University Of Strathclyde

Department of Mathematics and Statistics

Disease Mapping and Modelling

by

Oarabile Ruth Molaodi Bsc., MSc.

A thesis presented in the fulfilment of the requirements for the degree of Doctor of Philosophy

2009

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Acknowledgements

I will like to express my thanks and gratitude to Dr. Alison Gray and Professor Chris Robertson for dedicating their time to the supervision of this thesis. Thank you for your patience and help and may God's blessings be yours in abundance. Many thanks to Professor Stephen Marshall (EEE department) for his contribution on the idea of pseudo-colour maps.

To my mum and dad (Mr and Mrs Lempadi) who believed in my academic abilities at a very young age, thank you for your encouragements they kept me going. To my husband Eric, thank you for your prayers and support, and to my children Tshepo and Rejoice this work is dedicated to you.

My thanks to my church family (Struthers Pentecostal Church), Pastor Diana Rutherford and the crew, your love and prayers held me up and you made my stay in Scotland so sweet. All at STAMS, thank you for your encouragements and support I really enjoyed studying here. Ian, Lynne and Anne I appreciate all the help you offered.

I am thankful to Health Protection Scotland for providing data, and to the University of Botswana for funding this project.

Above all, thanks and honour and glory to Jesus Christ who gives strength, understanding and wisdom and sustains all things.

Abstract

The work of this thesis is in two parts. The first part (Chapters 2 to 4) focuses on the spatial techniques that are used in disease mapping to analyse disease data with the main focus being to compare maps over time. These methods are into two broad groups, non-parametric or interpolation methods, and modelling. For interpolation methods, nonparametric kernel regression and kriging methods are discussed, and for modelling, spatial, space-time and ecological models are discussed. The second part (Chapters 5 to 7) focuses on developing descriptive methods that can be used to compare two or more maps.

The first part of the work is a novel analysis of Scottish measles susceptibility data for pre-school and primary 1 and 2 school children for the period of 1999-2005. The spatial models and space-time models are both used to fit these data and compare the maps over time at both district and postcode sector levels. The interpolation methods are used as they can be helpful in comparing maps over time. Census variables obtained from the 2001 census data are used here. This enabled fitting of an ecological model to see if any of the variables can be useful in predicting measles susceptibility.

Since maps are similar to images, image analysis methods are adapted to help in the comparison of maps in the second part of the thesis. Other methods used are map-based methods, which are ratio maps, difference maps and pseudo-colour maps; use of plots of parameters that are obtained when fitting a model, plotting the overall mean, and the parameters for the unstructured (local) and structured (global) variation. Spatial autocorrelation methods, namely Moran's I, Geary's c and their correlograms are considered. Analogues of point process methods based on distances and used to test for complete randomness of a spatial point pattern are developed for use in comparing maps. Here the methods are based on differences rather than distances. The sensitivity and power of the methods is investigated using a simulation study. Methods that seemed to perform well are then used to compare patterns in susceptibility to measles over time, and also to compare NHS24 call uptake data for different disease syndromes.

List of Abbreviations

Abbreviations used in Chapter 5, 6, 7 and 8.

1. AVND-Average Neighbour Differences.
2. HVS - Human Visual System
3. IRD - Inter Region Differences.
4. MDD - Most Dissimilar Differences.
5. MDND - Most Dissimilar Neighbour Differences.
6. MOS - Mean Opinion Score.
7. MSD - Mean Square Difference.
8. MSDI- Most Similar Differences.
9. MSE - Mean Square Error.
10. MSND - Most Similar Neighbour Differences.
11. MSSIM - Mean Structural Similarity Index.
12. NwMSE - New Weighted Mean Square Error.
13. PSNR/PSNRR/PSNRM - Peak-to-Signal-Noise-Ratio.
14. SSIM/SSIMM/SSIMR - Structural Similarity Index.
15. SVD - Singular Value Decomposition.
16. UIQI - Universal Image Quality Index.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Epidemiology

Epidemiology deals with the occurrence of disease and with disease *aetiology* (the study of causes of disease). Last (1988) defines epidemiology as 'the study of the distribution and determinants of health related states or events in specified populations, and the application of this study to control of health problems'. Epidemiology normally concerns human populations and these populations can often be defined in geographical terms. Usually the population used is of a given area or country at a given time. In epidemiological analysis, the variation of population structure between geographical areas and time periods has to be taken into account (Beaglehole *et al.*, 1993).

Epidemiologists may try to establish whether over the years, there has been an increase or decrease in the prevalence or rate of the disease in question, whether a geographical area has a higher rate of disease than another area, and whether individuals with the disease in question has different characteristics from individuals without the disease. The characteristics may be demographic, e.g. age or gender; biological, e.g. blood levels of antibodies; social and economical, e.g. occupation; personal habits, e.g. diet; and genetic, e.g. blood groups. The concepts and methods used in epidemiology are derived from other disciplines such as biology, sociology and statistics (Lilienfeld and Stolley, 1994).

1.2 Spatial Epidemiology

Spatial epidemiology is the study of disease occurrence in relation to spatial or geographical locations. This simply deals with interpretation and analysis of geographical distribution of diseases.

One of the earliest analysis of distribution of disease is that of Snow (1854) who analysed the outbreak of cholera in relation to the location of the Broad Street water pump in London. Many of the statistical methods used in analysing geographical distribution of diseases were developed about a decade ago (Lawson *et al.*, 2003). This applies to both literature and statistical software development. This recent development shows a growing concern in society about environmental health issues in relation to the health of people, mainly because the statistical methods help in detection of sources of potential health hazards. This is seen as a fundamental issue in studies of environmental epidemiology (Diggle, 1993). The development of statistical methods also helps in the allocation of health resources by health services, as the methods allow more accurate depiction of the relationship between disease incidence and explanatory variables (see Chapter 2).

There are three main areas of application of the study of spatial distribution of disease. These are:

1. Disease Mapping
2. Ecological Analysis
3. Disease Clustering.

Disease Mapping: The focus of analysis here is to estimate the true *relative risk* (see Section 1.3.2) of disease across the geographical area of study or the disease map. Since the observed data may contain noise due to random variation, the purpose is to clean the noise from disease map and reveal the underlying structure. This can be done by smoothing and modelling (see Section 1.4). The aims were outlined by Lawson *et al.* (2000) as follows: (i) to describe the spatial variation in disease incidence for the formulation of aetiological hypotheses maps, (ii) to identify areas of unusual high risk so that action may be taken, and (iii) to provide a 'clean' map of disease risk in a region to allow better resource allocation and risk assessment.

Ecological Analysis: This focuses on the analysis of the spatial distribution of disease in relation to measured explanatory variables at the aggregated spatial level. In general, ecological analysis can be defined (in the sense of ecology) as focusing on explaining the spatial distribution of disease by the inclusion of explanatory variables. A study of this kind was done by Cook and Pocock (1983), who examined the relationship of cardiovascular incidence in the United Kingdom to a variety of variables which included water hardness, climate, location, socio-economical and genetic factors. Another study is by Donnelly (1995), who examined the respiratory health of school children and volatile organic compounds in the outdoor atmosphere. The general definition can include situations where the relationship between the case address locations and pollution hazards is through explanatory variables such as distance and direction from the hazard, thus individual data are directly related to the explanatory variables.

The problem with ecological analysis is *ecological bias*, which arises when the association between the variables at group/aggegated level is interpreted or applied as the association at individual level. The issue of ecological bias and methods based on individual level data have been discussed and developed by different authors, including Jackson *et al.* (2006), Steel *et al.* (2006) and Steel and Holt (1996). However, the advantage of ecological analysis is that it uses data already available at low cost, while individual level data are often unavailable for confidentiality reasons. It has been argued that the best method is to include both group level (to increase power to study small area variations) and individual level data (reducing ecological bias) when available in the analysis (Jackson *et al.*, 2006).

Disease Clustering: This concerns the analysis of disease clusters and their locations in a disease map. It takes a variety of forms:

General/non-specific clustering: The analysis here is concerned with assessing the whole disease map to see if it is clustered or not, i.e. whether there are clusters of regions with high or low disease rates. It does not determine where the clusters are formed but whether the map is clustered.

Specific/focused clustering: The analysis attempts to find out where the clusters are located in the map, if they exist. They may be in a fixed or known location, thus the relationship between disease incidence and the location of hazards may be as-

sessed. When locations of clusters are unknown, these are estimated from the data. In this case the clustering is called *non-focused clustering*. Ecological methods can be used in the analysis of focused clustering, while in non-focused clustering methods of analysis which allow estimation of the location of clusters and their form have to be constructed.

In this thesis we deal with the first area of spatial epidemiology, i.e. disease mapping. We will also move into ecological analysis.

1.2.1 Types of Spatial Data

There are two forms of mapped data which arise in studies of spatial distribution of disease, and these dictate which methods to use in analysing data. These are *case event data* and *region count data*. These are discussed by Lawson (2001).

Case event data: A study area/spatial window is defined, say W. For a fixed period of time, within W, the disease case events which occur at locations $x_i,\ i = 1, ..., m$, are recorded. These locations are usually residential addresses, e.g. street address, zip code (USA) or postcode (UK).

Count data: As in case event data, a study area W with m arbitrarily bounded subregions is defined, and these subregions may lie wholly or in part in W. Counts or totals of disease incidence, say $n_i,\ i = 1, ..., m$, in each subregion are recorded. Usually the m subregions are arbitrarily defined administrative regions/tracts. These are small areas within the study window, such as census tracts, health authority areas, counties or electoral districts. Count data are an aggregation of all the disease cases within a tract. Due to medical confidentiality, availability of case event data can be limited, but count data are commonly available from routine data sources such as government agencies. The methods of analysis used and developed in this project will be those for count data.

1.3 Disease Mapping

Firstly, we give some definitions. A *map*, as defined by MacEachren (1995) and Monmonier (1996), is a collection of spatially defined objects. This is a two-dimensional visual representation of Cartesian or polar coordinate locations of objects and some-

times their attributes; for example, a street map showing the streets and houses on the streets. In this case the houses may have attributes relating to the population of each household.

The term *disease* in disease mapping refers to the geographical distribution of disease within a population. This distribution can be expressed either as location addresses of individuals with the disease or total number of individuals with the disease in a small region/tract. The word *mapping* refers to the visual representation of the geographical distribution of the disease (Lawson and Williams, 2001).

1.3.1 Disease Maps

A *disease map* is a visual representation of a collection of disease objects. These disease objects may be residential locations of individuals or summary statistics for a group of individuals. These can be *mortality* (death rate due to the disease in question) or *morbidity* (rate of disease *prevalence/incidence*) maps. Prevalence of disease is the number of cases of disease in a population at a given point in time, and incidence of a disease is the number of new cases arising in a population in a given period of time (Beaglehole *et al.*, 1993).

There are three types of disease maps corresponding to certain types of data. These are:

- Dot maps used for case event data. See Figure 1.1.
- Choropleth maps for count data. Colour or shade of grey is used to represent different levels of disease rates in a study area. See Figure 1.2.
- Isopleth maps/contour maps, used to display continuous surfaces of measures made on counts or case event data. See Figure 1.3.

Below are the examples of a dot, choropleth and isopleth maps. These maps are produced using R, a public domain package for statistics and graphics available at http://cran.r-project.org/. The R codes used to produce Figures 1.1, 1.2 and 1.3 are given in Appendix C.

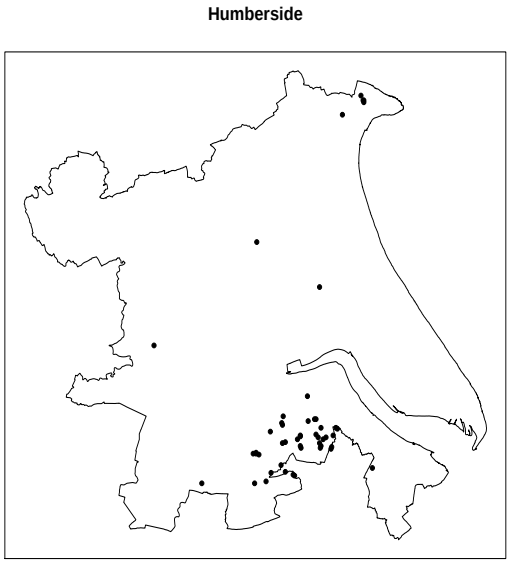

Figure 1.1: Humberside leukaemia and lymphoma case event map (1974-1986); data obtained from Lawson (2001).

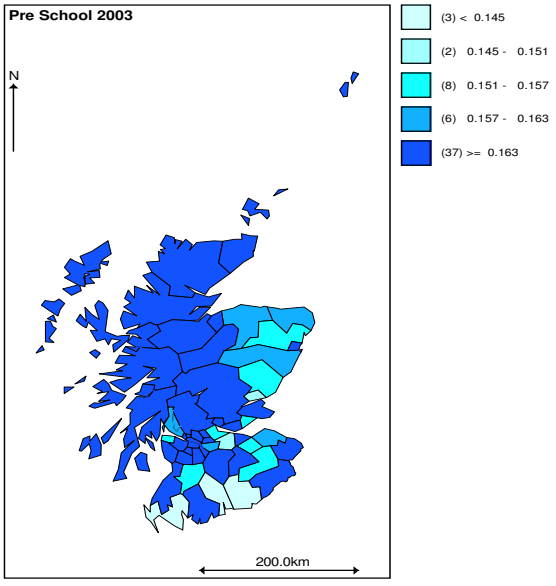

Figure 1.2: Choropleth map of susceptibility to measles in Scotland, 2003 pre-school (data from Health Protection Scotland).

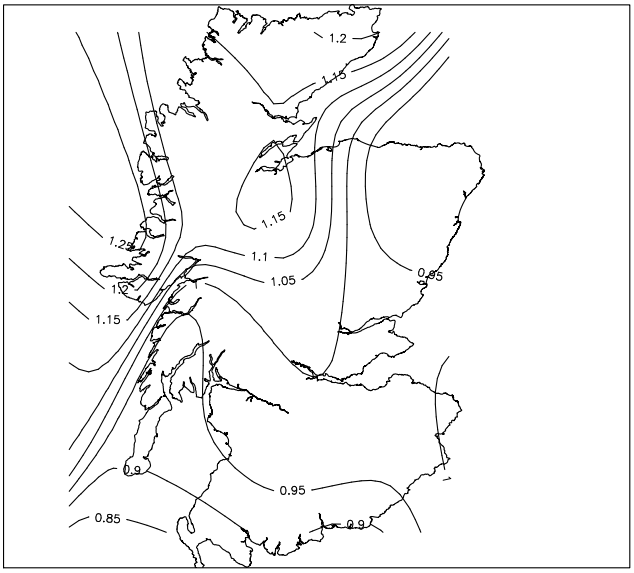

Figure 1.3: Isopleth map (using kernel smoothing; see Section 4.2) of susceptibility to measles in Scotland for 2002.

Figure 1.1 displays the case event map of childhood leukaemia and lymphoma in the north Humberside region of England for the period 1974-1986. Figure 1.2 displays measles susceptibility proportions for pre-school children in Scotland in 2003. Figure 1.3 is the isopleth map of the ratio of the observed/expected counts of measles susceptibility smoothed using kernel smoothing (see Chapter 4). The case event map (Figure 1.1) reveals clusters of leukaemia cases in the area. The choropleth map of measles susceptibility (Figure 1.2) shows spatial distribution of measles susceptibility, indicating areas of high and low susceptibility. The isopleth map of measles susceptibility (Figure 1.3) shows gradual change in measles susceptibility over space, and indicates high and low susceptibility areas.

Historical Overview of Disease Mapping

The outbreak of epidemic diseases in the eighteenth and nineteenth century influenced the use of disease maps to represent the geographical distribution of disease. It appears that the first dot maps were used in North America at the close of the eighteenth century showing the distribution of yellow fever (Howe, 1986; Stevenson, 1965). Seaman (1798) used two dot maps to show distribution of yellow fever in New

York. In 1820, Pascalis Ouviere used a dot map which was larger than that of Seaman, to show the distribution of yellow fever (Howe, 1986). The aim was to establish the factors that cause the disease. The concentration of disease was exhibited in one restricted area with specific environmental factors. This gave support to the local origin of the disease. Cartwright (1826) also used similar map focusing on Natchez in Mississippi, which suffered yellow fever in 1823. Seaman, Ouviere and Cartwright were anti-contagionists and they used the maps as a weapon of debate between them and the contagionists. Harty (1820) also used a map to mark the dates of the beginning of contagious fever from 1816-1818 in Ireland.

The outbreaks of cholera in the nineteenth century were also displayed in maps. The disease maps were used by Baker (1833) to accompany his 'Report of the Leeds Board of Health'. The areas where cholera prevailed were shaded red on the map, and based on the pattern of red areas he observed 'how exceedingly the disease prevailed in those parts of the town where there is a deficiency, often an extreme want, of sewage, drainage and paving' (Gilbert, 1958). Other maps of cholera were constructed by Ormerod (1848) about the Oxford situation, Shapter (1849) who published a book with a dot map representing the distribution of cholera deaths, and Petermann (1852) who produced a cholera map of the British Isles. He showed the districts that had the disease in 1831, 1832 and 1833. The most famous and celebrated cholera dot map is that of Snow (1854). He showed the distribution of cholera deaths in the Broad Street district of London in 1854 by placing a black dot at the victim's residence. He showed that cholera sufferers were those who drank from the Broad Street pump, and as a result the pump was closed and cholera cases ceased immediately. Several other maps of cholera were constructed after Snow's map, including that of Acland (1856). He used symbols of different shapes and colours to differentiate the cholera cases in 1832, 1849 and 1854 and to distinguish cases of cholera and those of choleraic diarrhoea. He found that altitude, drainage characteristics and the contamination of water resources influence cholera incidence. A cholera map similar to that of Snow was also constructed for cases of cholera in Hamburg and the adjoining suburb of Altona (Deneke, 1895). These also associated cholera with contaminated water supplied only to Hamburg.

Several other maps, apart from those of yellow fever and cholera were also published.

In 1839 Joseph-Francois Malgaine produced a map of hernia in France, to test aetiological theories relating to hernia. This map is considered to be the first statistical map of disease because of 'its use of transformation of raw data (the reciprocal transformation), its technique of agglomeration by departments, its use of shading to represent class intervals, and its overlaying of external environmental information of test hypotheses' (Glick, 1979). The other disease maps are those of Haviland (Glick, 1979). The maps were produced to display the distributions of heart disease, dropsy, cancer and phthisis in 1851 to 1860 in eleven registration divisions of England and Wales (Glick, 1979).

The next development of disease mapping was during the First World War. It was important to know which areas were affected by diseases. As a result two works were produced. These are 'Global Epidemiology' by Simmons *et al.* (1944-54) representing U.S. effort, and 'World Atlas for Epidemic Diseases' by Rodenwaldt and Jusatz (1961), representing German effort. Unlike the Americans, the Germans included maps (Glick, 1979). Many other atlases, other than those of world wars, were also produced. These include a disease atlas produced by May (1955), with seventeen maps that represented major infectious diseases of the world. This was a U.S. work similar to the German atlas. Howe produced atlases in 1963 with an update in 1971 (Howe (1963) and Howe (1971)). The work showed for the first time the differences in mortality experienced throughout the United Kingdom. Other atlases more detailed than that of Howe were those of Gardner *et al.* (1983) and Gardner *et al.* (1984). These were atlases of mortality for England and Wales which showed variations in distribution of diseases, and new hypotheses of causation were derived from the atlases. An atlas of cancer incidence (Kemp *et al.*, 1985) and an atlas of mortality (Lloyd *et al.*, 1987) were produced for the 56 government districts of Scotland (see Lawson and Williams, 2001).

Advantages of Using Disease Maps

'Maps provide an efficient and unique method of demonstrating distributions of phenomena in space. Though constructed primarily to show facts, to show spatial distributions with an accuracy which cannot be attained in pages of description or statistics, their prime importance is as research tools. They record observations in a succinct form; they aid analysis; they stimulate ideas and aid in the formation of working hypotheses; they make it possible to communicate findings' (Howe, 1971).

Disease maps are considered as an important contribution which a researcher can make (Copperthwaite, 1972). They display geographical distribution of disease. They can reveal unnoticed information which could not be detected from statistical tables, therefore they are an enhancement to tabular or verbal methods of communication.

Display of disease information on a map provides visual relationship of the disease and geographical location, which helps in understanding the impact of environment on health. High risk areas for a disease can be identified, which aids health authorities in allocation of health resources, by identifying which areas are in greatest need. For example, Figures 1.2 and 1.3 indicate that the north and north-west areas of Scotland are areas of the greatest need, as they have higher rates of susceptibility to measles in relation to other areas. Disease maps can display more than one factor at a time (*multiple-factor maps*). They can display the distribution of the disease and the topography of a region, therefore aetiological hypotheses can be established from the multiple-factor map. An example of a multiple-factor map is that of prevalence of Burkitt's lymphoma (Learmonth, 1972). The maps indicated that warm and relatively moist areas around Uganda and coastal regions of Kenya and Tanzania had high disease incidence in terms of tumours, while tumours were rare in the highlands of south west Uganda, central Kenya, northern Tanzania and central Tanzania. When analysing the maps it was concluded that the risk of tumour is related to the presence of malaria (Glick, 1979).

Constructing Disease Maps

Lawson (2001) and Lawson and Williams (2001) outline points to take into consideration when constructing disease maps:

Data: A decision has to be made on what kind of data are to be mapped as this affects the method of mapping. Therefore the form of data to be mapped must be chosen appropriately. It can be raw data, for example maps displaying observed counts of disease in regions; or data obtained by statistical processing. For example, producing isopleth maps requires further processing of disease data (see Chapter 4).

Geographical area: The geographical area of study for disease should be chosen or defined with great care. Unlike in other areas of applications of spatial epidemiology, the choice of geographical area in disease mapping may depend more on whether there

is available data to analyse than on the need to do the disease study. These data are commonly available at town level, which is too broad for epidemiological interest. Use of smaller areas allows identification of more localised risks but also means lower population figures and less stable estimates of rates.

Scale: The extent of aerial coverage possible and extent to which spatial structures are observed on the map are determined by the scale of the map. On constructing a disease map, the relationship between the map scale and the study/geographical area has to be established first, as the scale chosen should represent variation of the disease within the area of study. For example, consider Figure 1.2, the choropleth map of measles susceptibility in Scotland. If we were to zoom in (changing scale to increase resolution) to see more details, we would reduce the geographical area that can be mapped. Zooming in will allow smaller areas to be examined but this should be done at the same time as a change in resolution. Zooming out to a large aerial view (for example, the whole of the UK), the map will lose resolution and have less detail about the distribution of the disease.

Symbols: Inappropriate use of symbols can lead to the misrepresentation of data, therefore the use of symbols should be given consideration when used in a map. Symbols can be classified as point, line, colour and shading symbols with different sizes and shape. Point symbols are used in disease mapping, to represent individual events and a common size is used. When more than one disease is shown on the map, different sizes and shapes representing each disease should be used. Different sizes of symbols can also be used to show different measurements. For example, different symbol sizes can be plotted at the centroid of each tract in a study area to represent the weight of the number of counts at each tract, and the standardized mortality/morbidity ratio (SMR) (ratio of the observed count within a tract relative to the expected count) (see Section 1.3.2) can be represented this way. Line symbols are not commonly used in disease mapping except on contour plots to display a continuous surface (isopleth maps). A constant thickness of lines is used to show levels of constant effect/contour height. Usually contour intervals are specified, and fewer intervals leads to a smoother representation. Symbolisation can also be used to distort/exaggerate data in order to highlight important features. Monmonier (1996) gives examples of this. The use of symbols in this way should be carefully considered before it is used, as it can make

interpretation difficult.

Colour: Colours may be used in a choropleth map. There are problems associated with the use of colour as arbitrary use of colour intensities and types can produce distorted representation of disease distributions. Arbitrary scale differences between colour hues are not easy to interpret as differences in disease incidence (Lawson and Williams, 2001).

Interpretation of Maps

There are difficulties associated with the use of choropleth maps. The human eye is attracted to bright colours and colour changes, thus areas displaying bright colours may attract more attention than other areas. Change of colour can distort map appearance and influence the interpretation of the map. Use of different sizes and shapes of symbols can also distort interpretation of the map (Lawson and Williams, 2001). Snow (1854) used coffin symbols to depict cholera case addresses and these had an immediate emotive impact (Walter, 1993). The way objects are arranged on a map affects the interpretation. This is because a human eye can more quickly detect clusters (important in interpretation of disease maps) than other features (Ripley, 1981; Pickle and Herman, 1995). The choice of grey scaling can also affect interpretation, as the grouping is arbitrary (Berke, 2004). Irregular sizes and shapes of geographical areas make the interpretation difficult. Large less populated areas can visually dominate smaller areas which have a larger population and are more reliably estimated (Marshall, 1991). 'The visual impact of larger areas is higher and may dominate the map, leading to biased visual perception, whereas in human epidemiology it is the smaller, urban areas, and not the rural surroundings, that are primarily of interest due to population sizes' (Berke, 2004).

Lawson and Williams (2001) made basic recommendations to help in the interpretation of maps, and these relate to simplification of maps and the use of colours and symbols which represent the mapped data as clearly as possible. They suggested use of monochrome colour schemes, displaying a relative risk in each region, inclusion of statistical tables of the data used to produce the map, and production of an additional map of variability of data or estimates used.

To circumvent the problem of irregular area sizes, Oslon (1976) suggested re-drawing

a map so that areas are proportional to population sizes (a *cartogram*). This map does not maintain the contiguity of neighbouring areas but maintains their shapes. Schulman *et al.* (1988) proposed a cartogram algorithm maintaining contiguity of neighbouring areas but distorting their shapes. Non-contiguity cartograms can be more easily produced than the contiguity cartograms (Marshall, 1991). Dorling (2008) also argues for the use of worldmapper cartograms as they say more than the conventional map. The problems associated with choropleth maps can be avoided by smoothing estimates (see Section 1.4) and constructing isopleth maps which are easier to read than choropleth maps.

1.3.2 The Standardised Mortality/Morbidity Ratio (SMR)

In disease mapping the simplest form of displaying a disease distribution on a map is a crude representation of the observed rates, i.e. a map showing the number of counts/total in each location. This representation does not take into account the background population or the spatial distribution of population at risk of disease in question. To take into account the background population, a measure/estimate of a *relative risk* is obtained and mapped. Lawson (2001) defines *relative risk* as 'the measure of excess risk found in relation to that supported purely by the local population which is exposed or 'at risk". In disease mapping the estimate of the relative risk compares the observed with the expected disease incidence and it is referred to as the standardised mortality/morbidity ratio (SMR).

Let W be a study region with n tract/regions and O_i and E_i be number of observed counts and expected counts in the ith tract respectively, $i = 1, ..., n$. Then the standardised mortality/morbidity ratio in the ith tract is given by the following:

$$SMR_i = \frac{O_i}{E_i}\ , \quad i = 1, ..., n. \tag{1.1}$$

A value of SMR greater than 1 indicates a higher count than expected, less than 1 indicates fewer cases than expected and a value of 1 indicates that the observed and expected counts are equal. See also below.

The variance of the SMR is given by:

$$Var(SMR_i) = Var(\frac{O_i}{E_i}) = \frac{Var(O_i)}{{E_i}^2} = \frac{\theta_i E_i}{{E_i}^2} = \frac{\theta_i}{E_i}\ , \quad i = 1, ..., n, \tag{1.2}$$

if it is assumed that $O_i \sim Poisson(\theta_i E_i)$, a simplification which ignores spatial relationships but is commonly used; and where θ_i is the unknown relative risk for the ith region, estimated by SMR_i.

The expected counts can be calculated in two basic ways. Expected counts can be obtained as a product of death rates of a standard/reference population and the standard population of the study community, referred to as *indirect or external standardisation*. They can also be obtained from a product of death rates in the study community and the standard population in the study community, referred to as *direct standardisation* (Lawson and Williams, 2001). The standardised rates could be based on national, regional or study window total rates.

There are problems associated with the use of SMRs (Lawson, 2001; Lawson *et al.*, 2000). Since they are based on the ratios of observed to expected counts, any slight change in the expected values will give large changes in the estimates. A near zero expected value will result in a very large SMR for any observed count. Also the SMR does not differentiate between regions with zero observations. The variance of the SMR is proportional to $\frac{1}{E_i}$, therefore the variance in regions/tracts with a small population (and hence a small expected value) is large, and variance in tracts with a large population (hence a large expected value) is small. A variety of methods have been proposed to address the problems of SMRs and these methods will be discussed in the next Section. These methods can also be applied to spatio-temporal disease data. Looking at spatial and temporal epidemiological events helps to identify the disease trends and reduce the mistake of interpreting one-off SMRs as the level of health in a community (Lawson and Williams, 2001). Figure 1.2 is an example of a disease map based on the SMRs of measles susceptibility in Scotland.

1.4 Techniques of Disease Mapping of Count Data

Both non-parametric/informal and parametric methods have been proposed to improve relative risk estimation. The non-parametric methods discussed here use *smoothing* tools to reduce noise from the SMRs, based on interpolation methods/non-parametric methods. The parametric methods are based on *modelling* relative risk to produce estimates.

1.4.1 Smoothing

Smoothing allows interpolation of data values on a grid to neighbouring locations for which a data value is unavailable, so as to represent the disease variation continuously. There are different smoothing methods which can be used to interpolate SMRs onto a continuous surface. These methods include kernel smoothing, which was advocated by Breslow and Day (1987), and has been discussed by Bowman and Azzalini (1997). Brillinger (1990) used kernel smoothing for birth-rate data. Kernel smoothing has an advantage over other smoothing methods because it preserves the positivity of the SMRs, unlike other methods including kriging methods (Cressie, 1993). Kriging has been used by Carrat and Valleron (1992), Berke (2004), Lajaunie (1991), McNeill (1991), Oliver *et al.* (1992), Oliver *et al.* (1998), and Webster *et al.* (1994). For example, Berke (2004) suggested that negative interpolates in kriging can be avoided by choosing an appropriate method of kriging. He proposed using kriging on already smoothed estimates, using smoothing based on empirical Bayes/shrinkage estimates. Croner and De Cola (2001) used kriging to examine patterns over time in national public health data.

Other smoothing methods are discussed by Lancaster and Salkauskas (1986), Green and Silverman (1994), and Ripley (1981). Kelsall and Diggle (1998) proposed the use of generalized additive models which allow incorporation of covariates. Kelsall and Wakefield (2002) discussed approaches based on generalised linear modelling, similar to Diggle *et al.* (1998).

1.4.2 Modelling

Let O_i, E_i and n_i be the observed counts, expected counts and number of individuals at risk of disease in the ith region with relative risk θ_i respectively, $i = 1, ..., n$. A common model for the distribution of counts of a rare disease is the Poisson approximation to the Binomial distribution, i.e.

$$O_i \sim Pois(E_i\theta_i). \tag{1.3}$$

The likelihood of O_i is given as

$$L(O_i|\underline{\theta}) = \prod_{i=1}^{n} \frac{exp(-E_i\theta_i)}{O_i!}(E_i\theta_i)^{O_i}$$

$$= exp(-\sum_{i=1}^{n}(E_i\theta_i))\frac{\prod_{i=1}^{n}(E_i\theta_i)^{O_i}}{\prod_{i=1}^{n}O_i!} \propto exp(-\sum_{i=1}^{n}(E_i\theta_i))\prod_{i=1}^{n}(E_i\theta_i)^{O_i}$$

where $\underline{\theta} = (\theta_1, ..., \theta_n)$. The log-likelihood of O_i given θ_i will be given by

$$l(\underline{\theta}) = \sum_{i=1}^{n} O_i ln(E_i\theta_i) - \sum_{i=1}^{n} E_i\theta_i. \tag{1.4}$$

Differentiating the likelihood with respect to each of the θ_i, we obtain the maximum likelihood estimator of the relative risk θ_i as $\frac{O_i}{E_i}$, which is the SMR. In Section 1.3.2, we discussed the problems associated with mapping the SMRs and different models have been developed to help to deal with these problems. We review these models in Chapter 2.

1.5 Data

The data used in this thesis are based on the susceptibility to measles data obtained from a population database for all of Scotland. Measles, mumps and rubella (MMR) vaccine was introduced in the UK in 1988. This vaccine protects against measles, mumps and rubella. The first uptake (MMR1) is recommended for children at ages 13 months and the second uptake (MMR2) is recommended for the children between the ages of 3 years 4 months and 5 years (Friederichs *et al.*, 2006). In each birth cohort, for each postcode sector, the number of pre-school and primary school 1 and 2 school children susceptible to measles is provided as at 1st September each year. It is assumed that one dose of MMR vaccine has 90 % efficacy, i.e. 10 % of children who received one dose are still susceptible, and two doses gives combined efficacy of 99 %, i.e. 1 % of children who received both doses are still susceptible. However, MMR uptake has been decreasing since 1998 as a result of the association in the public mind of MMR vaccine with autism (Wakefield *et al.*, 1998) although this association has been disproved (Demicheli *et al.*, 2005; Baird *et al.*, 2008). Number of children susceptible in a postcode sector is estimated as number of children who have not received MMR vaccine+0.1 $\times$ number of children who received one dose+0.01$\times$ number of children who received a second dose.

The measles susceptibility data are available for seven different birth cohorts. We have pre-school children born between 1st March 1995-28th February 1997, 1st March 1996-28th February 1998, 1st March 1997-28th February 1999, 1st March 1998-28th February 2000, 1st March 1999-28th February 2001, 1st March 2000-28th February

2002, and 1st March 2001-28th February 2003, with susceptibility assessed as at 1st September 1999, 2000, 2001, 2002, 2003, 2004 and 2005, respectively. Primary 1 and 2 children were born between 1st March 1993-28th February 1995, 1st March 1994-28th February 1996, 1st March 1995-28th February 1997, 1st March 1996-28th February 1998, 1st March 1997-28th February 1999, 1st March 1998-28th February 2000, and 1st March 1999-28th February 2001, with susceptibility assessed as at 1st September 1999, 2000, 2001, 2002, 2003, 2004 and 2005 respectively. Five of the pre-school cohorts are also in primary 1 and 2 cohorts. These cohorts are those assessed for susceptibility at 1st September 1999, 2000, 2001, 2002 and 2003 in pre-school, corresponding to those assessed for susceptibility at 1st September 2001, 2002, 2003, 2004 and 2005 respectively in primary 1 and 2 children. All together we have 937 postcode sectors with data available.

To be able to perform an ecological analysis, we link the susceptibility to measles data to 2001 census data available at post code sector level from *http://census.ac.uk//casweb*. We select 11 census variables which are thought could be relevant to whether children in the local area get immunised or not. The data taken are in the form of census indicator variables relating to the percentage of people in households with no car, percentage of people in overcrowded households, percentage of unemployed males, and percentage of people in low social class households. The other data in the form of key statistics are the percentage of children aged 0-4 years in a postcode sector (target group for MMR vaccine), the percentage of households with dependent children, the percentage of people born in the EU but outside Britain (referred to as born other EU), the percentage of people born outside the EU (referred to as born elsewhere), the percentage of people with or without qualifications, the percentage of lone parents households and the percentage of people working in agriculture.

Analysis is also done at district level (large regions) to compare the results with those for postcode sectors (small regions). These larger districts are the 56 governmental districts of Scotland. The data for postcode sectors were combined to obtain data for 56 districts. Ethical approval was not required.

Descriptive Statistics and Maps of Measles Susceptibility

Table 1.1 and Table 1.2 show descriptive statistics for the proportions of measles susceptibility (raw data), for pre-school and primary school respectively, at district and postcode sector levels. For pre-school (Table 1.1) at district level the median/mean of proportions suggests that measles susceptibility decreased from 1999 to 2001, from 2001 to 2004 measles susceptibility increased and it decreased in 2005. The highest range is in 2003. At postcode sector level, the trend is similar to districts except that in 2004 susceptibility decreased and it increased again in 2005. The maps (Figures 1.4 and 1.5) show that generally there is increase in susceptibility over time with higher susceptibility in 2003 and 2004, and in 2005 susceptibility decreased. Regions with higher susceptibility were mostly those in the north and central areas.

For primary school 1 and 2 (Table 1.2), for both district and postcode sector levels median/mean of the proportions suggests a decrease in susceptibility from 1999 to 2003 and an increase from 2003 to 2005. At district level, the highest range is in 2005. The maps (Figures 1.6 and 1.7) show that most regions have lower susceptibility, especially in the southern regions, but susceptibility increased in 2005 mostly in the north and Highlands. In Chapters 3 and 4 this data will be smoothed to take into account variation due to small counts or population so that better interpretation can be made.

