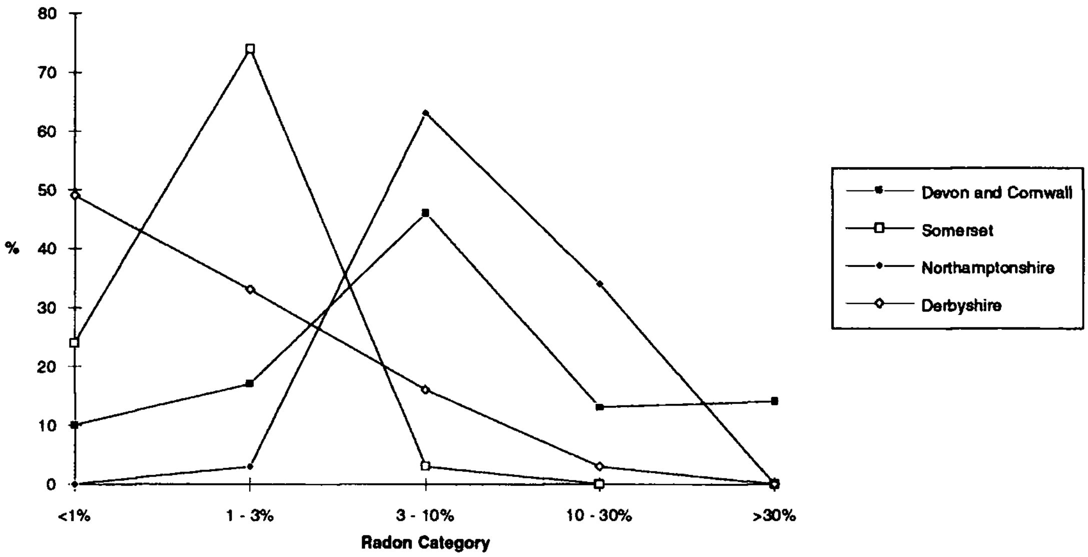
The effect of land capability on radon levels is not uniform in the five counties, but land capability grade does seem to exert some influence in all the Affected Areas. The graphs of land capability and radon for the four areas are problematic to compare and contrast because of the complex nature of the data (there are 25 classes in total), so reference is made to graphs incorporated in the text in the previous chapters (see Chapters 12.2, 12.3, 12.4 and 12.5). However, a tabular summary of the data is provided in Table 13.3 below.

Overall, there is an inverse relationship between the two variables, so that as radon levels increase, the capability of land for agriculture decreases. This pattern does not hold true for urban land which shows a variety of different relationships in the Affected Areas and is discussed separately (see below). The observed relationship is strongest in Devon and Cornwall (see Figure 12.2.6) most noticeably with radon levels of 1 - 3% and >30%. In Derbyshire, which has a more even spread of land capability, the relationship between the poorer land and radon levels 10 - 30% is strong, as is the association between Grades 1 to 3 and the lower radon levels (<1%).

In Somerset and Northamptonshire, the relationship between land capability and radon is affected by the dominance of certain categories. In Somerset, the two radon levels 1 - 3% (50%) and 3 - 10% (40%) together account for 90% of the area of the county and in Northamptonshire, Grade 3 land alone accounts for over nine tenths of the total area. Thus the correlation between radon and land capability is less marked than in Derbyshire and Devon and Cornwall. In Northamptonshire only tentative conclusions can be drawn from Figures 12.4.7 and 12.4.8 because of the lack of corroborating evidence from Grade 4 and 5 land.

The relationship between urban land and radon differs in all Affected Areas, although there are some similarities between the counties shown in the curves in Figure 13.4 below. Derbyshire shows a strong inverse linear relationship between urban land and radon levels. In this county urban land is closely associated with the lower radon levels (82% of urban land corresponds to radon levels between 0 - 3%), partly as a result of the location of the urban areas in the low lying valleys of the south of Derbyshire (see Chapter 12.5).

Figure 13.4 Urban Land and Radon Levels in the Affected Areas



The patterns between urban land and radon levels in Devon and Cornwall and Northamptonshire are related in that the two curves show distinct peaks in the 3 - 10% radon category. The association in Northamptonshire is more distinct but both show a close association with the middle radon levels. Devon and Cornwall however, also have a significant correlation with 10 - 30% and >30% that is not evident in the Northamptonshire data. Somerset shows a similar peak but is most closely associated with the lower radon levels of less than 3%. The proportion of urban land with low radon levels is around a quarter of the total area in Somerset but then rises to almost three-quarters (74%) with radon levels 1 - 3% before falling rapidly.

To summarise, the urban areas of Devon and Cornwall and Northamptonshire are characterised by relatively high radon levels, whereas in Somerset and Derbyshire radon levels are relatively low. Table 13.3 provides a statistical summary.

Table 13.3 Radon and Land Capability Summary Data*

<1%	**Grade 1 - 2**	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	5	3	3	0	10
Somerset	1	13	0	2	24
Northamptonshire	0	0.5	0	0	0
Derbyshire	36	27	6	0	49
1 - 3%	**Grade 1 - 2**	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	53	29	28	2	17
Somerset	69	50	26	44	73.5
Northamptonshire	3	20.5	0	0	3
Derbyshire	18	66	26	25	33
3 - 10%	**Grade 1 - 2**	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	6	22	27	6	46
Somerset	30	36	72	54	2.5
Northamptonshire	71	68	0	0	63
Derbyshire	46	7	35	47	16
10 - 30%	**Grade 1 - 2**	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	24	33	33	23	12
Somerset	0	1	2	0	0
Northamptonshire	26	11	0	0	34
Derbyshire	0	0	33	28	3
>30%	**Grade 1 - 2**	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
Devon and Cornwall	12	14	9	69	14
Somerset	0	0	0	0	0
Northamptonshire	0	0	0	0	0
Derbyshire	0	0	0.3	0	0

*percentages are a proportion of each Grade for each county

13.4 Social Factors in Northamptonshire - A comparison between the Affected Areas and with England as a whole

The impact of various social factors on radon levels was studied for the county of Northamptonshire. The results are discussed in detail in Chapter 12.4.4. It is interesting however, to view the social aspects of the country from a wider perspective. Table 13.4 below incorporates some data for Northamptonshire and presents a comparison with the other Affected Areas and with England as a whole. The overall population in the five counties (1991 Census) is 3,446,186 with Cornwall and Somerset both having less than half the total number of people in Devon and approximately half the population of Derbyshire. Derbyshire and Devon both have an estimated population in excess of 928,000 and Northamptonshire of over 578,000, whereas both Cornwall and Somerset have a population of less than 470,000. Together, the five Affected Areas accounted for 6.4% of the population of Great Britain in 1991 and 8.2% of the area. From analysing the social class distribution in the five counties and comparing it to the national average for England, it is clear that spatial patterns exist in social class at county level.

The lower population total in Cornwall is partly a result of the physical nature of the county. Farming and fishing are the main industries, together with tourism. The location of the county in the extreme south-western tip of the country does not make it favourable for the location of manufacturing industries, and primary industries can support only a low overall population density. Farming and fishing in Cornwall are on the whole traditionally family oriented and the methods employed are not manpower intensive, giving rise to sparsely populated rural areas and small fishing villages and market towns. The tourism industry (a major source of income for the county) is seasonal in its manpower requirements and in many places relies on people moving in from outside the county for the duration of the holiday season. The same is true of certain crops such as potatoes, soft fruit, bulbs and flowers, which have an intensive harvesting season. For these reasons, the actual population of Cornwall may often be more than the resident population at any one time. Overall the picture is one of a county dominated by the middle to lower classes, especially skilled manual workers (farmers, fishermen and the like), with a predominance of retired people who move into the south-west in order to live out their old age in a mild climate by the coast.

Devon is a large county in terms of area and population and is characterised by a variety of landscapes. The western part of the county is overshadowed by Dartmoor, a large area of open moorland which sustains only marginal livestock farming and some quarrying. In the far north-east is Exmoor, similarly marginal in nature though much smaller in area (Exmoor also extends into Somerset). The south and east of the county are characterised by rolling hills and valleys and a predominance of dairy farms. The towns of Plymouth and Exeter attract a variety of industries, especially Plymouth which has a major naval base and a working dockyard. Devon has the largest proportion of retired people out of the five counties, and at 30.2% the figure is significantly higher than the average for England which is 26.3% (see Table 13.5).

Somerset is largely an agricultural county but a large amount of the land is marginal in nature and the flat peaty Levels, Quantock and Mendip Hills and Exmoor cannot sustain either a large agricultural population or intensive farming methods. Industries associated with the Levels such as peat cutting and basket-weaving have declined over the last fifty years and there has been a flux of people out of the rural areas and into the main towns of Yeovil, Taunton, Bridgwater and the like, where various manufacturing industries (for example food and chemical industries) have located. Indeed the proximity of parts of Somerset to the south-west's regional capital Bristol, has encouraged more industry to locate in the county and the development of a good road and rail infrastructure has further enhanced the attractiveness of the county to inward investment. Partly as a result of this trend the proportion in the professional and managerial classes approximates to the national average.

Since the late eighteenth century, Derbyshire has been a traditional coal mining area, with roots in the textile industries that sprang up in the north of England in the early 1900s. The area is characterised by populated valleys and the steep rise of the Derbyshire Dales. The decline in coal mining in recent years has lead to out migration in parts of Derbyshire and an increase in unemployment in the region. The north-west of the county is less densely populated than the south and east due to the steep-sided valleys and moorland landscape of the Peak District, although some towns such as Buxton and Bakewell have prospered in the valleys. Sheep farming is the predominant activity in the upland region, although tourism generates seasonal income and is centred around 'honey-pot' locations such as Mam Tor and Derwent Reservoir, and the towns of Hathersage and Castleton. The lowland section of the county to the south and east is characterised by a high population density concentrated in the towns of Derby, Chesterfield, Ripley and the like. Industry is concentrated in these lowland towns. The county is dominated on the whole by the three lower classes, namely manual and partly skilled or unskilled workers and has the highest overall population density of the Affected Areas.

Northamptonshire is located at the southern end of the East Midlands and is able to sustain a high population density on its largely good quality arable land. Some industry has located in the county, especially following the movement of service industries and light manufacturing industries out of London over the last 20 years. The main urban areas of Northampton (the population of Northampton district increased from 155,536 in 1981 to 176,021 in 1991), Kettering, Daventry and Wellingborough are all traditionally market towns serving the surrounding agricultural population. Northamptonshire's proximity to the southern markets and the good transport network, such as the M1 and direct rail routes to London, means that it is well situated for easy access to the capital and the south-east. The county as a whole is dominated by the middle classes and it is the only one of the five Affected Areas to have a retired population below the national average.