Pre-School (Districts)							
Year	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Range
1999	0.112	0.135	0.144	0.147	0.158	0.214	0.102
2000	0.115	0.133	0.142	0.147	0.156	0.228	0.113
2001	0.111	0.129	0.139	0.142	0.153	0.207	0.096
2002	0.113	0.138	0.150	0.152	0.165	0.206	0.093
2003	0.122	0.159	0.172	0.176	0.186	0.285	0.163
2004	0.130	0.161	0.176	0.182	0.187	0.284	0.118
2005	0.119	0.150	0.163	0.167	0.176	0.236	0.117
Pre-School (Postcode Sectors)							
Year	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Range
1999	0.000	0.121	0.140	0.150	0.165	1.000	1.000
2000	0.000	0.122	0.141	0.150	0.167	1.000	1.000
2001	0.000	0.120	0.137	0.146	0.162	1.000	1.000
2002	0.000	0.129	0.148	0.157	0.174	1.000	1.000
2003	0.000	0.145	0.171	0.182	0.200	1.000	1.000
2004	0.000	0.000	0.148	0.131	0.190	1.000	1.000
2005	0.000	0.140	0.161	0.170	0.188	1.000	1.000

Table 1.1: Descriptive statistics for proportions susceptibile to measles, for pre-school for districts (above) and postcode sectors (below)

Primary School (Districts)							
Year	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Range
1999	0.033	0.053	0.068	0.072	0.087	0.151	0.118
2000	0.033	0.050	0.065	0.066	0.080	0.141	0.108
2001	0.031	0.047	0.059	0.064	0.076	0.137	0.106
2002	0.031	0.045	0.059	0.064	0.075	0.150	0.119
2003	0.030	0.044	0.056	0.061	0.072	0.127	0.097
2004	0.030	0.056	0.063	0.068	0.077	0.116	0.086
2005	0.042	0.066	0.078	0.082	0.091	0.172	0.130
Primary School (Postcode Sectors)							
Year	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	Range
1999	0.000	0.033	0.056	0.068	0.091	0.500	0.500
2000	0.000	0.036	0.056	0.069	0.083	1.000	1.000
2001	0.000	0.037	0.054	0.069	0.081	1.000	1.000
2002	0.000	0.037	0.055	0.071	0.080	1.000	1.000
2003	0.000	0.036	0.052	0.066	0.076	1.000	1.000
2004	0.000	0.044	0.061	0.075	0.083	1.000	1.000
2005	0.000	0.051	0.071	0.085	0.101	1.000	1.000

Table 1.2: Descriptive statistics for proportions susceptibile to measles, for primary school for districts (above) and postcode sectors (below)

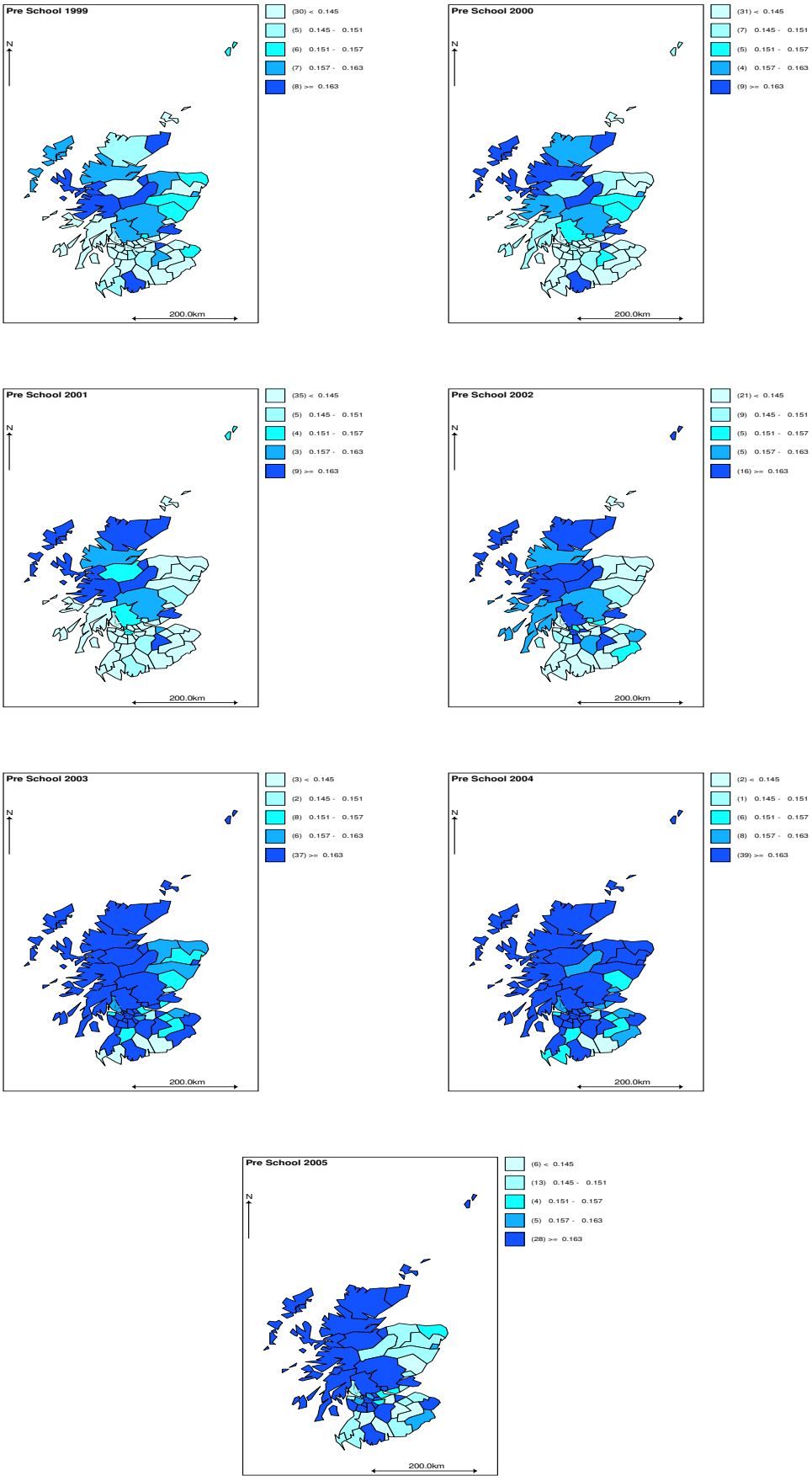

Figure 1.4: Maps of measles susceptibility proportions for districts, for 1999, 2000, 2001, 2002, 2003, 2004 and 2005 for pre-school children.

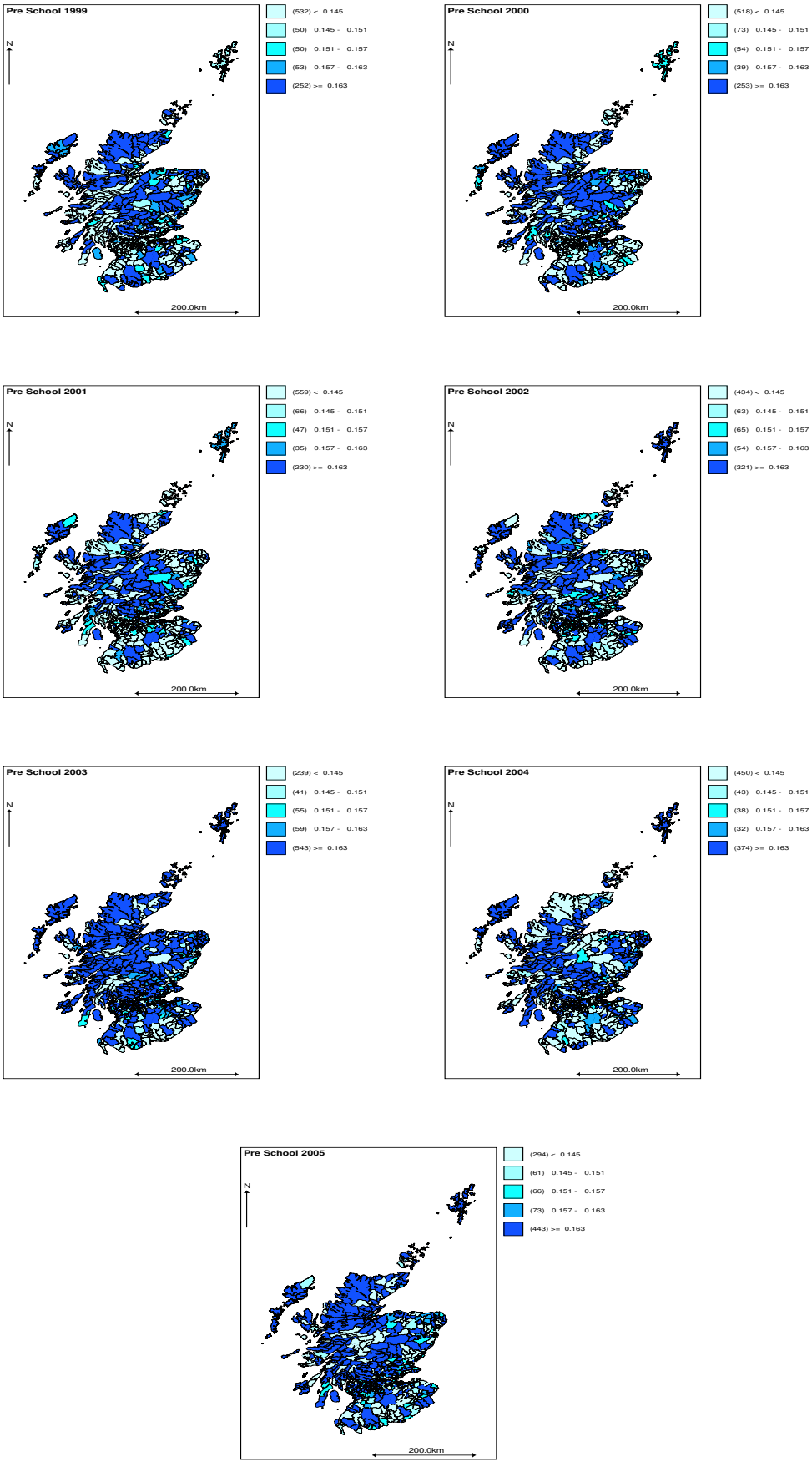

Figure 1.5: Maps of measles susceptibility proportions for postcode sectors, for 1999, 2000, 2001, 2002, 2003, 2004 and 2005 for pre-school children.

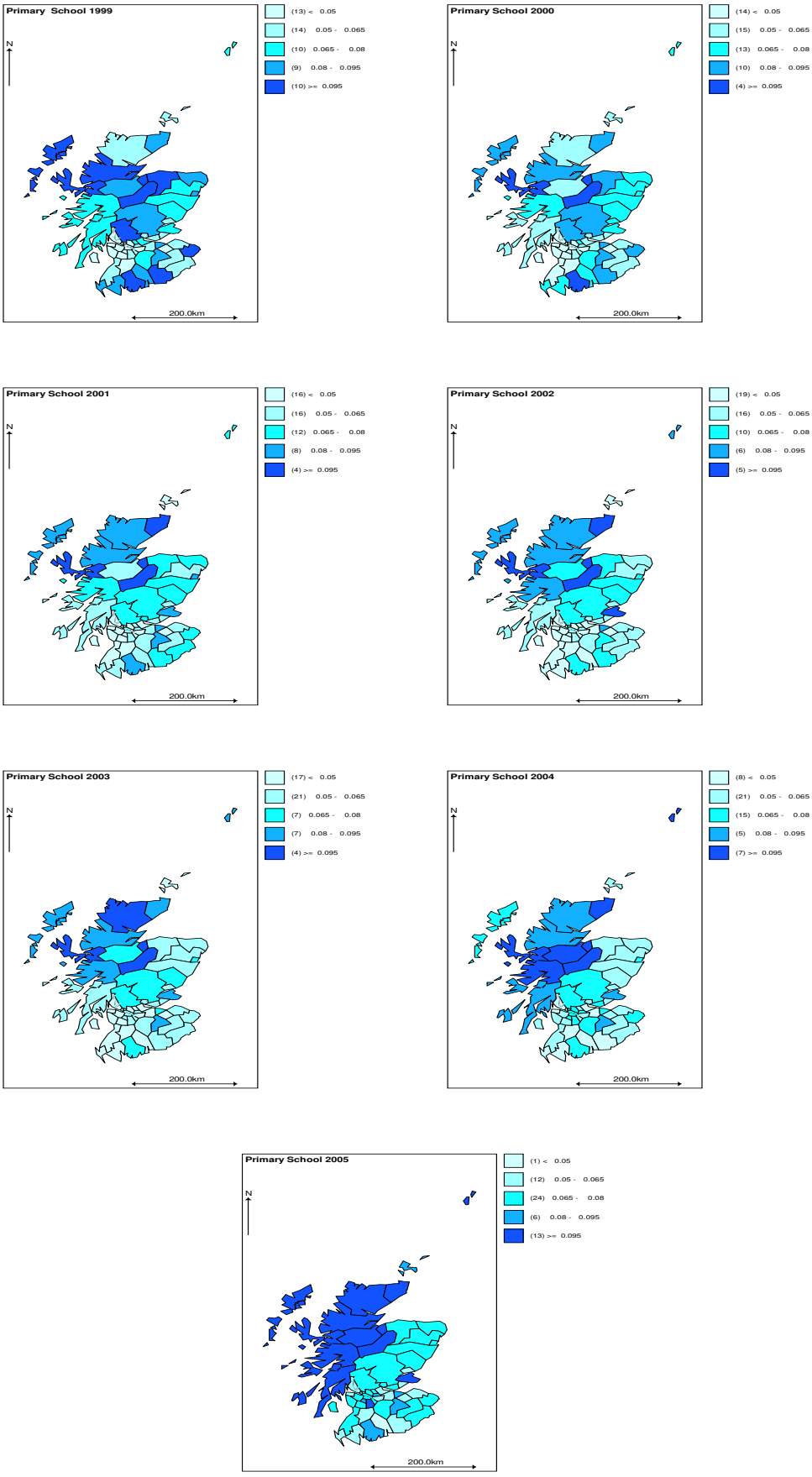

Figure 1.6: Maps of measles susceptibility proportions for districts, for 1999, 2000, 2001, 2002, 2003, 2004 and 2005 for primary school children.

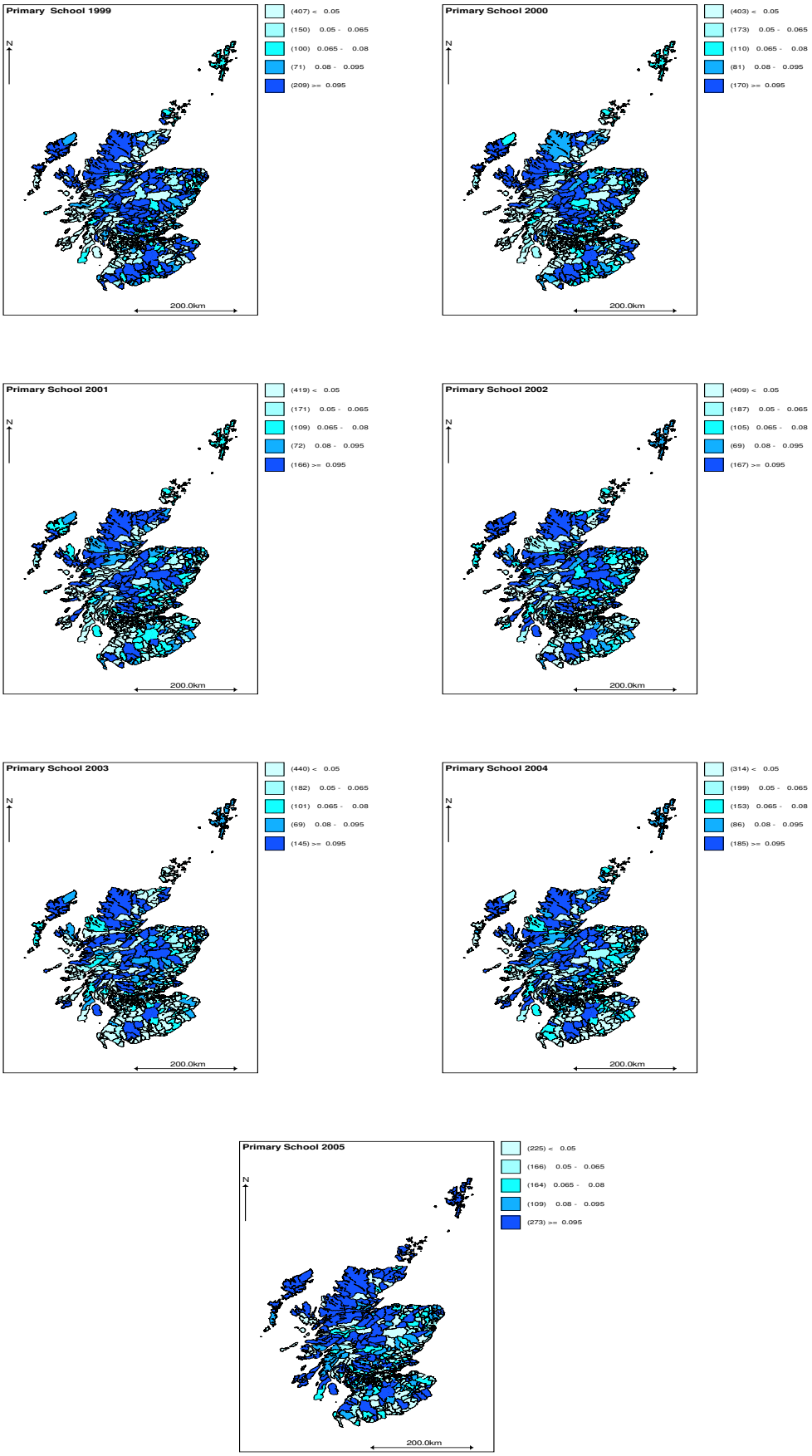

Figure 1.7: Maps of measles susceptibility proportions for postcode sectors, for 1999, 2000, 2001, 2002, 2003, 2004 and 2005 for primary school children.

Map of Names of Districts of Scotland

Figure 1.8 and 1.9 show a map of the 56 districts of Scotland and extracted central districts of Scotland respectively, with the district names given on Table 1.3. A postcode sector map is not presented here, as the postcode sectors are very small and not easy to see in the map.

Area Number	Area Name	Area Number	Area Name
1	Skye and Lochalsh	29	Perth and Kinross
2	Banff and Buchan	30	West Lothian
3	Caithness	31	Cumnock and Doon Valley
4	Berwickshire	32	Stewartry
5	Ross and Cromarty	33	Midlothian
6	Orkney	34	Stirling
7	Moray	35	Kyle and Carrick
8	Shetland	36	Inverclyde
9	Lochaber	37	Cunninghame
10	Gordon	38	Monklands
11	Western Isles	39	Dumbarton
12	Sutherland	40	Clydebank
13	Nairn	41	Renfrew
14	Wigtown	42	Falkirk
15	North East Fife	43	Clackmannan
16	Kincardine and Deeside	44	Motherwell
17	Badenoch and Strathspey	45	Edinburgh City
18	Ettrick and Lauderdale	46	Kilmarnock and Loudoun
19	Inverness	47	East Kilbride
20	Roxburgh	48	Hamilton
21	Angus	49	Glasgow City
22	Aberdeen City	50	Dundee City
23	Argyll and Bute	51	Cumbernauld and Kilsyth
24	Clydesdale	52	Bearsden and Milngavie
25	Kirkcaldy	53	Eastwood
26	Dunfermline	54	Strathkelvin
27	Nithsdale	55	Tweeddale
28	East Lothian	56	Annandale and Eskdale

Table 1.3: Table of names of the former 56 Districts of Scotland (www.gro_scotland.gov.uk) corresponding to the numbers on the map in Figure 1.8.

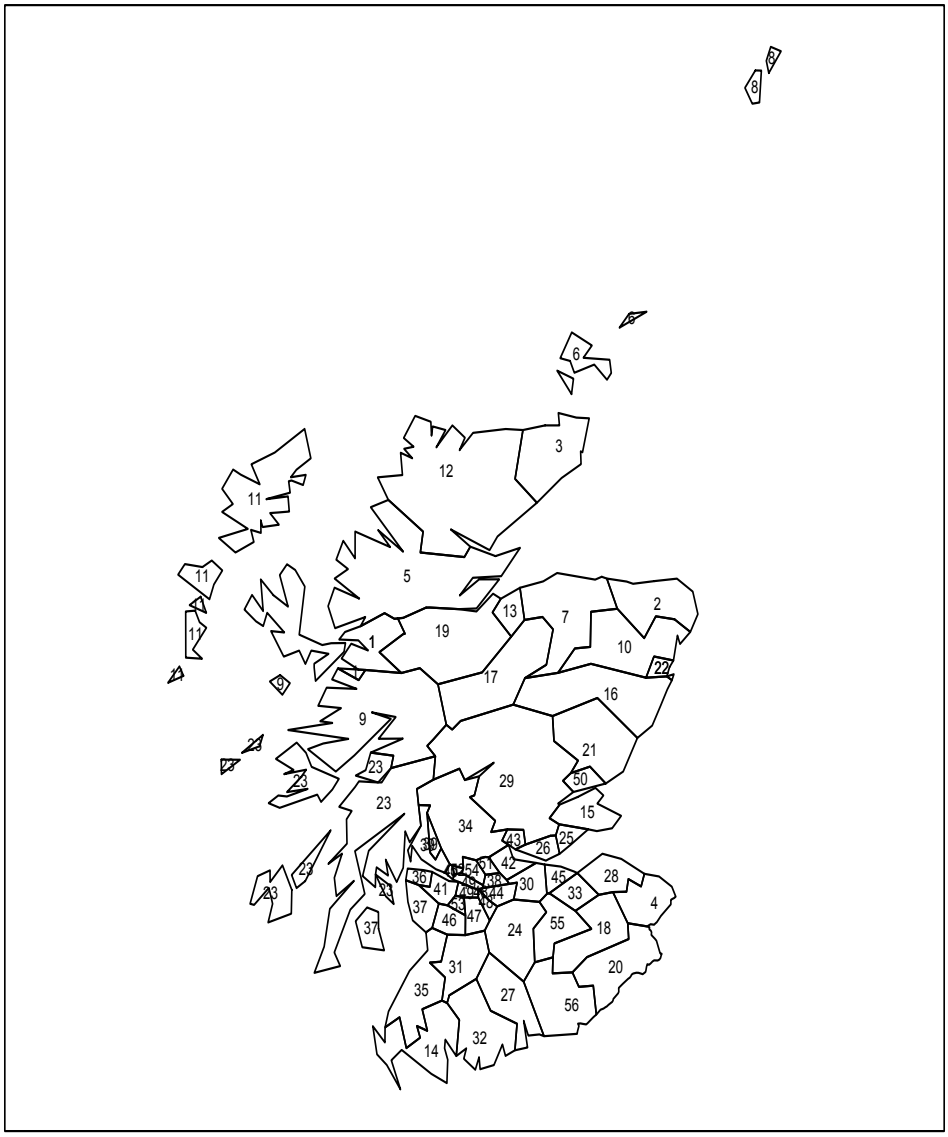

Figure 1.8: Map of the 56 Districts of Scotland.

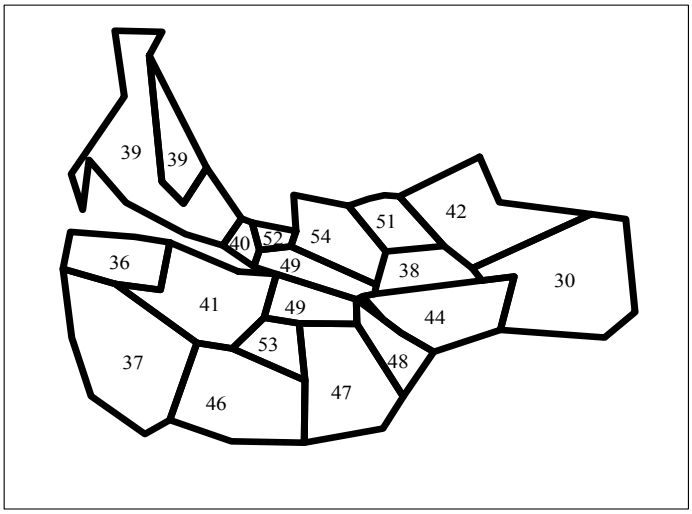

Figure 1.9: Map of Central Districts of Scotland (Enlarged).

1.6 Aim of Thesis

This thesis focuses on disease mapping. In the first part of the thesis we investigate and compare different approaches to the analysis and modelling of disease maps, using spatial data techniques. Here we have pre-school and primary 1 and 2 susceptibility to measles data available at different time points. The aims here are, using disease mapping models:

- To compare susceptibility to measles at pre-school and primary 1 and 2 school children. Children in primary 1 and 2 group are expected to have received a booster to MMR at ages between 3 years 4 months and 5 years, so it is expected that this group will have lower measles susceptibility than the pre-school group. Visual comparison of maps will enable us to see if all areas or regions change or decrease by the same amount or are there pockets where measles susceptibility does not decrease as much as in others.

- To look at time trends in the spatial distribution of measles susceptibility for pre-school and primary 1 and 2 school children. This may be able to show us whether there has been an increase or decrease or an increase followed by a decrease (or vice versa) in measles susceptibility over time. Also, for each group,

with the use of maps we will be able to see whether susceptibility in all areas changes by the same amount or not.

- To see if we can identify areas of high susceptibility at pre-school and primary 1 and 2 stages. These are important to public health as possible areas for a targeted vaccination campaign.

- Using the ecological model, to see if any of the census variables can be useful in the prediction of areas of high susceptibility.

The second part of the thesis focuses on developing methods for comparing disease maps, especially through time. Non-parametric smoothing methods, e.g. kriging, have been used to smooth data and compare trends over time, and these are explored to see whether they can be useful for these data. Maps are similar to images, thus some of the developed methods are based on methods which are used in image analysis to compare a reference image to a distorted image. Some methods will be based on methods used in the analysis of spatial point processes to test for complete spatial randomness of a spatial point pattern. Other methods that we consider are the use of ratio, difference and pseudo-colour maps (which use the primary colours red, green and blue to highlight different levels of susceptibility). Spatial autocorrelation methods and plots of parameters obtained from fitting models are also investigated.

1.7 Thesis Outline

The background and aim of this thesis are given above. In Chapter 2 we review the methods that have been used in modelling of relative risks, including space-time models and mapping of multiple diseases. In Chapter 3, we analyse the susceptibility to measles data (described in Section 1.5) using spatial models, space-time models and including census variables. Data are analysed at both 56 district and 937 postcode sector levels of Scotland and the two sets of results are compared.

Except for visual comparison of maps and the use of space-time models, so far in disease mapping methods have not been developed that can be used to compare extensively two or more maps, to be able to detect changes that have taken place, especially changes over time. As interpolation methods have been used to compare maps over time, these are discussed in Chapter 4 (mainly kernel smoothing and kriging). These

methods are applied to the susceptibility to measles data. In Chapter 5, new ideas for comparing disease maps are developed. These ideas include adapting some methods used in image analysis to compare a distorted image with a reference image, and analogues of point process methods which compare empirical cumulative distributions. In Chapter 6 the methods developed in Chapter 5 are tested on simulated data. As we have susceptibility to measles data available for 1999-2005, in Chapter 7 the methods that were found to be useful in Chapter 6 are used to compare maps of susceptibility to measles, for both pre-school and primary 1 and 2 school children, for 1999-2005, at both district and postcode sector level. We also use these methods to compare maps produced from proportions of calls to NHS24 attributed to cold/flu and fever, difficulty in breathing and cough, rash, and diarrhoea and vomiting syndromes (see Chapter 7). Overall conclusions and discussion, including some suggestions for further work, are given in Chapter 8.

Chapter 2

Models for Disease Mapping of Count Data

2.1 Introduction

In Chapter 1 we discussed SMRs and some of the problems associated with mapping them. The SMR is the maximum likelihood estimator of the relative risk (see Section 1.4.2). In this chapter we look at different models that have been proposed and used in disease mapping to address some of the problems relating to mapping of SMRs. In Chapter 3 spatial data are analysed over a number of years, therefore we need to review models which might be useful. We review the Poisson-Gamma model, lognormal/logistic model, mixture models, linear Bayes method, space-time models and how some of these models can be extended for mapping two or more diseases. Different methods have been used to estimate parameters in these models, and these methods are also reviewed here.

In most of the models, the focus of attention is on modelling the relative risks θ_i. One simple model is to construct a regression model that includes covariates, i.e. a log-linear model of relative risk with fixed effects (Lawson *et al.*, 2003). The log-linear model is used to ensure that the relative risks are positive. For example, we may assume that there is a spatial trend or long range variation, and in this case the tract centroid coordinates or functions of the coordinates can be used as covariates. Thus we have

$$\theta_i = exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}), \tag{2.1}$$

where $exp(\beta_0)$ represents the overall rate in the whole study area, β_1 and β_2 are linear parameters, and x_{1i} and x_{2i} are the coordinates of the centroid of the ith region, i=1,...,n. The model can be extended to include higher order trend surfaces, by including powers of the coordinates. The model could also include covariates measured in each region or tract, which could be found from national census data, relating to socio-economic measures such as deprivation indicators for each region. For example, the percentage of people unemployed can give a measure for increased disease risk due to correlation with poor housing and ill-health. The model can be specified in general as

$$\theta_i = exp(\mathbf{x_i}\beta), \tag{2.2}$$

where $\mathbf{x}$ is an $n \times p$ matrix consisting of observations on $p-1$ covariates and an intercept term, β is a p x 1 parameter vector $(\beta_0, \beta_1, ..., \beta_{p-1})$ and $\mathbf{x_i}$ denotes the ith row of $\mathbf{x}$. However, fitting the above simple model does not guarantee that the disease map will be clean of all noise. There may be unobserved effects, usually termed *random effects*, which have been discussed by many authors both in statistical methodology and epidemiology. These include Mantol *et al.* (1981), Tsutakwa (1988), Breslow and Clayton (1993), Bernardinelli *et al.* (1995b), Lawson (2001), Lawson and Clark (2002), and Richardson (2003).

Random effects are extra variation components which are estimable within a map (Lawson, 2003). These can occur as a result of individuals in a region who may have different susceptibilities to a disease of interest (*frailty effect*), or be associated with spatial variation not accounted for by covariates in the analysis. There may be extra Poisson variation or overdispersion, which occurs due to variation in disease rates exceeding their expected level under the Poisson model. One source of this overdispersion is when there are large numbers of zero counts, as in rare diseases. Overdispersion can occur due to *unstructured/uncorrelated heterogeneity* or *structured heterogeneity* (due to correlation between spatial regions), sometimes referred to as *spatial clustering* or spatial autocorrelation (Bernardinelli *et al.*, 1995b) or to a combination of both. Unstructured heterogeneity is a form of independent and spatially uncorrelated variation. This occurs as a result of differences between spatial units. Structured heterogeneity implies that there is spatial autoccorrelation between the regions. Regions which share a boundary or are neighbours to each other tend to have similar disease rates, thus spatial autocorrelation may arise as a result of clustering of a disease in question or by

the existence of unobserved or frailty effects (Lawson *et al.*, 2003). All these variations should ideally be included within the chosen model.

2.2 Models

In this section, we review and discuss some common models that have been proposed and used in disease mapping. These are two broad groups of models: non-spatial and spatial. Spatial models take into account that neighbouring regions tend to have the same disease rate, i.e. existence of spatial autocorrelation, while non-spatial models ignore the existence of spatial autocorrelation. We also describe some methods which are used to estimate the parameters in the models.

2.2.1 Poisson-Gamma Model

Clayton and Kaldor (1987) proposed the Poisson-Gamma model in which the relative risks were assumed to be Gamma distributed. Let O_i, E_i, θ_i be the observed count, expected count and relative risk parameter in the ith region respectively, $i = 1, ..., n$. Then the observed counts are assumed to be distributed as

$$O_i \sim Pois(E_i\theta_i), \quad i = 1, ..., n. \tag{2.3}$$

Let the relative risks θ_i be independent and identically distributed. If the prior distribution is such that

$$\theta_i \sim \Gamma(\nu, \alpha), \tag{2.4}$$

with shape parameter ν and scale parameter α, yielding mean and variance, $\frac{\nu}{\alpha}$ and $\frac{\nu}{\alpha^2}$, respectively, then the posterior distribution for θ_i will be Gamma distributed also, where

$$p(\theta_i|O_i) \sim \Gamma(O_i + \nu, E_i + \alpha).$$

Then the posterior expectation is

$$E(\theta_i|O_i, \nu, \alpha) = \frac{O_i + \nu}{E_i + \alpha} = w_i\text{SMR}_i + (1 - w_i)\frac{\nu}{\alpha}, \tag{2.5}$$

where

$w_i = \frac{E_i}{E_i+\alpha}$ and $\text{SMR}_i = \frac{O_i}{E_i}$ (Lawson *et al.*, 2003).

The marginal likelihood of O_i given θ_i will be negative Binomial with mean $\frac{E_i\nu}{\alpha}$ and variance $\frac{E_i\nu}{\alpha} + \frac{{E_i}^2\nu}{\alpha^2}$. Estimation can be done using empirical or full Bayesian estimation (see Section 3.4). When the observed counts O_i are large, the Bayes estimates will be close to the SMRs, as the w_i will be approximately 1. With O_i small, the Bayes estimates will tend to be closer to the overall mean risk $\frac{\hat{\nu}}{\hat{\alpha}}$, as the w_i will be closer to zero.

In the case where covariates are to be included, this model is restrictive but Clayton and Kaldor (1987) suggested that area level or ecological covariates $\underline{z}_i$ can be incorporated by allowing distinct values (α_i) for the scale parameters of the distributions of each θ_i, and assuming a log-linear model such that

$$E[log(\theta_i)] = log(\frac{\nu}{\alpha_i}) = \underline{z}_i^T \underline{\beta} \tag{2.6}$$

so that the relative risks are now distributed as $\theta_i \sim \Gamma(\nu, \alpha_i)$. The Poisson-Gamma model allows for overdispersion, but cannot cope with spatial correlation.

2.2.2 Lognormal and Logistic Model

A lognormal model was considered by Clayton and Kaldor (1987). This model was further developed by Besag *et al.* (1991), and referred to as the Besag, York and Mollié (BYM) model. This model assumes that the log relative risk is the sum of two independent components. Taking the observed counts to be Poisson distributed as in (2.3), the log relative risks are modelled as

$$log(\theta_i) = \alpha + u_i + v_i, \tag{2.7}$$

with $\theta_i = exp(\alpha + u_i + v_i)$, where α is the overall mean, u_i is the structured heterogeneity (spatial autocorrelation or spatial clustering), v_i is the unstructured heterogeneity (which measures overdispersion in the individual region), and both u_i and v_i are assumed to be independent. We note that when the rate of the disease is not very small (i.e. the disease is not very rare), assuming the observed counts to be Poisson distributed will not always be appropriate. In such a case the observed counts are assumed to be Binomially distributed, and instead of the lognormal model, the logistic

model will be obtained as

$$O_i \sim Bin(n_i, p_i), \quad logit(p_i) = ln(\frac{p_i}{1-p_i}) = \alpha + u_i + v_i. \tag{2.8}$$

For the logistic model n_i and p_i are number of people and the unknown probability of disease respectively in the ith region , $i = 1, 2, ...n$, with $p_i = \frac{e^{\alpha+u_i+v_i}}{1+e^{\alpha+u_i+v_i}}$. The parameters α, u_i and v_i are as described in the lognormal model.

In the model, unstructured heterogeneity is assumed to be normally distributed as

$$v_i \sim N(0, \tfrac{1}{\tau_v{}^2}).$$

For the structured heterogeneity, a spatial correlation structure is used. Several Gaussian Markov random field models have been used (Best *et al.*, 2005 and Ugarte *et al.*, 2006), but the most commonly used is the conditional autoregressive (CAR) model proposed by Besag *et al.* (1991) and used as follows:

$$(u_i|u_j, i \neq j, \frac{1}{\tau_u^2}) \sim N(\bar{u_i}, \frac{1}{\tau_i^2}), \quad i = 1, ..., n, \quad j = 1, ..., n, \tag{2.9}$$

where

$$\bar{u_i} = \frac{1}{\sum_{j,j\neq i} w_{ij}} \sum_{j,j\neq i} u_j w_{ij}, \quad \tau_i{}^2 = \frac{\tau_u{}^2}{\sum_{j,j\neq i} w_{ij}}, \tag{2.10}$$

and the weights are such that

$$w_{ij} = 1 \text{ if } i \text{ and } j \text{ are labels of adjacent regions, and 0 otherwise,} \tag{2.11}$$

hence taking account of the neighbourhood structure of the regions. Parameters $\tau_v{}^2$ and $\tau_u{}^2$ are the inverse variances of the random effects and they control the variability of v_i and u_i respectively. The lognormal model can incorporate covariates so that (2.7) becomes

$$log(\theta_i) = \alpha + u_i + v_i + \sum_{s=1}^{p} \beta_s x_{is} \tag{2.12}$$

with $\theta_i = exp(\alpha + u_i + v_i + \sum_{s=1}^{p} \beta_s x_{is})$, where $\beta_s, s = 1, ..., p$, are the coefficients of the covariates and x_{is} is the value of covariate s in the *ith* region, $i = 1, ..., n$. For the logistic model, when the covariates are in the model the model (2.8) becomes:

$$O_i \sim Bin(n_i, p_i), \quad logit(p_i) = \alpha + u_i + v_i + \sum_{s=1}^{p} \beta_s x_{is} \tag{2.13}$$

where n_i and p_i are number of people and the unknown probability of disease respectively in the ith region, $i = 1, 2, ...n$, and the $\beta_s, s = 1, ..., p$ are the coefficients of covariates $x_1, ..., x_p$. Therefore,

$$p_i = \frac{e^{\alpha+u_i+v_i+\sum_{j=1}^{p}\beta_j x_{ij}}}{1+e^{\alpha+u_i+v_i+\sum_{j=1}^{p}\beta_j x_{ij}}}.$$

Unlike the Poisson-Gamma model, the log-normal and logistic models can easily incorporate the spatial correlation and include covariates. Estimation can be carried out using empirical or full Bayesian estimation.

2.2.3 Mixture Models

Concerns (Lawson *et al.*, 2000) that the use of parametric prior models for the relative risks could over-smooth the relative risks have led to the development of non-parametric maximum likelihood (NPML) estimation for the prior of the relative risk. This was first proposed by Clayton and Kaldor (1987), and developed by Schlattmann and Böhning (1993) and Heisterkamp *et al.* (1993). These models are mixture models which detect discontinuities in the map (differences in relative risk between neighbouring regions (Knorr-Held and Rasser, 2000)). However, they do not take account of the spatial autocorrelation. The risk variation is modelled by a mixture of components not by a global model (Lawson *et al.*, 2000).

The model assumes that small regions in the whole study area can be grouped into discrete homogenous relative risk classes/components $C_1, ..., C_K$ with constant relative risk α_k within the kth class (Böhning *et al.*, 2000, and Lawson and Clark, 2000). There are K distinct levels of overall risk, $\alpha_1, ..., \alpha_K$ and probability p_k of belonging to class k, $\quad k = 1, ..., K$ such that

$$\theta_i \sim f_{\theta_i}(\alpha_1, ..., \alpha_k), \quad \text{and} \quad P(\theta_i = \alpha_k) = p_{ik} \tag{2.14}$$

Following Leyland and Davies (2005), for a mixture of Poisson distributions the conditional likelihood is

$$f(O_i|\underline{\alpha}) = \sum_{k=1}^{K} \frac{p_k e^{-\alpha_k E_i}(\alpha_k E_i)^{O_i}}{O_i!}, \ i = 1, ..., n. \tag{2.15}$$

By Bayes theorem, the probability that the ith region belongs to a class k is estimated as

$$\hat{p}_{ik} = P(w_{ik} = 1|O_i, E_i) = \frac{\hat{p}_k e^{-\hat{\alpha}_k E_i}(\hat{\alpha}_k E_i)^{O_i}}{\sum_{k=1}^{K} \hat{p}_k e^{-\hat{\alpha}_k E_i}(\hat{\alpha}_k E_i)^{O_i}}, \ i = 1, ..., n, \tag{2.16}$$

where the ith region is assigned to the class for which $\hat{p}_{ik}$ is largest, w_{ik} is a latent/hidden random variable and $w_{ik} = 1$ if the ith region belongs to class k and 0

otherwise. The posterior empirical Bayes estimates are then weighted averages given by

$$\hat{\theta}_i = E(\theta_i|O_i, E_i) = \frac{\sum_{k=1}^{K} \hat{\alpha}_k \hat{p}_k e^{-\hat{\alpha}_k E_i} (\hat{\alpha}_k E_i)^{O_i}}{\sum_{k=1}^{K} \hat{p}_k e^{-\hat{\alpha}_k E_i} (\hat{\alpha}_k E_i)^{O_i}} \quad \text{(Heisterkamp } et \text{ } al.\text{, 1993).} \quad (2.17)$$

Biggeri *et al.* (2003) formalised the proposed approach using a pseudo-likelihood to derive a transitional non-parametric maximum pseudo-likelihood (TNPMPL) estimator. This method incorporates spatial autocorrelation. For each area, there are K risks α_{ik} such that (2.15) becomes

$$f(O_i|\alpha_{ik}, k = 1, ..., K) = \sum_{k=1}^{K} \frac{p_k e^{-\alpha_{ik} E_i} (\alpha_{ik} E_i)^{O_i}}{O_i!} \quad (2.18)$$

with the logarithm of α_{ik} modelled in terms of SMRs in the regions contiguous to i (Leyland and Davies, 2005) as

$$log(\alpha_{ik}) = \beta_k + \phi_k log(\frac{\sum_j w_{ij} O_i}{\sum_j w_{ij} E_i}), \quad (2.19)$$

with weights w_{ij} given by (2.11), and here also the posterior empirical Bayes estimates are the weighted average of support points in each region as in (2.17) but in this case weights are region specific. For both NPML and TNPML, the estimation is done via the EM algorithm (see Section 2.5.1).

Models by Schlattmann and Bhning (1993) ignore the location of regions, i.e. members of a mixture class may be located over the whole region. Knorr-Held and Rasser (2000), proposed a spatial partition model which considered the location of regions. Here the relative risks θ_i are assumed to be constant over a set of one or more contiguous regions. The clusters $C_k, k = 1, .., K$ are a partition of regions each with relative risk α_k which is constant among the regions belonging to that cluster. Some regions are cluster centres. Cluster centres (K of them) are chosen at random. The regions are then allocated to a cluster if the cluster centre is closest to it in terms of the minimal number of boundaries that have been crossed to reach it, and regions that have same distance to two or more clusters are assigned to the cluster with the smallest index position (Knorr-Held and Rasser, 2000). The observed counts O_i are assumed to be Poisson distributed with mean $\alpha_j E_i$, and in region i of cluster j

$$log(\alpha_j) \sim N(\mu, \sigma^2), \quad j = 1, ..., K \quad (2.20)$$

with hyperparameters μ and σ^2, and prior for the number of clusters $K \sim Unif(1, c_{max})$ or a geometric distribution. In general c_{max} may be set to be number of regions (Best

et al., 2005). Reversible jump MCMC is used to fit the cluster model.

Lawson and Clark (2002) also proposed a spatial mixture model that allows both smoothness and discontinuities on the map. This model assumes that the log-relative risks can be decomposed into three additive components, rather than the two used in Besag *et al.* (1991), as

$$log(\theta_i) = \alpha + v_i + p_i u_i + (1 - p_i)\varphi_i \tag{2.21}$$

where v_i is a component representing unstructured heterogeneity (measuring overdispersion in an individual region), and $v_i \sim N(0, \frac{1}{\tau_v{}^2})$, $p_i \sim beta(a, a)$. The two mixing components are $u_i,\ i = 1, ..., n$, a spatial correlation, and $\varphi_i,\ i = 1, ..., n$, a component that models discrete jumps. When $p_i = 1$ we obtain the BYM model, and $p_i = 0$ for every i gives a pure jump model (Lawson et al., 2000; Lawson and Clark, 2002). For u_i, the usual CAR model is adopted, and for the jump component the prior is

$$\pi(\varphi_1, ..., \varphi_n) \propto \frac{1}{\sqrt{\lambda}} \exp(-\frac{1}{\lambda} \sum_{i\ j} \|\varphi_i - \varphi_j|),$$

The neighbours for this prior are defined as in the lognormal model and the prior uses total absolute differences between neighbours (Lawson *et al.*, 2003).

The authors observed that the maps produced from this model were visually closer to those obtained from the SMRs than those from the BYM model. This suggests the possibility of using models that do not over-smooth the map while allowing both smoothing and jumping in the relative risks (Lawson and Clark, 2002).

2.2.4 Linear Bayes Method

Marshall (1991) used a linear Bayes method. If the prior distribution of θ_i has mean m_i and variance v_i, the best linear Bayes estimator of the relative risk θ_i (in the sense of total squared error loss) is

$$\hat{\theta}_i = m_i + \frac{v_i}{\frac{m_i}{E_i} + v_i}(\text{SMR}_i - m_i). \tag{2.22}$$

When E_i is large the estimate tends to the SMR, and for small E_i the estimate will tend towards the prior mean. Assuming constant mean and variance, i.e. $m_i = m$ and $v_i = v$, and using the method of moments, the global estimates (estimates shrunk towards the global mean) were obtained as

$$\hat{m} = \frac{\sum_i O_i}{\sum_i E_i} \tag{2.23}$$

and

$$\hat{v} = \frac{1}{\sum_i E_i} \sum_i E_i(\text{SMR}_i - \hat{m})^2 - \frac{\hat{m}}{\sum_i E_i}. \tag{2.24}$$

Since the estimate of $\hat{v}$ can be negative, the convention is adopted that $\hat{\theta}_i = \hat{m}$ if $\hat{v} < 0$.

Marshall (1991) obtained the local estimates (considering spatial correlation), by using only the regions that are neighbours to a region to obtain the prior mean and variance, i.e.

$$\hat{m}_i = \frac{\sum_j w_{ij} O_j}{\sum_j w_{ij} E_j},$$

where $w_{ij} = 1$ if region i and j are neighbours, and 0 otherwise, and

$$\hat{v_i} = \frac{1}{\sum_j w_{ij} E_j} \sum_j w_{ij} E_j (\text{SMR}_j - \hat{m}_j)^2 - \frac{\hat{m}_j}{\sum_j w_{ij} E_j} \quad \text{(Leyland } \textit{et al.}\text{, 2005)}.$$

2.2.5 Space-time Modelling

In the above sections we have considered methods of smoothing when data are available at one time point. When data are available at a number of different time points, space-time models can be used to model the risk in space and time. These models consider spatial smoothing, temporal smoothing and spatio-temporal interaction. Space-time modelling may help in identifying time trends, and in producing maps at different time points. It may be helpful in disease surveillance to monitor spatial patterns over time, revealing regions which have consistently high relative risks, so these can be investigated further (MacNab and Dean, 2001). Most space-time models used in disease mapping are an extension of the Besag *et al.* (1991) (BMY) model.

Bernardinelli *et al.* (1995a) proposed a space-time model as follows. Let O_{ik} be the observed counts for the ith region, $i = 1, ..., n$, and the kth time interval, $k = 1, ..., T$. Similarly E_{ik} and θ_{ik} are the expected count and relative risk in region i at time point k. Then

$$O_{ik}|\underline{\theta} \sim Pois(E_{ik}\theta_{ik}), \quad log(\theta_{ik}) = \alpha + u_i + v_i + \beta * t_k + \delta_i * t_k, \tag{2.25}$$

where α is a constant (overall mean rate), u_i is the structured heterogeneity, v_i is the unstructured heterogeneity, $\beta * t_k$ is the linear trend term in t_k, where t_k, $k = 1, ..., T$ is the kth time point, T is the total number of time points, and δ_i is the random space-time interaction effect. The conditional autoregressive (CAR) model is used

for the spatial correlation terms u_i and for the space-time interaction terms δ_i. The uncorrelated heterogeneity is assumed to have the Normal distribution

$$v_i \sim N(0, {\tau_v}^2).$$

The parameters α, β, and the inverse parameters of τ_u, τ_v, and the δ_i are assigned priors.

Waller *et al.* (1997) proposed a model where the hierarchical specification by Besag *et al.* (1991) is applied to each time point k separately, $k = 1, .., T$. In this model

$$O_{ik}|\underline{\theta} \sim Pois(E_{ik}\theta_{ik}), \quad log(\theta_{ik}) = \alpha + {u_i}^k + {v_i}^k, \tag{2.26}$$

where O_{ik}, E_{ik} and θ_{ik} are the observed count, expected count and relative risk in region i at time point k. This model allows the spatial patterns at each time point to be completely different. As in the log-normal model, the conditional autoregressive (CAR) model is used for the spatial correlation at each time point k, and the uncorrelated heterogeneity term at each time point k is

$${v_i}^k \sim N(0, {\tau_{vk}}^2).$$

Taking the observed counts to be Binomially distributed, this space-time model becomes

$$O_{ik}|\underline{\theta} \sim Bin(n_{ik}, p_{ik}), \quad logit(p_{ik}) = \alpha + {u_i}^k + {v_i}^k, \quad k = 1, ..., T \tag{2.27}$$

where n_{ik} and p_{ik} are number at risk and probability of risk in region i at time point k respectively.

One space-time model that is based on a generalised additive mixed model (GAMM), is a semi-parametric mixed effects model. This additive extension of generalised linear mixed models was proposed and used by MacNab and Dean (2001) and MacNab and Dean (2002). The model incorporates spatial random effects and both fixed and random temporal effects.

Let O_{it} be the observed count and μ_{it} be the conditional expectation of O_{it} given the random spatial and temporal effects, for the ith region $i = 1, ..., n$ and year t, $t = 1, ..., T$. Then the logarithm of the conditional mean count for the ith region is taken as

$$log(\mu_{it}) = log(n_{it}) + log(m) + S_0(t) + \theta_i + S_i(t), \tag{2.28}$$

where n_{it} is population at risk in region i at time point t, m is the overall mean rate, $S_0(t)$ is the fixed mean rate trend at time t over all local regions considered, the θ_i are the random spatial effects (area specific relative risks which may be spatially correlated) and $S_i(t)$ is the random temporal trend effect for the ith region at time t. MacNab and Dean (2001) and (2002) used a conditional autoregressive model to model the random spatial effects and used a cubic B-spline for the fixed $S_0(t)$ and random $S_i(t)$ temporal effects. $S_0(t)$ is then given by

$$S_0(t) = \sum_{k=1}^{K} \beta_{0,k} p_k(t) \tag{2.29}$$

and

$$S_i(t) = \sum_{k=1}^{K} \beta_{i,k} p_k(t), \tag{2.30}$$

where $\beta_{0,k}$ and $\beta_{i,k}$ are estimates of the fixed and random effects respectively, and $p_k(t)$ (p not a probability here) is the kth B-spline basis function evaluated at time t.

A special case of (2.28) replaces $S_i(t)$ by an approximating random linear trend such that

$$log(\mu_{it}) = log(n_{it}) + log(m) + S_0(t) + \theta_i + \beta_i t. \tag{2.31}$$

This model accommodates a non-linear trend modelled by a spline. Then $log(m)+S_0(t)$ is the linear local region relative risk trend and the values of $\theta_i + \beta_i t$ are area-temporal effects which measure the departure of local region relative risks from the overall spline rate (MacNab and Dean, 2002). This model may represent a generalised linear model with

$$E(\underline{O}|\underline{b}) = \underline{\mu} = exp(log(m) + X\underline{a} + Z\underline{b}) \tag{2.32}$$

where $\underline{a}$ is a $(K+1) \times 1$ vector of fixed effects given by $\underline{a} = (log(m), \beta_{0,1}, ..., \beta_{0,K})^T$, and $\underline{b}$ is a vector of random effects of length $2n$ given by $\underline{b} = (\theta_1, ..., \theta_n, \beta_{1,1}, ..., \beta_{n,K})^T$ with $\underline{b} \sim MVN(0, \Sigma)$. The design matrix $\underline{X}$ of the fixed effects will be $nT \times (K+1)$, i.e.

$$\underline{X} = \begin{pmatrix} 1 & p_1(t_1) & ... & p_K(t_1) \\ & ... & ... & ... & ... \\ 1 & p_1(t_1) & ... & p_K(t_1) \\ & ... & ... & ... & ... \\ 1 & p_1(t_T) & ... & p_K(t_T) \\ & ... & ... & ... & ... \\ 1 & p_1(t_T) & ... & p_K(t_T) \end{pmatrix}$$

and $\underline{Z}$ is an $nT \times 2n$ matrix given by

$$\underline{Z} = \begin{pmatrix} Z_0 & Z_1 \\ Z_0 & Z_2 \\ ... & ... \\ Z_0 & Z_T \end{pmatrix}$$

where Z_0 is an $n \times n$ identity matrix, Z_l is a diagonal matrix of dimension n with all diagonal elements equal to t_l, $l = 1, ..., T$, and n=number of regions.

The GAMM model can be extended to include covariates. MacNab and Dean (2001) and (2002) used Penalised-quasi likelihood (PQL) for estimation of the parameters but an MCMC approach (see Section 2.6) can also be used. The advantage of the GAMM model is the use of lower order B-splines, a simple design matrix for generalised mixed models (GLMM) analysis, and computationally simple estimation, but the disadvantage is lack of invariance to change of basis when using random coefficients with splines (MacNab and Dean, 2001).

2.2.6 Mapping Multiple Diseases

In disease mapping, models for mapping two (bivariate models) or more diseases have also been developed and could be used to compare maps at two time points. These are joint models which are extensions of those used in the case of a single disease. Since many diseases share common risk factors, for example smoking and alcohol consumption, joint mapping of two or more related diseases may provide a way to borrow strength across diseases as well as across nearby regions and provide better estimates of risk (Best et al., 2005). When interest is in a rare disease, joint modelling of a rare disease with a more common and related disease may help in increasing precision in the estimates of the rare disease. The main aim of joint mapping is to find similarities and dissimilarities in the spatial distribution of disease risk (Dabney and Wakefield, 2005). Identifying similar patterns in the spatial variation of the related diseases may give more convincing evidence of common risk factors than in the analysis of a single disease (Held et al., 2005). For example, in a case when the risk factors of a disease are known but unmeasured at an area level, the similarity of relative risk rates to those of a second disease may imply that the diseases have common (or subset) risk factors, and dissimilarity of spatial patterns may imply that the risk factors are not a

cause of the second disease (Dabney and Wakefield, 2005).

In the case of two diseases, let O_{1i} and O_{2i} be observed counts from disease 1 and 2 respectively in region i. The observed counts are distributed as

$$O_{1i} \sim Pois(\theta_{1i}E_{1i}), \qquad O_{2i} \sim Pois(\theta_{2i}E_{2i}), \tag{2.33}$$

where E_{1i} ,E_{2i} are expected counts for the two diseases and θ_{1i} and θ_{2i} are the unknown relative risks for the two diseases in region i, $i = 1, ..., n$.

Knorr-Held and Best (2001) proposed a shared component model for the joint spatial analysis of a disease, with the aim of identifying shared and disease specific varying patterns of risk (Best *et al.*, 2005). The relative risk is split into three components and modelled as

$$log(\theta_{1i}) = \alpha_1 + \rho_i\delta + \beta_{1i}, \quad log(\theta_{2i}) = \alpha_2 + \frac{\rho_i}{\delta} + \beta_{2i}, \tag{2.34}$$

where α_1 and α_2 are overall levels of relative risk for disease 1 and 2 respectively, ρ_i is a shared component (representing unmeasured risk factors) shared by both diseases, β_{1i} and β_{2i} are components specific to diseases 1 and 2 respectively, and δ allows for a different gradient in the shared component for each disease.

The three components ρ_i, β_{1i} and β_{2i}, are assumed to be independent, with each one following a spatial prior distribution. In principle, any spatial prior distribution can be used (Best et al., 2005). Knorr-Held and Best (2001) used the spatial partition model by Knorr-Held and Rasser (2000), and the Best *et al.* (2005) used the Besag *et al.* (1991) spatial prior for each component.

The advantages of the shared model are the flexibility of a spatial prior and the ability to estimate and map the shared and specific components of the risk separately. Held *et al.* (2005) extended this model to more than two diseases.

Dabney and Wakefield (2005) proposed a proportional mortality model. Let the observed counts be modelled as (2.33). The log-relative risk is given by

$$log(\theta_{1i}) = \alpha_1 + u_{1i} + v_{1i}, \quad log(\theta_{2i}) = \alpha_2 + u_{2i} + v_{2i}, \tag{2.35}$$

where u_{1i} and u_{2i} are spatially structured effects and v_{1i} and v_{2i} are unstructured effects. Let the sum of the two disease counts in area i be $M_i = O_{1i} + O_{2i}$. Then

Dabney and Wakefield (2005) assume

$$O_{1i}|M_i \sim Bin(M_i, p_i) \tag{2.36}$$

with logit function

$$logit(p_i) = log(\frac{E_{1i}}{E_{2i}}) + \alpha + {u_i}^* + {v_i}^*, \tag{2.37}$$

where

$$p_i = \frac{\exp(log(\frac{E_{1i}}{E_{2i}}) + (\alpha_1 - \alpha_2) + (u_{1i} - u_{2i}) + (v_{1i} - v_{2i}))}{1 + \exp(log(\frac{E_{1i}}{E_{2i}}) + (\alpha_1 - \alpha_2) + (u_{1i} - u_{2i}) + (v_{1i} - v_{2i}))}, \tag{2.38}$$

and $\alpha = \alpha_1 - \alpha_2$ and ${u_i}^*$ can be thought of as $u_{1i} - u_{2i}$, so that values of the ${u_i}^*$ capture similarity and dissimilarity between the spatial random effects of each disease. That is, large values will indicate that the two disease differ significantly (in terms of spatial component) in region i, thus a map of the spatial components will highlight similarities and dissimilarities between the two diseases. The unstructured ${v_i}^*$ can also be taken as $v_{1i} - v_{2i}$ and a map of these values will indicate the regions where the unobserved unstructured effects are similar or dissimilar between the two diseases. Area level covariates can be included in the model.

Space-time modelling of two diseases or two sub-groups of one disease was considered by Richardson *et al.* (2006) as follows. Let O_{1it} and O_{2it} be observed counts from diseases 1 and 2 in region i and time $t, t = 1, ..., T$. Then observed counts are distributed as

$$O_{1it} \sim Pois(\theta_{1it}E_{1it}), \qquad O_{2it} \sim Pois(\theta_{2it}E_{2it}), \tag{2.39}$$

where E_{1it}, E_{2it} are expected counts for the two diseases, and θ_{1it} and θ_{2it} are unknown relative risks for the two diseases in region i, $i = 1, ..., n$, at time t. The log relative risks are modelled in the most general model as

$$log(\theta_{1it}) = \alpha_1 + \rho_i\delta + \xi_t\kappa + \varsigma_{it} + \varepsilon_{1it}, \quad log(\theta_{2it}) = \alpha_2 + \frac{\rho_i}{\delta} + \frac{\xi_t}{\kappa} + \varsigma_{it} + \vartheta_i + \eta_t + \varepsilon_{2it} \tag{2.40}$$

where ρ_i is the shared spatial pattern, ϑ_i is the disease 2 differential from the shared spatial pattern (disease-space interaction in region i), and a spatially correlated (CAR) prior is used for both of these, ξ_t is a shared time trend, η_t is the disease 2 differential at time t from the shared time trend (disease-time interaction), and a first order random walk prior is used for ξ_t and η_t. This random walk prior is a one-dimensional version of the CAR prior with adjacency weight matrix defining the temporal neighbours of time t as periods $t - 1$ and $t + 1$. The terms δ and κ are coefficients representing relative

effects of the shared risk terms, and ς_{it} is the space-time interaction for region i and time t. The authors chose to assume $\varsigma_{it} \sim N(0, \tau_\varsigma)$. The terms ε_{1it} and ε_{2it} account for overdispersion and the authors chose a zero-mean multivariate normal distribution for the distribution of $(\varepsilon_{1it}, \varepsilon_{2it})^T$. This is an extension of the Knorr-Held and Best (2001) model.

2.3 Comparison of Models

The Poisson-Gamma model is simple to use, but it is unable to cope with spatial correlation between relative risks in neighbouring areas. The BYM model can accommodate both spatial correlation and covariates. Lawson *et al.* (2000) concluded in their analysis comparing models that the Poisson-Gamma and BYM models outperform the mixture models. Unlike the mixture models, the Poisson-Gamma and BYM models are based on smoothing methods which often smooth over large discontinuities in the risk surface, but these jumps may be important to help in allocation of resources (Lawson and Clark, 2002; Lawson *et al.*, 2003).

For the mixture models, the NPML approach performs very well in identifying regions with extreme high relative risks but seems to fail when spatial autocorrelation is present (Militino *et al.*, 2001; Biggeri *et al.*, 2003). However, the TNPML approach addresses this and Biggeri *et al.* (2003) found that, unlike NPML, TNPMPL gives estimates that are closer to the parametric models. The disadvantage of TNPMPL is that it does not allow for incorporation of covariates. Compared to the mixture models proposed by Lawson and Clark (2002), the NPML and TNPML approaches do not consider that other regions in the map may have smooth relative risk transitions.

The strength of using non-parametric mixture models is the ability to divide a region into clusters, making the production of a disease map simple. These clusters are presented in a map legend with no need to categorise relative risks into percentiles as in parametric methods (Militino *et al.*, 2001). The disadvantage is that it is difficult to obtain standard errors of estimates provided by the EM algorithm (Dempster *et al.*, 1977).

Comparing the non-spatial and spatial model, even though non-spatial models perform well they do not incorporate spatial autocorrelation, which leads to the increase

of autocorrelation in the residuals (Lawson *et al.*, 2000).

In Chapter 3, the Poisson-Gamma model and BYM model are compared using the measles data described in Section 1.5. The analysis is then done extensively with the use of BYM models and the space-time model of Waller *et al.* (1997).

2.4 Estimation of Parameters

The most common method of estimation in disease mapping is the use of Bayesian estimation for hierarchical models. The first level of the model is commonly the Poisson model for observed counts, i.e.

$$O_i|\underline{\theta} \sim Pois(E_i\theta_i), \tag{2.41}$$

and the second level models the extra Poisson variation through a *prior* distribution, often Gamma. This distribution is given by the investigator, based on prior belief concerning the behaviour of the relative risk parameters, and is parameterised by *hyperparameters*. Let $L(\underline{O}|\underline{\theta})$ be the likelihood of the observed counts given the relative risks and $g(\underline{\theta}|\underline{\lambda})$ be the prior distribution of the relative risks given hyperparameters $\underline{\lambda}$, where $\underline{O} = (O_1, O_2, ..., O_n)$, $\underline{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ are vectors of observed counts and relative risks respectively. Then

$$p(\underline{\theta}|\underline{O}, \underline{\lambda}) \propto L(\underline{O}|\underline{\theta})g(\underline{\theta}|\underline{\lambda}) \tag{2.42}$$

is the marginal *posterior* distribution of the relative risk describing the behaviour of the parameters (relative risks) after the data are observed and prior assumptions have been made. The relative risks $\underline{\theta}$ may be estimated from the posterior distributions as the *posterior mean* or *mode*. When there are many data the likelihood contributes more to the relative risk estimation and when the data are fewer then the prior distribution dominates.

2.4.1 Empirical and Full Bayesian Estimation

The earlier Bayesian hierachical disease methods focused on *empirical Bayes methods* to estimate the relative risk. These methods use two-level hierarchical models and they seek to approximate the posterior distribution. The estimates of the hyperparameters

$\underline{\lambda}$ are the maximum likelihood estimates derived from the marginal likelihood of $\underline{\lambda}$ given by

$$L(\underline{\lambda}) = \int L(\underline{O}|\underline{\theta})g(\underline{\theta}|\underline{\lambda})d\underline{\theta}. \tag{2.43}$$

In the empirical Bayes setting to estimate the parameters in the model it is common to use penalised quasi-likelihood (PQL) methods (see Ugarte *et al.* (2006) and Section 2.5.2). PQL is straightforward and computationally simple, and performs very well provided the expected count is not less than two (Leroux, 2000). The EM algorithm (Dempster *et al.*, 1977) is also commonly used (Section 2.5.1).

Recently, full Bayesian estimation of relative risk has become available due to successful applications of Markov Chain Monte Carlo (MCMC; see Section 2.6) methods of posterior sampling and its implementation in software packages. In the full Bayesian approach, the hyperparameters of the prior distribution of the relative risks are given a distribution (referred to as a *hyperprior* ($f(\underline{\lambda})$). Estimation is based on the posterior distribution of the relative risk given the data.

In general, the joint posterior distribution of the relative risks and the hyperparameters $\underline{\lambda}$ given the observed data is

$$p(\underline{\theta}, \underline{\lambda}|\underline{O}) \propto L(\underline{O}|\underline{\theta})g(\underline{\theta}|\underline{\lambda})f(\underline{\lambda}) \tag{2.44}$$

with the marginal posterior distribution of $\underline{\theta}$ given the observed data given by

$$m(\underline{\theta}|\underline{O}) = \int p(\underline{\theta}, \underline{\lambda}|\underline{O})d\underline{\lambda}. \tag{2.45}$$

The criticism of the empirical Bayesian estimation is that, as it uses estimates of hyperparameters, uncertainty is not allowed for in the hyperparameters (Leyland and Davies, 2005). Bernardinelli *et al.* (1992) pointed out that the estimates produced by empirical Bayesian estimation are inexact, as they are conditional on a point estimate for the smoothing parameter, while the full Bayesian method gives approximate estimates.

2.4.2 Choice of the Prior Distribution for Variance Parameters

Choosing prior distributions is an important part of Bayesian analysis. These give information about the unknown parameters, and combined with the probability distribution of the data, this gives the posterior distribution which is used to make decisions

concerning the unknown parameter (Gelman, 2002). Therefore, when assigning a prior distribution one should consider the kind of information available and the properties of the posterior distribution obtained.

In full Bayesian modelling, a challenge is in choosing the hyperprior distribution for the variance (σ^2) parameters. When the sample is large, the choice of hyperprior will have less effect since the data will dominate the hyperprior, but for small sample sizes the choice of hyperprior becomes very important. In practice, a sensitivity analysis should be carried out by comparing reasonable choices of hyperpriors, to investigate the influence of the hyperprior on the relative risks and other parameters.

Some of the proposed prior distributions are uniform distributions (Gelman *et al.*, 2003) and inverse-Gamma distributions (Spiegelhalter *et al.*, 1994, 2003). These are noninformative and improper prior distributions. A uniform distribution can be assigned to $log(\sigma)$ as this will be working with a logarithm of a positive parameter, but this gives an improper posterior distribution. Thus an option is to define $\sigma \sim U(0, A)$ which yields a proper posterior distribution as $A \to \infty$ (Gelman, 2006). For a noninformative but proper prior, Gelman (2006) recommends approximating the uniform distribution for σ by a uniform on a wide range i.e. $\sigma \sim U(0, 100)$ or a half-normal distribution centered at zero with high standard deviation.

The inverse-gamma (ϵ, ϵ) is the most commonly used prior distribution on the variance σ^2, with ϵ set to low values such as 1 or 0.01 or 0.001. This prior distribution gives an improper posterior distribution when the limit $\epsilon \to 0$, therefore reasonable values of ϵ must be used. Gelman (2006) demonstrated that when low values of σ are possible (near zero), then the inference is very sensitive to ϵ, and thus does not recommend use of the inverse-gamma prior. In the case that more prior information is required, Gelman (2006) recommends working with more flexible prior distributions that behave well near zero rather than the inverse-gamma prior. He recommends half-t family distributions such as a half-Cauchy prior for σ with a high value of the scale parameter; for example half-Cauchy (25).

2.5 Techniques Used in Empirical Bayesian Estimation

In this section some of the techniques used in empirical Bayesian estimation are briefly discussed, namely the EM algorithm and penalised quasi-likelihood.

2.5.1 Estimation and Maximisation (EM) Algorithm

To compute the maximum likelihood estimates of the hyperparameters ($\underline{\lambda}$) from the marginal likelihood (2.43), the EM algorithm can be used (Dempster *et al.*, 1977). This algorithm computes the estimates through an iteration between an estimation step (E-step) and a maximisation step (M-step). Using the log-likelihood, at the E-step the estimate of the marginal posterior distribution of $\underline{\theta}$ (2.42) is evaluated at a particular set of values $\underline{\lambda}^{(p)}$ after p iterations, thus, up to constant terms not involving $\underline{\theta}$,

$$log[p(\underline{\theta}|\underline{O}, \underline{\lambda} = \underline{\lambda}^{(p)})] = log[L(\underline{O}|\underline{\theta})] + log[g(\underline{\theta}|\underline{\lambda} = \underline{\lambda}^{(p)})]. \tag{2.46}$$

Let $q(\underline{\lambda}^{'}|\underline{\lambda}) = E[log(p(\theta|\underline{\lambda}^{'}))|\underline{O}, \underline{\lambda}]$, assumed to exist for all pairs $q(\underline{\lambda}^{'}, \underline{\lambda})$.

At the E-step $q(\underline{\lambda}|\underline{\lambda}^{(p)})$ is computed. The M-step maximises (2.46) over the hyperparameters $\underline{\lambda}$ (Leyland and Davies, 2005), i.e. determines $\underline{\lambda}^{(p+1)}$ which maximises $q(\lambda|\lambda^{(p)})$.

2.5.2 Penalised Quasi-likelihood

Penalised quasi-likelihood (PQL) (Breslow and Clayton, 1993) is an approximation technique for generalized linear mixed models which uses weighted least-squares estimation for estimating fixed effects parameters and likelihood equations from an approximating normal model for estimating variance components. Following the formulation of Dean *et al.* (2004), let O_{ij}, E_{ij} and θ_{ij} be the observed count, expected count and relative risk for the ith region and jth age group, $i = 1, ..., n,\ j = 1, ..., J$, then

$$O_{ij} \sim Pois(E_{ij}\theta_{ij}), \quad i = 1, ..., n, \quad j = 1, ..., J \tag{2.47}$$

where $E_{ij} = n_{ij}a_j$, n_{ij} is the population in ith region and jth group and a_j is a fixed age effect.

Decomposing region by age group we have:

$$log(\theta_{ij}) = m_i + w_{ij}, \quad i = 1, ..., n, \quad j = 1, ..., J \tag{2.48}$$

where m_i represents the structured and unstructured heterogeneity, $\underline{m} \sim N(\underline{0}, \underline{D}_u)$, $\underline{D}_u = \sigma_m^2(\phi \underline{Q}^{-1} + (1-\phi)\underline{I}_u)$, $\underline{Q}$ is an $n \times n$ matrix determined by the neighbourhood structure, $\underline{I}_u$ is the $n \times n$ identity matrix, and ϕ is the relative weight between structured and unstructured variation ($\phi = 1$ means no unstructured heterogeneity, $\phi = 0$ means no structured heterogeneity). The distribution of $m_i|\underline{m}_{-i}$ is taken as $N(\bar{m}_{\delta_i}, \frac{\sigma_m^2}{\delta_i})$, where $\underline{m}_{-i}$ is the set of the random effects excluding the ith one, and $\bar{m}_{\delta_i}$ is the mean of the random effects corresponding to the δ_i regions in the neighbourhood of region i. The term $w_{ij} \sim N(0, \sigma_w^2)$ is the age-region interaction term independent from m_i.

In terms of a generalised linear mixed model (2.47) and (2.48) can be expressed generally as

$$E(\underline{O}|\underline{b}) = \underline{\mu}^b = g^{-1}(\underline{\alpha} + (\underline{X})(\underline{\beta}) + (\underline{Z})(\underline{b})) \tag{2.49}$$

where $\underline{\alpha}$ is the offset parameter vector given here by the $\log(n_{ij})$, $\underline{b}$ is a vector of random effects given by $(\underline{u}^T, \underline{w}^T)^T$, $(\underline{Z})(\underline{b}) = \underline{Z}_1\underline{u} + \underline{Z}_2\underline{w}$, where $\underline{Z}_1$ and $\underline{Z}_2$ are $IJ \times J$ and $IJ \times IJ$ design matrices respectively and $\underline{Z}_2$ is an identity matrix. Here $\underline{\beta}$ is the vector of fixed effects and is given by $\underline{\beta} = (log(a_{ij}))$, $i = 1, ..., n$, $j = 1, ..., J$ and $\underline{X}$ is the corresponding design matrix ($IJ \times J$). The function $g^{-1}(\cdot)$ is the inverse of the link function, so the linear predictor is $\eta = g(\underline{\mu}^b)$. Thus for the log-linear predictor $\eta = log(\underline{\mu}^b)$, g^{-1} is the exponential function, and $\eta_{ij} = log(n_{ij}) + log(a_j) + m_i + w_{ij}$.