There are clearly some similarities between the social structures of the three counties in the south-west, especially between Cornwall and Devon. Northamptonshire however has a quite different class structure influence by its environment and position in the country. The social structure in Derbyshire is related to the traditional industries of the county and its marginal landscape and differs in most respects from the other Affected Areas. All five counties have some unique characteristics however and the differences, as well as the similarities, in terms of social class have been outlined in more detail below (see Table 13.5).

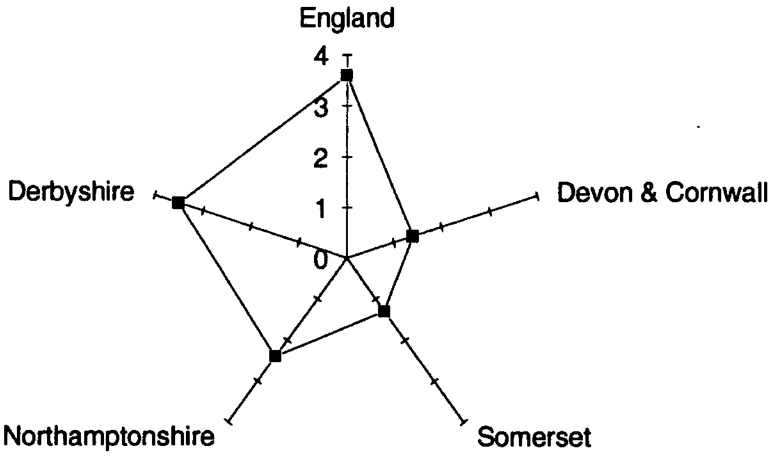
Table 13.4 General Census Statistics for the Affected Areas and for England (1991 Census)

	Devon & Cornwall	**Somerset**	**Northamptonshire**	**Derbyshire**	**ENGLAND**
Persons present 1981	1,390,985	427,167	528,443	910,141	46,226,100
Persons present 1991	1,486,881	461,647	570,598	918,648	46,382,051
% of 1991 total for GB	2.7	0.9	1.1	1.7	85.6
Area (hectares)	1,027,427	345,207	236,697	262,858	13,036,268
Residents per hectare	1.4	1.3	2.4	3.5	3.6
% aged under 16	18.8	19.2	21.5	19.6	20.1
% of households consisting only of pensioners	29.2	28.7	22.9	25.3	24.8*
% aged 75 and over	9.2	8.6	6.3	6.9	7.1

* Great Britain (Source: OPCS 1993a,c,d,e,f)

A graph of the population densities in the four areas and in England is shown in Figure 13.5 below. The population density in Derbyshire approximates to the average for England and is one third higher than the next highest county, Northamptonshire. The average for Great Britain is just 2.4 persons per hectare, a figure which serves to highlight the sparsely populated nature of Scotland and Wales.

Figure 13.5 Residents per Hectare in the Affected Areas; a comparison with England (1991 Census)



The results of the analysis between radon and social aspects of the population in Northamptonshire (see Chapter 12.4.4) show that high population density at ward level is associated with high radon levels. If this correlation holds true throughout the Affected Areas, then it would be expected that Derbyshire, with the highest population density would also have high radon levels. However, this association is not substantiated by the results of the analysis between urban areas and radon levels in Derbyshire (see Figure 13.4), which show that urban areas are more likely to be associated with the lower radon levels. In Devon and Cornwall however, which have low average population densities, urban areas are almost entirely associated with areas having more than 3% of homes above the Action Level. There appears, therefore, to be no pattern between the counties with regards to population density and the proportion of homes above the Action Level. This finding illustrates the fact that associations which exist at one spatial level (*e.g.* wards in Northamptonshire) do not necessarily exist at another (*e.g.* at county level for the other Affected Areas). This phenomenon is known as the aggregation problem (Openshaw, 1984) and has been discussed in more detail in Chapter 10.3.2. The sensitivity of spatial statistics to zonal definition has been effectively demonstrated by Openshaw and Taylor (1982).

A summary of the differences and similarities between the counties in terms of social class is discussed below, with reference to Table 13.5. A description of the classes is provided in Chapter 4 and Chapter 12.4.4.

Table 13.5 Social Class Statistics for the Affected Areas and for England
(1991 Census - % of households in each class)

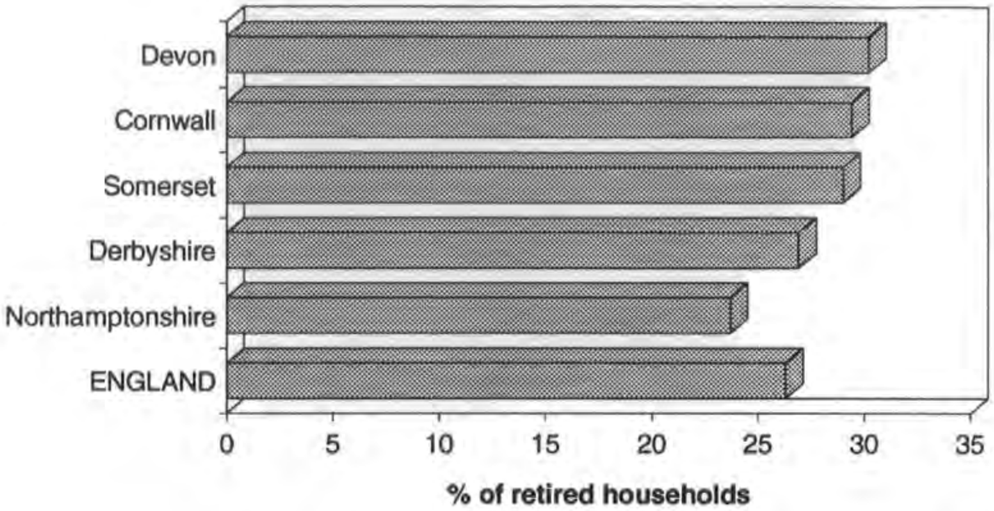
County	**I**	**II**	**IIIN**	**IIIM**	**IV**	**V**	**Retired**	**Other***	**Inad. described**
Devon	3.2	17.7	7.8	15.6	8.1	2.8	30.2	3.0	11.7
Cornwall	2.7	16.7	7.0	16.3	8.0	3.2	29.4	3.6	13.1
Somerset	3.8	19.6	7.4	17.2	8.7	2.8	29.0	1.9	9.6
Derbyshire	4.1	17.5	7.1	19.9	9.0	3.2	26.9	0.9	11.6
Northants	3.7	20.5	8.6	20.8	9.8	2.9	23.7	1.3	8.4
ENGLAND	**4.3**	**19.5**	**8.6**	**16.6**	**8.5**	**2.8**	**26.3**	**1.6**	**11.8**

* Armed forces/on a government training scheme (Source: OPCS 1993a,c,d,e,f)

Social class was analysed with respect to radon levels at district level for the county of Northamptonshire. Analysis of the overlays identified that, in Northamptonshire, low radon levels are more likely to be associated with higher proportions in social classes I and II and lower proportions in IV and V. Northamptonshire is the only Affected Area to have above the national average in classes I and II, although Somerset approximates to the national average. Cornwall has the lowest percentage of the Affected Areas in class I with only 2.7% of the population. With the exception of Northamptonshire, all the Affected Areas have fewer than the average proportion in class IIIN (Cornwall has only 7% in this category), but Somerset, Northamptonshire and Derbyshire have a higher than average proportion in class IIIM and class IV (partly-skilled occupations). Class V (unskilled occupations) is the lowest social class and on average accounts for only 2.8% of the households in England. The Affected Areas all have similar proportions in this class.

The three counties in the south-west of the country, Cornwall, Devon and Somerset, have the highest percentage of retired households (29.4%, 30.2% and 29.0% respectively). This is due to the well established migration of retired people to the south and west of England. The proportion of retired people in Derbyshire approximates to the national average and Northamptonshire has fewer than the average with just 26.3%.

Figure 13.6 A Comparison of the Percentage of Households Consisting Only of Pensioners* in the Affected Areas (1991 Census)



* Women aged 60 years and over and men aged 65 years and over

The proportion of households consisting only of pensioners was analysed with respect to radon levels in Northamptonshire. The results from this county highlight a relationship between high radon levels (>10%) and areas with at least one in four homes consisting only of pensioners. Although Northamptonshire has fewer pensioners than the national average, they tend to live in and around the urban areas, where radon levels are generally higher. The counties with the highest proportion of households consisting only of pensioners are Devon and Cornwall, which also have the greatest proportion of people aged 75 years and over (9.6% and 8.8% respectively). In general, the south-west of England (including Somerset) has an aged population structure, with the highest proportion of dependants living on the south coast and in the south-west, compared to anywhere else in the United Kingdom. The observed relationship at county level between high radon levels in Devon and Cornwall and a high proportion of pensioners is not necessarily indicative of a causal relationship. The influence of geology for example, in these counties is more likely to the cause of elevated radon levels. In the absence of any other findings regarding radon levels and population age structure, it is therefore problematic to extrapolate correlations for the other Affected Areas based on the results identified for the county of Northamptonshire. Indeed any results found in Northamptonshire are specifically related to the social and physical characteristics of Northamptonshire alone. A graph showing the proportion of pensioners in the Affected Areas is included for the purposes of comparison (see Figure 13.6).

Whilst it is possible to highlight differences and similarities between the counties in terms of radon, land capability and social factors, it is problematic to identify a single cause for any variations that are identified. Differences in the natural and social characteristics of a county relate specifically to the spatial scale, time and location under investigation and conclusions drawn from analyses can not be confidently applied to another scale, time or location. The comparison does however, serve to highlight aspects of similarity and continuity between the Affected Areas, which could be used as a basis for further investigation. It also highlights aspects of the counties that differ, which can be equally as important in assessing why patterns exist in national and local radon levels. A summary of the main findings from this chapter can be found in Chapter 14.

14. Summary

The aim of this thesis - to assess the geographical differences and similarities of radon Affected Areas in England - was furthered by the identification of key data sources which were mapped in conjunction with radon data using ARC/INFO. Part One of the thesis is concerned with the availability of the key datasets, namely radon data, land capability data and census data. The radon data were supplied by the NRPB on diskette for the five Affected Areas. The land capability data were digitised from MAFF source maps and integrated with the radon data in ARC/INFO. It is envisaged that in the future the bulk of geographical data at both global and national scales will be derived from earth observation from space. Thus the trend in land use data provision, away from conventional paper map sources towards digital data compiled from satellite imagery, is part of a much wider move within geography as a whole. In the United Kingdom, the publication (and subsequent release of a digital version in 1994) of the Land Cover Map of Great Britain has significantly advanced the availability of land cover data in this country. Similarly, the development of digital data dissemination for census data over the last fifteen years has greatly facilitated the use of these data within GIS. There were a number of new questions in the 1991 Census, among them the postcode of households. The future of spatial data provision seems set to lie in the increased use of postcodes, which offer more spatial detail (and a greater scope for aggregation) than census zones (enumeration districts, wards and the like). In addition, the re-representation of existing spatial data, for example the generation of 'surfaces' for derived socio-economic variables, represent a fundamental re-modelling away from a zonal to a surface-based form or representation. Population geographers are increasingly concerned with new methods of displaying discrete data, for example Martin and Bracken (1993) offered population-weighted centroids.