The integrated quasi-likelihood function is given by (Dean *et al.*, 2004)

$$|\underline{D}|^{-\frac{1}{2}} \int exp[-\frac{1}{2}\sum_{i,j} d_{ij}(O_{ij}, \mu_{ij}^b) - \frac{1}{2}\underline{b}^T \underline{D}^{-1}\underline{b}]d\underline{b} \tag{2.50}$$

where $d(O, \mu) = -2\int_O^\mu \frac{O-u}{u}du$, $\text{var}(\underline{b}) = \underline{D}$, where

$$\underline{D} = \begin{pmatrix} \underline{D}_u & \underline{0} \\ \underline{0} & \sigma_w^2 I_w \end{pmatrix}$$

I_w is the identity matrix of dimension IJ, and the covariance matrix $\underline{D}$ depends upon an unknown vector $\underline{\xi}$ of variance components. The PQL is obtained by taking a quadratic expansion of the exponent in (2.45) about its maximising value before integration.

To use iterated weighted least squares to estimate the fixed effects, Breslow and Clayton (1993) define the response $\underline{Y}$ to be $Y_{ij} = \eta_{ij} - \alpha + (O_{ij} - \mu_{ij})g^{'}(\mu_{ij})$, where $g^{'}(\mu_{ij}) = \frac{1}{\mu_{ij}}$, $i = 1, ..., n,\ j = 1, ..., J$. The associated normal model is

$$\underline{Y} = (\underline{X})(\underline{\beta}) + (\underline{Z})(\underline{b}) + \varepsilon, \tag{2.51}$$

where $\varepsilon \sim N(\underline{0}, \underline{W}^{-1})$, $\underline{W} = diag[var(O_{ij}|\underline{b})[g^{'}(\mu_{ij})]^2]^{-1}$, with $\hat{\beta} = (\underline{X}^T\underline{V}^{-1}\underline{X})^{-1}\underline{X}^T\underline{V}^{-1}\underline{Y}$, where $(\underline{X}^T\underline{V}^{-1}\underline{X})^{-1}$ is the estimated asymptotic variance and $\underline{V} = \underline{W}^{-1}+(\underline{Z})(\underline{D})(\underline{Z}^T)$. Here $\underline{V} = \underline{W}^{-1} + (\underline{Z}_1)(\underline{D}_u)(\underline{Z}_1^T) + \sigma_w^2\underline{I}_w$ and $W = diag(\mu_{ij})$. The random effects $\underline{b}$ are estimated as empirical Bayes estimates of the posterior mean, and are given by $\hat{\underline{b}} = (\underline{D})(\underline{Z}^T)\underline{V}^{-1}(\underline{Y} - \underline{X}\hat{\underline{\beta}})$.

The REML equations (Harville, 1977) below (Dean et al., 2004) are used to estimate variance components,

$$\frac{1}{2}[(\underline{Y} - \underline{X}\hat{\underline{\beta}})^T\underline{V}^{-1}\frac{\partial \underline{V}}{\partial \underline{\xi}_r}\underline{V}^{-1}(\underline{Y} - \underline{X}\hat{\underline{\beta}}) - tr(\underline{P}\frac{\partial \underline{V}}{\partial \underline{\xi}_r})] = 0, \quad r = 1, 2, 3, \tag{2.52}$$

where $\underline{P} = \underline{V}^{-\frac{1}{2}}(\underline{I} - \underline{H})\underline{V}^{-\frac{1}{2}}$, and $\underline{H} = \underline{V}^{-\frac{1}{2}}\underline{X}(\underline{X}^T\underline{V}^{-1}\underline{X})^{-1}\underline{X}^T\underline{V}^{-\frac{1}{2}}$ is the hat matrix.

The asymptotic variance of $\hat{\underline{\xi}}$ is ζ^{-1}, with the components of ζ given by
$\zeta_{rs} = \frac{1}{2}tr[\underline{P}(\frac{\partial \underline{V}}{\partial \underline{\xi}_r})\underline{P}(\frac{\partial \underline{V}}{\partial \underline{\xi}_s})], \quad r, s = 1, 2, 3.$

Given the initial estimates and initially fixing the variance components, PQL firstly solves for the $(\hat{\underline{\beta}}, \hat{\underline{b}})$ using the equations above, then the variance components are then updated and the process is repeated iteratively until convergence of both the mean and variance parameters.

2.6 Markov Chain Monte Carlo Methods

In Bayesian disease mapping when it is not possible to obtain the estimated parameters directly from the posterior distribution (due to the complexity of the models), posterior sampling algorithms referred to as Markov Chain Monte Carlo Methods (MCMC) are used. These methods use iterative simulation to sample parameters from a distribution that becomes closer and closer to the posterior distribution, say $g(\underline{\theta}|\underline{O})$, where $\underline{\theta} = (\theta_1, \theta_2, ...\theta_n)$ and $\underline{O} = (O_1, O_2, ..., O_n)$ are the vectors of relative risks to be estimated and observed counts respectively ,(n is the number of regions). Some of the many authors who discuss these methods include Robert and Casella (1999), Gilks *et*

al. (1996), and Casella and George (1992).

The distribution of the sampled parameters depends on the last value sampled, hence the samples form a *Markov chain*. A Markov chain is a (time) sequence of random variables, say $\underline{\theta}_0, \underline{\theta}_1, ...$, such that the next state, $\underline{\theta}_{t+1}$, $t \geq 0$, is sampled from a distribution $P(\underline{\theta}_{t+1}|\underline{\theta}_t)$ which does not depend on the history of the chain $(\underline{\theta}_0, \underline{\theta}_1, ..., \underline{\theta}_{t-1})$ but only on the current state $\underline{\theta}_t$. $P(\underline{\theta}_{t+1}|\underline{\theta}_t)$ is referred to as the *transition kernel* of the chain (Gilks *et al.*, 1996). If the chain is run for a long time, then it will eventually forget its initial state and will converge to a stationary distribution which does not depend on time t and the starting state $\underline{\theta}_0$. As t increases, the samples obtained are dependent samples from the stationary equilibrium distribution.

The Markov chain has to be constructed in such a way that the stationary distribution is the required posterior distribution. This requires constructing the correct transition probabilities, to obtain the posterior distribution $g(\underline{\theta}|\underline{O})$ as the equilibrium distribution. The convergence of the chain to a stationary distribution, which is assumed to be the posterior distribution, should be assessed (see below).

The following algorithms are used for the construction of the Markov chain:

1. the Metropolis algorithm, and (its extension) the Metropolis-Hastings algorithm, and more commonly

2. the Gibbs Sampler algorithm.

2.6.1 Metropolis and Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm (Hastings, 1970) is a generalisation of a Metropolis algorithm (Metropolis *et al.*, 1953). For the Metropolis-Hastings algorithm at each time t, the next state $\underline{\theta}_{t+1}$ is obtained by sampling $\underline{\theta}^{'}$ from a *proposal distribution* $q(\cdot|\underline{\theta}_t)$ which may depend on $\underline{\theta}_t$. The proposal function must be an irreducible and aperiodic transition function (Tierney, 1995; Roberts, 1995) for suitable convergence. Different choices of proposal function may be used.

The proposal $\underline{\theta}^{'}$ is accepted with probability $\alpha(\underline{\theta}_t, \underline{\theta}^{'})$, where

$$\alpha(\underline{\theta}, \underline{\theta}^{'}) = min\{1, \frac{g(\underline{\theta}^{'}|\underline{O})q(\underline{\theta}|\underline{\theta}^{'})}{g(\underline{\theta}|\underline{O})q(\underline{\theta}^{'}|\underline{\theta})}\}. \tag{2.53}$$

If $\underline{\theta}^{'}$ is accepted, then the next state becomes $\underline{\theta}_{t+1} = \underline{\theta}^{'}$, otherwise the chain does not move, i.e. $\underline{\theta}_{t+1} = \underline{\theta}_t$.

The Metropolis algorithm considers only proposal distributions which are symmetric, having the form $q(\underline{\theta}|\underline{\theta}^{'}) = q(\underline{\theta}^{'}|\underline{\theta})$. The acceptance probability in (2.53) reduces to $\alpha(\underline{\theta}, \underline{\theta}^{'}) = min\{1, \frac{g(\underline{\theta}^{'}|O)}{g(\underline{\theta}|O)}\}$.

2.6.2 Gibbs Sampling

Gibbs sampling is a special and very popular case of Metropolis-Hastings, proposed by Geman and Geman (1984). The proposal distribution is generated from the conditional distribution of θ_i given all other $\theta_j, i = 1, ..., n, \quad j = 1, ..., n, \quad j \neq i$, and the proposed value is accepted with probability 1.

Let $q(\theta_j|\theta_j^{'}) = p(\theta_j|\theta_{-j}^{t-1})$ if $\theta_{-j} = \theta_{-j}^{t-1}$, and zero otherwise, where $p(\theta_j|\theta_{-j}^{t-1})$ is the conditional distribution of θ_j given all other θ values (denoted θ_{-j}) at time $t-1$, $j = 1, ..., n$. Then

$$\frac{q(\underline{\theta}|\underline{\theta}^{'})}{q(\underline{\theta}^{'}|\underline{\theta})} = \frac{g(\underline{\theta}^{'}|O)}{g(\underline{\theta}|O)},$$

hence (2.53) gives $\alpha(\underline{\theta}, \underline{\theta}^{'}) = 1$.

2.6.3 Metropolis-Hastings Algorithm versus Gibbs Sampling

There are disadvantages and advantages associated with each algorithm. The Gibbs sampling requires evaluation of a conditional distribution, and at each iteration it samples a new single value for each $\theta_i, i = 1, ..., n$, and does not provide block updates of parameters. The Metropolis-Hastings algorithm does not need to evaluate a conditional distribution, but it does not guarantee the acceptance of a new value, and it provides block updates of parameters (Lawson *et al.*, 2003).

2.6.4 Assessing Convergence

When using MCMC methods, the convergence of the Markov chain to a posterior distribution has to be assessed. Robert and Cassela (1999) and Chen *et al.* (2000) review methods to assess this convergence, by checking distributional properties of the samples generated. The chain has to be run for an initial *burn-in-period* (period before convergence is reached) until it can be assumed to have reached convergence.

After convergence, the chains are run to produce more accurate posterior estimates. The burn-in estimates are usually discarded and not used in calculating the estimate of the parameter(s) of interest.

The accuracy of the estimates can be assessed using the Monte Carlo standard error of the mean. This is the standard deviation of the difference between the mean of the sampled values and the true posterior means. A rule of thumb is that the Monte Carlo standard error (MCSE=$\frac{Sd}{n}$), where Sd is the standard deviation of the chain values and n is the number of iterations, should be less than 5% of the standard deviation of the parameter estimate (Lawson *et al.*, 2003).

When running a *single chain*, convergence can be assessed by methods which look for stabilisation of the posterior probability in a time series. Brooks-Draper and Raftery-Lewis diagnostics can also be used (Lawson *et al.*, 2003). The Raftery-Lewis single chain diagnostic estimates the number of iterations required to estimate accurate quantiles, while the Brooks-Draper diagnostic estimates the number of iterations needed to quote the mean estimate to a given number of significant figures with accuracy. These diagnostics are used in packages such as MLwiN. Trace plots simply plot the value of the estimated parameter across iterations.

Single chain methods can be used to assess multiple chains but there are also methods which can only be used for multiple chains. Gelman and Rubin (1992), Brooks and Gelman (1998) and Robert and Casella (1999) proposed the use of the Gelman-Rubin statistic, used in packages such as WinBUGS. This statistic is based on running parallel chains starting from different values and computes the ratio of between chain runs to within chain runs. The output is a plot where the width of the central 80% interval of the pooled runs is given by a green line, the blue line gives the average width of the 80% interval of the within runs, and the ratio of the pooled/within is a red line (Lawson *et al.*, 2003), (See Section 3.2.1 for an example). Convergence is said to have been achieved if the pooled/within ratio has converged to 1 and both the pooled and within interval lines have converged to a stable value.

There is a debate about whether to run one single chain for a long time or multiple chains with different starting points. However the multiple chains have an advantage as they provide evidence for convergence across different subspaces (Lawson *et al.*,

2003).

2.6.5 Assessing Goodness-of-fit

In Bayesian and hierarchical modelling (Lawson *et al.*, 2003), the Bayesian Information Criterion (BIC) is widely used to help in model selection, and can be estimated from the output of an MCMC algorithm. For a model with log-likelihood $l(\underline{\theta})$, let p be the number of linearly independent parameters, n be the number of data points and

$$\hat{l}(\underline{\theta}) = \frac{1}{K}\sum_{i=1}^{K} l(\underline{\theta}_i)$$

be the averaged log-likelihood over K posterior samples $\underline{\theta}_i$ of $\underline{\theta}$, $i = 1, ..., k$. Then the Bayesian Information Criterion value will be given by

$$2\hat{l}(\underline{\theta}) - p\ \ln(n).$$

Given two models, BIC attempts to identify the "true" model, and the model with the lowest value of BIC is the preferred model.

Spiegelhalter *et al.* (2002) have proposed another model choice criterion, called the Deviance Information Criterion (DIC). This is based on the principle that DIC is given by goodness of fit and complexity of the model. Following Spiegelhalter *et al.* (2002), the goodness of fit is measured through deviance and is given by

$$D(\underline{\theta}) = -2logL(\underline{O}|\underline{\theta})$$

where $\underline{O}$ are the observed data. The complexity is measured by the estimate of the effective number of parameters as

$$p_D = E_{\underline{\theta}|\underline{O}}(D) - D(E_{\underline{\theta}|\underline{O}}(\underline{\theta})) = \overline{D} - D(\overline{\underline{\theta}}),$$

which is the posterior mean deviance minus the deviance evaluated at the posterior mean of the parameter. Thus, the DIC is then given by

$$DIC = D(\overline{\underline{\theta}}) + 2p_D = D(\overline{\underline{\theta}}) + 2(\overline{D} - D(\overline{\underline{\theta}})) = \overline{D} + p_D. \tag{2.54}$$

The model with the smallest DIC is estimated to be the best model that fits a set of data.

Analysis of residuals also helps in the assessment of model goodness of fit. A residual

is defined as a standardised difference between the observed value and the fitted model value. Standardisation is based on a measure of the variability of the difference between the two values (Lawson *et al.*, 2003). The ith unstandardised residual is defined as

$$r_i = O_i - \hat{O}_i \tag{2.55}$$

and the standardised residual as

$$r_i = \frac{(O_i - \hat{O}_i)}{\sqrt{var(O_i - \hat{O}_i)}}, \quad i = 1, ..., n. \tag{2.56}$$

When the observed counts are Poisson distributed, we have

$$r_i = \frac{(O_i - E_i\hat{\theta}_i)}{\sqrt{Var(O_i - E_i\hat{\theta}_i)}}. \tag{2.57}$$

The Pearson chi-squared residual sum of squares (RSS) can also be used to assess goodness of fit. This is the sum of squares of the residuals in (2.57).

The following Bayesian residual was described by Carlin and Louis (1996), as

$$r_i = O_i - \frac{1}{K}\sum_{k=1}^{K} E(O_i|\theta_i^{(k)}), \tag{2.58}$$

where $E(O_i|\theta_i)$ is the expected value from the posterior predictive distribution and $\theta_i^{(k)}$ is a set of k parameter values sampled from the posterior distribution by the MCMC algorithm. When a constant region risk rate is assumed, then a residual which averages the posterior samples will be given by

$$r_i = O_i - \frac{1}{K}\sum_{k=1}^{K} E_i\theta_i^{(k)}. \tag{2.59}$$

Spiegelhalter *et al.* (1996) suggested obtaining (2.56) at each iteration of a posterior sampler and averaging over the converged sample.

2.7 Conclusion

In this chapter we have outlined and discussed the models which are used in disease mapping of count data, together with the different methods that can be used to estimate parameters in these models. The models discussed are those used when data are available at one time point and different time points (space-time models), and also

models used in mapping two or more diseases.

Some models like the Poisson-Gamma and NPML models do not incorporate spatial correlation, while models like the lognormal and logistic models and TNPML do easily allow for spatial correlation. As neighbouring regions tends to have similar disease counts, it is better to take into account the spatial correlation. Fitting a non-spatial model will not take this into account, therefore, spatial models may be preferred over non-spatial models. Also, inclusion of covariates in a model will help with accounting for unobserved effects, if they exist, thus models that allow for this may be preferred.

It is important to realise that in trying to circumvent the problems of mapping the SMRs, we also should not over-smooth the relative risks, as this may result in wrong interpretation of the distribution of the disease in question. Considering jumps in the relative risks surface is of great importance also. The Poisson-Gamma and BYM models are based on smoothing methods, while the non-parametric mixture models NPML and TNPML detect discontinuities in the map. Thus, it may be helpful to put these two concepts together, as in the model proposed by Lawson and Clark (2002) (described in Section 4.2.3).

As for estimation of parameters, for empirical Bayesian estimation PQL is considered to be straightforward and simple. Breslow and Clayton (1993) suggest that PQL is a very useful approximation method except for the analysis of very small counts. The PQL method was compared to maximum likelihood estimation by Leroux (2000), who recommends maximum likelihood estimators only for very large sample sizes.

Even though full Bayesian estimation is used widely due to the availability of software that can handle such complexity, Leyland and Davies (2005) argue that empirical Bayesian estimation still has its place.

In the next chapter, using susceptibility to measles data described in Section 1.5, we choose to fit models that are commonly used to these data. The Poisson-Gamma and lognormal models are fitted using empirical and full Bayesian estimation, and the two estimation methods are compared. The log-normal and logistic models are compared to see which one will fit the data better and are used to compare maps over time. The space-time model of Waller *et al.* (1997) is also fitted to the data and used to compare maps over time. Inclusion of covariates is also considered.

Chapter 3

Analysis of Measles Data

3.1 Introduction

Measles, mumps and rubella vaccine (MMR) uptake decreased since 1998 (Friederichs *et al.*, 2006) as a result of a postulated association with autism and bowel disease claimed by Wakefield *et al.* (1998). The association has since been disproved by some authors, including Demicheli *et al.* (2005) and Baird *et al.* (2008). There are public health concerns with falling MMR vaccine rates and so it is of interest to examine trends in MMR uptake over time and also regional variation in uptake rates. Recently, there have been confirmed cases of measles in England and Wales, with cases in 2008 nearly 40% higher than in 2007 (www.hpa.org.uk/hpr/archives/2009/hpr0509.pdf).

In this chapter we analyse the measles susceptibility data described in Section 1.5. Increases in measles susceptibility estimates were predicted from the decreases in MMR uptake (Friederichs *et al.*, 2006). For each of the pre-school and primary 1 and 2 groups, the analysis is done at both district and postcode sector level. Children in the primary 1 and 2 school group should have received an MMR booster, and so are expected to have lower susceptibility than pre-school children. The analysis will enable us to compare susceptibility for these two groups.

For each group, pre-school and primary 1 and 2 children, susceptibility to measles is compared over time for the period 2000-2005. This analysis will be able to show time trends in the spatial distribution of measles susceptibility, and thus inform if there has been an increase or decrease over time. As susceptibility maps will be produced here, the spatial distribution of measles susceptibility will be visually compared over

time, and the areas in which susceptibility has changed or not changed over time will be able to be seen. If susceptibility has changed in some areas, it should be possible to see whether different areas have changed by the same amount or not. As it is important for public health to know which areas have high susceptibility, the aim here is also to see if these areas can be identified, as this may help with targeted vaccination campaigns. Here visual comparison of maps is used but in Chapter 7 the descriptive methods developed for this purpose will be used.

The census area level data, described in Section 1.5, is also used here to see if any of the variables can be used to predict which areas have high susceptibility. Friederichs *et al.* (2006) used the Scotland measles data, and found that the decrease in MMR uptake was associated with increase in deprivation. Previous studies in England and Wales have highlighted high educational attainment, deprivation and population density as being negatively associated with MMR vaccine uptake (Wright and Polack, 2006). A recent study for the whole of the UK has also shown that a lower uptake of the MMR vaccine is associated with children who live in households with other children, i.e. being the three or more in the family, lone parent households, households with highly educated mother (AS/A level, degree or above), households with unemployed or self-employed mother, and households with mother under 20 years or over 34 years when she gave birth to cohort child (Pearce *et al.*, 2008).

Friederichs *et al.* (2006), analysed measles data for Scotland pre-school and primary 1 and 2 school children, to assess MMR uptake across Scotland from its introduction in 1988 to 2005. A linear regression model was used to analyse the data and to examine effects of deprivation on susceptibility to measles. The study we do here is different from Friederichs *et al.* (2006) study in the sense that it is the first study to use the spatial model, space-time model and spatial ecological model, to analyse Scottish measles susceptibility data. Also maps of susceptibility to measles are produced here, thus allowing the spatial distribution of measles susceptibility to be visualised for each year and compared over time, which has not been done before. Friederichs *et al.* (2006) used only the Carstairs index to examine effects of deprivation, whereas here the four components of deprivation are used, i.e. percentage of people in households with no car, percentage of people in households with low social class, percentage of people in overcrowded households and percentage of unemployed males. These com-

ponents are kept separate in the analysis to check if they had similar effects. Also other variables are included: percentage of children aged between 0-4, percentage of people born in other European Union countries (other than UK), percentage of people born elsewhere (other than EU), percentage of lone parent households, and percentage of people working in agriculture, percentage of people with no qualifications, and percentage of people with high qualifications (first degree, higher degree and professional qualifications).

In the next section, the software used here is described. Using one data set of measles susceptibility, an illustration of how the model fitting is done in disease mapping is presented based on some of the models reviewed in Chapter 2 (which are considered to be commonly used). This will also allow a comparison of empirical and full Bayesian modelling. As our interest is in fitting a full Bayesian model to the whole data set, the full Bayesian lognormal and logistic models are compared to see which one will fit the data better, and the best fitting model will then be used to fit the rest of the data.

3.2 Model fitting

The software packages used here are R (http://www.r-project.org) for data management, empirical Bayesian modelling, and selection of census variables and WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs), for full Bayesian modelling. The WinBUGS codes for the models fitted in this chapter are based on the ones given in Lawson *et al.* (2003), and given in Appendix D. WinBUGS uses MCMC methods to sample from the posterior distribution and allows mapping of the fitted parameters, including relative risks and residuals, using a spatial module called GeoBUGS.

Firstly, to see how some of the models fit to the data, the Poisson-Gamma model, lognormal model of Besag *et al.* (1991) and the logistic model (Section 3.2.2) were fitted to the pre-school susceptibility to measles data for 1999 at district level. This also will allow the comparison of these models in fitting these data. The empirical and full Bayesian modelling based on the Poisson-Gamma and lognormal model are compared. The full Bayesian lognormal model and full Bayesian logistic model were also compared to select a model that fits the data best. In R we use the package DCluster, which contains the functions *empbaysmooth* and *lognormalEB* which pro-

duce empirical Bayesian estimates (EB) for a Poisson-Gamma and log-normal model respectively. The estimates for the two models are all based on the models proposed by Clayton and Kaldor (1987) and described in the previous chapter.

3.2.1 Poisson-Gamma and Log-normal Models

Firstly, the ratio of observed counts/expected counts and their standard errors were obtained, for the raw susceptibility data for 1999 pre-school, for 56 districts. The number of expected children susceptible to measles was obtained as

$$E_i = N_i * \frac{\sum O_i}{\sum N_i}, \quad i = 1, 2, ..., n,$$

where N_i and O_i are total number of children and observed number of children (product of total number and proportion susceptible) susceptible to measles in the ith district respectively and n is the number of districts.

The empirical and full Bayesian Poisson-Gamma and log-normal (without the spatial term) models were fitted. In the full Bayesian modelling, two chains with different initial values were run. For the Poisson-Gamma model the first 2000 iterations were discarded as burn-in and each chain was run for a further 2000 iterations. The parameters monitored were α and β of the Gamma distribution and the relative risks θ_i. The parameters α and β were assigned an exponential hyperprior. For the log-normal model, the first 4000 iterations were discarded as burn-in and chains were run for a further 4000 iterations. The parameters monitored were the relative risk parameters (θ_i), the overall mean (α) and the inverse variance of the uncorrelated heterogeneity, tau.v (τ_v). When specifying the prior distributions for the parameters in the log-normal model, Gamma(0.1, 0.001) priors were used for the inverse variance (τ_v). This was assigned with an understanding that the hyperprior will give a large variance, therefore it will be relatively flat over a large range, thus will have little influence on the likelihood of the data.

Figures 3.1 and 3.2 show examples of how convergence was monitored for α and β in the Poisson-Gamma model. The trace plots (Figure 3.1) of the estimated parameters show that the parameter estimates have converged, i.e. simulation has stabilised. The Gelman and Rubin plots (Figure 3.2) show that the parameters converged after 2000 iterations, as indicated by the pooled/within chain variance ratio (red) line having

converged to 1 and both the pooled (green) and within (blue) interval widths lines having converged to a stable value.

Table 3.1 shows the minimum, maximum, range and mean of the ratios of observed/expected counts, and the empirical and full Bayesian estimates of relative risks obtained from the Poisson-Gamma and log-normal models. Smoothing with both models increased the minimum (from 0.512), reduced the maximum (from 2.342) and reduced the range of the ratio of observed/expected (from 1.830), as would be expected. For the Poisson-Gamma model, the empirical Bayesian range is 1.189 and the full Bayesian range is 1.161. For the log-normal model the range for the empirical estimates is 1.280 and for the full Bayesian estimates the range is 1.356.

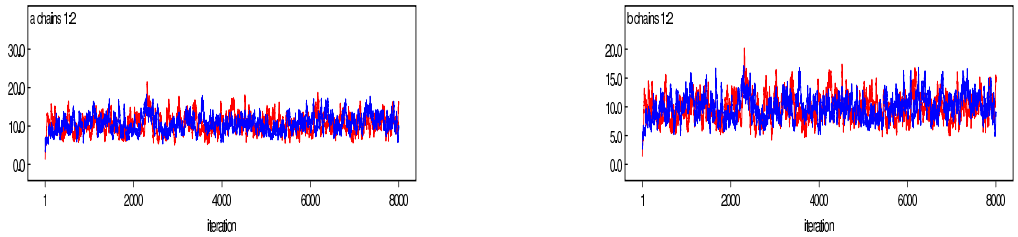

Figure 3.1: Trace plots diagnostic for α (a chains) and β (b chains), where red and blue are the traces for chains 1 and 2.

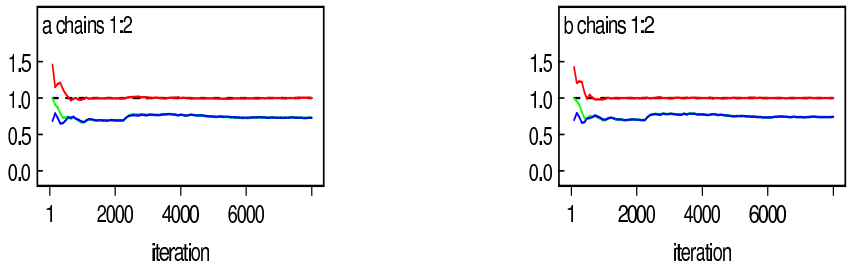

Figure 3.2: Plots of Gelman and Rubin convergence diagnostic for α (a chains) and β (b chains), where the pooled/within chain variance ratio line is red, pooled chain variance line is green and within chain variance line is blue.

The range of susceptibility rates produced by the Poisson-Gamma model is slightly greater for empirical Bayes than full Bayesian estimation, while for the log-normal

model the range is higher for the full Bayesian method than the empirical Bayesian method. Also, the mean of the smoothed rates produced by each model is smaller than the mean of the ratios (observed/expected). Therefore, smoothing the rates has helped remove random variability from the data due to small observed and expected counts.

Figure 3.3 shows maps of the ratios (observed/expected) and of their standard errors (top), maps of the empirical and full Bayesian estimates from the Poisson-Gamma model (middle), and empirical and full Bayesian estimates from the log-normal model (bottom). Cut-off points are chosen in such a way that all maps use the same ranges of values so that they are comparable. The darker colour in the key indicates a higher susceptibility rate, while lighter colour indicates a lower susceptibility rate. The standard error map shows that most regions have standard errors of 0.05-0.10 and these are mostly those with high susceptibility rates, and regions in the central area with low susceptibility rates have lower variation. The maps produced from the estimates obtained by smoothing are smoother, with some extreme rates removed. The maps produced by both empirical and full Bayesian methods have the same spatial pattern. It can be observed that high susceptibility is found in the central regions and some districts in the south. Table A.1 in Appendix A shows the ratio (observed/expected) and the smoothed susceptibility rates from fitting the four models. For most districts the empirical and full Bayesian models give the same or very close results.

Table 3.1 gives parameters for the Poisson-Gamma and lognormal models for empirical and full Bayesian models. For the Poisson-Gamma models, the values of the parameters are slightly larger for the full Bayesian than for the empirical Bayesian model. For the lognormal models, the values of the parameters are slightly larger for the empirical Bayesian model than for the full Bayesian model. The RSS values indicate that there is no difference between fitting the Poisson-Gamma and lognormal empirical Bayesian models, and there is no difference between fitting the Poisson-Gamma and lognormal full Bayesian models.

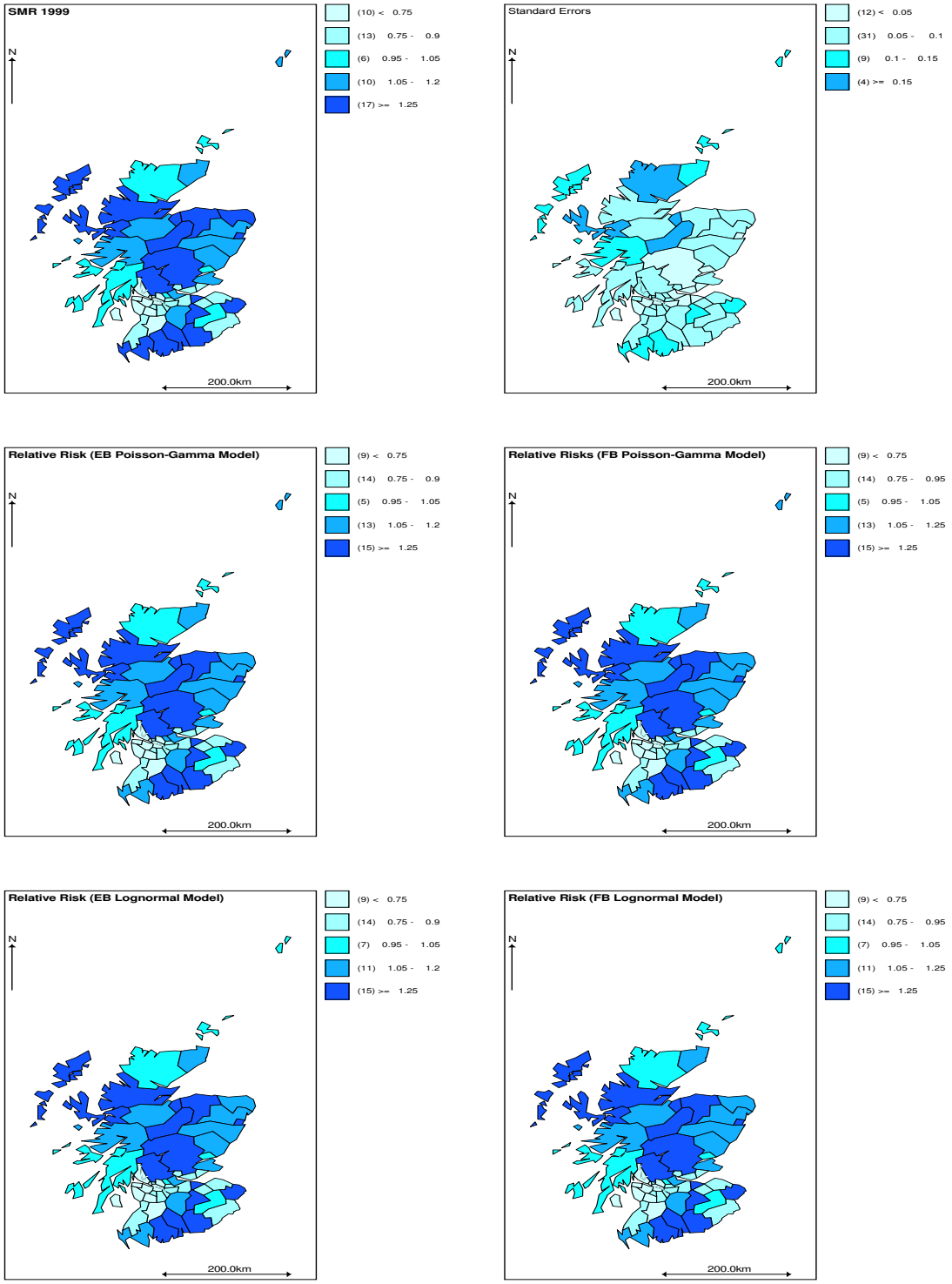

Figure 3.3: Maps of raw susceptibility to measles rates (observed/expected) (top left), standard errors (top right), empirical Bayesian (middle right) for Poisson-Gamma model and full Bayesian (middle left), and empirical Bayesian for log-normal (bottom right) and full Bayesian for log-normal (bottom left), for estimates of susceptibility to measles for pre-school 1999.

	SMR	PG(EB)	PG(FB)	LN(EB)	LN(FB)
Minimum	0.5115	0.5806	0.5822	0.5885	0.5745
Maximum	2.3420	1.7700	1.7430	1.8680	1.9300
Range	1.8305	1.1894	1.1608	1.2795	1.3555
Mean	1.1050	1.0770	1.076	1.0800	1.0950
a		9.86	10.73		
b		9.16	9.99		
α				0.0367	0.025
σ_v^2				0.098	0.102
RSS		55.06	59.26	55.00	59.27

Table 3.1: Table of minimum, maximum, range and mean of observed/expected ratios using the empirical (EB) and full Bayesian (FB) estimates of Poisson-Gamma (PG) and log-normal (LN) models for pre-school 1999.

However, the RSS indicate that the empirical models fit these data better than the full Bayesian model, we choose to pursue full Bayesian modelling as we have access to WinBugs which will enable us to fit the models and producing maps easily.

In our analysis we are interested in fitting a model with covariates, to see if they have any effects on susceptibility to measles, thus we choose to pursue the use of the lognormal model. Firstly in the next section the lognormal model is compared with the logistic model to see which of the two models will fit the data better.

The interest in this thesis is to use full Bayesian modelling to analyse the susceptibility to measles data, therefore the lognormal and logistic model are compared based on the full Bayesian analysis.

3.2.2 Comparing Log-normal and Logistic Models

The full Bayesian logistic model (2.8) in Appendix D was fitted to the 1999 pre-school data and compared with the full Bayesian log-normal model (2.7) to see which of the two models fits these data better. The overall goodness of fit measures used to compare the models are the residual sum of squares (RSS), and the residual maps.

RSS for both the models is almost the same, 60.58 (lognormal) and 60.68 (logistic), indicating that there is little difference in the way these models fits this measles susceptibility data. The maps of the standardised residuals in Figure 3.4 show that the

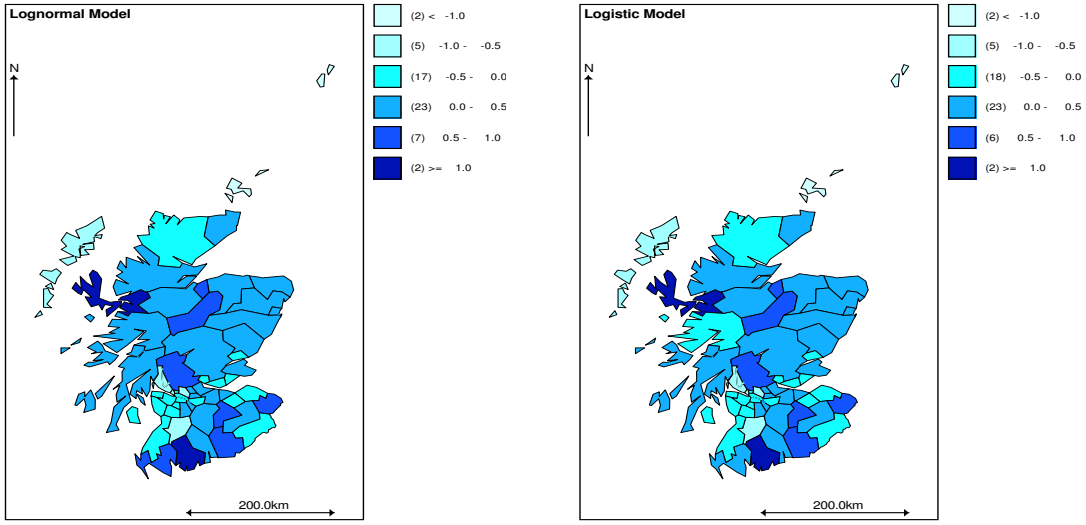

Figure 3.4: Maps of standardised residuals of the log-normal (left) and logistic (right) models for pre-school 1999.

two models fit the data very similarly, with only Lochaber district having a residual in a different class when comparing the two maps. The range of the residuals for the log-normal model is 2.570, with minimum -1.489 and maximum 1.081. The range of the residuals for the logistic model is 2.588 with minimum -1.489 and maximum 1.099. The DIC for lognormal model is 439.92 and for logistic model is 436.26. The change in DIC is -3.66, indicating that the logistic model might fit these data better than the lognormal model.