The use of data that have been aggregated for artificially imposed areal units such as grid squares or wards, inevitably introduce some error. The problems of the ecological fallacy and the MAUP are ever prevalent in aggregated data, especially when the user has no control over the method or scale of aggregation. As areas used to display the data become larger, the level of generalisation increases and, since the areal units are often irregularly shaped, the effect of the areal partition is directionally variable and is very difficult to quantify (Langford and Unwin, 1994). Openshaw and Taylor (1982) effectively demonstrated the sensitivity of spatial statistics to zonal definition (see Chapter 10.3.2). In addition to these inherent sources of error, some error was introduced into the analysis during the digitising process and as a result of discrepancies between the census ward boundaries and ward data.

The identification of several questions involving radon data, land capability data and 1991 Census data, served as a useful framework for the analysis of the findings in Part Three. The questions that were posed are as follows:

1. What spatial patterns exist in the radon data set?
2. What spatial patterns exist in the land capability data set?
3. What spatial patterns and relationships exist between radon gas and land capability?

In addition (in Northamptonshire):

4. What is the relationship between population density and radon levels?
5. What is the relationship between social class and radon levels?
6. What is the relationship between the proportion of households containing only pensioners and radon levels?

Once the land capability, census and radon data had been digitally captured and integrated in ARC/INFO the *identity* command was issued to overlay the different datasets. This enabled any spatial relationships between the data to be identified and to some extent quantified using statistics produced automatically in ARC/INFO. Although the most accurate way of establishing correlations between social information, land capability zones and indoor radon data is to examine site-specific data this does not facilitate the assessment of geographical differences and similarities at the county level. Data at the ward, district and five kilometre grid square level were integrated with the land capability zones and, through the use of exploratory data analysis, conclusions were drawn about the human and physical characteristics shared by the Affected Areas. The results of the overlays have been discussed in Chapters 12 and 13 using descriptive statistics (*e.g.* percentages, modes, distributions) rather than adopting the inferential method. The plethora of maps and graphs included in Chapters 12 and 13 portray the results of the overlays in a way which is clear and thought provoking. A distinct advantage of including few statistical analyses is that the data can be analysed in what is essentially their raw format and the reader is open to make his or her own conclusions. Some similarities between the Affected Areas have been identified from the results of the overlays, as well as some noteworthy differences. The reader is referred to Chapter 13 for a more detailed assessment of the principal geographical differences and similarities of the Affected Areas. The following provides a summary of the main findings.

The general trend, identified from the initial results obtained from Devon and Cornwall, suggests that as land capability for agriculture decreases (from Grade 1 to 5), the level of radon in the home increases. The close correlation between land capability and bedrock (*e.g.* Grade 5 land and granite in Devon and Cornwall; Grades 1 and 2 and New Red Sandstone in east Devon) means that land capability will undoubtedly appear to influence indoor radon levels because of its association with the underlying bedrock. In general, higher radon levels in all the Affected Areas are associated with the poorer land capability classes (*i.e.* Grades 4 and 5), although the relationship is strongest in Devon, Cornwall and Derbyshire. Results in Northamptonshire are strongly influenced by the dominance of the Grade 3 category (which accounts for 92% of the county area), indeed only tentative conclusions can be drawn from the results of the land capability and radon overlay because of the lack of any Grade 4 or 5 land. In Somerset the relationship is affected by the dominance of the radon categories 1 - 3% and 3 - 10%, which together account for 90% of the total area of the county.

The relationship between urban land and radon levels does not follow the general trend identified above (probably because of the unique impact of urbanisation on the land), nor does it conform to any pattern. The urban areas of Devon, Cornwall and Northamptonshire are characterised by relatively high radon levels, namely grid squares with more than 3% of homes estimated to be above the Action Level. On the other hand, radon levels in urban areas in Somerset and Derbyshire are low. In Derbyshire for example, 82% of the urban land is associated with areas where fewer than 3% of homes exceed the Action Level.

Social class data (classes I to V) from the 1991 Census were mapped in conjunction with the radon data for the seven districts of Northamptonshire. The relationship between radon and social class is most clearly illustrated in social classes I and V (the two extremes). Results show that polygons in the overlay with an above average proportion in class I are more likely to have a low percentage of homes above the Action Level. On the other hand, polygons with above average proportions in class V are slightly more likely to be associated with areas having more then 3% of homes above the Action Level. In summary, areas with a relatively low proportion of homes above the Action Level are more likely to have a high proportion in social classes I and II and a low proportion in classes IV and V. At the county level, Northamptonshire and Somerset have approximately average proportions in classes I and II with respect to social class in England. Cornwall and Devon have a similar overall social structure to each other, and share with Derbyshire an affinity with the lower social classes. The three counties all have lower proportions than the average for England in the higher social classes (I, II and IIIN) and higher than average proportions in the lower social classes (IIIM, IV and V). In Devon, Cornwall and Somerset retired households account for more than 29% of all households (compared to the English average of 26.3% and a figure for Northamptonshire of just 23.7%). Results in Northamptonshire (at ward level), between radon levels and the proportion of households consisting only of pensioners, identified a relationship between areas of high radon (>10% of homes above the Action Level) and areas where 25% or more of households contain only pensioners.

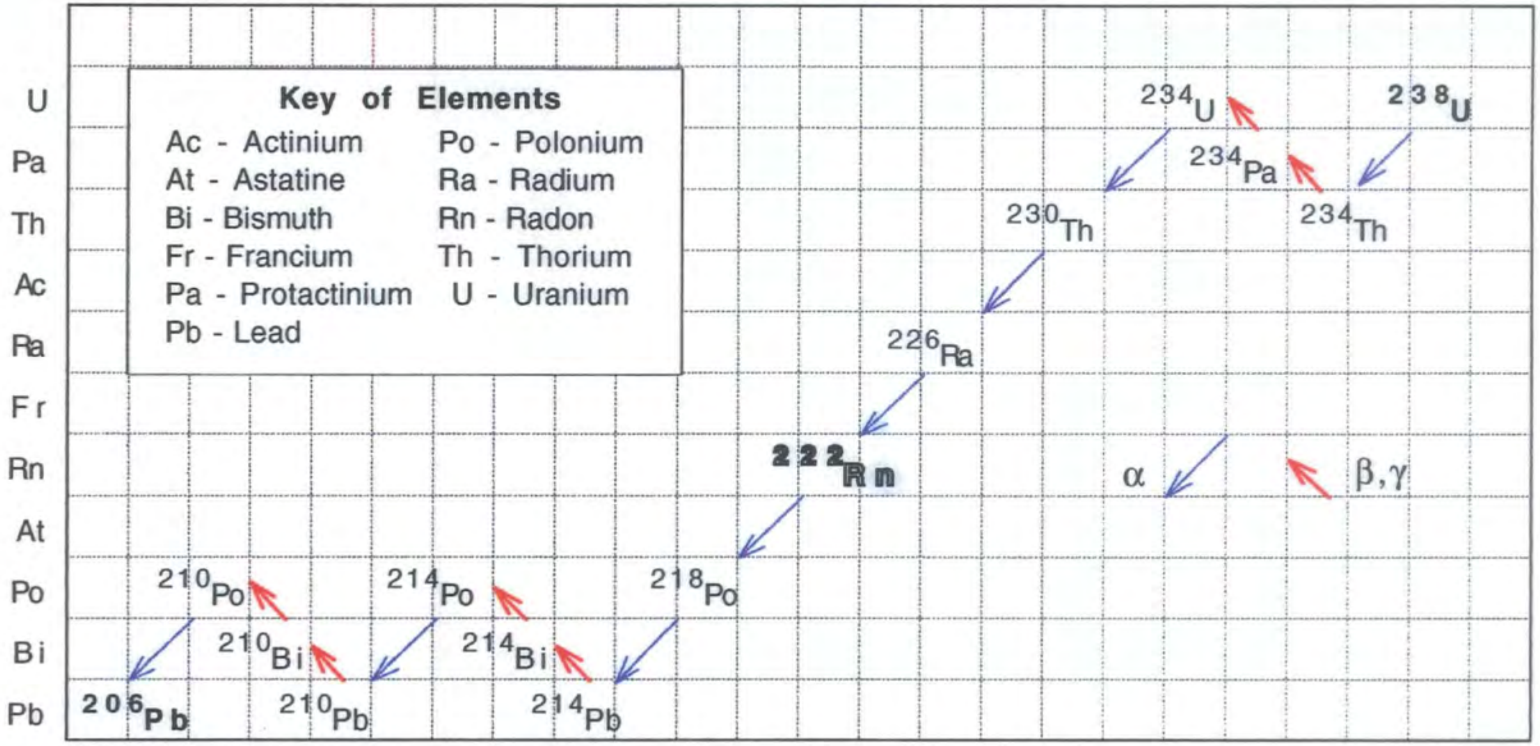
Population density at both ward and district level was analysed for Northamptonshire. A relationship was found to exist between high radon levels (>10% of homes above the Action Level) and high population density (more than 50 persons per hectare), at ward level. The converse is also true, so that low population densities are, on the whole, more likely to be associated with areas with lower radon levels. This result agrees with the correlation identified above between urban areas and radon levels in Northamptonshire. Population density for the seven districts of Northamptonshire produced similar overall results to the analysis at ward level, except for a significant difference between the two sets of data within the radon category

3 - 10%. The ward data show a strong correlation between the radon category 3 - 10% and densities of less than 20 people per hectare, whereas the district data show an equally strong association with both the lowest population densities (0 - 4 people per hectare) and the highest (20 - 24 people per hectare). This is likely to be a result of the dominance of the 3 - 10% radon category in the county as a whole, coupled with the poor resolution of the population density data at district level. With regards to the other Affected Areas, there appears to be no pattern between radon and population density at the county level (for example, as shown above urban areas show markedly different relationships).