In the next section we choose to use the logistic model, especially since we have N_i and O_i available in our datasets. Also susceptibility p_i is not small in all areas, especially for pre-school, so assuming a Poisson distribution will not always be appropriate. The space-time model of Waller *et al.* (1997) is also used and ecological analysis is done using the logistic model.

These models were fitted in WinBUGS, and Appendix D gives the WinBugs code. For all the analyses convergence was checked using time series plots/traces and the Gelman and Rubin convergence diagnostic. When specifying the prior distributions for the parameters, Gamma(0.1, 0.001) priors were used for the inverse variances and noninformative N(0,1E-05) priors were used for the regression coefficients and for the intercept (i.e. a flat prior).

3.3 Analysis of Measles Data for 56 districts

In this section, susceptibility to measles data for pre-school and primary 1 and 2, for 2000-2005, is analysed at district level. The logistic model (2.8) is fitted separately to each of the 6 years in each group, and the space-time model (2.27) is also fitted to each group of the pre-school and primary 1 and 2 children. The analysis will enable us to compare measles susceptibility over time for each group and to compare susceptibility to measles of pre-school children to primary 1 and 2 school children. Also the difference between fitting a space-time model and a spatial model to each time point will be assessed.

3.3.1 Comparing Maps Over Time at District Level

The logistic model (2.8) and space-time model (2.27) were fitted to both the pre-school and primary 1 and 2 school data at district level. Convergence is different for each model and for pre-school and primary 1 and 2 school children, and even in each group, time points may have different burn-in periods. For the logistic model, on average 6,000 (pre-school) and 4,000 (primary 1 and 2) iterations were run before convergence was achieved and the parameter estimates are based upon a further 6,000 and 4,000 iterations for pre-school and primary 1 and 2 school children respectively. For the space-time model, a burn-in of 12,000 (pre-school) and 10,000 (primary 1 and 2) iterations was needed and estimates are based upon a further 12,000 and 10,000 iterations respectively.

For pre-school, Figures 3.5 and 3.6 show maps of susceptibility obtained by fitting the logistic model (2.8) to each time period for each of these groups, and the space-time model (2.27), respectively. The limits on the maps were chosen such that all the maps have the same groups, so that maps can be compared. For all the maps we tried to balance a 5th in each group over all 6 maps and retain interpretable boundaries. Looking at the logistic model (Figure 3.5), generally the maps show that susceptibility is increasing over time, with 2003 and 2004 having the highest number of districts with high susceptibility, and susceptibility decreased in 2005. In 2000 susceptibility is less than 15% in central and south Scotland and greater than 16% in the north and west of Scotland. The number of regions with susceptibility to measles of less than 15% increased in 2001; this includes regions in the north east.

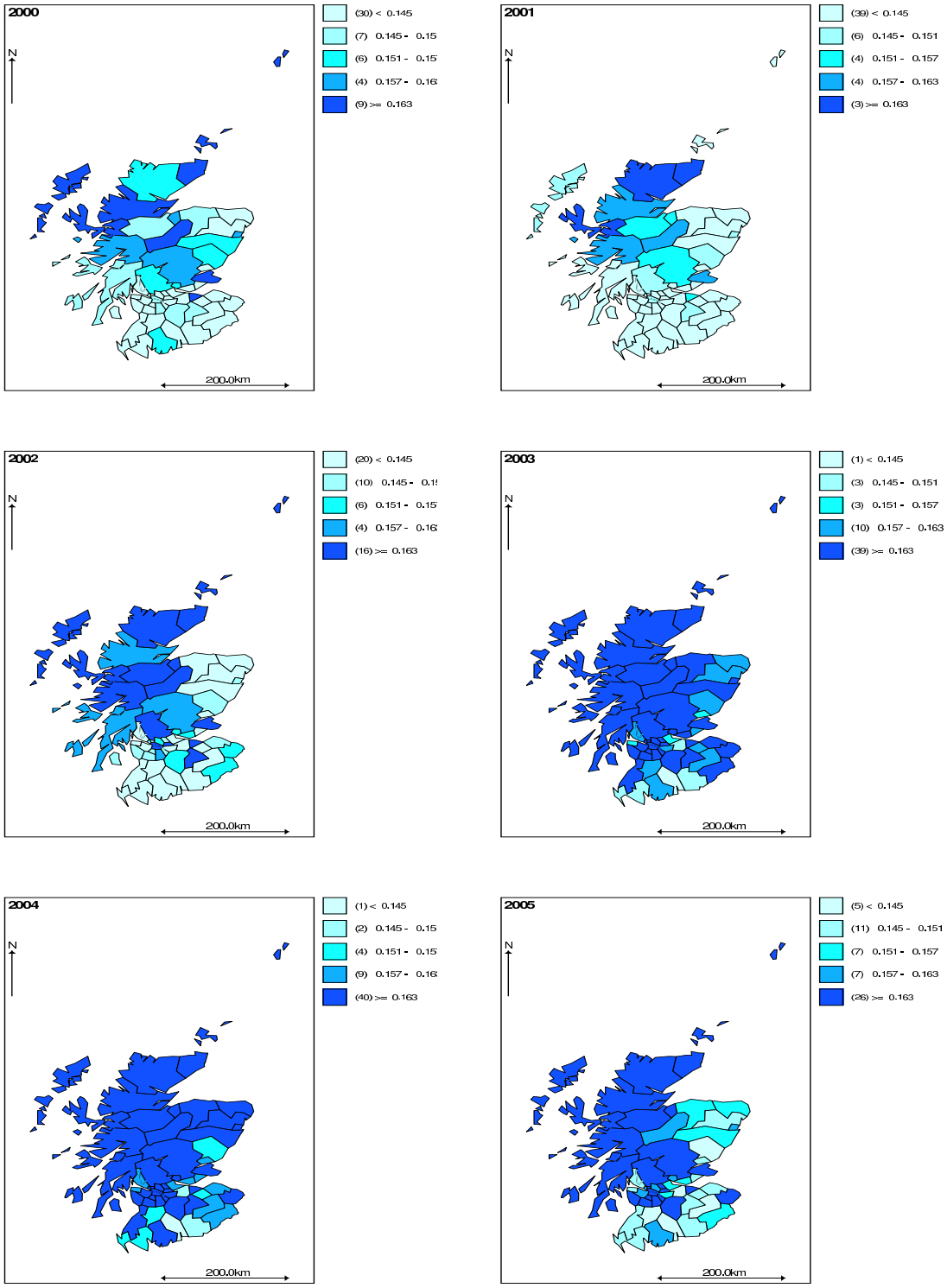

Figure 3.5: District level maps of estimated probabilities of pre-school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model).

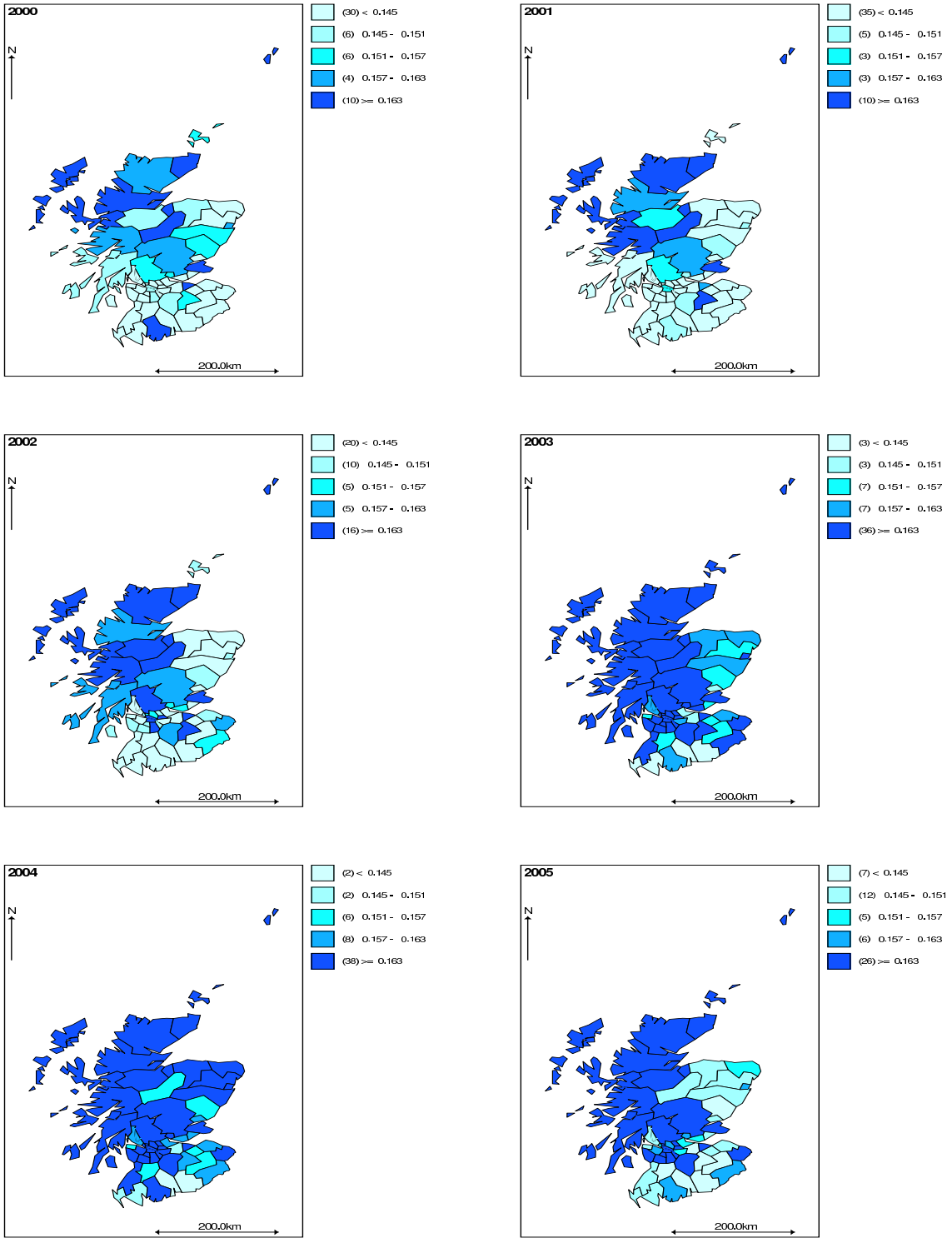

Figure 3.6: District level maps of estimated probabilities of pre-school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (space-time model).

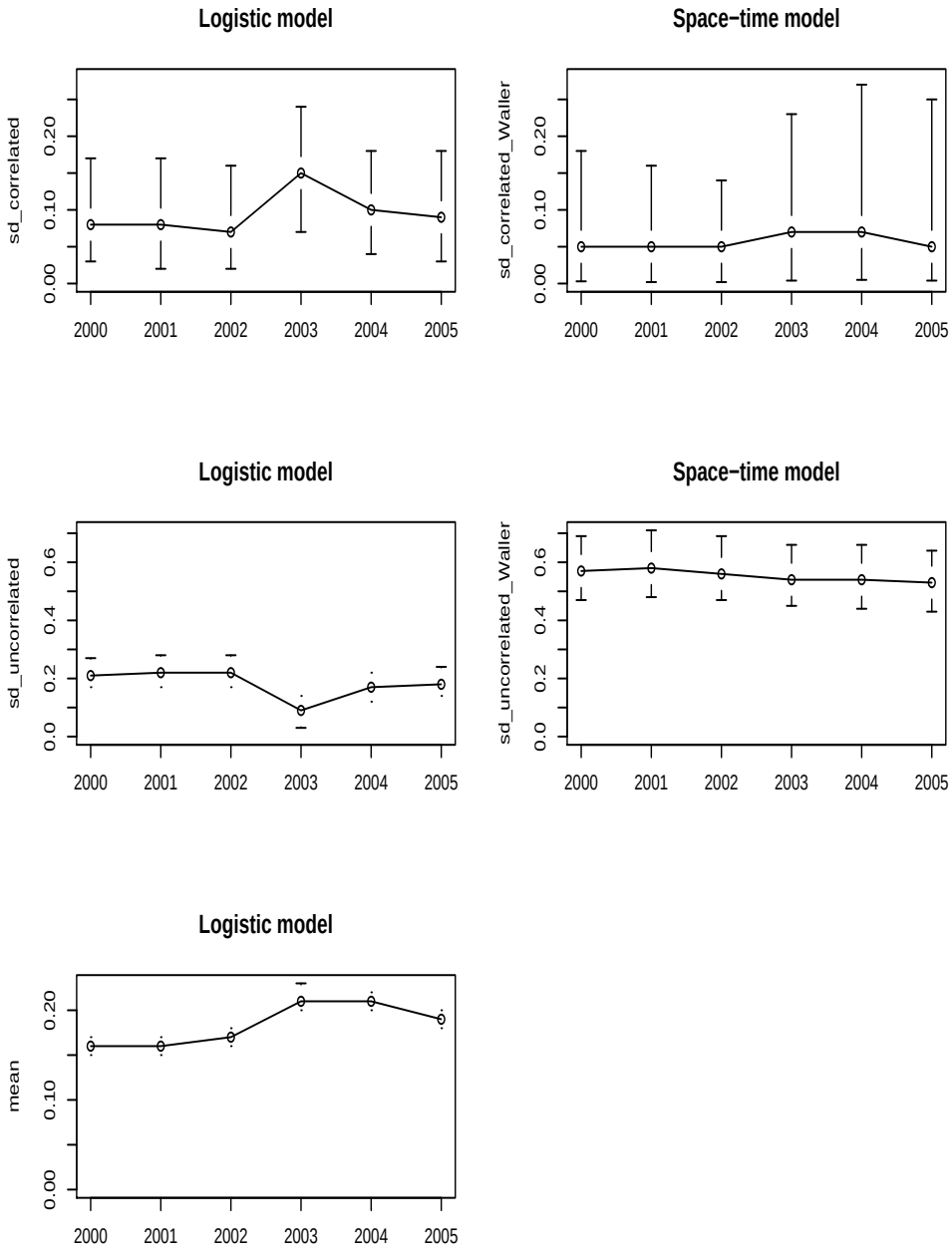

Figure 3.7: Plots of variability components and mean with 95% credible interval bars, for logistic and space-time Waller *et al.* (1997) models for pre-school children at district level, against years 2000-2005.

Year	Pre-School					
	Logistic model				Space-time model	
	α	proportion	σ_u	σ_v	σ_u	σ_v
2000	−1.83 (-1.89,-1.77)	0.16 (0.15,0.17))	0.08 (0.03,0.17)	0.21 (0.17,0.27)	0.05 (0.003,0.18)	0.57 (0.47,0.69)
2001	−1.86 (-1.92,-1.80)	0.16 (0.15,0.17)	0.08 (0.02,0.17)	0.22 (0.17,0.28)	0.05 (0.002,0.16)	0.58 (0.48,0.71)
2002	−1.78 (-1.84,-1.71)	0.17 (0.16,0.18)	0.07 (0.02,0.16)	0.22 (0.17,0.28)	0.05 (0.002,0.14)	0.56 (0.47,0.69)
2003	−1.56 (-1.60,-1.53)	0.21 (0.20,0.23)	0.15 (0.07,0.24)	0.09 (0.03,0.14)	0.07 (0.004,0.23)	0.54 (0.45,0.66)
2004	−1.57 (-1.62,-1.52)	0.21 (0.20,0.22)	0.10 (0.04,0.18)	0.17 (0.12,0.22)	0.07 (0.005, 0.27)	0.54 (0.44,0.66)
2005	−1.67 (-1.73,-1.62)	0.19 (0.18,0.20)	0.09 (0.03,0.18)	0.18 (0.14,0.24)	0.05 (0.004,0.25)	0.53 (0.43,0.64)
Year	Primary School					
	Logistic model				Space-time model	
	α	proportion	σ_u	σ_v	σ_u	σ_v
2000	−2.75 (-2.84,-2.66)	0.06 (0.06,0.07)	0.23 (0.09,0.39)	0.31 (0.23,0.40)	0.27 (0.05,0.45)	0.21 (0.13,0.31)
2001	−2.80 (-2.87,-2.71)	0.06 (0.06,0.07)	0.22 (0.10,0.35)	0.29 (0.21,0.38)	0.27 (0.14,0.42)	0.19 (0.11,0.28)
2002	−2.76 (-2.81,-2.70)	0.06 (0.06,0.07)	0.29 (0.17,0.44)	0.16 (0.07,0.25)	0.29 (0.16,0.45)	0.17 (0.05,0.27)
2003	−2.85 (-2.93,-2.78)	0.06 (0.05,0.06)	0.22 (0.11,0.36)	0.27 (0.19,0.35)	0.24 (0.11,0.40)	0.23 (0.16,0.32)
2004	−2.71 (-2.78,-2.63)	0.07 (0.06,0.08)	0.18 (0.04,0.32)	0.25 (0.18,0.33)	0.21 (0.09,0.35)	0.18 (0.10,0.25)
2005	−2.47 (-2.52,-2.41)	0.08 (0.08,0.09)	0.21 (0.10,0.35)	0.21(0.14,0.28)	0.21 (0.06,0.30)	0.32 (0.25,0.41)

Table 3.2: Pre-school and primary 1 and 2 posterior means with lower and upper credible intervals in the brackets, for overall mean level (α), proportion ($\frac{e^{\alpha}}{1+e^{\alpha}}$), standard deviations due to correlated heterogeneity (σ_u) and uncorrelated heterogeneity (σ_v) for logistic and space-time (Waller *et al.*, 1997) models for the 56 districts.

Thus susceptibility is lower in most regions in 2001 than 2000. In 2002, susceptibility started to increase. The number of regions with less than 15% susceptibility decreases, and pockets of increased measles susceptibility can be observed in the south and the Borders regions. There is a global increase in susceptibility in 2003, with susceptibility greater than 16% in the western regions, and in 2004 isolated regions with low susceptibility can be seen in the south and Borders. Susceptibility decreased in 2005 but only in the north east and the Borders.

High susceptibility is also observed in the urban areas of Glasgow, Edinburgh, Aberdeen and Dundee, especially from 2002 to 2005. As for the cities of Edinburgh and Aberdeen, high susceptibility has been seen from 2000. The space-time model (Figure 3.6) maps give similar interpretation to the logistic model maps (Figure 3.5), but in 2001 the space-time model map shows that there were more regions with high measles susceptibility (greater than 16%) than as shown by the logistic model. In 2004 and 2005, the space-time model shows that Badenoch and Strathspey district had a lower measles susceptibility than as shown by the logistic model. Such differences may occur

as the space-time model has only one α parameter, thus this model may give similar results to the logistic model when α does not vary over time, as in the case of primary 1 and 2 school children (see Figures 3.8 and 3.9).

Table 3.2 gives values of the model parameters and plots of these parameters are shown in Figure 3.7. For the logistic model and pre-school, the years 2000 to 2002 have similar overall proportion and structured standard deviations, and only a very slight increase in unstructured standard deviation, indicating a slight change in spatial variation. Spatial variation changed very significantly in 2003. This year had the smallest unstructured variation, indicating that the regions became more similar with high susceptibility and the highest structured standard deviation as clusters of higher rates increased. The logistic model shows that the overall proportion ranges from 0.16-0.21.

The credible intervals suggest that there was no real difference in overall proportion from 2000 to 2002, but the overall proportion increased from 2002 and decreased in 2005, with 2003 and 2004 being similar and having the highest overall proportions. This corresponds with what is shown by the maps, i.e the number of regions with high susceptibility increased over time, and is higher in 2003 and 2004 with a decrease in 2005. The structured standard deviation ranges from 0.07-0.15 (logistic model) and 0.05-0.07 (space-time model), and the unstructured standard deviation ranges from 0.09-0.22 (logistic model) and 0.53-0.58 (space-time model). For both models and both standard deviations, the credible intervals are wide. For the structured standard deviation, for both models the credible interval suggests no real differences over time. For unstructured standard deviation, for the space-time model the credible intervals suggest no real differences but for the logistic model the credible intervals suggest that the 2003 unstructured standard deviation was different from other years (lower).

With the exception of 2003, from both models unstructured standard deviations are larger than structured standard deviations, indicating that variation in susceptibility rates is more due to general differences between regions than local clustering.

For primary 1 and 2, Figures 3.8 and 3.9 show maps of susceptibility obtained by fitting the logistic model (2.8) to each time period for each of these groups, and the space-time model (2.27) respectively. The two models give similar maps. As expected,

each year has fewer regions with higher measles susceptibility and lower probabilities of children susceptible than in the pre-school group. Generally susceptibility increased most from 2004 to 2005. In 2000 susceptibility is less than 5% in central Scotland and there are isolated pockets of susceptibility greater than 9.5%. Table 3.3 and Figure 3.10 show overall proportion ranging from 0.06-0.08 (logistic model), unstructured standard deviation ranging from 0.16-0.31 for the logistic model and 0.17-0.32 for space-time model, and structured standard deviation ranging from 0.18-0.29 for logistic model and 0.21-0.29 for the space-time model. The two models give similar results. Year 2000 had the greatest unstructured variability, indicating more variation in susceptibility between regions than in other years.

In 2001, relative to 2000, susceptibility in northern regions over 9.5% tended to decrease. For both models unstructured standard deviation is slightly less than in 2000 but still indicating variation in susceptibility between regions. These two years have the same overall proportion, which is the smallest among all years, indicating in general for these two years susceptibility rates were lower compared to other years, and almost similar local standard deviations. In year 2002 the central regions still have susceptibility of less than 5%. There is a decrease in susceptibility in the southern, Borders and north east regions, with pockets of increased susceptibility in the north west and central regions. This year (2002) has the lowest unstructured standard deviation, indicating regions are more similar than in other years, and the highest local standard deviation, indicating that there is more clustering of susceptibility rates than in other years. In 2004 we observe increased susceptibility in the central regions, and in 2005 an increase in susceptibility can be observed with susceptibility greater than 9.5% in the west and north. Also here some of the urban regions, Aberdeen and Edinburgh, had high susceptibility rates for each year. An increased overall proportion in 2004 and 2005 indicated the increase in high susceptibility rates.

For primary 1 and 2, 2005 has highest unstructured variability for the space-time model, while for the logistic model 2000 had the highest unstructured variation. Credible intervals are wider for the standard deviation parameters and they do not suggest any major differences over time, unlike for pre-school 2003 (logistic model). Also structured and unstructured deviation are roughly equal, unlike for pre-school. There is very little change for α over time, though there is a suggestion of an increasing trend.

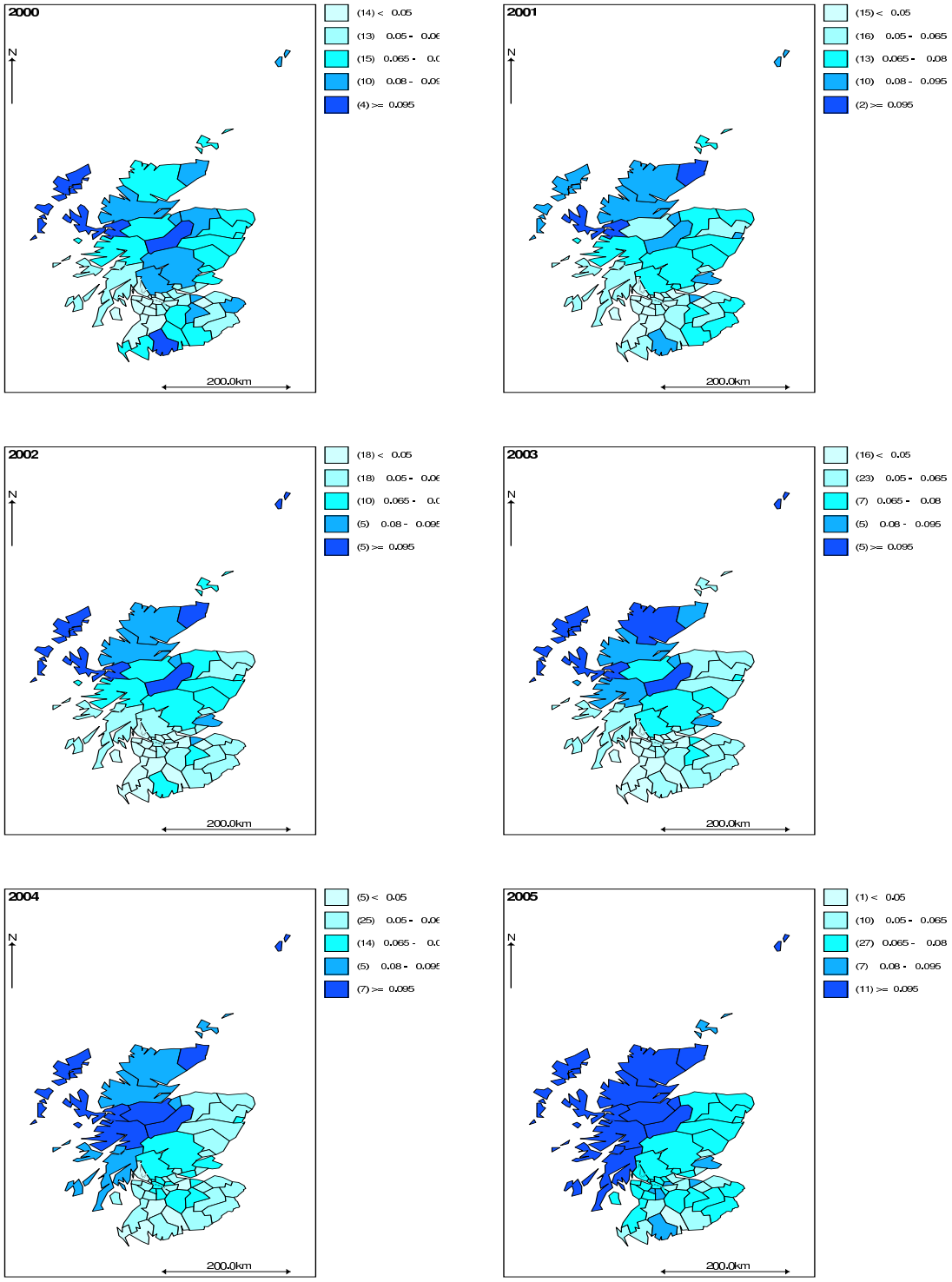

Figure 3.8: District level maps of estimated probabilities of primary 1 and 2 school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model).

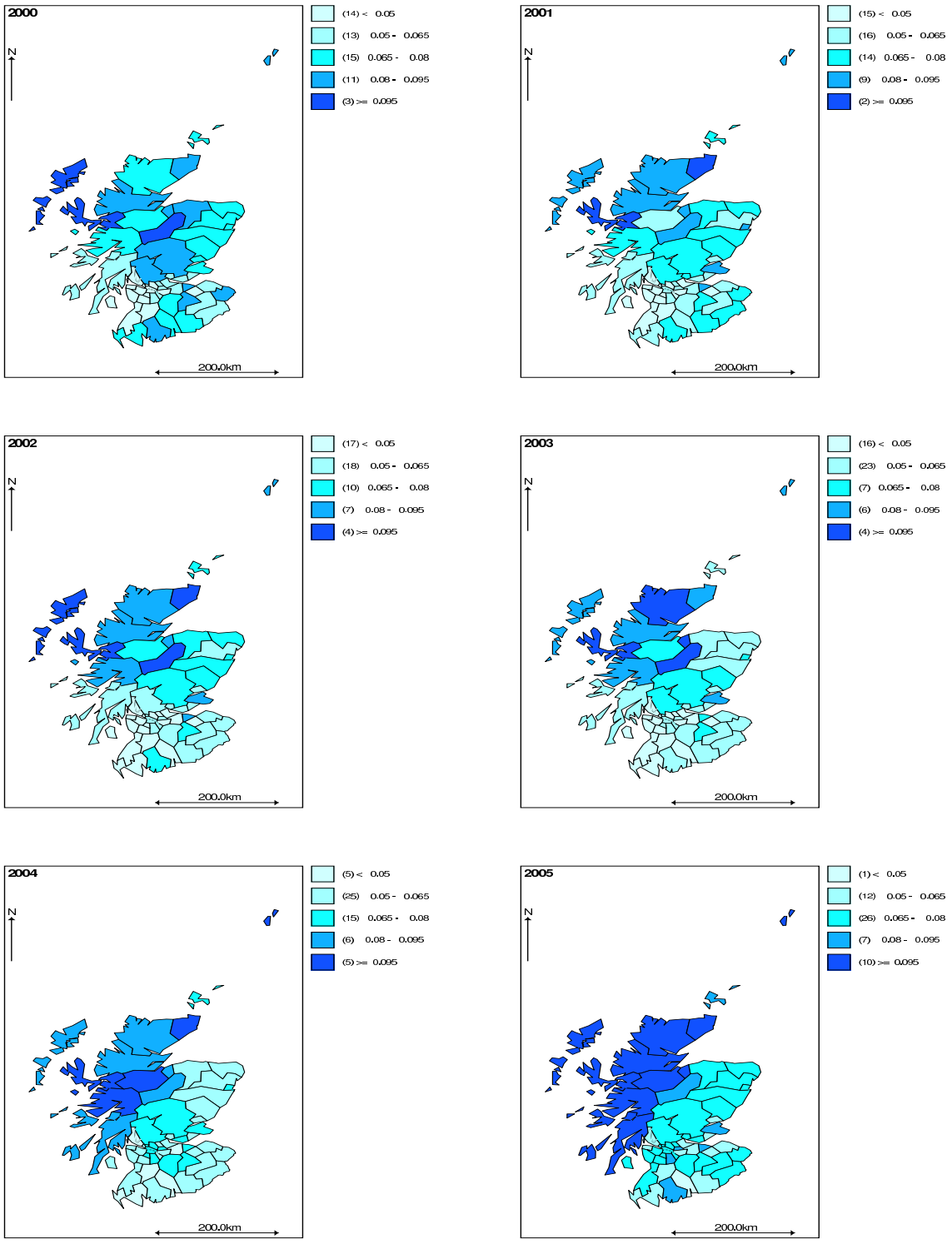

Figure 3.9: District level maps of estimated probabilities of primary 1 and 2 school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (space-time model).

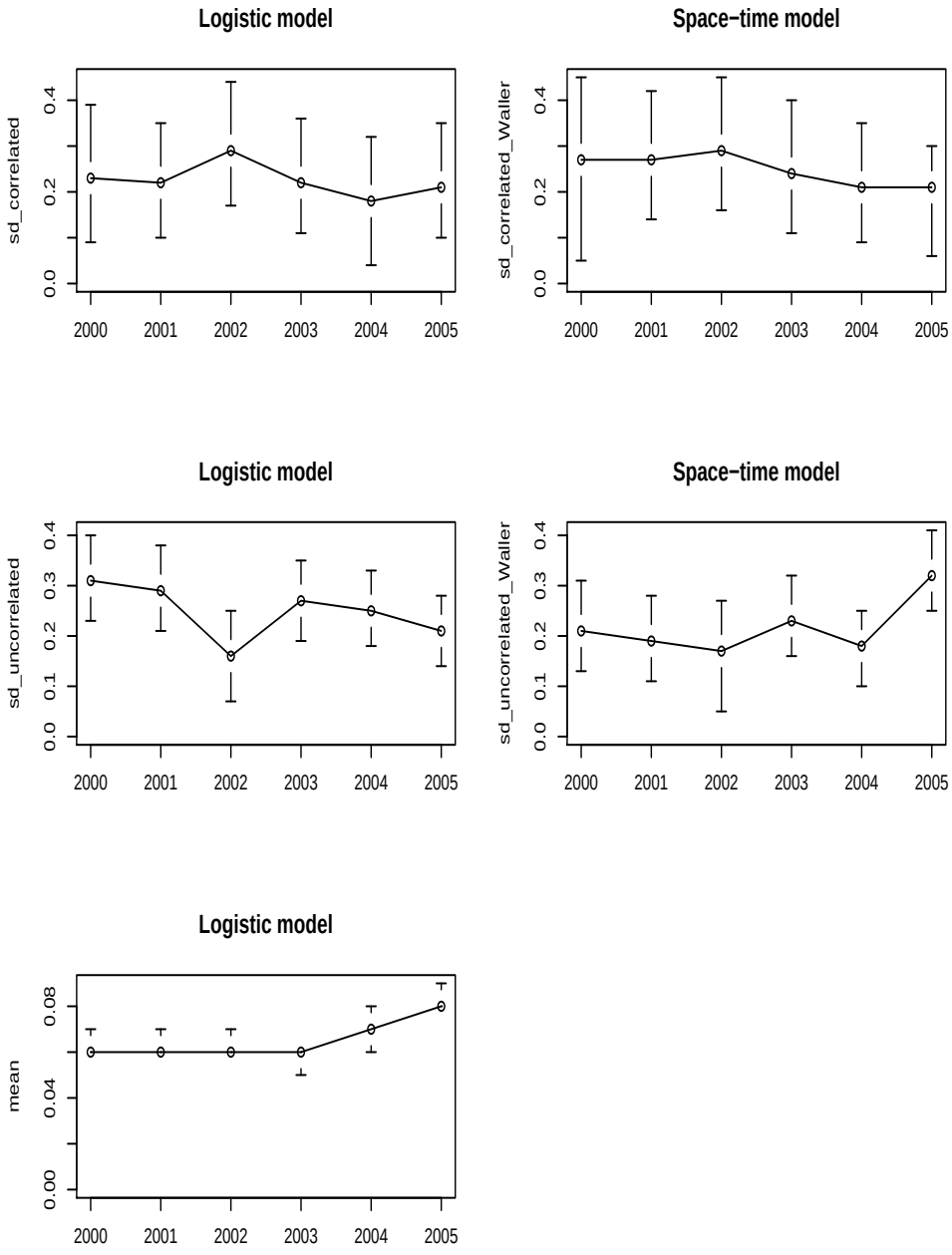

Figure 3.10: Plots of variability components and mean with 95% credible interval bars, for logistic and space-time Waller *et al.* (1997) models for primary 1 and 2 children at district level, against years 2000-2005.

The difference observed between the models may be due to the fact that when fitting the space-time model, this model contains only one α. Because of that, the space-time model may only really be appropriate if α does not vary over time, as in the case of primary 1 and 2. Otherwise as in the case of pre-school, we obtain higher values of unstructured variation than those obtained from the logistic model. Therefore in this case it will be better to use the logistic model.

Comparing pre-school and primary 1 and 2 school children, the overall proportion for pre-school is higher than it is for primary 1 and 2 school years, indicating that pre-school children have higher susceptibility rates than primary 1 and 2, as expected. Susceptibility is increasing significantly over time in pre-school compared to primary 1 and 2 school children, as can be observed from the maps of pre-school getting darker more quickly (regions with high susceptibility increasing) over time than in the case of primary 1 and 2. Urban districts of Edinburgh, Glasgow, Dundee and Abeerden are indicated in both groups to have high susceptibility rates.

Generally, districts are relatively less similar in primary 1 and 2 than in pre-school, indicated by higher standard deviations in primary 1 and 2 than pre-school. Some birth cohorts are in both pre-school and primary 1 and 2 school groups. Susceptibility was estimated in 2000, 2001, 2002 and 2003 for pre-school, corresponding to 2002, 2003, 2004 and 2005 for primary 1 and 2 school groups. The 2003 pre-school/ 2005 primary 1 and 2 birth cohort has the highest susceptibility in both groups.

The next section focuses on comparing maps at postcode level sector. The results can then be compared to the district level results to see if there are any differences or similarities.

3.4 Comparing Maps Over Time at Postcode Sector Level

It is expected that the same regions shown at district level to have high/low susceptibility will be observed at postcode sector level, thus giving the same interpretation of the maps, but, as the postcode sectors are smaller, there is more potential for them to be spatially clustered than districts. Also there is more sampling variation

at postcode sector level than at district level. For the postcode sector analysis, for the logistic model, a burn-in period of 8,000 (pre-school) and 9,000 (primary 1 and 2) iterations was used before convergence was achieved, and the parameter estimates are based upon a further 10,000 iterations for both pre-school and primary 1 and 2. For the space-time model, a burn-in of 12,000 iterations was needed for both pre-school children and results are based on a further 12,000 iterations.