Indoor radon levels depend on numerous physical and social parameters, not least local variations in radium concentrations, ground permeability, soil moisture content, weather conditions, details of house construction and habits of occupants. This thesis has not aimed to tackle the influence of all of these factors but by choosing to analyse one aspect of the physical environment and one aspect of the social environment, some geographical differences and similarities between the Affected Areas have been identified. It is precisely because of the combination of factors listed above that any simple correlation identified between radon and another aspect of the physical or social environment will be predisposed to uncertainty and will be compounded by the implicit complexity of the radon hazard. Thus the relationships identified in this thesis require further testing using site-specific data in controlled conditions before any firm conclusions can be drawn as regards the spatial characteristics of the gas.

Appendices

Appendix 1 Radioactive Decay Series of 238-Uranium



(Source: NRPB)

Appendix 2 County Council Data Sets

Cornwall County Council, Truro

All 1981 and 1991 census data for the county is held on computer. A separate mapping package enables the data to be thematically mapped down to enumeration district level. Enumeration district, ward and district maps are held in both hard copy and on computer.

Aerial photographs are periodically commissioned for the whole county. The latest available photographs were taken in the late 1980s.

There is no central land use survey data bank. Any data on land use change is held by the district councils, if at all. This situation is unlikely to change in the near future.

There is no GIS currently in use.

Devon County Council, Plymouth

All 1981 and 1991 census data is available for the county in hard copy format. There are selected datasets only, from the two censuses held on a mainframe computer. A mapping package is used to output thematic maps for parishes and wards. Parish and ward maps are also available as hard copy.

There is an up to date set of aerial photographs for the whole county flown in 1992 and 1993 at 1:250,000. These are currently being converted into digital format at Devon County Council, but the process is slow and will take some time. These will eventually produce and up to date source of land use data for the county.

Land use data - see above

There is currently no GIS in use, but there are plans to buy one in the near future.

Somerset County Council, Taunton

Paper maps of the 1971 boundaries are available. Census data for the 1981 and 1991 censuses are held in both hard copy format and on computer. Digital maps down to enumeration district are available for both censuses. Thematic census maps can be output using the GIS Atlas. The data is held on a stand-alone computer, running various packages including SASPAC.

Aerial photographs for the whole county are available for 1971 (black and white), 1981 (black and white) and 1992 (colour) and were purchased from Geonex. The 1992 photographs are at 1:12,000.

Land use maps of the county are available in hard copy and have been produced from the aerial photographs.

The GIS 'Atlas' is used.

Northamptonshire County Council, Northampton

The OPCS county reports are available in hard copy format for 1981 and 1991 censuses. Maps from the 1981 and 1991 censuses down to enumeration districts are held in both hard copy and also in digital format in the GIS MapInfo which is used to produce thematic maps of the county. The Council are registered to use NOMIS census data.

There are no up to date aerial photographs available for the county.

The Council have bought an Ordnance Survey copyright license and hold OS maps of the county in digital format. These are used as the primary source of land use data. The MAFF Land Capability Classification maps are available.

The GIS 'MapInfo' is used.

Derbyshire County Council, Matlock

Census data for the 1981 and 1991 censuses are held on a mainframe computer which runs SASPAC software. 1991 maps are available in digital format and were purchased from OPCS. 1981 maps are only available in paper format. There has been a move towards digital data and it is possible that there will be more data available in digital format in the future. There is no 'in-house' mapping package but there is some liaison with a grant-maintained school that has a mapping package capable of mapping census data.

The most recent aerial photographs are approximately 15 years old.

There is no GIS currently in use.

Appendix 3 Revised Land Capability Figures for Somerset

The impact of land capability digitising errors on the results of the overlay between land capability and radon data (square metres)

Somerset - Digitised Data (square metres)

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
<1%	3,951,780	291,039,611	0	4,451,633	20,001,395	319,444,419
1 - 3%	280,745,566	1,149,369,692	104,142,329	102,539,837	61,403,736	1,698,201,160
3 - 10%	123,021,578	841,574,272	290,359,295	123,698,711	2,117,864	1,380,771,720
10 - 30%	0	32,769,459	9,562,556	0	0	42,332,015
>30%	0	0	0	0	0	0
TOTAL	407,718,924	2,314,753,034	404,064,180	230,690,181	83,522,995	3,440,749,314

Somerset - Errors corrected

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
<1%	3,951,780	291,039,611	0	4,451,633	20,001,395	319,444,419
1 - 3%	280,745,566	1,149,323,692	104,182,329	102,539,837	61,409,736	1,698,201,160
3 - 10%	123,021,578	841,554,272	290,379,295	123,698,711	2,117,864	1,380,771,720
10 - 30%	0	32,769,459	9,562,556	0	0	42,332,015
>30%	0	0	0	0	0	0
TOTAL	407,718,924	2,314,687,034	404,124,180	230,690,181	83,528,995	3,440,749,314

% Difference

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
<1%	0	0	0	0	0	0
1 - 3%	0	0	0.001	0	0.002	0.003
3 - 10%	0	0.0006	0	0	0	0.0006
10 - 30%	0	0	0	0	0	0
>30%	0	0	0	0	0	0
TOTAL	0	0.0006	0.001	0	0.002	0.0036

Appendix 4 Extract From the 1991 Census Form: Questions 14 - 16

lease read A below, tick the box that applies and follow the instruction by the box ticked.

A Did the person have a paid job last week (any of the boxes 1, 2, 3 or 4 ticked at question 13)?	YES ☐ Answer questions 14, 15, 16, 17 and 18 about the main job last week, then go on to question 19 NO ☐ Answer B	YES ☐ Answer questions 14, 15, 16, 17 and 18 about the main job last week, then go on to question 19 NO ☐ Answer B
B Has the person had a paid job within the last 10 years?	YES ☐ Answer questions 14, 15 and 16 about the most recent job, then go on to question 19 NO ☐ Go on to question 19	YES ☐ Answer questions 14, 15 and 16 about the most recent job, then go on to question 19 NO ☐ Go on to question 19
Hours worked per week How many hours per week does or did the person usually work in his or her main job? **Do not count** overtime or meal breaks.	Number of hours worked per week ☐	Number of hours worked per week ☐
Occupation Please give the full title of the person's present or last job and describe the main things he/she does or did in the job. **At a**, give the full title by which the job is known, for example: 'packing machinist'; 'poultry processor'; 'jig and tool fitter'; 'supervisor of typists'; 'accounts clerk'; rather than general titles like 'machinist'; 'process worker'; 'supervisor' or 'clerk'. Give rank or grade if the person has one.	**a** Full job title	**a** Full job title
At b, write down the main things the person actually does or did in the job. If possible ask him/her to say what these things are and write them down. **Armed Forces** — enter 'commissioned officer' or 'other rank' as appropriate at **a**, and leave **b** blank. **Civil Servants** — give grade at **a** and discipline or specialism, for example: 'electrical engineer'; 'accountant'; 'chemist'; 'administrator' at **b**.	**b** Main things done in job	**b** Main things done in job
Name and business of employer (if self-employed give the name and nature of the person's business) **At a**, please give the name of the employer. Give the trading name if one is used. Do not use abbreviations.	**a** Name of employer	**a** Name of employer
At b, describe clearly what the employer (or the person if self-employed) makes or does (or did). **Armed Forces** — write 'Armed Forces' at **a** and leave **b** blank. For a member of the Armed Forces of a country other than the UK — add the name of the country. **Civil Servants** — give name of Department at **a** and write 'Government Department' at **b**. **Local Government Officers** — give name of employing authority at **a** and department in which employed at **b**.	**b** Description of employer's business	**b** Description of employer's business

Appendix 5 Summary Overlay Data for the Affected Areas

Results of the overlays between land capability and radon data for the Affected Areas (square metres)

Devon and Cornwall

Sum area of polygons						
	Grade 1 - 2	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**	**TOTAL**
<1%	23,968,322	191,179,206	25,000,000	0	24,865,223	265,012,751
1 - 3%	264,423,468	2,142,017,181	255,808,196	22,405,473	42,398,995	2,727,053,313
3 - 10%	31,898,701	1,612,256,319	247,453,223	67,441,823	112,376,714	2,071,426,780
10 - 30%	121,963,461	24,10,383,243	297,449,646	274,120,328	30,582,845	3,134,499,523
>30%	57,763,111	1,054,766,417	80,400,853	828,240,281	35,053,749	2,056,224,411
TOTAL	500,017,063	7,410,602,366	906,111,918	1,192,207,905	245,277,526	10,254,216,778

Mean polygon area					
	Grade 1 - 2	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
<1%	2,996,040	10,062,063	25,000,000	0	6,216,306
1 - 3%	5,876,077	14,473,089	12,181,343	2,000,684	2,649,937
3 - 10%	2,899,882	14,656,876	10,758,836	5,187,833	7,023,545
10 - 30%	4,878,538	15,451,175	9,914,988	5,832,347	1,798,991
>30%	4,813,593	11,985,755	5,742,918	14,790,005	1,947,430

Somerset

Sum area						
	Grade 1 - 2	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**	**TOTAL**
<1%	3,951,780	291,039,611	0	4,451,633	20,001,395	319,444,419
1 - 3%	280,745,566	1,149,369,692	104,142,329	102,539,837	61,403,736	1,698,201,160
3 - 10%	123,021,578	841,574,272	290,359,295	123,698,711	2,117,864	1,380,771,720
10 - 30%	0	32,769,459	9,562,556	0	0	42,332,015
>30%	0	0	0	0	0	0
TOTAL	407,718,924	2,314,753,034	404,064,180	230,690,181	83,522,995	3,440,749,314

Mean area					
	Grade 1- 2	**Grade 3**	**Grade 4**	**Grade 5**	**Urban**
<1%	1,975,890	12,653,896	0	1,112,908	4,000,279
1 - 3%	5,973,310	11,849,172	4,733,742	8,544,986	2,923,987
3 - 10%	5,348,764	13,149,598	7,445,110	4,581,439	2,117,864
10 - 30%	0	16,384,729	9,562,556	0	0
>30%	0	0	0	0	0

Northamptonshire

Sum area

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
<1%	0	9,988,080	0	0	0	9,988,080
1 - 3%	2,307,227	441,604,843	0	0	3,331,162	447,243,232
3 - 10%	52,921,208	1,466,780,011	0	0	65,856,140	1,585,557,359
10 - 30%	19,387,247	237,567,816	0	0	35,517,287	292,472,350
>30%	0	0	0	0	0	0
TOTAL	74,615,682	2,155,940,750	0	0	104,704,589	2,335,261,021

Mean area

	Grade 1- 2	Grade 3	Grade 4	Grade 5	Urban
<1%	0	4,994,040	0	0	0
1 - 3%	3,845,038	16,984,802	0	0	832,791
3 - 10%	1,960,045	18,108,395	0	0	3,873,891
10 - 30%	2,769,607	13,198,212	0	0	5,919,548
>30%	0	0	0	0	0

Derbyshire

Sum area

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban	TOTAL
<1%	57,951,311	159,503,312	66,808,576	0	87,392,142	371,655,341
1 - 3%	29,251,084	387,319,548	318,199,985	97,059,515	58,057,476	889,887,608
3 - 10%	74,688,895	38,254,148	426,034,469	186,521,425	28,203,499	753,702,436
10 - 30%	0	0	395,753,576	110,221,335	4,604,322	510,579,233
>30%	0	0	6,971,598	0	0	6,971,598
TOTAL	161,891,290	585,077,008	1,213,768,204	393,802,275	178,257,439	2,532,796,216

Mean area

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
<1%	4,829,276	6,380,132	7,423,175	0	5,826,143
1 - 3%	3,656,386	12,494,179	9,642,424	10,784,391	2,419,062
3 - 10%	4,336,112	9,563,537	12,910,135	1,167,589	2,563,954
10 - 30%	0	0	16,489,732	5,511,067	2,302,161
>30%	0	0	6,971,598	0	0

All Affected Areas - mean polygon area

	Grade 1 - 2	Grade 3	Grade 4	Grade 5	Urban
<1%	2,450,302	8,522,533	8,105,794	278,227	4,010,682
1 - 3%	4,837,703	13,950,311	6,639,377	5,332,515	2,206,444
3 - 10%	3,636,201	13,869,602	7,778,520	2,734,215	3,894,814
10 - 30%	1,912,036	11,258,529	8,991,819	2,835,854	2,505,175
>30%	1,203,398	2,996,439	3,178,629	3,697,501	486,858

Appendix 6 Summary Overlay Data for Northamptonshire

Results of the overlays between social data and radon data in Northamptonshire (square metres)

Radon and population density overlay for wards

Number of polygons				
people per hectare	**<1%**	**1 - 3%**	**3 - 10%**	**>10%**
0.0 - 0.9	4	51	192	35
1.0 - 2.9	0	24	106	18
3.0 - 19.9	0	8	56	21
20 - 49.9	0	1	22	25
>50	0	1	4	8
Sum area				
	<1%	**1 - 3%**	**3 - 10%**	**>10%**
0 - 0.9	9,429,400	365,613,042	993,264,736	17,535,261
1.0 - 2.9	0	84,472,598	369,006,049	54,081,806
3.0 - 19.9	0	17,428,949	188,727,814	31,367,815
20 - 49.9	0	240,892	29,334,335	24,063,800
>50	0	171,250	2,551,606	5,841,806
Mean area				
	<1%	**1 - 3%**	**3 - 10%**	**>10%**
0 - 0.9	2,357,350	7,168,883	5,173,254	5,010,075
1.0 - 2.9	0	3,519,692	3,481,189	3,004,545
3.0 - 19.9	0	2,178,619	3,370,140	1,493,705
20 - 49.9	0	240,892	1,333,379	962,552
>50	0	171,250	637,902	730,226

Radon and population density overlay for districts

Number of polygons				
people per hectare	**<1%**	**1 - 3%**	**3 - 10%**	**>10%**
0 - 4	4	29	106	23
5 - 9	0	7	6	0
10 - 19	0	0	0	0
20 - 24	0	0	6	5
Sum area				
	<1%	**1 - 3%**	**3 - 10%**	**>10%**
0 - 4	9,429,400	421,290,658	1,502,691,638	264,678,179
5 - 9	0	46,636,077	25,705,638	0
10 - 19	0	0	0	0
20 - 24	0	0	54,487,258	26,029,672
Mean area				
	<1%	**1 - 3%**	**3 - 10%**	**>10%**
0 - 4	2,357,350	14,527,264	14,176,336	11,507,747
5 - 9	0	6,662,297	4,284,273	0
10 - 19	0	0	0	0
20 - 24	0	0	9,081,210	5,205,934

Radon and social class I overlay for districts

Number of polygons

% in class I	<1%	1 - 3%	3 - 10%	>10%
1 - 1.9	0	7	6	0
2 - 2.9	0	0	8	3
3 - 3.9	2	6	15	10
4 - 4.9	0	12	14	5
5 - 5.9	2	11	75	10

Sum area

	<1%	1 - 3%	3 - 10%	>10%
1 - 1.9	0	466,360,767	25,705,638	0
2 - 2.9	0	0	112,022,081	50,806,201
3 - 3.9	90,889	31,778,374	184,667,221	104,601,080
4 - 4.9	0	196,819,988	266,098,139	36,136,591
5 - 5.9	9,338,511	192,692,297	994,391,455	99,163,978

Radon and social class II overlay for districts

Number of polygons

% in class II	<1%	1 - 3%	3 - 10%	>10%
10 - 14	0	7	6	0
15 - 19	2	6	23	13
20 - 24	0	12	14	5
25 - 29	2	11	75	10

Sum area

	<1%	1 - 3%	3 - 10%	>10%
10 - 14	0	46,636,077	25,705,638	0
15 - 19	90,889	31,778,374	296,689,302	155,407,281
20 - 24	0	196,819,988	266,098,139	36,136,591
25 - 29	9,338,511	192,692,297	994,391,455	99,163,978

Radon and social class IIIN overlay for districts

Number of polygons

% in class IIIN	<1%	1 - 3%	3 - 10%	>10%
6.1 - 7	0	12	14	5
7.1 - 8	0	3	44	4
8.1 - 9	4	21	54	14
9.1 - 10	0	0	6	5

Sum area

	<1%	1 - 3%	3 - 10%	>10%
6.1 - 7	0	196,819,988	266,098,139	36,136,591
7.1 - 8	0	52,703,948	572,612,163	18,753,594
8.1 - 9	9,429,400	218,402,799	689,686,974	209,787,993
9.1 - 10	0	0	54,487,257	26,029,671

Radon and social class IIIM overlay for districts

Number of polygons				
% of class IIIM	<1%	1 - 3%	3 - 10%	>10%
18 - 19.9	2	11	75	10
20 - 21.9	2	6	23	13
22 - 23.9	0	12	14	5
24 - 25.9	0	7	6	0
Sum area				
	<1%	1 - 3%	3 - 10%	>10%
18 - 19.9	9,338,511	192,692,297	994,391,455	99,163,978
20 - 21.9	90,889	31,778,374	296,689,302	155,407,281
22 - 23.9	0	196,819,988	266,098,139	36,136,591
24 - 25.9	0	46,636,077	25,705,638	0

Radon and social class IV overlay for districts

Number of polygons				
% in class IV	<1%	1 - 3%	3 - 10%	>10%
7 - 8.9	2	9	53	9
9 - 10.9	2	20	59	19
11 - 12.9	0	0	0	0
13 - 14.9	0	0	0	0
15 - 16.9	0	7	6	0
Sum area				
	<1%	1 - 3%	3 - 10%	>10%
7 - 8.9	90,889	84,482,322	702,792,126	97,325,002
9 - 10.9	9,338,511	336,808,337	854,386,769	193,382,848
11 - 12.9	0	0	0	0
13 - 14.9	0	0	0	0
15 - 16.9	0	46,636,077	25,705,638	0

Radon and social class V overlay for districts

Number of polygons				
% in class V	<1%	1 - 3%	3 - 10%	>10%
2 - 2.4	2	20	45	11
2.5 - 2.9	2	9	53	9
3 - 3.4	0	0	14	8
3.5 - 3.9	0	7	6	0
Sum area				
	<1%	1 - 3%	3 - 10%	>10%
2 - 2.4	9,338,511	336,808,337	687,877,431	116,546,976
2.5 - 2.9	90,889	84,482,322	702,792,126	97,325,002
3 - 3.4	0	0	166,509,338	76,835,873
3.5 - 3.9	0	46,636,077	25,705,638	0

Radon and % of households consisting only of pensioners for wards

Number of polygons				
% of households	**<1%**	**1 - 3%**	**3 - 10%**	**>10%**
<15	0	6	18	8
15 - 19.9	2	16	45	10
20 - 24.9	2	46	209	34
25 - 29.9	0	11	82	35
>=30	0	6	26	20
Sum area				
	<1%	**1 - 3%**	**3 - 10%**	**>10%**
<15	0	16,974,906	58,002,329	6,432,500
15 - 19.9	9,338,511	155,182,482	177,280,962	8,703,432
20 - 24.9	90,889	229,319,183	978,799,852	92,268,099
25 - 29.9	0	4,608,738	311,133,268	125,212,716
>=30	0	19,641,422	57,668,128	58,091,101

References

References

ALLEN, RE (1990). *The Concise Oxford Dictionary of Current English.* 8th ed. Oxford University Press, Oxford.

ARNHEIM, R (1969). *Visual Thinking.* University of California Press, Berkeley.

AUSLIG (1994). Spot News. *Acres Update* February. 2-4.

BALCHIN, WGV (1984). Data Sources for Land Use Survey. *Land Use Policy* **1**(1), 4-13.

BALCHIN, WGV (1985). The Development of Land Use Maps. *Land Use Policy* **2**(1), 3-15.

BALL, K (1993a). *Radon - Dangers and Benefits of a Radioactive Gas.* British Geological Survey, Nottingham.

BALL, K (1993b). Radioactive Gas. *Earthworks* **4**.

BALL, TK, CAMERON, DG, COLMAN, TB, ROBERTS, PD (1991). Behaviour of Radon in the Geological Environment: A Review. *Quarterly Journal of Engineering Geology* **24**, 169-182.

BALL, TK, MILES, JCH (1993). Geological and Geochemical Factors Affecting the Radon Concentration in Homes in Cornwall and Devon, UK. *Environmental Geochemistry and Health* **15**(1), 27-36.

BARKE, M, HENDERSON, P, MORPHET, C (1993). Mapping the Census. *Geographical Magazine* **65**(7), 36-39.

BARR, CJ (1990). Mapping the Changing Face of Britain. *Geographical Magazine* **62**(9), 44-47.

BARR, R (1993). Census Geography: A Review. In: *The 1991 Census User's Guide.* (Eds: Dale, A, Marsh, C) HMSO, London, 70-76.

BERG, LD (1993). Between Modernism and Postmodernism. *Progress in Human Geography* **17**(4), 490-507.

BIBBY, JS, MACKNEY, D (1969). Land Use Capability Classification. *Soil Survey Technical Monograph* No 1.

BLAKEMORE, M, RYBACZUK, K (1987). Digital Mapping. In: *World Mapping Today*. (Eds: Parry, RB, Perkins, CR) Butterworths, London, 46-52.

BLAKEMORE, M, RYBACZUK, K (1993). Information Supply, Availability, Costs, Ethics and Freedom. Challenges for 'Open' Europe. *Mapping Awareness and GIS in Europe* **7**(1), 20-23.