Figure 3.11 and Figure 3.12 show maps of susceptibility obtained by fitting the logistic model (2.8) to each time period for each of these groups, and the space-time model (2.27) respectively. For pre-school, as at district level, generally susceptibility to measles is increasing over time followed by a decrease in 2005. In 2000 susceptibility is less than 15% in central and south Scotland and greater than 16% in the north and the Highlands. In 2001, there is an increase in number of regions in the north east with lower susceptibility of less than 15% and the number of postcode sectors with greater than 16% susceptibility (higher susceptibility) decreases in the north. These postcode sectors in the north again have an increase in susceptibility in 2002. In 2003 a global increase in susceptibility is observed and this is increased further in 2004, but isolated pockets of postcode sectors with low susceptibility can be seen here. Susceptibility decreased in 2005, especially in the north east, south and the Borders. The urban areas of Glasgow, Edinburgh, Aberdeen and Dundee have high susceptibility rates, as observed in the districts.

Unlike the logistic model, the maps obtained from the space-time model (Figure 3.12) show that susceptibility decreased in 2002, followed by an increase from 2003, while for the logistic model susceptibility decreased in 2001, followed by an increase from 2002. The space-time model indicates that 2000 and 2001 have more regions with susceptibility greater than 16% than are shown by the logistic model. The logistic model shows that 2003 has more (almost twice as many) regions with susceptibility greater than 16% than are shown by the space-time model.

Table 3.3 and Figure 3.13 show values and plots of parameters obtained from the models. For the logistic model, the overall proportion ranges from 0.17-0.21. As the number of postcode sectors with low susceptibility increased in 2001, the overall proportion here is lower than in 2000.

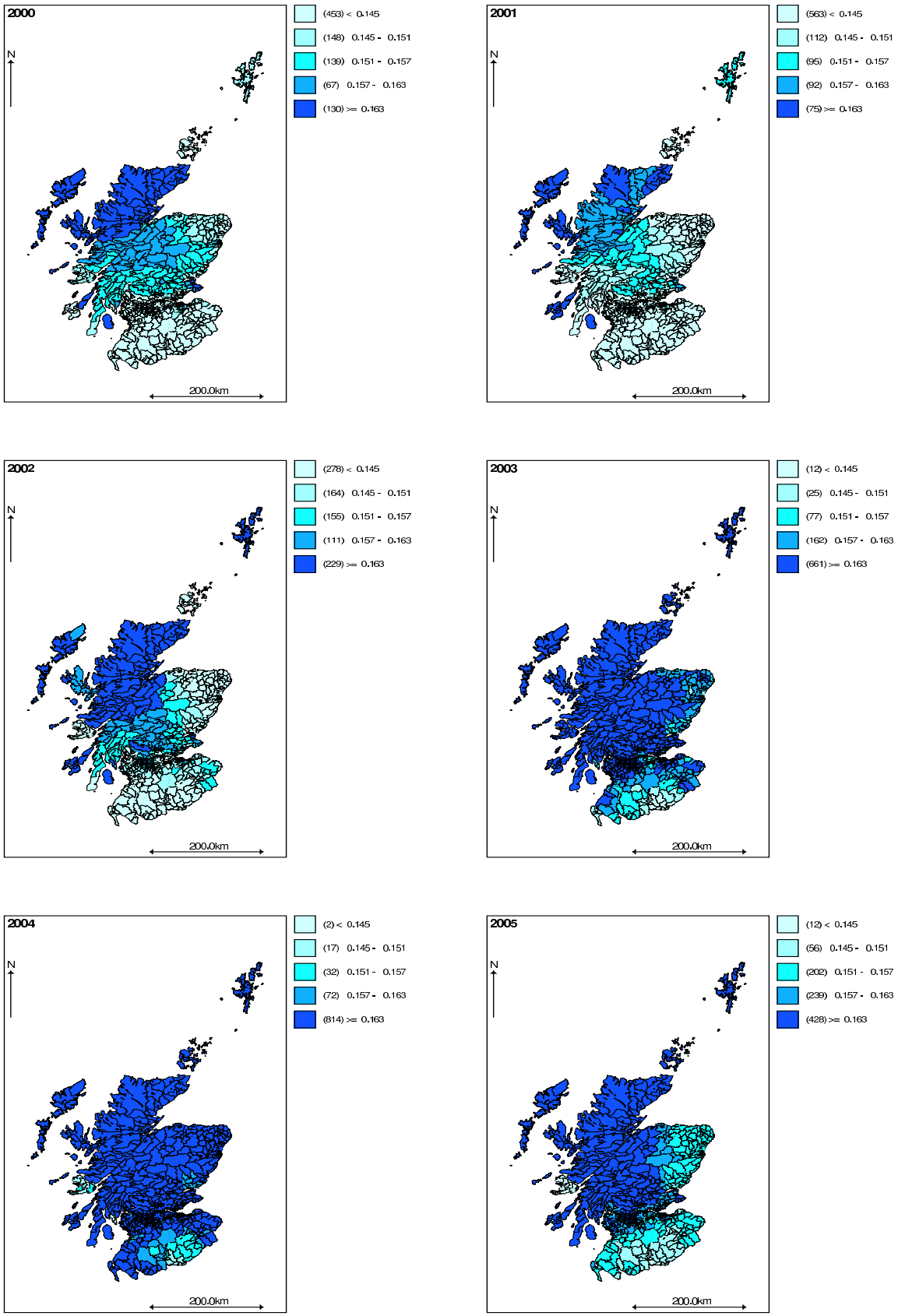

Figure 3.11: Postcode sector level maps of estimated probabilities of pre-school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model).

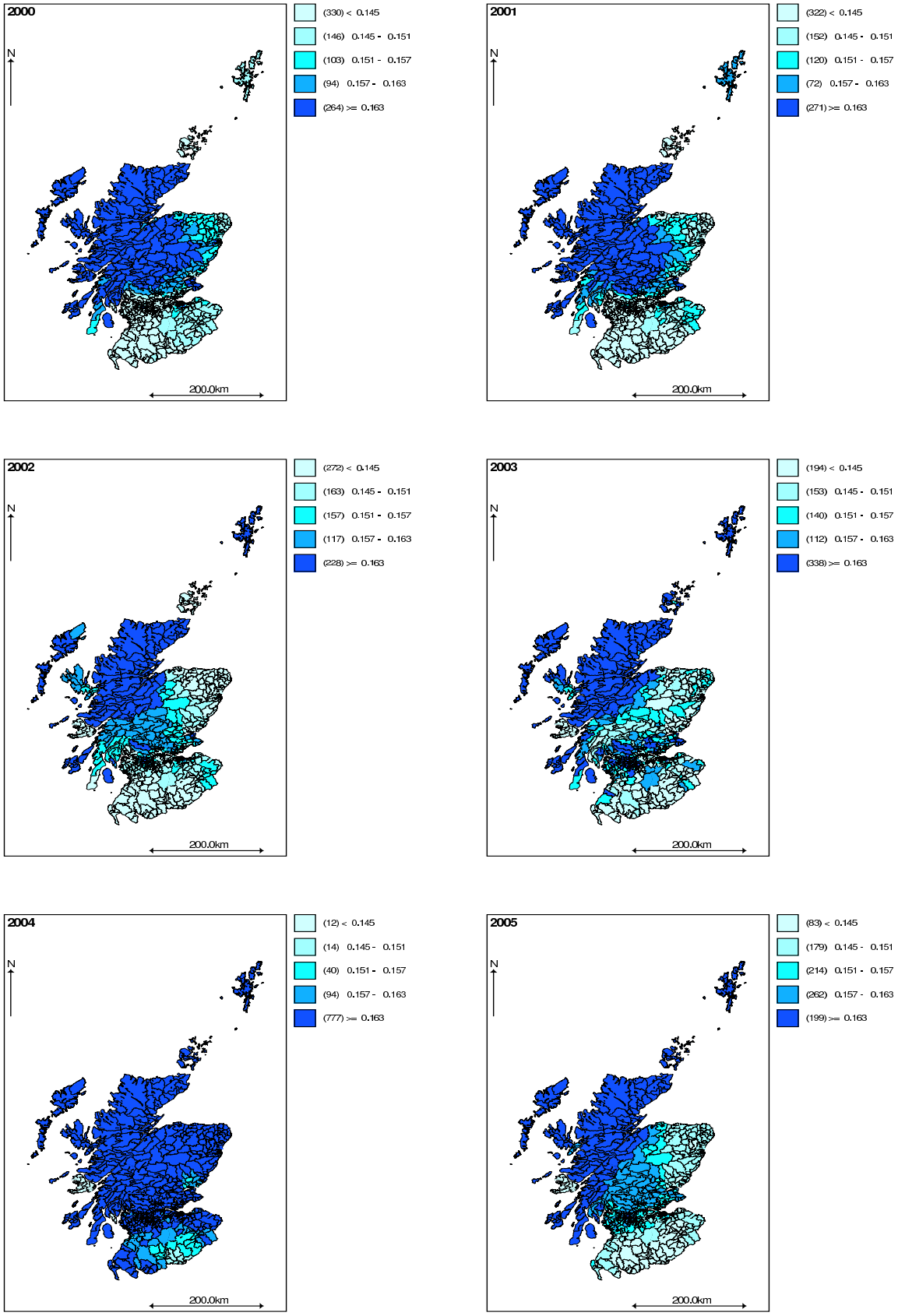

Figure 3.12: Postcode sector level maps of estimated probabilities of pre-school children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (space-time model).

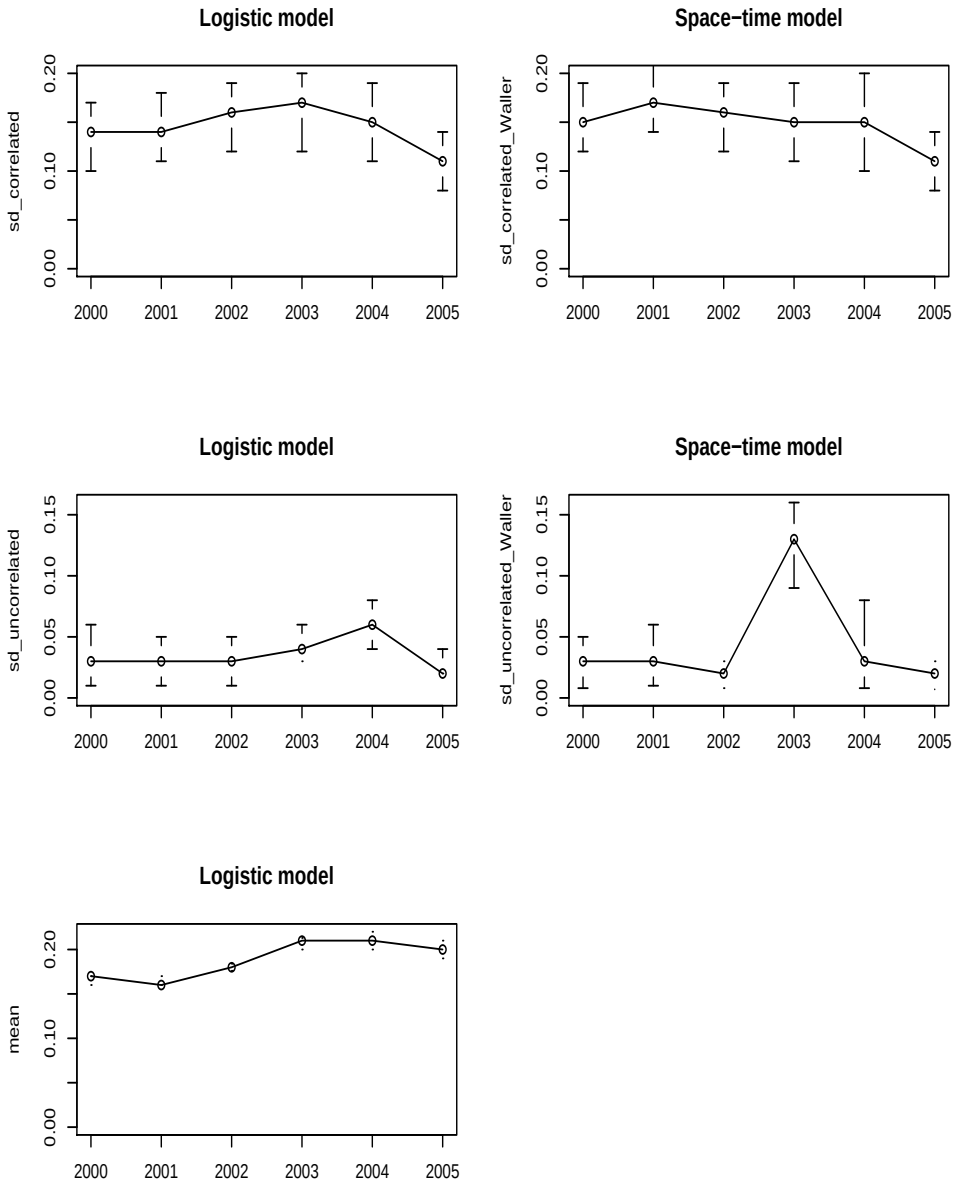

Figure 3.13: Plots of variability components and mean with 95% credible interval bars, for pre-school children, for logistic and space-time models, for 937 postcode sectors, against years 2000-2005.

Year	Pre-School					
	Logistic model				Space-time model	
	α	proportion	σ_u	σ_v	σ_u	σ_v
2000	−1.77(-1.78,-1.75)	0.17 (0.16,0.17)	0.14 (0.10,0.17)	0.03 (0.01,0.06)	0.15 (0.12,0.19)	0.03 (0.008,0.05)
2001	−1.80 (-1.82,-1.78))	0.16 (0.16,0.17)	0.14 (0.11,0.18)	0.03 (0.01,0.05)	0.17 (0.14,0.21)	0.03 (0.01,0.06)
2002	−1.72 (-1.74,-1.70)	0.18 (0.176,0.184)	0.16 (0.12,0.19)	0.03 (0.01,0.05)	0.16 (0.12,0.19)	0.02 (0.008,0.03)
2003	−1.55 (-1.57,-1.52)	0.21 (0.20,0.213)	0.17 (0.12,0.20)	0.04 (0.03,0.06)	0.15 (0.11,0.19)	0.13 (0.09,0.16)
2004	−1.57 (-1.62,-1.53)	0.21 (0.20,0.22)	0.15 (0.11,0.19)	0.06 (0.04,0.08)	0.15 (0.10,0.20)	0.03 (0.008,0.08)
2005	−1.63 (-1.65,-1.61)	0.20 (0.19,0.21)	0.11 (0.08,0.14)	0.02 (0.02,0.04)	0.11 (0.08,0.14)	0.02 (0.007,0.03)
Year	Primary School					
	Logistic model				Space-time model	
	α	proportion	σ_u	σ_v	σ_u	σ_v
2000	−3.21 (-3.45,-2.87)	0.041(0.03,0.06)	0.54 (0.44,0.62)	0.09 (0.01,0.18)	0.39 (0.30,0.47)	0.29 (0.24,0.33)
2001	−2.77 (-2.81,-2.73)	0.063 (0.060,0.065)	0.47 (0.36,0.56)	0.18 (0.09,0.24)	0.37 (0.30,0.46)	0.28 (0.24,0.33)
2002	−2.75 (-2.78,-2.71)	0.064 (0.062,0.066)	0.40 (0.31,0.50)	0.16 (0.03,0.22)	0.34 (0.27,0.41)	0.24 (0.20,0.29)
2003	−2.81 (-2.84,-2.77)	0.060 (0.058,0.063)	0.37 (0.28,0.46)	0.20 (0.14,0.25)	0.33 (0.25,0.40)	0.26 (0.22,0.31)
2004	−2.65 (-2.68,-2.62)	0.071 (0.068,0.073)	0.40(0.32,0.48)	0.11(0.02, 0.17)	0.30 (0.24,0.37)	0.26 (0.26,0.31)
2005	−2.48 (-2.51,-2.45)	0.084 (0.081,0.086)	0.40 (0.34,0.46)	0.05 (0.01,0.12)	0.25 (0.17,0.33)	0.37 (0.32,0.41)

Table 3.3: Pre-school and primary 1 and 2 posterior means, with 95% credible intervals in brackets, for overall mean level (α), proportion ($\frac{e^{\alpha}}{1+e^{\alpha}}$), standard deviations due to correlated heterogeneity (σ_u) and uncorrelated heterogeneity (σ_v) for logistic and space-time (Waller *et al.*, 1997) models for 937 postcode sectors.

There is an increase from 2002 to 2003, as it can be seen from the maps that high susceptibility rates increased. The highest overall proportion is in 2003 and 2004, as shown by the maps that most postcode sectors have higher susceptibility rates than in other years, and in 2005 there was a slight decrease in overall proportion as postcode sectors with high susceptibility rates decreased. The credible intervals for the overall proportions suggests there was no real differences in overall proportion for 2000 to 2002 and in 2003 there was an increase in overall proportions and this remained similar in 2004, and there was no real difference in 2005.

For both models, the structured standard deviation ranges from 0.11-0.17 with wider credible intervals suggesting no major differences over time for structured variation. The unstructured standard deviations ranges from 0.02-0.06 for the logistic model and 0.02-0.13 for the space-time model. For the space-time model the credible intervals suggests a change in 2003, while for both models for other years there are no major differences.

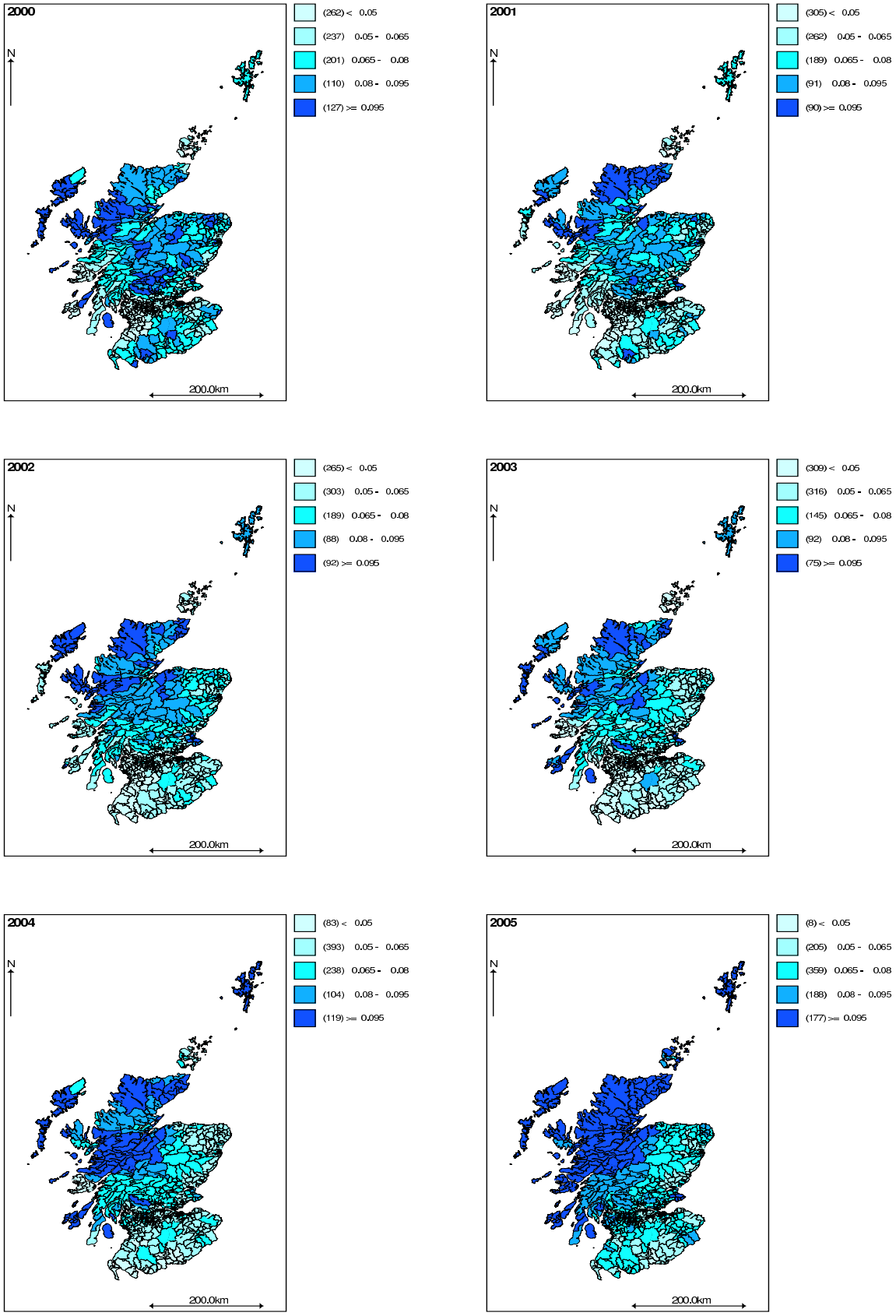

Figure 3.14: Postcode sector level maps of estimated probabilities of primary 1 and 2 children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model).

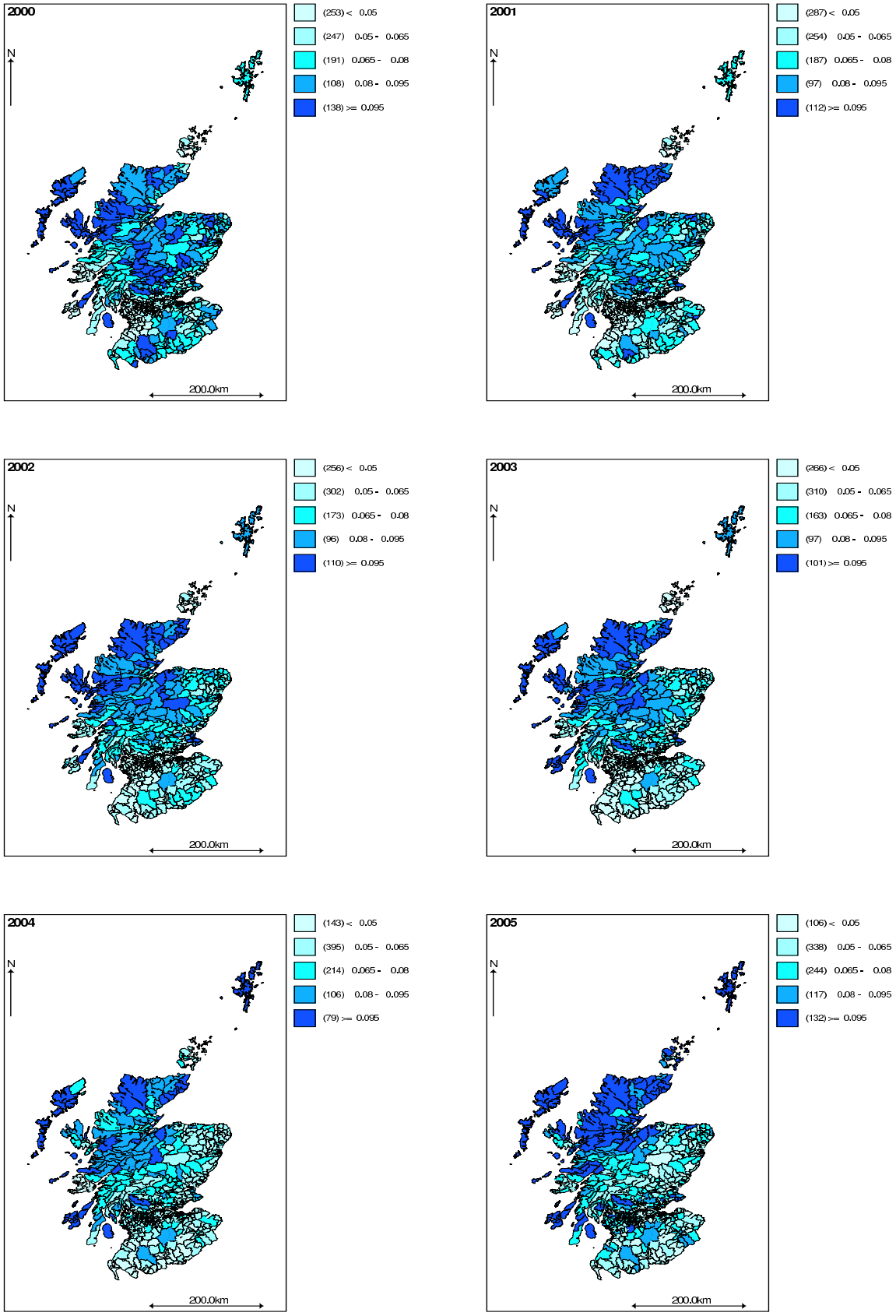

Figure 3.15: Postcode sector level maps of estimated probabilities of primary 1 and 2 children susceptible to measles for 2000, 2001, 2002, 2003, 2004 and 2005 (space-time model).

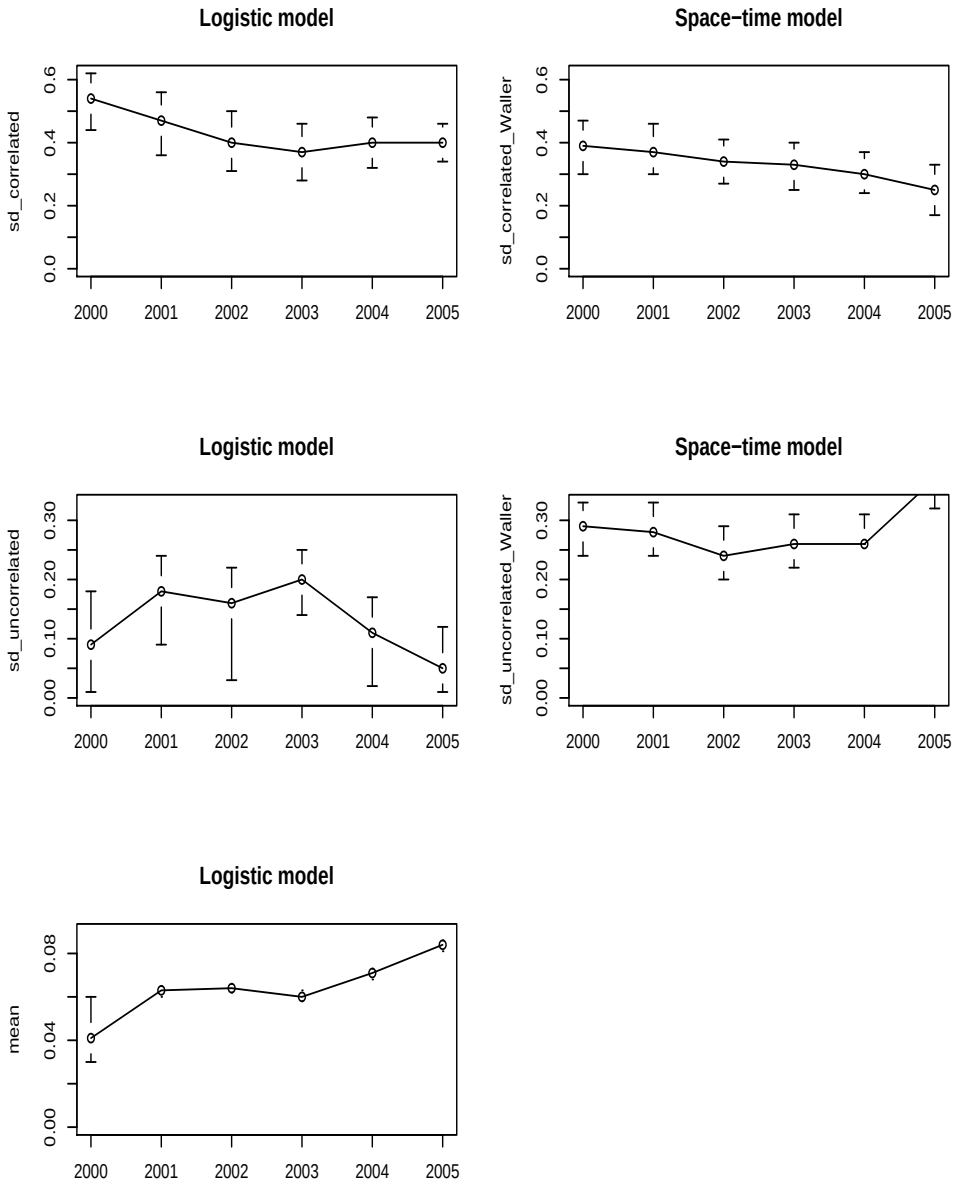

Figure 3.16: Plots of variability components and mean with 95% credible interval bars, for primary 1 and 2 children, for logistic and space-time models, for 937 postcode sectors, against years 2000-2005.

Overall, for each year, the value of the structured standard deviation is higher than for the unstructured variation, thus variability in susceptibility is more due to local clustering of rates than variation between postcodes.

For primary 1 and 2 school children, the maps of susceptibility rates are shown in Figures 3.14 and 3.15 for the logistic model (2.8) and space-time model (2.27) respectively. It is observed that in 2000 the central postcode sectors had lower susceptibility of less than 5%, and there are isolated pockets of high susceptibility greater than 9.5%. In 2001 there is a reduction in the number of postcode sectors with high susceptibility rates, but still pockets of these can be observed. The number of postcode sectors with low susceptibility rates, less than 5%, increased in the central and the west regions.

There is a slight increase in the number of regions with high susceptibility rates in the north in 2002, with low susceptibility rates increasing in the south. In 2003, there is a slight increase of low susceptibility rates in the southern regions, with isolated pockets of high susceptibility. Susceptibility increases in most districts in the north and north west in 2004 and low susceptibility increased in the north east. In 2005 more regions have increased high susceptibility rates, with susceptibility greater than 9.5% in the north and west.

Table 3.4 and the plots of parameters in Figure 3.16 show that generally the overall proportion (ranging from 0.041-0.084), is increasing over the years, but in 2003 there is a decrease in overall proportion. This agrees with what is shown by the maps, as in 2003 the number of regions with high susceptibility rates, i.e greater than 9.5%, was lower than in all other years. Overall proportion increased from 2004 to 2005, and this is shown by an increase in the number of postcode sectors with high susceptibility rates in this year. The structured standard deviation ranges from 0.37-0.54 for the logistic model and 0.25-0.39 for the space-time model, and for both models the credible intervals are wider, therefore there might be no big change between the years. The unstructured standard deviation decreases initially over time with an increase from 2004 and 2005. This agrees with what the maps show, i.e. a decrease of clustering of postcode sectors with similar rates and in 2004 to 2005 there was increased evidence of clustering of postcode sectors with higher rates. The unstructured standard deviation ranges from 0.25-0.39 for the logistic model and 0.24-0.37 for the space-time model. This does not show much of a trend but the logistic model shows that there

was a decrease in 2004 and 2005, while the space-time model indicates an increase between these two years. Looking at the credible intervals, they suggest that there was no real difference in overall proportions from 2000-2003 and from 2003-2005 there was an increase in overall proportions. For the structured and unstructured standard deviations, the credible intervals are wide and suggest no real differences over time.

As for districts, for both pre-school and primary 1 and 2 school children, maps given by the space-time model (2.27) are similar to those of the logistic model (2.8). The logistic model gives higher values for the structured standard deviations than the space-time model and lower values for the unstructured standard deviations than the space-time model. From both models it is observed that variation is due to structured variation more than unstructured variation.

Comparing pre-school and primary 1 and 2 school children, as for districts, the overall proportion for pre-school is higher than the overall proportion for primary 1 and 2, indicating that pre-school children have higher susceptibility rates than primary 1 and 2, as expected. The 2003 (pre-school) and 2005 (primary 1 and 2) birth cohorts have the highest susceptibility in both groups. Susceptibility is increasing more noticeably over time in pre-school children than in primary 1 and 2 school children, as shown by the regions in the maps becoming darker quickly for pre-school children than primary 1 and 2. For both pre-school and primary 1 and 2 children, and for both models, the values of the structured standard deviation are higher than the unstructured standard deviation. This shows that regional variation in susceptibility is due more to clustering than general variation. Primary 1 and 2 has higher values for both standard deviations and models than pre-school. Thus, susceptibility is more similar in postcode sectors in pre-school children than in primary 1 and 2 children.

3.5 Comparison of District and Postcode Sector Levels

Comparing maps fitted at districts (Figures 3.5-3.6 and 3.8-3.9) and postcode sector (Figures 3.11-3.12 and 3.14-3.15) levels, the interpretation is similar especially for primary 1 and 2 school children (Figures 3.8-3.9 and Figures 3.14-3.15), and for pre-school logistic model (Figures 3.5 and 3.11), except that at postcode sector level,

there is more clustering than at district level. For pre-school, the space-time model at district level (Figure 3.6) show that unlike at postcode sector level (Figure 3.12), susceptibility was lower in 2000 and 2001 and increased in 2002. For postcode sectors, susceptibility was higher in most regions in 2000 and 2001 and decreased in 2002. Tables 3.3 and 3.4 give estimates of the structured and unstructured standard deviations for the districts and 937 postcode sectors respectively. For pre-school and for both models, structured standard deviation (σ_u) is much higher at postcode sector level. The unstructured standard deviation (σ_v) is lower at postcode level. For primary 1 and 2, for both models structured standard deviation is again higher at postcode level than at district level. The unstructured standard deviation is lower at postcode level than at district level for the logistic model, while the space-time model gives different results, i.e. structured standard deviation is lower at district level than postcode sector level.

The structured standard deviation tends to be larger at postcode sector level than at district level i.e. there tends to be more spatial clustering of postcode sectors than districts. This is because districts are larger and each of them has neighbours which are large, and they have less variation in susceptibility rates than postcode sectors. Therefore, there is more potential for the susceptibilities of the postcode sectors to be spatially clustered than the larger districts. Districts are larger, and postcode sectors are smaller and more likely to be binomial in distribution for susceptibility. Districts are a combination of postcode sectors, and so by adding them up we accumulate extra binomial dispersion giving unstructured variability, which is why the districts have larger unstructured variation than postcode sectors.

The next section focuses on fitting an ecological model to the pre-school and primary 1 and 2 school susceptibility to measles data. Thus the relationship between susceptibility and census variables is explored, both at district and postcode sector level.

3.6 Ecological Variable Selection

The census variables considered here are percentage of people in households with no car, percentage of people in overcrowded households, percentage of unemployed males, percentage of people in households with low social class, percentage of children aged

0-4, percentage of households with dependent children, percentage of individuals born in other European countries, percentage of individuals born outside EU (elsewhere), percentage of lone parent households, percentage of individuals working in agriculture and percentage of individuals with or without qualifications (five levels). Firstly the correlation between each of the explanatory variables was investigated. The percentage of households with dependent children correlated with percentage of children aged 0-4, thus only the percentage with children aged 0-4 years was included in the model as this is the age of vaccination. Also because of correlation of education levels, only two categories were included in the model. These were no qualifications, or high qualifications (i.e. first degree, high degree or professional qualification). Altogether, 11 census variables were used. These were percentage of people in households with no car, percentage people in overcrowded households, percentage of unemployed males, percentage of people in households with low social class, percentage of children aged 0-4, percentage of people born in other European countries, percentage of people born outside EU (elsewhere), percentage of lone parent households, percentage of people working in agriculture, percentage of people with high qualifications and percentage of people with no qualifications. Scatter plots were used initially to assess the linearity of the variables with $logit(p)$ (see equation (3.1)) and they were linear.

For the ecological analysis, we fitted the logistic CAR model with covariates model (2.13) using WinBUGS. Since we have 11 census variables, fitting all these in WinBUGS is very time consuming (convergence is very slow) and there were problems with computer memory, especially when trying to use a DIC tool, to compare fitted models and choose significant variables. (In some cases the computer may hang up, and here WinBugs was run on a Windows XP Pentium 4 computer with 3GHz and 512MB of RAM).

In order to avoid these problems, firstly we fitted a generalised linear mixed model (logistic regression) below:

$$O_i \sim Bin(n_i, p_i), \quad logit(p_i) = \alpha + v_i + \sum_{j=1}^{p} \beta_j x_{ij}, \tag{3.1}$$

where n_i and p_i are number of people and the unknown probability of children susceptible to measles in the ith region respectively, α is a constant, $v_i \sim N(0, \sigma_v)$ is the random effect for region i and the $\beta_j s, j = 1, ..., p$, are the coefficients of covariates $x_1, ..., x_p$. Thus

$$p_i = \frac{e^{\alpha+v_i+\sum_{j=1}^{p}\beta_j x_{ij}}}{1+e^{\alpha+v_i+\sum_{j=1}^{p}\beta_j x_{ij}}}.$$

The penalised-quasi-likelihood method was used to fit the model to each year's/cohort's data in R, using function *glmmPQL* in library MASS, for both pre-school and primary 1 and 2 school children at each of district and postcode sector levels. By using backward elimination, census variables significant at the 5% level were obtained.