BLAKEMORE, M, SINGH, G (1992). *Cost-Recovery Charging for Government Information. A False Economy?* Gurmukh Singh and Associates Ltd., London.

BLYTHE, A (1992). Radon - The Local Authority Role. *Radiation Protection Dosimetry* **42**(3), 203-210.

BOARD, C (1968). Land Use Surveys: Principles and Practice. In: *A Memorial Volume to Sir Dudley Stamp*: IBG Special Publication No. 1 (Eds: Embleton, C, Coppock, JT). Institute of British Geographers, London, 29-41.

BRACKEN, I, HIGGS, G, MARTIN, D, WEBSTER, C (1989). *A Classification of Geographical Information Systems Literature and Applications.* CATMOG 52. Geo Books, Norwich.

BROOKINS, DG (1990). *The Indoor Radon Problem*. Colombia University Press, New York.

BROWN, A (Ed.) (1992). *The UK Environment*. HMSO, London.

BUDD, JTC (1989). National Countryside Monitoring Scheme. In: *Rural Information for Forward Planning. ITE Symposium* No. 21 (Eds: Bunce, RGH, Barr, CJ). Institute of Terrestrial Ecology, Grange Over Sands.

BUNCE, RGH, BARR, CJ, FULLER, RM (1992). Integration of Methods for Detecting Land Use Change, with Special Reference to Countryside Survey 1990. In: *Land Use Change: The Causes and Consequences*. 1st ed. (Ed: Whitby, MC). Institute of Terrestrial Ecology & HMSO, London, 69-78.

BURRILL, A, TERRES, J (1994) Agricultural Information Systems for Europe. In: *GIS Europe* **3**(6), 22-25.

BURROUGH, PA (1986). *Principles of Geographical Information Systems for Land Resources Assessment.* Clarendon Press, Oxford.

BURROUGH, PA (1990). Development of Intelligent Geographical Information Systems. *International Journal of Geographical Information Systems* **6**(1), 1-11.

BUTTENFIELD, BP, MCMASTER, RB (1991). *Map Generalisation: Making Rules for Knowledge Representation.* Longman Scientific & Technical, Harlow.

CADOUX-HUDSON, J, HEYWOOD, DI (Eds.) (1991). *Geographical Information: AGI Yearbook.* Taylor & Francis and Miles & Arnold, London.

CARVER, S (1991). Error Modelling in GIS: Who Cares? In: *Geographic Information: AGI Yearbook* (Eds: Cadoux-Hudson, J, Heywood, DI). Taylor & Francis and Miles & Arnold, London, 228-234.

CHRISMAN, NR (1983). The Role of Quality Information in the Long-term Functionality of a Geographical Information System. In: *Automated Cartography: International Achievements and Challenges* (Ed: Weller, BS). Proceedings of the 6th International Symposium on Automated Cartography, 303-312.

CHRISMAN, NR (1989). Modelling Error in Overlaid Categorical Maps. In: *Accuracy of Spatial Databases* (Eds: Goodchild, M, Gopal, S). Taylor & Francis, London, 21-34.

CHRISMAN, NR (1990). The Accuracy of Map Overlays: A Reassessment. In: *Introductory Readings in GIS* (Eds: Peuquet, DJ, Marble, DF). Taylor and Francis, London, 309-320.

CLARK, A (1992). 1991 Census: Data collection. *Population Trends* **70**, 22-27.

CLARK, G (1982). *The Agricultural Census - United Kingdom and United States.* Geo Books, Norwich.

CLIFF, A, HAGGETT, P (1989). Plotting Disease. *Geographical Magazine* **61**(6), 26-29.

CLIFF, AD, HAGGETT, P (1988). *Atlas of Disease Distributions.* Blackwell, Oxford.

CLIFF, AD, ORD, JK (1973). *Spatial Autocorrelation*. Pion, London.

COPPOCK, JT (1964). *An Agricultural Atlas of England and Wales*. Faber & Faber, London.

COPPOCK, JT (1978). Land Use. In: *Reviews of United Kingdom Statistical Sources* (Ed: Maunder, WF). Pergamon Press, Oxford.

DAHLBERG, RE (1986). Combining Data From Different Sources. *Surveying and Mapping* **46**(2), 141-149.

DALE, A (1993). Fieldwork and Data Processing. In: *The 1991 Census User's Guide* (Eds: Dale, A, Marsh, C). HMSO, London, 84-110.

DALE, A (1993). The Content of the 1991 Census: Change and Continuity. In: *The 1991 Census User's Guide* (Eds: Dale, A, Marsh, C). HMSO, London, 16-51.

DALE, A, MARSH, C (Eds.) (1993). *The 1991 Census User's Guide*. HMSO, London.

DALE, PF (1986). Problems and Research Needs in the Presentation of Land Information. *Surveying and Mapping* **46**(2), 151-155.

DAVIDSON, DA (1992). *The Evaluation of Land Resources*. 2nd ed. Longman Scientific & Technical, Harlow.

DEANS, T (1961). A Galena-Wulfenite-Uraniferous-Asphalite Horizon in the Magnesian Limestone of Nottinghamshire. *Journal of the Mineralogical Society* **32**, 705-715.

DE MAN, EWH (1990). Establishing a Geographical Information System in Relation to its Use: A Process of Strategic Choices. In: *Introductory Readings in GIS* (Eds: Peuquet, DJ, Marble, DF). Taylor & Francis, London, 325-340.

DENHAM, C (1993). Census Geography: An Overview. In: *The 1991 Census User's Guide* (Eds: Dale, A, Marsh, C). HMSO, London, 52-69.

DEPARTMENT of ENVIRONMENT (Ed.) (1987). *Handling Geographic Information (The Chorley Report)*. HMSO, London.

DEWDNEY, JC (1985). *The UK Census of Population 1981*. CATMOG 43. Geo Books, Norwich.

DIXON, DW, GREGORY, T (1992). Reduction of Radon Exposure in Buildings. *Radiological Protection Bulletin* **131**, 15-19.

DOLL, R (1992). Risks from Radon. *Radiation Protection Dosimetry* **42**(3), 149-155.

DOUGLAS, DH (1992). The Proper Distribution of Tasks to Digitise a Clean Boundary Network. *Cartographica* **29**(2), 37-45.

DRAGANIC, IG, DRAGANIC, ZD, ADLOFF, JP (1991). *Radiation and Radioactivity: On Earth and Beyond*. CRC Press Inc., New York.

DUVAL, JS, OTTON, JK (1990). Radium Distribution and Indoor Radon in the Pacific Northwest. *Geophysical Research Letters* **17**, 801-804.

EMBLETON, C, COPPOCK, JT (Eds.) (1968). *A Memorial Volume to Sir Dudley Stamp*: IBG Special Publication No. 1. Institute of British Geographers, London.

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE INC. (1987). *PC Arc/Info Starter Kit*. Environmental Systems Research Institute Inc., Redlands (USA).

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE INC. (1993). ESRI Next Generation GIS Software. *Mapping Awareness and GIS in Europe* **7**(9), 33-36.

EVANS, IS (1980). *British Census Data and Their Use in Quantitative Geography*. Working Paper 17, University of Durham Census Research Unit.

EVANS, IS (1981). Census Data Handling. In: *Quantitative Geography: a British View*. (Eds: Wrigley, N, Bennett, RJ) Routledge & Kegan Paul, London, 46-59.

FISHER, PF, WILKINSON, GG (1987). Recent Developments and Future Trends in Geo-Information Systems. *The Cartographic Journal* **24**, 64-70.

FITZPATRICK-LINS, K, JOHNSON, TL, OTTON, JK (1990). Radon Potential Defined by Exploratory Data Analysis and GIS. *US Geological Survey Bulletin* **1908**, E1-E10.

FLOWERDEW, R, OPENSHAW, S (1987). *A Review of the Problems of Transferring Data From one Set of Areal Units to Another Incompatible Set*. North East Regional Research Laboratory - Report 87/0.

FOSTER, P, SHAND, PJ (Eds.) (1990). *AGI Yearbook*. Taylor & Francis and Miles & Arnold, London.

FOTHERINGHAM, AS, KNUDSEN, DC (1987). *Goodness-of-Fit Statistics*. CATMOG 48. Geo Books, Norwich.

FOX, DJ (1991). Mapping from Satellite Remote Sensing - Past, Present and Future. *The Cartographic Journal* **28**, 55-58.

FREEMAN, D, GREEN, D, HASSELL, D, PATERSON, K (1993). Getting Started With GIS. *Teaching Geography* **18**(2), 57-60.

FRESHNEY, EC, TAYLOR, RT (1985). Southwest England. In: *Geological Structure of Great Britain, Ireland and Surrounding Seas* (Poster; Ed: Dunning, FW). Geological Society of London, London.

FULLER, R, GROOM, G (1993). The Land Cover Map of Great Britain. *GIS Europe* **2**(9), 25-28.

FULLER, RM, SHEAIL, J, BARR, CJ (1994). The Land of Great Britain, 1930-1990: A Comparative Study of Field Mapping and Remote Sensing Techniques. *The Geographical Journal* **160** (2), 173-184.

GARNSWORTHY (1990). Tradeable Information Initiative. In: *AGI Yearbook 1990* (Eds: Foster, P, Shand, PJ). London, 106-108.

GATRELL, AC, DUNN, CE, BOYLE, PJ (1991). The Relative Utility of the Central Postcode Directory and Pinpoint Address Code in Applications of GIS. *Environment and Planning A* **23**, 1447-1458.

GEODATA INSTITUTE (1991). *GIS Survival Guide*. Southampton University.

GOLDSMITH, E, HILDYARD, N (Eds.) (1990). *The Earth Report 2*. Mitchell Beazley, London.

GOODCHILD, MF (1986). *Spatial Autocorrelation.* CATMOG 47. Geo Books, Norwich.

GOODCHILD, MF, GOPAL, S (Eds.) (1989). *The Accuracy of Spatial Databases*. Taylor & Francis, London.

GREEN, BMR (1991). *Radon Hazards and Remedies*. National Radiological Protection Board, Chilton.

GREEN, BMR, LOMAS, PR, BRADLEY, EJ, WRIXON, AD (1989). *Gamma-Radiation Levels Outdoors in Great Britain*. NRPB - R191. National Radiological Protection Board, Chilton.

GREEN, BMR, LOMAS, PR, LUYKX, F (1991). Natural Radiation Atlas of Europe. *Radiation Protection Dosimetry* **36**(2-4), 85-89.

GREEN, BMR, LOMAS, PR, O'RIORDAN, MC (1991). *Radon in Dwellings in Scotland*. NRPB - M324. National Radiological Protection Board, Chilton.

GREEN, BMR, LOMAS, PR, O'RIORDAN, MC (1992). *Radon in Dwellings in England*. NRPB - R254. National Radiological Protection Board, Chilton.

GREEN, P (1992). Radon - Whose responsibility? *Radiation Protection Dosimetry* **42**(3), 251-256.

GUNDERSEN, LCS (1992). Hidden Hazards of Radon - Scanning the Country for Problem Locations. *Earth* **1**(6), 54-61.

HAGGETT, P , CLIFF, AD, FREY, A (1977). *Locational Analysis in Human Geography.* 2nd ed. Edward Arnold, London.

HAINING, R (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. 2nd ed. Cambridge University Press, Cambridge.

HAKIM, C (1982). *Secondary Analysis in Social Research.* Allen & Unwin, London.

HEALEY, M (Ed.) (1991). *Economic Activity and Land Use: The Changing Information Base for Local and Regional Studies*. Longman Scientific & Technical, Harlow.

* For HAMPEL *et al.* (1986) reference see pg. 230

HEYWOOD, I, CORNELIUS, S (1988). *Managing Environmental Radioactivity Monitoring Data: A GIS Approach.* North East Regional Report - Report 88/7.

HEYWOOD, I, OPENSHAW, S (1989). Radioactive Britain. *Geographical Magazine* **61**(6), 30-32.

HILTON, K, MCMORROW, J, STEWART, NJ (Eds.) (1988). *Spaceview UK*. The Remote Sensing Society, Nottingham.

HILTON, N (1968). An Approach to Agricultural Land Classifications in Great Britain. In: *A Memorial Volume to Sir Dudley Stamp*: IBG Special Publication No. 1 (Eds: Embleton, C, Coppock, JT). Institute of British Geographers, London, 127-142.

HOBSON, J (1992). The Government Radon Programme. *Radiation Protection Dosimetry* **42**(3), 165-168.

HORNSBY, M (1993). £3 Million Survey Reveals Decline of this Green and Pleasant Land. *The Times* (18 November 1993), 5.

HUGHES, JS, O'RIORDAN, MC (1993). *Radiation Exposure of the UK Population - 1993 Review*. NRPB - R263. National Radiological Protection Board, Chilton.

HUNTING SURVEYS AND CONSULTANTS LTD. (Ed.) (1986). *Monitoring Landscape Change*. Hunting Surveys and Consultants Ltd., Borehamwood.

IRVINE, J, MILES, I, EVANS, J (Eds.) (1979). *Demystifying Social Statistics*. Pluto Press Ltd., London.

JOHNSON, RJ, GREGORY, D, SMITH, DM (Eds.) (1994). *The Dictionary of Human Geography*. 3rd ed. Blackwell, Oxford.

JOHNSTON, RJ (Ed.) (1985). *The Future of Geography*. Methuen, London.

JOHNSTON, RJ (1987). *Geography and Geographers: Anglo-American Human Geography Since 1945*. 2nd ed. Edward Arnold, London.

JONES, RC (1981). *Reconnaissance Geochemical Maps of Parts of south Devon and Cornwall*. Mineral Reconnaissance Programme Report No. 44. British Geological Survey, London.

KAY, H, MAY, S (1992). *Pictures and Pixels: A Teaching Pack of Satellite Images and Aerial Photographs*. National Remote Sensing Centre Ltd, Farnborough.

KITMITTO, K (1994). CHEST the ticket for academic software: provision of spatial sciences software and datasets to higher education in Britain. *Mapping Awareness* **8**(3), 14-15.

KLECKNER, RL (1981). A National Program of Land Use and Land Cover Mapping and Data Compilation. In: *Planning Future Land Use* **42** (Eds: Peterson, G, Beatty, M). ASA Special Publication, 7-13.

LANGFORD, M, UNWIN, DJ (1994). Generating and Mapping Population Density Surfaces Within a Geographical Information System. *The Cartographic Journal* **31**, 21-26.

LANGRAN, G (1989). A Review of Temporal Database Research and its Use in GIS Applications. *International Journal of GIS* **3**, 215-232.

LAURINI, R, THOMPSON, D (1992). *Fundamentals of Spatial Information Systems*. A.P.I.C. Series No. 37. Academic Press Ltd., London.

LEE, TR (1992). The Publics Perception of Radon. *Radiation Protection Dosimetry* **42**(3), 257-262.

LIEVESLEY, D (1994). Regional and Local Statistics: Statistics User's Council Annual Conference 16 November 1993. *Mapping Awareness* **8**(3), 20-21.

LIEVESLEY, D, MASSER, I (1993). An Overview of Geographic Information in Europe: Part One. *Mapping Awareness and GIS in Europe* **7**(10), 9-12.

LIEVESLEY, D, MASSER, I (1994). An Overview of Geographic Information in Europe: Part Two. *Mapping Awareness and GIS in Europe* **8**(1), 33-35.

LO, CP, SHIPMAN, RL (1990). A GIS Approach to Land-Use Change Dynamics Detection. *Photogrammetric Engineering and Remote Sensing* **56**, 1483-1491.

MCRAE, SG, BURNHAM, CP (1981). *Land Evaluation*. Clarendon Press, Oxford.

MACDOUGALL, EB (1975). The Accuracy of Map Overlays. *Landscape Planning* **2**, 23-30.

MACINTYRE, S, MACIVER, S, SOOMAN, A (1993). Area, Class and Health: Should We Be Focusing on Places or People? *Journal of Social Policy* **22**(2), 213-234.

MAFFINI, G, ARNO, M, BITTERLICH, W (1989). Observations and Comments on the Generation and Treatment of Error in Digital GIS Data. In: *Accuracy of Spatial Databases* (Eds: Goodchild, M, Gopal, S). Taylor & Francis, London, 55-67.

MAGUIRE, DJ (1989). *Computers in Geography*. Longman, Essex.

MAGUIRE, DJ, GOODCHILD, MF, RHIND, DW (Eds.) (1991). *GIS: Principles and Applications*. Longman Scientific & Technical, Harlow.

MARSH, C (1993). General Issues. In: *The 1991 Census User's Guide* (Eds: Dale, A, Marsh, C). HMSO, London, 155-164.

MARSH, C (1993). Privacy, Confidentiality and Anonymity in the 1991 Census. In: *The 1991 Census User's Guide* (Eds: Dale, A, Marsh, C). HMSO, London, 111-128.

MARSH, C, ARBER, S, WRIGLEY, N, RHIND, DW, BULMER, M (1988). The View of Academic Social Scientists on the 1991 Census of Population: A report of the Economic and Social Research Council Working Group. *Environment and Planning A* **20**, 851-889.

MARTIN, D (1989). Mapping Population Data from Zone Centroid Locations. *Transactions of the Institute of British Geographers* **14**, 90-97.

MARTIN, D (1991). *GIS and their Socio-Economic Applications*. Routledge, London.

MARTIN, D (1992). Postcodes and the 1991 Census: Issues, Problems and Prospects. *Transactions of the Institute of British Geographers* **17**, 350-357.

MARTIN, D (1993). *The UK Census of Population 1991*. CATMOG 56. Enviornmental Publications, Norwich.

MARTIN, D, BRACKEN, I (1993). The Integration of Socio-Economic and Physical Resource Data for Applied Land Management Information Systems. *Applied Geography* **13**, 45-53.

MASSER, I, BLAKEMORE, M (Eds.) (1991). *Handling Geographic Information: Methodology and Potential Applications*. Longman, Harlow.

MATHER, PM (1992). Remote Sensing and the Detection of Change. In: *Land Use Change: The Causes and Consequences*. 1st ed. (Ed: Whitby, MC). Institute of Terrestrial Ecology & HMSO, London, 60-67.

MAUNDER, WF (Ed.) (1978). *Reviews of United Kingdom Statistical Sources*. Pergamon Press, Oxford.

* For McLaren & Tulip (1991) reference see pg. 230

MEDYCKYJ-SCOTT, D (1991). "What have I done now?". In: *Geographic Information: AGI Yearbook* (Eds: Cadoux-Hudson, J, Heywood, DI). Taylor & Francis and Miles & Arnold, London, 235-241.

METTERS, J (1992). Setting Radon in Context. *Radiation Protection Dosimetry* **42**(3), 159-164.

MEYER, G, VON MEYER, N (1993). Data Conversion: It Isn't Just For Breakfast Anymore. *ACSM Bulletin* **142**, 37.

MILES, JCH (1991). Indoor Radon and Lung Cancer: Reality or Myth? *Radiological Protection Bulletin* **120**, 25-28.

MILES, JCH, CLIFF, KD, GREEN, BMR, MUIRHEAD, CR (1992a). Natural Radiation Environment. *Radiological Protection Bulletin* **130**, 17-19.

MILES, JCH, GREEN, BMR, LOMAS, PR, (1992b). *Radon Affected Areas: Derbyshire, Northamptonshire and Somerset.* Reprinted from Documents of the NRPB, Vol. 3, No. 4, Chilton.

MILES, JCH, GREEN, BMR, LOMAS, PR, CLIFF, KD (1990). *Radon Affected Areas: Cornwall and Devon.* Reprinted from Documents of the NRPB, Vol. 1, No. 4, Chilton.

MINISTRY OF AGRICULTURE, FISHERIES AND FOOD (1976). *Agricultural Land Classification of England and Wales: the Definition and Identification of Sub-grades within Grade 3.* Technical Report No. 11/1. MAFF, London.

MINISTRY OF AGRICULTURE, FISHERIES AND FOOD (1988). *Agricultural Land Classification of England and Wales.* MAFF, London.

MINISTRY OF AGRICULTURE, FISHERIES AND FOOD, SCOTTISH OFFICE AGRICULTURE AND FISH DEPARTMENT, DEPARTMENT OF AGRICULTURE FOR NORTHERN IRELAND, WELSH OFFICE AGRICULTURE DEPARTMENT (1991). *The Digest of Agricultural Census Statistics - United Kingdom June 1991.* HMSO, London.

MOHAN, JF (1983). *Linking and Analysing Census and Medical Data*. Working Paper **21**, University of Durham Census Research Unit.

MONKHOUSE, FJ, WILKINSON, HR (1963). *Maps and Diagrams.* Methuen, London.

MORAWSKA, L, PHILLIPS, CR (1993). Dependence of the Radon Emanation Coefficient on Radium Distribution and Internal Structure of the Material. *Geochimica et Cosmochimica Acta* **57**, 1783-1797.

MULLER, JC (Ed.) (1991). *Advances in Cartography*. 1st ed. International Cartographic Association & Elsevier Applied Science, London.

NATIONAL RADIOLOGICAL PROTECTION BOARD (1993). *Radon Update Leaflet*. NRPB, Chilton.