Fitting models with significant census variables in WinBUGS, for districts, a burn-in period of 8,000 (pre-school) and 10,000 (primary 1 and 2) iterations was used before convergence was achieved and the parameter estimates are based upon a further 12,000 iterations for both pre-school and primary 1 and 2. For postcode sectors a burn-in period of 12,000 (pre-school) and 15,000 (primary 1 and 2) iterations was used before convergence was achieved and the parameter estimates are based upon a further 15,000 (pre-school) and 17,000 primary (1 and 2) iterations.

3.6.1 District Level

At district level, for pre-school, when all census variables are in the model (Table A.2 in Appendix A), the only significant variable at 5% level is the percentage of people in households with low social class for 2003, 2004 and 2005 (and none for 2000-2002). After backward elimination, the census variables found to be significant at the 5% level (in Table A.3) are percentage of people in households with no car for the years 2002 and 2005, percentage of people in households with low social class for 2001-2005, percentage of male unemployed for 2001, 2003 and 2004, and percentage of individuals born in other EU countries for 2000 and 2001. For primary 1 and 2 school children, when all census variables are in the model (Table A.4) the significant variables are percentage of people in overcrowded households for 2000 and 2005, percentage of individuals born in other EU for 2000, 2001 and 2005, and percentage of individuals with no qualification for 2005 only. The significant variables (after backward elimination) are shown in Table A.5, namely percentage of people in overcrowded households for 2000, 2001, 2002 and 2005, percentage of children aged 0-4 years for 2002 and 2003, percentage of individuals born in other EU countries for 2000, 2001, 2002, 2003 and 2005, percentage of people in households with no car for 2004, percentage of people in households with low social class for 2003 and 2004, and percentage of individuals with no qualifications for 2005 only.

For both pre-school and primary 1 and 2 groups, susceptibility increases as percentage of households with low social class, percentage of individuals born in other EU, and percentage of people in households with no car, increase. For pre-school alone, susceptibility increases as percentage male unemployment increases. For primary 1 and 2 alone, susceptibility increases with increasing percentage of people who are not qualified, and decreases with increase in percentage of people in overcrowded households and percentage of children aged 0-4 years.

3.6.2 Postcode Sector Level

At postcode sector level, for pre-school, when all census variables are in the model (Table A.6), the significant variables were the percentage of individuals born in other EU countries for 2000, 2001 and 2003, percentage of individuals with no qualifications for 2000, percentage of individuals with high qualifications for 2000, 2001, 2003 and 2004, percentage of individuals working in agriculture and percentage of people in households with no car for 2000, 2001, 2002, 2003 and 2005, percentage people in overcrowded households for 2000 and 2001, percentage of individuals unemployed for 2001, 2002, 2003 and 2005, and percentage of people in households with low social class for 2000, 2001, 2002, 2003 and 2005. After backward elimination (Table A.7) the census variables found to be significant at the 5% level were percentage of individuals born in other EU countries for 2000-2003, percentage of individuals working in agriculture for 2000-2003 and 2005, percentage of people in households with no car for 2000-2003 and 2005, percentage of people in overcrowded households for 2001 and 2002, percentage of individuals unemployed for 2001, 2002, 2003 and 2005, percentage of people in households with low social class for 2001-2003 and 2005, and percentage of individuals with high qualifications only significant for 2004. For primary 1 and 2, when all census variables are in the model (Table A.8), the significant variables were the percentage of individuals born in other EU countries for 2000-2004, percentage of individuals with no qualifications for 2002, percentage of individuals with high qualifications for 2002, percentage of individuals working in agriculture, percentage of people in households with no car, percentage of people in overcrowded households and percentage of people in households with low social class for 2000-2005, and percentage of individuals unemployed for 2004 and 2005. After backward elimination (Table A.9), the significant variables at 5% were percentage of individuals born in other EU countries, percentage

of individuals working in agriculture, percentage of people in households with no car, percentage of people in overcrowded households and percentage people in households with low social class, for all years. Percentage of individuals with no qualifications and high qualifications are only significant in 2002, with percentage of individuals unemployment significant for 2003, 2004 and 2005.

In the next section, those variables found to be significant are fitted using a spatial logistic model at both district and postcode sector level.

3.7 Ecological analysis of Measles Data at District Level

In this section, we investigate the relationship between the spatial variation of susceptibility to measles and census variables, over the period 2000-2005, for both pre-school and primary 1 and 2 school children at district level. The logistic model (2.13) was fitted for each year/cohort with census variables found to be significant in the 3 or more years when using penalised-quasi-likelihood.

These census variables were percentage of male unemployment and percentage of people in households with low social class for pre-school children, and percentage of people born in other EU countries and percentage of people in overcrowded households for primary 1 and 2 children. When fitting the model with census variables, there was a problem with model convergence i.e. simulation stabilising when using these variables. To solve this problem, each of the census variables was standardised by subtracting its mean and dividing by its standard deviation before fitting in WinBUGS, and this overcame the lack of convergence.

Maps of the levels of the census variables at district level in Figure 3.17 show that percentage of people born in other EU countries is in general higher (at least 1.4%) in Inverness, Aberdeen, North East Fife, Edinburgh, Midlothian and East Lothian. Percentage of people in overcrowded households is lower in the northern districts and higher in the western districts and central. Most districts (40 out of 56) have low percentage of unemployment male, while percentage of people in households with low social class is highest in the north and some districts in the south and major cities.

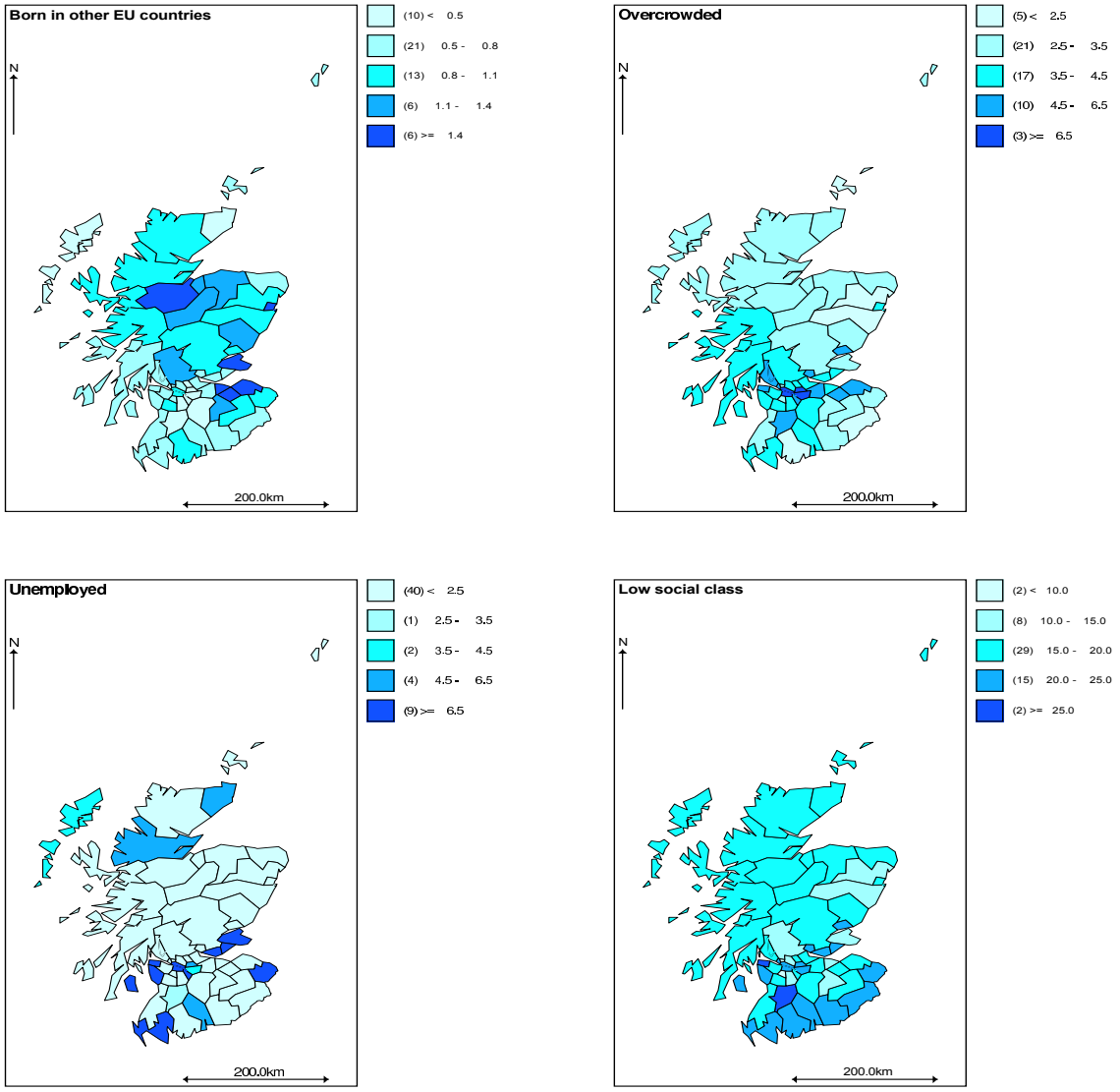

Figure 3.17: Maps of the levels of the census variables used in the logistic CAR model for 56 districts.

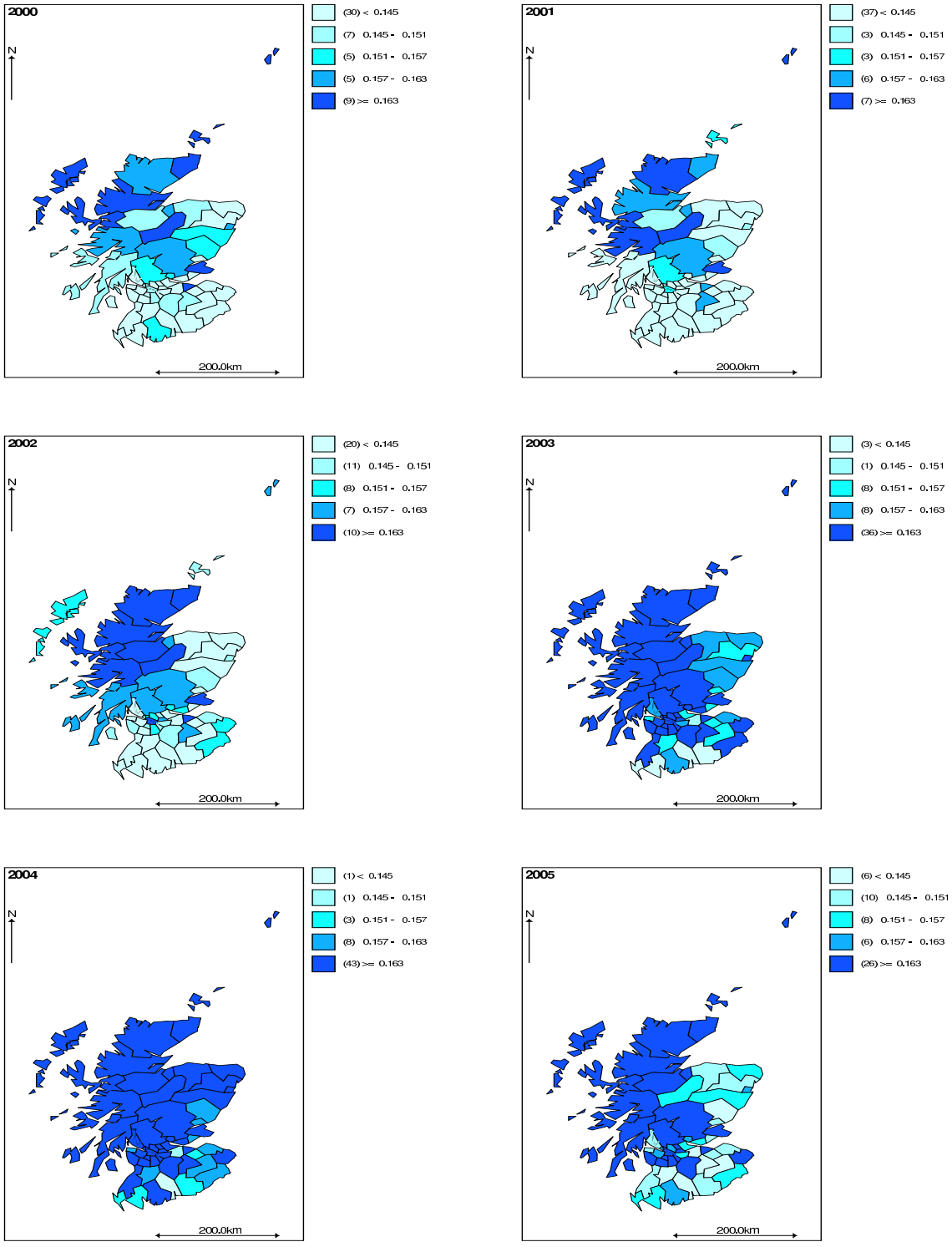

Figure 3.18: District level maps of estimated probabilities of children susceptible to measles (pre-school) for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model with census variables).

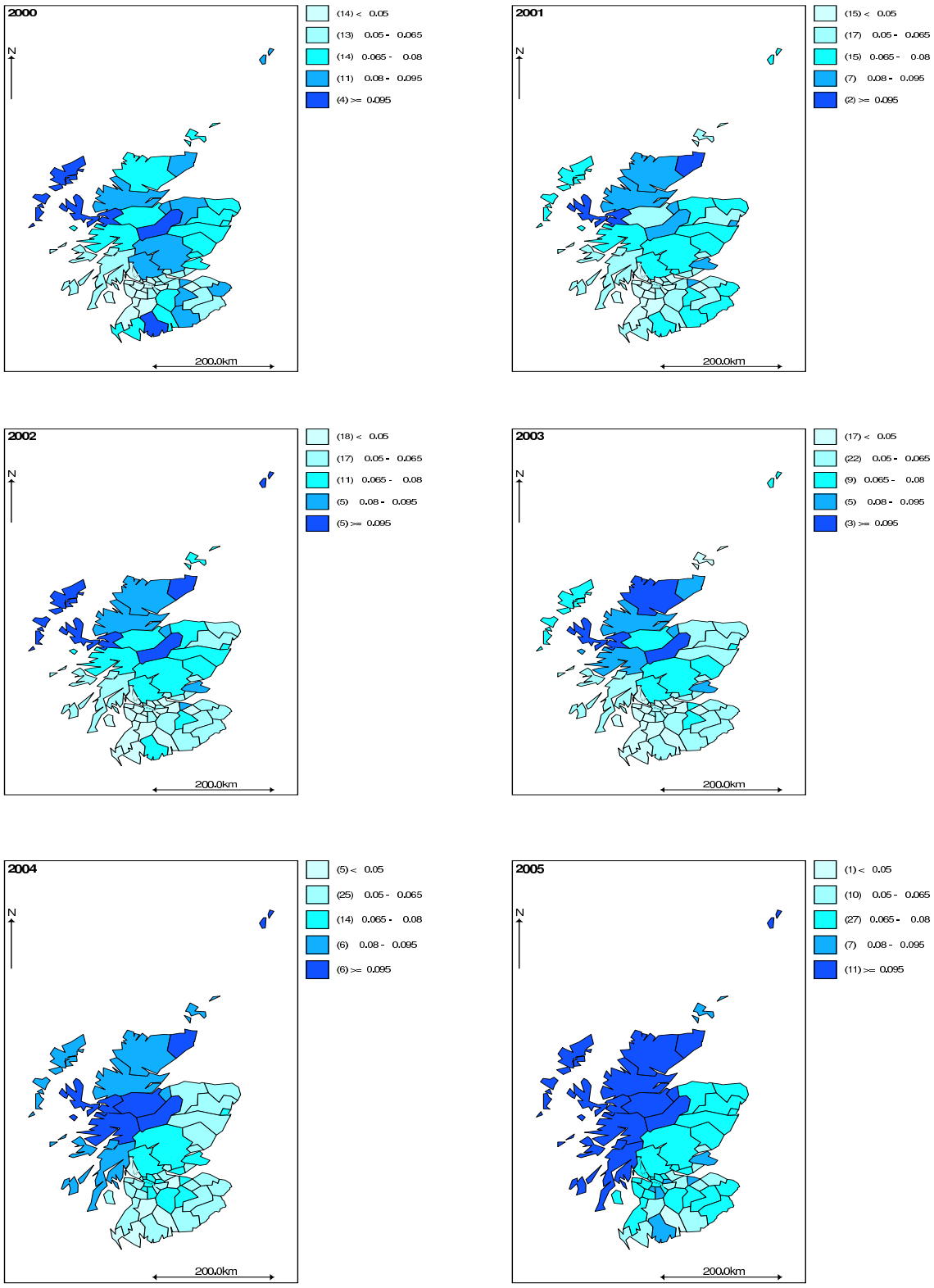

Figure 3.19: District level maps of estimated probabilities of children susceptible to measles (primary 1 and 2) for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model with census variables).

The maps of susceptibility to measles in Figure 3.18 (pre-school) and 3.19 (primary 1 and 2) were produced to see how they compare with the maps without census variables (Figures 3.5 and 3.8). The same spatial patterns are observed for maps without and maps with census variables. This shows that inclusion of these variables does not explain the spatial differences at district level.

Tables 3.4 (pre-school) and 3.5 (primary 1 and 2) show results obtained by fitting the logistic CAR model (2.13) at district level. The signs of the parameter estimates for the census variables do not change much over time (i.e. direction of relationship is the same, negative/positive for all the years). Exceptions are percentage of people in households with low social class in 2005 for both pre-school and primary 1 and 2, and percentage of people in overcrowded households in 2003 and 2004 for primary 1 and 2.

	2000			2001			2002		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-1.83	-1.91	-1.75	-1.86	-1.94	-1.78	-1.72	-1.76	-1.68
σ_u	0.09	0.03	0.17	0.08	0.03	0.17	0.11	0.03	0.20
σ_v	0.22	0.17	0.28	0.23	0.17	0.29	0.08	0.03	0.12
% unemployed	0.002	-0.01	0.01	0.003	-0.01	0.01	0.002	-0.003	0.008
%low social class	-0.0001	-0.005	0.004	-0.0007	-0.006	0.004	-0.002	-0.004	0.0008
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-1.61	-1.69	-1.54	-1.52	-1.56	-1.47	-1.68	-1.75	−1.62
σ_u	0.10	0.03	0.21	0.12	0.06	0.20	0.09	0.03	0.18
σ_v	0.20	0.15	0.26	0.10	0.05	0.15	0.19	0.14	0.24
% unemployed	0.01	-0.005	0.01	0.003	-0.006	0.01	0.004	−0.01	0.01
%low social class	-0.0006	-0.005	0.004	-0.001	-0.004	0.002	0.0004	−0.004	0.0053

Table 3.4: Posterior mean parameter estimates for significant census variables (log odds ratios) and standard deviation due to uncorrelated and correlated heterogeneity with their lower (LCL) and upper (UCL) 95% credible limits, for pre-school children for 2000-2005, for 56 districts.

For pre-school, the higher the percentage of unemployment males, the higher the probability of susceptibility. Susceptibility decreases with increase in percentage of people in households with low social class, except for 2005 when susceptibility seems to increase with increase in percentage of people in households with low social class. For primary 1 and 2 children, susceptibility increases with increase in percentage born in

other EU countries. Susceptibility decreases with increase in percentage of people in overcrowded households for 2000, 2001 and 2005, but for 2003 and 2004 susceptibility increases with increasing percentage of people in overcrowded households. The magnitude of the effects is very small in all cases. Overall, the 95% credible intervals for these census variables for both pre-school and primary 1 and 2 groups at district level all include zero, suggesting that at district level these variables are not significantly different from zero, i.e. in no case is there a significant effect.

As the explanatory variables in general have a small effect, there are very little changes in the structured and unstructured variation when variables are in the model. For pre-school (Tables 3.2 and 3.4), inclusion of variables in the model increased the clustering standard deviation in 2000, 2002 and 2004 and decreased it in 2003 and 2005, with the effect in 2001 remaining the same. The unstructured standard deviations increased in 2000, 2001, 2003 and 2005 and decreased in 2002 and 2004. We note for 2002 that the inclusion of explanatory variables decreased very much the unstructured standard deviation from 0.22 to 0.11, indicating that explanatory variables accounted for regional variation in measles susceptibility while in 2003 the unstructured variation increased significantly.

	2000			2001			2002		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-2.70	-2.84	-2.56	-2.74	-2.85	-2.64	-2.74	-2.87	-2.61
σ_u	0.22	0.07	0.40	0.30	0.17	0.47	0.25	0.12	0.39
σ_v	0.31	0.23	0.41	0.17	0.07	0.27	0.25	0.18	0.34
% born other EU	0.02	-0.02	0.06	0.002	-0.03	0.03	0.02	-0.01	0.06
% overcrowded	-0.002	-0.02	0.02	-0.0003	-0.02	0.01	0.004	-0.01	0.02
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-2.78	-2.89	-2.68	-2.63	-2.75	-2.51	-2.45	-2.56	-2.33
σ_u	0.25	0.14	0.40	0.16	0.054	0.31	0.21	0.08	0.36
σ_v	0.19	0.12	0.28	0.25	0.18	0.33	0.21	0.15	0.29
% born other EU	0.008	-0.02	0.04	0.03	-0.008	0.06	0.006	-0.03	0.04
% overcrowded	0.001	-0.01	0.02	0.003	-0.01	0.02	-0.003	-0.02	0.01

Table 3.5: Posterior mean parameter estimates for significant census variables (log odds ratios) and standard deviation due to uncorrelated and correlated heterogeneity with their lower (LCL) and upper (UCL) 95% credible limits, for primary school children for 2000-2005, for 56 districts.

For primary 1 and 2 (Tables 3.3 and 3.5), there was a decrease in structured variation

2000, 2002 and 2004, and an increase in 2001, 2003 and 2005. Inclusion of variables increased the unstructured standard deviation in 2000, 2001, 2003 and 2005, and decreased it in 2002 and 2004 for pre-school, while for primary 1 and 2 it remained the same for 2000, 2004 and 2005, decreased for 2001 and 2003 and increased in 2002.

The following section focuses on fitting the models to smaller areas (postcode sectors) than districts. It is expected that there will be less ecological bias at postcode sector level as regions are smaller.

3.8 Ecological Analysis of Measles Data at Postcode Sector Level

To describe the relationship between the spatial variation of susceptibility and census variables at postcode sector level, the logistic CAR model (2.13) was again fitted for each year/cohort with census variables found to be significant in 3 or more years when using penalised-quasi-likelihood. For both pre-school and primary 1 and 2, these census variables were: the percentage of people born in other EU countries, the percentage of people working in agriculture, the percentage of people in households with no car, the percentage of people in overcrowded households, the percentage of male unemployed and the percentage of people households with low social class. For pre-school, we note that even though percentage overcrowded was significant in only two of the years 2000-2005 (2000 and 2001), this was included for consistency and comparability. As for districts (Section 3.7), these census variables were standardised before fitting them in the model.

Figure 3.20 shows maps of levels of the census variables which were included in the model. For the percentage of people born in other EU countries most postcode sectors have less than 2.5% of people in this category. For percentage of people working in agriculture, it can also be seen that most postcode sectors with $5\% - 10\%$ of the population working in agriculture are in the northern part of Scotland. For the percentage of people in households with no car, most postcode sectors in the north have 5-10% of people in households without cars, while percentage of people in households with no car is higher (at least 20%), in the south and central. Most postcode sectors have less than 5% of people in overcrowded households. About 33 postcode sectors have at

least 20% of male unemployed and close to half of the postcode sectors has $5\% - 10\%$ of male unemployed. Most postcode sectors have at least 15% of people in households with low social class, and these are found all over Scotland.

The maps of susceptibility in Figures 3.21 and 3.22 for pre-school and primary 1 and 2 children respectively, for the logistic model with census variables, show the same spatial pattern over time as maps with no census variables, except that for each group and each year the number of postcode sectors with high susceptibility is decreased when the census variables are in the model, thus there is slight evidence that the explanatory variables explain some of the spatial differences in susceptibility to measles.

Tables 3.6 and 3.7 give the results of the fitted logistic model with significant census variables for pre-school and primary 1 and 2 respectively. It is observed that the signs for the estimated coefficients of the census variables do not change over the years, indicating that the directions of the relationships between the census variables and susceptibility do not change over time, except for percentage of people in overcrowded households in 2004 and 2005 for pre-school when the effect is almost zero.

The effect of having a larger percentage of the population born in other EU countries, larger percentage of people working in agriculture, larger percentage of people in households with no car, and larger percentage of unemployed males is always positive for both pre-school and primary 1 and 2 school children. An increase in the percentage of each of these variables is associated with an increase in susceptibility. For percentage of the population born in other EU countries, there is an effect for all the cohorts except for 2005 as the credible interval for that year spans zero, For the percentage of people working in agriculture, there is an effect for all cohorts except for pre-school in 2004. For percentage of people in households with no car and percentage unemployed males, there is no effect for 2004 pre-school for both, and no effect for percentage of unemployed males for primary 1 and 2 for 2000 and 2001.

The effect of having a large percentage of people in overcrowded households and large percentage of people in households with low social class is always negative, except for percentage of people in overcrowded households in 2004 and 2005 for pre-school, where this might as well be zero as the credible interval spans zero. Thus an increase in the percentage of these variables results in a decrease in susceptibility.

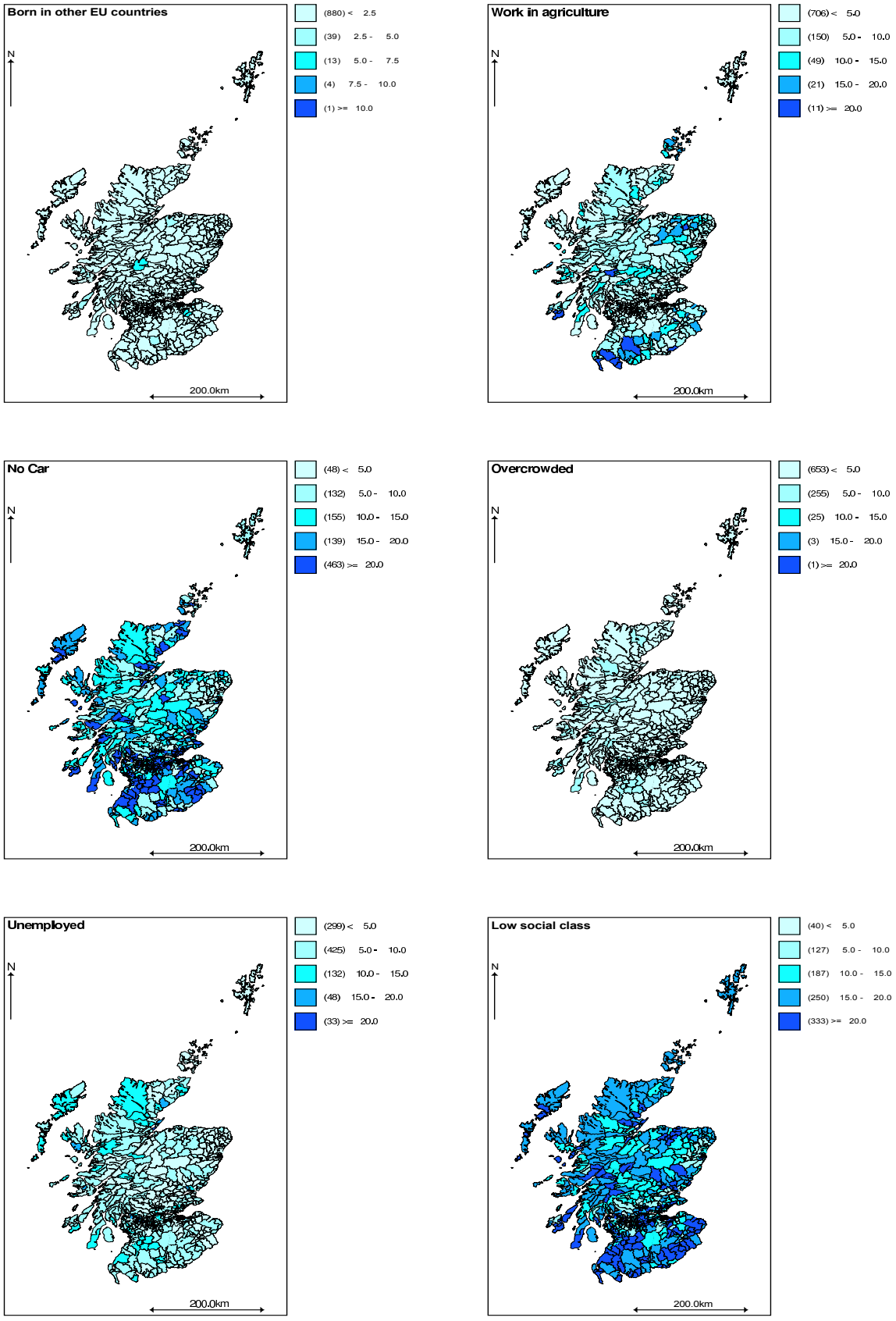

Figure 3.20: Maps of levels of census variables used in the logistic CAR model, for 937 postcode sectors.

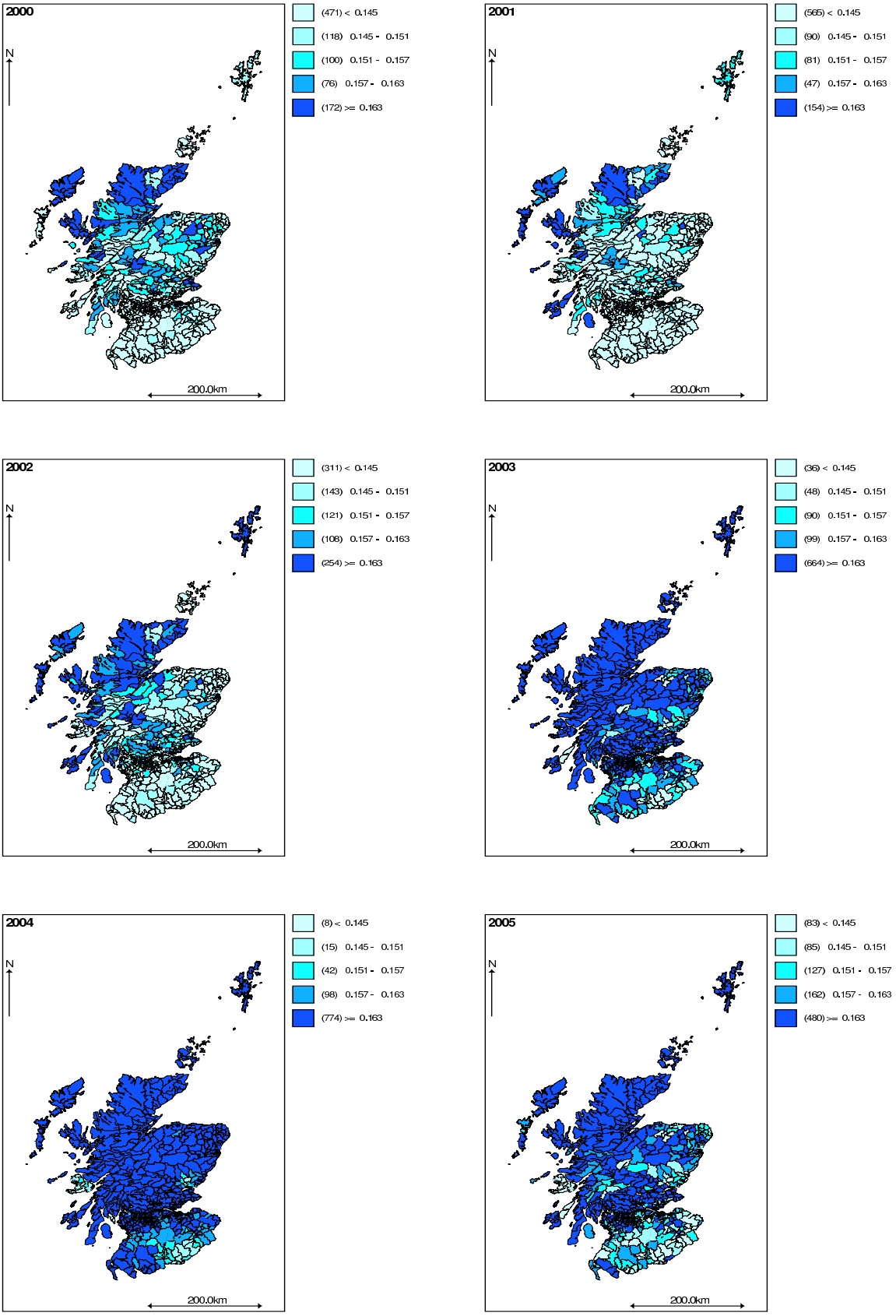

Figure 3.21: Postcode sector level maps of estimated probabilities of children susceptible to measles (pre-school) for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model with census variables).

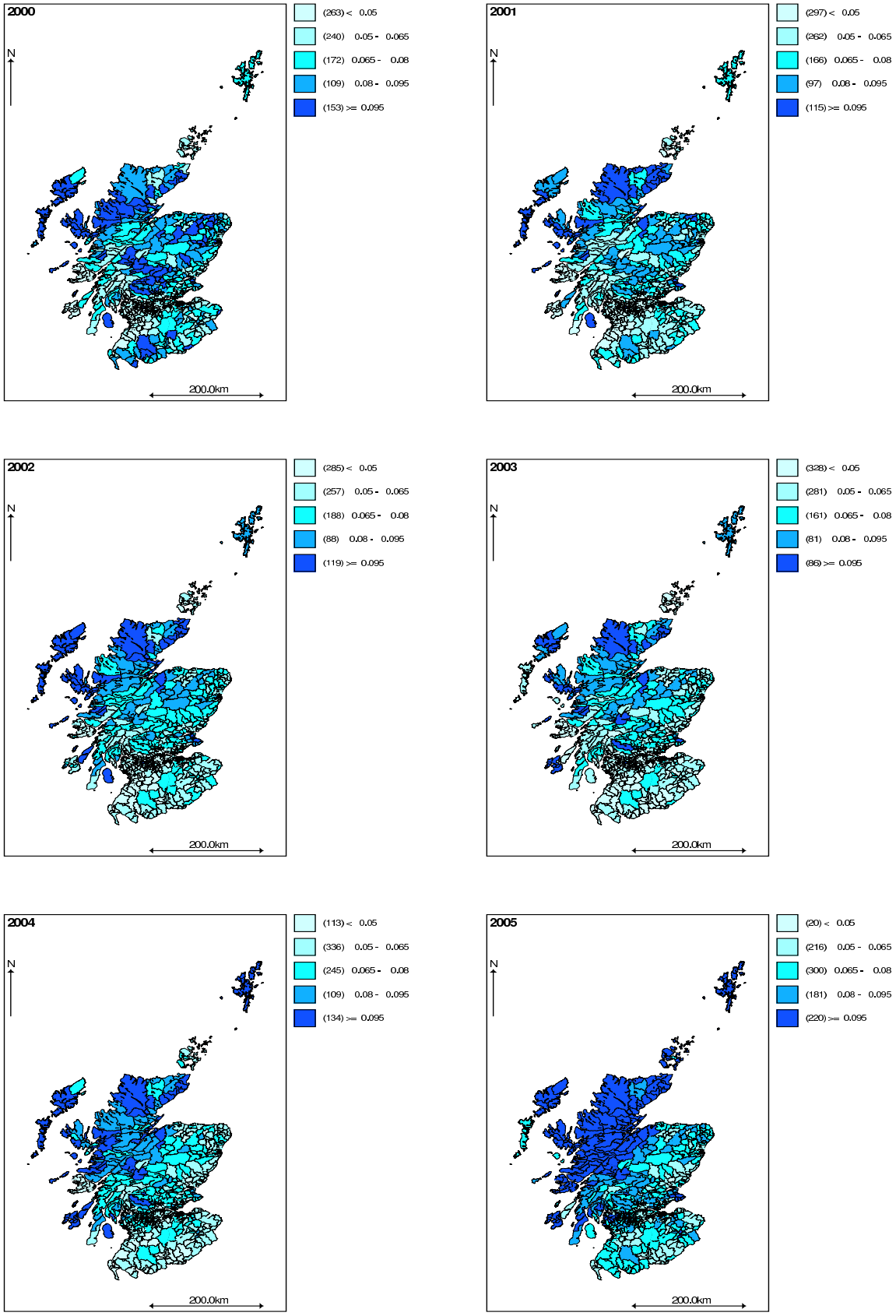

Figure 3.22: Postcode sector level maps of estimated probabilities of children susceptible to measles (primary 1 and 2) for 2000, 2001, 2002, 2003, 2004 and 2005 (logistic model with census variables).