O'CARROLL, P, RODRIGUEZ-BACHILLER, A, GLASSON, J (1994). *Directory of Digital Data Sources in the United Kingdom.* Working Paper No. 149. Oxford Brookes University School of Planning, Oxford.

OLSEN, RW (1990). Digital Spatial Data Technology and Applications. *US Geological Survey Bulletin* **1908**, D1-D10.

OPCS (1990). *1981 Census - User Guide 50: Small Area Statistics.* HMSO, London.

OPCS (1992). *1991 Census - National Monitor for Great Britain.* CEN 91 CM 56. HMSO, London.

OPCS (1993a). *1991 Census - Ward and Civil Parish Monitor for Cornwall and Isles of Scilly.* WCP8. HMSO, London.

OPCS (1993b). *1991 Census - Sex, Age and Marital Status: Great Britain.* Topic Monitor CEN 91 TM SAM. HMSO, London.

OPCS (1993c). *1991 Census - Ward and Civil Parish Monitor for Devon.* WCP11. HMSO, London.

OPCS (1993d). *1991 Census - Ward and Civil Parish Monitor for Somerset.* WCP36. HMSO, London.

OPCS (1993e). *1991 Census - Ward and Civil Parish Monitor for Derbyshire.* WCP10. HMSO, London.

OPCS (1993f). *1991 Census - Ward and Civil Parish Monitor for Northamptonshire.* WCP30. HMSO, London.

OPCS (1993g). *1991 Census - County Report: Northamptonshire - Part 2.* HMSO, London.

OPCS (1994). *Mid-1993 Population Estimates for England and Wales.* Monitor PP1 94/2. HMSO, London.

OPCS, GRO (S) (1991). *1991 Census - User Guide 3: Local Statistics and Small Area Statistics Prospectus.* HMSO, London.

OPENSHAW, S (1986). *The Modifiable Areal Unit Problem.* CATMOG 46. Geo Books, Norwich.

OPENSHAW, S (1989). Learning to Live with Errors in Spatial Databases. In: *Accuracy of Spatial Databases* (Eds: Goodchild, M, Gopal, S). Taylor & Francis, London, 263.

OPENSHAW, S, TAYLOR, PJ (1981). The MAUP. In: *Quantitative Geography: a British View* (Eds: Wrigley, N, Bennett, RJ). Routledge & Kegan Paul, London, 60-69.

OPENSHAW, S, TAYLOR, PJ (1982). A Million or So Correlation Coefficients: Three Experiments of the MAUP. In: *Statistical Applications in the Spatial Sciences* (Ed: Wrigley, N). Pion, London.

OWEN, L (1993). Radon - A New Environmental Health Hazard. *Geography* **78**(2), 194-198.

PARRY, ML, HOSSELL, JE, WRIGHT, LJ (1992). Land Use in the United Kingdom. In: *Land Use Change: The Causes and Consequences.* 1st ed. (Ed: Whitby, MC). Institute of Terrestrial Ecology & HMSO, London, 7-14.

PARRY, RB, PERKINS, CR (Eds.) (1987). *World Mapping Today*. Butterworths, London.

PATTIE, N (1993). Farmwatch UK. *GIS Europe* **2**(10), 22-24

PERKINS, CR, PARRY, RB (Eds.) (1990). *Information Sources in Cartography*. Bowker-Sauer, London.

PETERSON, G, BEATTY, M (1981). *Planning Future Land Uses*. ASA Special Publication 42, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America: Madison.

PEUQUET, DJ, MARBLE, DF (Eds.) (1990). *Introductory Readings in GIS*. Taylor & Francis, London.

PLANT, J, OSTLE, D, MILLER, JM (1983). *Natural Levels of Uranium in Britain - Economic and Environmental Significance*. IGS Report 83/1 32, HMSO, London.

PLANT, JA (1978). Regional Geochemical Maps of Uranium in Northern Scotland: Environmental and Economic Considerations. In: *Geological Aspects of Uranium in the Environment: Miscellaneous Papers* **7**. Geological Society of London, London, 21-36.

PRITCHARD, T (1994). United Kingdom GIS Market Set for Significant Growth. *Mapping Awareness* **8**(6), 18-19.

RALPHS, M (1993). Data Quality Issues and GIS: A Discussion. *Mapping Awareness and GIS in Europe* **7**(7), 39-41.

* For RAPER *et al.* (1992) reference see pg. 230

REES, P (1992). The ESRC/UFC-ISC 1991 Census of Population Initiative: Delivering the Data of the Decade.

REES, P (1992). A Guide to Using the 1991 Census of Population Small Area Statistics and Local Base Statistics.

REIMER, GM (1990). The Occurrence and Transport of Radon in the Natural Environment. *Geophysical Research Letters* **17**, 799.

RHIND, DW (1981). Geographical Information Systems in Britain. In: *Quantitative Geography: A British View*. 1st ed. (Eds: Wrigley, N, Bennett, RJ). Routledge & Kegan Paul, London, 17-35.

RHIND, DW (1991). Counting the people: The role of GIS. In: *Geographical Information Systems* **2** (Eds: Maguire, DJ, Goodchild, MF, Rhind, DW). Longman, London.

RHIND, DW (1992). Data Access, Charging and Copyright and their Implications for GIS. *International Journal of Geographical Information Systems* **6**(1), 13-30.

RHIND, DW (1993). Maps, Information and Geography: A New Relationship? *Geography* **78**(2), 150-159.

RHIND, DW, HUDSON, R (1980). *Land-Use*. Methuen, New York.

ROSE, AW, HAWKES, HE, WEBB, JS (1979). *Geochemistry in Mineral Exploration*. 2nd ed. Academic Press, London.

ROWLEY, J (1993). Standards: Progress (?) and Outlook. *Mapping Awareness and GIS in Europe* **7**(6), 13-15.

SALINI, G (1992). The 1992 Sievert Lecture - Ethical Issues in Radiation Protection. *Health Physics* **63**(2), 139-148.

SAYER, A (1984). *Method in Social Science: A Realist Approach*. Hutchinson, London.

SCHROEDER, GL, KRANER, HW, EVANS, RB (1965). Diffusion of Radon in Several Naturally Occurring Soil Types. *Journal of Geophysical Research* **70**, 471-472.

SINCLAIR, G (1992). *The Lost Land: Land Use Change in England 1945 - 1990*. Council for the Protection of Rural England, London.

SLONECKER, T, GROSKINSKY, B (1993). GPS Applications in the U.S. Environmental Protection Agency. *Surveying and Land Information Systems* **53**(2), 117-124.

SMITH, W (1992). Keynote Address : AGI 1991 Conference, Birmingham. *Mapping Awareness and GIS in Europe* **6**(1), 9-14.

SMITH, W (1993). Taking stock of GIS. *Mapping Awareness and GIS in Europe* **7**(6), 9-12.

STAMP, LD (1932). The Land Utilisation of Britain. *Nature* **129**, 709-711.

STEINITZ, C (1993). The changing face of GIS from 1965 to 1993. *GIS Europe* **2**(7), 38-40.

STRAIN, P, ENGLE, F (1994). *Portrat Unseres Planeten: Satellitenbild-Atlas*. Westermann, Braunshweig.

SUTHERLAND, DS (1994). Radon Workshop - Geology, Environment, Techniques. *Geoscientist* **4**(2), 27-30.

TAMMEMAGI, HY, SMITH, NL (1975). A Radiogeologic Study of the Granites of SW England. *Journal of the Geological Society of London* **131**, 145.

* For TAYLOR (1991) reference see pg. 230

TOBLER, WR (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* **46** (suppl.), 234-240.

TOWNSHEND, JRG (Ed.) (1981). Image Analysis and Interpretation for Land Resources Survey. In: *Terrain Analysis and Remote Sensing* (Ed: Townshend, JRG). Allen & Unwin, London, 56-108.

UNEP (1992). The GRID meta-database. *GRID News* **5**(4), 1.

UNWIN, DJ (1981). *Introductory Spatial Analysis*. Methuen, London.

UNITED STATES GEOLOGICAL SURVEY (1992). *The Geology of Radon*. Public Information Factsheet. US Department of Interior & US Geological Survey, Washington.

VEREGIN, H (1989). Error Modelling for the Map Overlay Operation. In: *Accuracy of Spatial Databases* (Eds: Goodchild, M, Gopal, S). Taylor & Francis, London, 3-18.

VISVALINGAM, M (1988). Issues Relating to Basic Spatial Units. *Mapping Awareness* **2**, 42-45.

WALLACE, W (1971). *The Logic of Science in Sociology*. Aldine de Gruyter, New York.

WATTS, P (1990). Safe as Houses? *Green Magazine* **1**, 63-66.

WEBB, GAM (1992). Exposure to Radon. *Radiation Protection Dosimetry* **42**(3), 191-196.

WEBSTER, R, OLIVER, MA (1990). *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, New York.

WELLER, BS (Ed.) (1983). *Automated Cartography: International Achievements and Challenges*. Proceedings of the 6th International Symposium on Automated Cartography.

WHITBY, MC (Ed.) (1992). Land Use Change: The Causes and Consequences. *Institute of Terrestrial Ecology Symposium* no. 72.

WOLFF, SP (1991). Radon and Socio-Economic Indicators. *Lancet* **337**, 1476.

WRIGLEY, N (Ed.) (1982). *Statistical Applications in the Spatial Sciences*. Pion, London.

WRIGLEY, N (1990). ESRC and the 1991 Census. *Environment and Planning A* **22**, 573-582.

WRIGLEY, N, BENNETT, RJ (Eds.) (1981). *Quantitative Geography: A British View*. Routledge & Kegan Paul, London.

WRIXON, AD, GREEN, BMR, LOMAS, PR, MILES, JCH, CLIFF, KD, FRANCIS, EA, DRISCOLL, CMH, JAMES, AC, O'RIORDAN, MC (1988). *Natural Radiation Exposure in UK Dwellings*. NRPB - R190 (abridged). National Radiological Protection Board, Chilton.

Addendum

HAMPEL, FR, RONCHETTI, EM, ROUSSEEUW, PJ, STAHEL, WA (1986). *Robust Statistics*. Wiley, New York.

MCLAREN, R, TULIP, J (1991). Spatial Data Overview: UK vs the Rest of the World. In: *Geographical Information 1991: AGI Yearbook* (Eds: Cadoux-Hudson, J, Heywood, DI). Taylor & Francis and Miles & Arnold, London, 65-68.

RAPER, JF, RHIND, DW, SHEPHERD, JW (1992). *Postcodes: The new geography*. Longman Scientific & Technical, Harlow.

TAYLOR, DRF (Ed.) (1991). *Geographic Information Systems: The microcomputer and modern cartography*. Series: Modern Cartography Volume 1. Pergamon, Oxford.