	2000			2001			2002		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	−1.75	−1.77	−1.73	−1.79	−1.81	−1.77	−1.71	−1.73	−1.69
σ_u	0.08	0.06	0.11	0.08	0.05	0.11	0.09	0.06	0.13
σ_v	0.03	0.02	0.06	0.03	0.02	0.06	0.03	0.02	0.05
% born other EU	0.03	0.01	0.05	0.03	0.005	0.05	0.02	0.00001	0.04
% working in agriculture	0.04	0.02	0.07	0.04	0.02	0.07	0.04	0.01	0.06
% no car	0.15	0.10	0.19	0.16	0.11	0.20	0.13	0.08	0.18
% overcrowded	−0.02	−0.05	0.01	−0.02	−0.05	0.008	−0.005	−0.04	0.03
% unemployed	0.05	0.001	0.10	0.06	0.02	0.11	0.08	0.03	0.12
% low social class	−0.13	−0.17	−0.10	−0.14	−0.18	−0.11	−0.15	−0.19	−0.11
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	−1.53	−1.55	−1.51	−1.59	−1.63	−1.56	−1.63	−1.65	−1.60
σ_u	0.11	0.08	0.15	0.14	0.09	0.19	0.07	0.05	0.10
σ_v	0.03	0.02	0.05	0.04	0.02	0.08	0.03	0.02	0.05
% born other EU	0.02	0.002	0.04	0.04	0.01	0.08	0.02	−0.005	0.04
% working in agriculture	0.03	0.005	0.06	0.007	−0.03	0.04	0.04	0.009	0.06
% no car	0.10	0.05	0.15	0.04	−0.02	0.09	0.09	0.04	0.13
% overcrowded	−0.01	−0.04	0.02	0.002	−0.03	0.04	0.004	−0.02	0.03
% unemployed	0.103	0.06	0.15	0.02	−0.03	0.06	0.06	0.02	0.11
% low social class	−0.16	−0.20	−0.13	−0.05	−0.09	−0.009	−0.15	−0.19	−0.12

Table 3.6: Posterior mean parameter estimates for significant census variables (log odds) and standard deviation due to uncorrelated and correlated heterogeneity with their lower (LCL) and upper (UCL) 95% credible limits for pre-school children, for 2000-2005, for 937 postcode sectors.

For pre-school and for primary 1 and 2 in 2000, the effect of percentage of people in overcrowded households is not very large as this variable might as well be zero looking at the 95% credible limits for all the years.

Percentage of people born in other EU countries, percentage of people working in agriculture, percentage of people in households with no car, percentage of individuals unemployed, percentage of people in overcrowded households, and percentage of people in households with low social class have standard deviations of 1.11%, 4.44%, 16.20%, 4.86%, 2.65% and 7.50% respectively (calculated from the census data).

As the census variables were standardised (Section 3.8), for pre-school children, as percentage of the population born in other EU countries in a sector changes by 1 standard deviation, log odds of susceptibility changes by 0.02 to 0.04 across the years (Table 3.6). For a unit change in percentage of the population born in other EU countries

	2000			2001			2002		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	−3.30	−3.58	−3.01	−2.76	−2.80	−2.72	−2.74	−2.77	−2.70
σ_u	0.38	0.28	0.47	0.31	0.23	0.39	0.26	0.19	0.34
σ_v	0.12	0.03	0.20	0.19	0.14	0.24	0.17	0.11	0.22
% born other EU	0.04	0.003	0.07	0.04	0.004	0.08	0.04	0.007	0.08
% working in agriculture	0.14	0.10	0.19	0.11	0.06	0.15	0.08	0.04	0.13
% no car	0.32	0.23	0.41	0.33	0.24	0.42	0.27	0.18	0.35
% overcrowded	−0.04	−0.10	0.02	−0.07	−0.13	−0.01	−0.07	−0.13	−0.02
% unemployed	0.08	−0.003	0.16	0.07	−0.01	0.16	0.11	0.03	0.19
% low social class	−0.31	−0.39	−0.24	−0.29	−0.36	−0.22	−0.26	−0.33	−0.19
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	−2.78	−2.82	−2.75	−2.63	−2.67	−2.60	−2.46	−2.49	−2.43
σ_u	0.23	0.16	0.32	0.30	0.22	0.39	0.33	0.27	0.39
σ_v	0.19	0.15	0.24	0.11	0.03	0.19	0.04	0.02	0.09
% born other EU	0.05	0.02	0.09	0.05	0.02	0.09	0.03	-0.005	0.06
% working in agriculture	0.09	0.05	0.14	0.06	0.02	0.11	0.05	0.01	0.10
% no car	0.24	0.16	0.33	0.18	0.09	0.26	0.13	0.05	0.21
% overcrowded	−0.06	−0.12	−0.01	−0.06	−0.11	−0.004	−0.07	−0.12	−0.02
% unemployed	0.12	0.04	0.21	0.15	0.07	0.22	0.14	0.06	0.21
% low social class	−0.27	−0.34	−0.20	−0.26	−0.32	−0.19	−0.23	−0.30	−0.17

Table 3.7: Posterior mean parameter estimates for significant census variables (log odds) and standard deviation due to uncorrelated and correlated heterogeneity with their lower (LCL) and upper (UCL) 95% credible intervals for primary 1 and 2 school children, for 2000-2005, for 937 postcode sectors.

it is predicted that the log odds of susceptibility changes by a factor of $\frac{0.02}{1.11} = 0.0180$ to $\frac{0.04}{1.11} = 0.0360$ and the odds by $e^{0.0180} = 1.0182$ to $e^{0.0360} = 1.0367$. For a 1 standard deviation change in percentage of people working in agriculture, the log odds of susceptibility changes by 0.007 to 0.04. Thus for a unit change in percentage of people working in agriculture the log odds of susceptibility changes by $\frac{0.007}{4.44} = 0.00158$ to $\frac{0.04}{4.44} = 0.00901$ and the odds by $e^{0.00158} = 1.00158$ to $e^{0.00901} = 1.00905$. For a 1 standard deviation increase in percentage of people in households with no car, log odds of susceptibility changes by 0.04 to 0.16. For a unit change in percentage of people in households with no car the log odds of susceptibility is expected to change by $\frac{0.04}{16.20} = 0.00247$ to $\frac{0.16}{16.20} = 0.00988$ and the odds by $e^{0.00247} = 1.00247$ to $e^{0.00988} = 1.00993$. As percentage of people in overcrowded households changes by 1 standard deviation, log odds of susceptibility changes by -0.005 to 0.004. Thus for a unit change in percentage of people in overcrowded households, log odds of sus-

ceptibility changes by $\frac{-0.005}{2.65} = -0.00189$ to $\frac{0.004}{2.65} = 0.00151$ and odds changes by $e^{-0.00189} = 0.998$ to $e^{0.00151} = 1.00151$. As percentage of males unemployed increases by 1 standard deviation, log odds of susceptibility changes by 0.02 to 0.10. Thus for a unit change in percentage unemployed males the log odds of susceptibility changes by $\frac{0.02}{4.86} = 0.00412$ to $\frac{0.10}{4.86} = 0.0206$ and the odds by $e^{0.00412} = 1.00413$ to $e^{0.0206} = 1.0208$. As percentage of people in households with low social class changes by 1 standard deviation, log odds susceptibility changes by -0.16 to -0.05. Thus for a unit change in percentage of people in households with low social class, log odds of susceptibility changes by $\frac{-0.16}{7.5} = -0.0213$ to $\frac{-0.05}{7.5} = -0.00667$ and odds changes by $e^{-0.0213} = 0.979$ to $e^{-0.00667} = 0.993$.

For primary 1 and 2 school children, for a 1 standard deviation increase in percentage of the population born in other EU countries, log odds changes by 0.03 to 0.05 across the years (Table 3.7). For a unit change in percentage of the population born in other EU countries, the log odds of susceptibility changes by $\frac{0.03}{1.11} = 0.0270$ to $\frac{0.05}{1.11} = 0.0450$ and the odds by $e^{0.0270} = 1.0274$ to $e^{0.0450} = 1.0460$. As percentage of people working in agriculture increases by 1 standard deviation, log odds of susceptibility changes by 0.05 to 0.14. Thus for a unit change in percentage of people working in agriculture, the log odds of susceptibility changes by $\frac{0.05}{4.44} = 0.0113$ to $\frac{0.14}{4.44} = 0.0315$ and the odds by $e^{0.0113} = 1.0114$ to $e^{0.0315} = 1.0320$. For a 1 standard deviation increase in percentage of people in households with no car, log odds of susceptibility changes by 0.13 to 0.33. For a unit change in percentage of people in households with no car, the log odds of susceptibility is expected to change by $\frac{0.13}{16.20} = 0.00802$ to $\frac{0.33}{16.20} = 0.0204$ and the odds by $e^{0.00802} = 1.00805$ to $e^{0.0204} = 1.0206$. As percentage of people in overcrowded households increases by 1 standard deviation, log odds of susceptibility changes by -0.07 to -0.04. Thus for a unit change in percentage of people in overcrowded households, log odds of susceptibility changes by $\frac{-0.07}{2.65} = -0.0264$ to $\frac{-0.04}{2.65} = -0.0151$ and odds changes by $e^{-0.0264} = 0.974$ to $e^{-0.0151} = 0.985$. As percentage of males unemployed increases by 1 standard deviation, log odds of susceptibility changes by 0.007 to 0.15. Thus for a unit change in percentage of males unemployed the log odds of susceptibility changes by $\frac{0.007}{4.86} = 0.00144$ to $\frac{0.15}{4.86} = 0.0309$ and the odds by $e^{0.00144} = 1.00144$ to $e^{0.0309} = 1.0314$. As percentage of people in households with low social class increases by 1 standard deviation, log odds of susceptibility changes by -0.31 to -0.2. For a unit change in percentage of people in households with low social class, log odds of

susceptibility changes by $\frac{-0.31}{7.5} = -0.0396$ to $\frac{-0.23}{7.5} = -0.0294$ and odds changes by $e^{-0.0396} = 0.961$ to $e^{-0.0294} = 0.971$. Table 3.8 shows a summary of these changes in the log odds and odds ratios.

Percentage of people in households with no car and percentage of people in households with low social class have higher absolute values of the coefficients than other census variables in the model for both pre-school and primary 1 and 2 children. This reflects the work of Friederichs *et al.* (2006). The relationship between susceptibility and these variables is much stronger than the relationship with other variables. For all

	Pre-School		Primary School	
Variable	Log Odds	Odds	Log Odds	Odds
% born other EU	0.018 to 0.036	1.018 to 1.037	0.027 to 0.045	1.027 to 1.046
% working in agriculture	0.002 to 0.009	1.002 to 1.009	0.011 to 0.032	1.011 to 1.032
% no car	0.002 to 0.010	1.002 to 1.010	0.008 to 0.020	1.008 to 1.021
% overcrowded	−0.002 to 0.002	0.998 to 1.002	−0.026 to −0.015	0.974 to 0.985
% unemployed	0.004 to 0.021	1.004 to 1.021	0.001 to 0.031	1.001 to 1.031
% low social class	−0.021 to −0.007	0.979 to 0.993	−0.040 to −0.029	0.961 to 0.971

Table 3.8: Range of effects on log odds and odds for a unit change in each of the census variables.

cohorts in both pre-school and primary 1 and 2 children, the magnitude of coefficients is higher (in absolute value) in primary 1 and 2 children than in pre-school children. This suggests that the relationship between susceptibility and all the census variables in the model is stronger in primary 1 and 2 than in pre-school children.

It is observed from Tables 3.4 (no census variables), and Tables 3.7 and 3.8 (census variable), that the inclusion of census variables in the model decreases the variability due to clustering for all years for both pre-school (0.11-0.16 reduced to 0.07-0.14) and primary 1 and 2 (0.37-0.54 reduced to 0.23-0.38) children, but the variability due to clustering pattern over the years remains the same as for the model without census variables. Inclusion of variables does not change the unstructured variability for pre-school but there is a slight increase for primary 1 and 2.

3.9 Comparing District and Postcode Sector Levels

Ecological analysis at postcode sector level (small areas) may be expected to be closer to analysis at individual level when compared to district level (large areas). Thus at postcode sector level there might be less ecological bias but more measurement error than at district level.

The census variables which were fitted at district level were also fitted at postcode sector level (Tables 3.10 and 3.11) for a direct comparison of postcode sectors and districts (Tables 3.5 and 3.6). These census variables were percentage of unemployment males and percentage of people in low social class households for pre-school children, and percentage of people born in other EU countries and percentage of people in overcrowded households for primary 1 and 2 children.

For pre-school, the higher the percentage unemployed males the higher the susceptibility at both district and postcode sector level. Susceptibility decreases with increase in percentage of people in low social class households at both district and postcode sector level, except for 2005 at district level when susceptibility seems to increase with increase in percentage of people in low social class households.

	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-1.76	-1.78	-1.74	-1.80	-1.82	-1.78	-1.71	-1.74	-1.69
σ_u	0.12	0.09	0.16	0.12	0.09	0.15	0.13	0.09	0.16
σ_v	0.03	0.02	0.06	0.03	0.02	0.06	0.03	0.02	0.05
% unemployed	0.13	0.09	0.17	0.15	0.11	0.18	0.15	0.11	0.19
% low social class	-0.11	-0.14	-0.08	-0.12	-0.15	-0.08	-0.12	-0.15	-0.09
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-1.54	-1.56	-1.52	-1.57	-1.62	-1.53	-1.63	-1.65	-1.61
σ_u	0.13	0.10	0.17	0.14	0.09	0.19	0.08	0.05	0.11
σ_v	0.03	0.02	0.05	0.04	0.02	0.09	0.03	0.02	0.05
% unemployed	0.15	0.12	0.19	0.04	0.003	0.08	0.11	0.08	0.15
% low social class	-0.14	-0.18	-0.11	-0.05	-0.09	-0.02	-0.13	-0.16	-0.10

Table 3.9: Posterior mean parameter estimates for significant census variables (log odds ratios) and standard deviation due to uncorrelated and correlated heterogeneity, with their lower (LCL) and upper (UCL) 95% credible limits for pre-school children for 2000-2005 and 937 postcode sectors.

	2000			2001			2002		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-2.94	-3.11	-2.84	-2.78	-2.81	-2.73	-2.74	-2.78	-2.71
σ_u	0.47	0.35	0.58	0.42	0.32	0.51	0.36	0.28	0.44
σ_v	0.13	0.02	0.23	0.20	0.13	0.26	0.18	0.10	0.24
% born of other EU	0.06	0.02	0.10	0.06	0.02	0.10	0.06	0.02	0.10
% overcrowded	-0.06	-0.11	-0.02	-0.06	-0.11	-0.02	-0.06	-0.10	-0.02
	2003			2004			2005		
Parameter	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL
α	-2.80	-2.84	-2.77	-2.64	-2.68	-2.61	-2.47	-2.50	-2.45
σ_u	0.33	0.26	0.41	0.37	0.29	0.45	0.38	0.32	0.44
σ_v	0.21	0.15	0.26	0.12	0.03	0.19	0.05	0.02	0.11
% born of other EU	0.07	0.03	0.11	0.07	0.03	0.10	0.04	0.009	0.08
% overcrowded	-0.06	-0.10	-0.02	-0.07	-0.11	-0.03	-0.10	-0.13	-0.06

Table 3.10: Posterior mean parameter estimates for significant census variables (log odds ratios) and standard deviation due to uncorrelated and correlated heterogeneity, with their 95% lower (LCL) and upper (UCL) credible limits for primary 1 and 2 school children for 2000-2005 and 937 postcode sectors.

For primary 1 and 2 children, susceptibility increases with increase in percentage of people born in other EU countries at both district and postcode sector level. Susceptibility decreases with increase in percentage of people in overcrowded households at postcode sector level, while at district level it varies over the years, i.e. susceptibility decreases with increase in percentage of people in overcrowded households for 2000, 2001 and 2005, but for 2003 and 2004 susceptibility increases with increasing percentage of people in overcrowded households.

For all years in both pre-school and primary 1 and 2 groups, the magnitude of all coefficients is higher (in absolute value) at postcode sector level than district level, thus the relationship between susceptibility and the census variables is stronger at postcode sector level than district level. In fact, the 95% credible intervals for these census variables for both pre-school primary 1 and 2 groups at district level all include zero, suggesting that the effects of these variables are not significantly different from zero, but this is not the case for postcode sectors.

3.10 Discussion

In this chapter we first used the pre-school 1999 data set to illustrate the fitting of the empirical and full Bayesian methods, using the commonly used disease models, i.e. the Poisson-Gamma model and the lognormal model. These data were also used to compare the log-normal and logistic models which both included the spatial term. There was little difference in the way the two models fitted the data, the logistic model was chosen for analysis of the data throughout the chapter.

In model fitting we compared the empirical and full Bayesian modelling (for both Poisson-Gamma and log-normal models) based on the 1999 pre-school data. The results indicated that for these data smoothing with empirical Bayes is not very different from full Bayesian smoothing.

To compare susceptibility over time, and compare susceptibility to measles of pre-school to that of primary 1 and 2 school children, at district level (larger regions) and postcode sector (smaller regions) level, a logistic model was fitted to each of the pre-school and primary 1 and 2 school children groups, at each time point (2000-2005). The space-time model (Waller *et al.* (1997)) was also fitted to each of pre-school and primary 1 and 2 school children groups. For pre-school and primary 1 and 2, an ecological model, i.e. a logistic model including census variables, was fitted at each time point, at district and postcode sector level, to determine which census variables will help predict higher susceptibility to measles. As the interest is in seeing if the explanatory variables have the same effect in the different years, the space-time model (Waller *et al.* (1997)) that includes census variables was not fitted. Fitting of models at district and postcode sector level enabled comparison of measles susceptibility at these two levels.

Comparing susceptibility over time, for both district and postcode sector levels, susceptibility was found to be increasing over the years for pre-school, with 2003 and 2004 being the highest, and a decrease was seen in 2005. For primary 1 and 2 school children, the change in susceptibility was not very much until a higher increase in 2005 (this year is the same birth cohort as for pre-school 2003). For both district and postcode sectors, there is greater geographical variation for primary 1 and 2 school children than pre-school children. Since postcode sectors are smaller than districts,

more clusters can be observed, thus the variation due to clustering is higher at postcode level.

The census variables which were considered are the four components of deprivation, which are: percentage of people in households with no car, percentage of people in overcrowded households, percentage of people in households with low social class and percentage of unemployed males. These were kept separate in the analysis to check if these components had similar effects, and we also included other variables, i.e. percentage of children aged between 0-4, percentage of lone parent households, percentage of people born in other European Union countries (other than UK), percentage of people born elsewhere (other than EU), and percentage of people working in agriculture, percentage of people with no qualifications, and percentage of people with high qualifications. At district level, very few census variables were found to be significant compared to postcode sector level, and their effect was nonsignificant, i.e. the credible interval spanned zero.

At postcode sector level, some of the variation in susceptibility is explained by people born in other EU countries (immigration) and people working in agriculture (rurality), and all four components of deprivation (people with no car, overcrowded households, low social class and unemployment). The effects of these components are different, i.e. susceptibility increases with higher percentage of people in households with no car and percentage of unemployed males, and decreases with higher percentage of people in overcrowded households and percentage of people in low social class households. Generally there are consistent results of explanatory variables over the years, thus data could have been pooled together by fitting a space-time model, but this model will give only one estimate and the year differences are random effects. Also this may not work well for years such as 2004 for pre-school, which has only one significant variable.

Both area and individual level factors may influence the high susceptibility rates, but the weakness of this analysis is that it considered only area level factors, and so care should be taken when interpreting the results, i.e. not to interpret them as individual effects. The other weakness is that the census data are at 2001 which is unchanging for 2005 and of less relevance than in 2000, 2001 and 2002.

This analysis identified regions with high susceptibility, and these are the more rural areas, highly deprived areas and areas with higher number of immigrants. These areas may be at risk of measles outbreak and are the areas that mostly need to be targeted when campaigning for measles, mumps and rubella (MMR) vaccine uptake.

The analysis here has shown that comparing maps over time is essential. For example, public health practitioners would wish to identify pockets of areas with changes in susceptibility from one year to the next. Thus the next chapter will focus on using interpolation methods to compare maps over time. Systematic methods that can be used to compare maps over time will be developed in Chapter 5.

Chapter 4

Smoothing

4.1 Introduction

This chapter looks at some of the most commonly used smoothing or interpolation methods used for smoothing count data to produce isopleth maps (continuous surface maps displayed using contours). Choropleth maps are used to show spatial and temporal patterns (see Chapter 3), but isopleth maps can help strengthen the ability to visually communicate patterns over time (Croner and De Cola, 2001) and aid in map comparison. We therefore consider them in this chapter before considering in Chapter 5 methods specifically developed to compare maps. A disadvantage of choropleth maps is that the grey scale grouping is arbitrary and the choice can affect the interpretation, so they must be interpreted with caution (see Chapter 3). Also, the patchy map may be difficult to interpret due to varying size and shape of the regions, so isopleth maps can be a solution to these problems (Berke, 2004).

In this chapter we focus on isopleth maps and consider various ways to produce them. These may be compared with the choropleth maps in Chapter 3 produced from the same data, of susceptibility to measles of pre-school children obtained from Health Protection Scotland. This has been described in Section 1.5.

There are a number of methods that can be used to interpolate count data. These methods originated with point process data where location is random and here location is fixed. These methods include non-parametric kernel regression smoothers (Waller and Gotway, 2004; Bowman and Azzalini, 1997; Hardle, 1990; Wand and Jones, 1995), and kriging (Cressie, 1993; Waller and Gotway, 2004), which is a geostatistical method

of smoothing, trend polynomial surfaces (Ripley, 1981), splines (De Boor, 1978) and distance-weighted methods (Ripley, 1981). The trend polynomial, splines and distance weighted methods make assumptions that do not take advantage of the spatial structure (Carrat and Valleron, 1992). Here we consider non-parametric kernel regression smoothers, widely used and simpler to use, and kriging, which takes advantage of the spatial structure.

Section 4.2 is a review of nonparametric kernel regression methods and Section 4.3 is a review of kriging methods. In Section 4.4 kernel smoothing and ordinary kriging are used to smooth data and produce isopleth maps. Firstly, pre-school measles susceptibility data at district level is used and methods compared. One method is then chosen and used to smooth primary 1 and 2 data at district level and both pre-school and primary 1 and 2 data at postcode sector level.

4.2 Nonparametric Kernel Regression

Following Waller and Gotway (2004), let $Z_1, Z_2, ...Z_n$ be data from a population with probability distribution $f(\underline{Z}|\underline{\theta})$, available at locations $s_1 = (x_1, y_1), ..., s_n = (x_n, y_n)$ respectively. To estimate the unknown parameters $\underline{\theta} = (\theta_1, ..., \theta_n)$, Brillinger (1990) maximises the weighted log-likelihood of the data

$$\sum_{j=1}^{n} w_{ij} log(f(z_j|\underline{\theta})),$$

where w_{ij} is a weight dependent on distance between the centroids of locations s_i and s_j. For different problems and different distributions we obtain different estimators. Brillinger (1990) obtained an estimator of $\underline{\theta}$ for the case when data are normally distributed with mean θ_i and variance σ^2 as

$$\hat{\theta}_i = \frac{\sum_{j=1}^{n} w_{ij} Z_j}{\sum_{j=1}^{n} w_{ij}},\ i = 1, ..., n. \tag{4.1}$$

This is a locally weighted mean at centroid i, obtained as a result of minimising

$$\sum_{j=1}^{n} w_{ij}(Z_j - \theta_i)^2.$$

For rates $r_i = \frac{Z_i}{n_i}$, where Z_i is Poisson distributed with mean and variance $n_i\theta$, n_i is the population size in region i and θ is the risk of disease, the locally smoothed rate is given by

$$\hat{r}_i = \frac{\sum_{j=1}^{n} w_{ij} Z_j}{\sum_{j=1}^{n} w_{ij} n_j},\ i = 1, ..., n. \tag{4.2}$$

This is a ratio of two smoothers, one smoother applied to the Z_i and one to the n_i. This is advantageous if we are concerned with smoothing variation in population which may occur as a result of sampling or counting errors (Waller and Gotway, 2004). Kafadar (1996) also proposed the use of a ratio smoother.

Different choices of weights (w_{ij}) give different estimators. Here we consider weights based on kernel functions (Waller and Gotway, 2004). Then using weights

$$w_{ij} = K(\tfrac{s_i - s_j}{h})$$

in (4.1) gives the estimator at location s_i as

$$\hat{\theta}_i = \frac{\sum_{j=1}^{n} K(\frac{s_i - s_j}{h}) Z_j}{\sum_{j=1}^{n} K(\frac{s_i - s_j}{h})}, \tag{4.3}$$

while (4.2) becomes

$$\hat{r}_i = \frac{\sum_{j=1}^{n} K(\frac{s_i - s_j}{h}) Z_j}{\sum_{j=1}^{n} K(\frac{s_i - s_j}{h}) n_j}, \tag{4.4}$$

where $s_i = (x_i, y_i)$ and $s_j = (x_j, y_j)$ are the centroid locations for region i and j, h is the smoothing parameter and $K(\cdot)$ is the kernel function, a bivariate probability density function which is symmetric about the origin and integrates over 1 for the whole domain, e.g. Gaussian (Wand and Jones, 1995). Larger values of h lead to smoother maps, while smaller values lead to less smooth maps. Equation (4.3) is often referred to as the *Nadaraya-Watson kernel estimator*.

Observations near the edges of the study region tend to have fewer local neighbours than interior regions. The resulting smoothed values near the edges are obtained by averaging over smaller number of neighbours than the interior values (Waller and Gotway, 2004). This is referred to as an *edge effect or boundary bias*. The kernel estimators discussed above have the disadvantage of edge effects, thus the bias of this estimator will be less in the interior than at the exterior or edges of the study area. The local linear regression smoothers can provide boundary bias correction (Fan and Gijbels, 1992; Simonoff, 1996).

A local polynomial regression estimator at location s_i minimises with respect to $\beta_m, m = 0, ..., p$

$$\sum_{j=1}^{n} (Z_j - \beta_0 - \beta_1(s_i - s_j) - ... - \beta_p(s_i - s_j)^p)^2 K(\frac{s_i - s_j}{h}), \tag{4.5}$$

where $K(\cdot)$ is a kernel function and h is the bandwith/smoothing parameter, and p is the order of the polynomial. Following Simonoff (1996), define the design matrix $\underline{X}$ as

$$\underline{X} = \begin{pmatrix} 1 & (s_i - s_1) & ... & (s_i - s_1)^p \\ & ... & ... & ... & ... \\ 1 & (s_i - s_n) & ... & (s_i - s_n)^p \end{pmatrix}$$

and the weight matrix as

$$\underline{W} = \frac{1}{h} diag[K(\tfrac{s_i - s_1}{h}), ..., K(\tfrac{s_i - s_n}{h})].$$

Then when $\underline{X}^T(\underline{W})(\underline{X})$ is invertible,

$$\underline{\hat{\beta}} = (\underline{X}^T(\underline{W})(\underline{X}))^{-1}\underline{X}^T(\underline{W})(\underline{Z}).$$

The estimator of β_0 in (4.5) is then given by

$$\hat{\beta}_0 = \underline{y_1}^T(\underline{X}^T(\underline{W})(\underline{X}))^{-1}\underline{X}^T(\underline{W})(\underline{Z}) = (\underline{A})(\underline{Z}) \tag{4.6}$$

where $(\underline{y}_1)$ is the $(p+1) \times 1$ vector having the value 1 in the 1st entry and zero elsewhere.

When p=1, (4.5) becomes

$$\sum_{j=1}^{n}(Z_j - \beta_0 - \beta(s_i - s_j))^2 K(\frac{s_i - s_j}{h}). \tag{4.7}$$

The local linear estimator of θ_i at location s_i is obtained by finding $\hat{\beta}_0$ and $\hat{\beta}$ which minimise (4.7). This can be given an explicit formula (Simonoff, 1996; Bowman and Azzalini, 1997) as

$$\hat{\theta}_i = \frac{1}{nh}\sum_{j=1}^{n}\frac{\{a_2(s_i;h) - a_1(s_i;h)(s_i - s_j)\}K(\frac{s_i - s_j}{h})Z_j}{a_2(s_i;h)a_0(s_i;h) - a_1(s_i;h)^2} \tag{4.8}$$

where $a_r(s_i;h) = \frac{1}{nh}\sum_{j=1}^{n}(s_i - s_j)^r K(\frac{s_i - s_j}{h})$, r= 0, 1 or 2.

The Nadaraya-Watson estimator can be obtained the same way by omitting the β term in (4.7).

4.3 Kriging

Kriging is a geostatistical method used in the earth sciences (Cressie, 1993). This method can be used in disease mapping to interpolate relative risk estimates onto a continuous surface and produce isopleth maps. These maps strengthen the ability to visualise disease patterns over time. Authors who have used kriging in disease mapping include Berke (2004), Croner and De Cola (2001) and Carrat and Valleron (1992).

Kriging has an advantage over other interpolation methods such as splines, distance-weighted methods and trend polynomial surface, as it takes into account the spatial structure of the data (Croner and De Cola (2001); Carrat and Valleron (1992)). The disadvantages are that kriging can produce negative interpolates and heterogeneous variances in the estimates (Berke, 2004). Berke (2004) suggested that empirical Bayes smoothing of the relative risk before using kriging can help address the problem of heterogeneous variances, as empirical Bayes smoothing can shrink unstable estimates to the local or global mean and, by borrowing strength from the neighbours, stabilise the effect on variance. The problem of negative interpolates can also be addressed by choosing an appropriate kriging method (Berke, 2004).

Let $\underline{Z} = (Z_1, Z_2, ..., Z_n)^T$ be the spatial data observed at locations s_i, with coordinates (x_i, y_i), $i = 1, 2, ..., n$. Then kriging assumes that $\underline{Z}$ can be modelled as a stochastic process, within the spatial linear model below. This model takes into account the spatial correlation (Berke (2004); Cressie (1993)):

$$Z(\underline{s}) = \mu(\underline{s}) + \delta(\underline{s}), \quad \delta(\underline{s}) \sim N(0, \Sigma), \tag{4.9}$$

where $\mu(\underline{s}) = E(Z(\underline{s}))$ is the mean structure, $\underline{s} = (s_1, ..., s_n)^T$ is the set of locations, $\delta(\underline{s})$ is the observation at location s of the correlated error process and Σ is a spatially structured variance-covariance matrix. The structure of the spatial variation is estimated by the use of a semi-variogram. The semi-variogram is defined as

$$\gamma(h) = \tfrac{1}{2}E(Z(s+h) - Z(s))^2,$$

where s is a location and $s+h$ is a location at distance h from s. The classical estimator of the semi-variogram for a given lag or distance h is

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{s_i, s_j} (Z(s_i) - Z(s_j))^2, \tag{4.10}$$

where the summation is over the $N(h)$ pairs of locations s_i and s_j at distance h from each other. For irregular spatial location points, we sum over pairs s_i and s_j such that s_i and s_j are between $h - \frac{h}{2}$ and $h + \frac{h}{2}$ distance apart. The computations of $\gamma(h)$ are repeated for $2h, 3h, 4h, ..., kh$. Plotting $\hat{\gamma}(h)$ against h gives the empirical semi-variogram.

There are different types of semi-variogram models which can be fitted to an empirical semi-variogram. The basic models are the linear, spherical and exponential models (Cressie, 1993). The expression for a linear model is:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + b_1 h, & h \neq 0. \end{cases}$$

The spherical model is:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + C_1(1.5\frac{h}{C_2} - 0.5(\frac{h}{C_2})^3), & 0 < h < C_2, \\ C_0 + C_1, & h \geq C_2. \end{cases}$$

The exponential model is:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + C_1(1 - e^{\frac{-h}{C_2}}), & h \neq 0. \end{cases}$$

If $\gamma(h) \to C_0 > 0$ as $h \to 0$, C_0 is called a *nugget effect/variance* (the amount by which the variance differs from zero). This measures the small scale variance, due to measurement error (which occurs when measurement is done several times and different results obtained). $C_0 + C_1$ is called the *sill* and this describes the level of the semi-variogram where the variance no longer increases (the level of an asymptote). There is no sill in the linear model. The spherical model reaches an asymptote sharply, whereas the exponential model levels out slowly. C_2 is called the *range of influence*, and this is the distance between two locations beyond which observations appear independent, i.e. variance no longer increases. This is the distance at which the sill is reached.

When the semi-variogram is estimated without the nugget effect, kriging leads to direct interpolation at the locations, giving predicted residuals equal to model residuals. Predicting with the nugget effect gives smaller residuals, thus a smoother prediction surface (Berke, 2004). The semi-variogram estimates parameters of the theoretical model through a weighted or ordinary least squares technique (Cressie, 1993).

There are different types of kriging which exist as a result of the underlying assumptions and the aims of the analysis (Cressie (1993) and Waller and Gotway (2004)). For example, as a result of the knowledge and estimation of the spatial mean $\mu(s)$ in (4.9), there is *simple kriging* (a linear prediction assuming a known mean), *ordinary kriging* (a linear prediction with a constant unknown mean function) and *universal kriging* (a linear prediction with non-stationary mean). We discuss ordinary and universal kriging (commonly used methods) below.

4.3.1 Ordinary Kriging

In ordinary kriging (Waller and Gotway, 2004; Cressie, 1993) prediction assumes a constant unknown mean and known semi-variogram, which still has to be fitted. If $\underline{Z} = (Z_1, Z_2, ..., Z_n)^T$ is the spatial data observed at location s_i, $i = 1, 2, ..., n$, the value of $\underline{Z}$ at location s_0 can be estimated from the nearest sampling values by model assumption (4.9) and the linear formula

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i); \quad \sum_{i=1}^{n} \lambda_i = 1. \tag{4.11}$$

The estimation of the $\lambda_i s$ is based on using the semi-variogram $\gamma(h)$ and two properties, namely unbiasedness ($E[\hat{Z}(s_0)] = E[Z(s_0)]$) which is guaranteed by the constraint ($\sum_{i=1}^{n} \lambda_i = 1$) and minimising mean squared prediction error (defined as $E[\hat{Z}(s_0) - Z(s_0)]^2$).

The $\lambda_i s$ are found by the method of Lagrange multipliers (Waller and Gotway, 2004). We find $\lambda_1, ..., \lambda_n$ and the Lagrange multiplier ϕ that minimise the function

$$E[(\sum_{i=1}^{n} \lambda_i Z(s_i) - Z(s_0))^2] - 2\phi(\sum_{i=1}^{n} \lambda_i - 1), \tag{4.12}$$

where the second term is minimised when $\sum_{i=1}^{n} \lambda_i = 1$, ensuring unbiasedness.

The constraint $\sum_{i=1}^{n} \lambda_i = 1$ implies that

$$[\sum_{i=1}^{n} \lambda_i Z(s_i) - Z(s_0)]^2 = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j (Z(s_i) - Z(s_j))^2 + \sum_{i=1}^{n} \lambda_i (Z(s_0) - Z(s_i))^2. \tag{4.13}$$

Taking expectations of both sides of (4.13) we obtain

$$E[(\sum_{i=1}^{n} \lambda_i Z(s_i) - Z(s_0))^2] = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j E[(Z(s_i) - Z(s_j))^2] + \sum_{i=1}^{n} \lambda_i E[(Z(s_0) - Z(s_i))^2]. \tag{4.14}$$

so that (4.12) becomes

$$-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(s_i-s_j)+2\sum_{i=1}^{n}\lambda_i\gamma(s_0-s_i)-2\phi(\sum_{i=1}^{n}\lambda_i-1). \tag{4.15}$$

Equation (4.15) is minimised by differentiating with respect to $\lambda_1, ..., \lambda_n$ and ϕ in turn and setting the partial derivatives equal to zero to obtain a system of linear equations below, referred to as *ordinary kriging equations*:

$$\sum_{i=1}^{n}\lambda_i\gamma(s_i-s_j)+\phi=\gamma(s_0-s_i), i=1,2,...,n, \quad \sum_{i=1}^{n}\lambda_i=1. \tag{4.16}$$

To obtain the $\lambda_i s$ and ϕ, the system of linear equations above is solved and the results used in equation (4.11) to give the ordinary kriging predictor, where the weights of $\hat{Z}(s_0)$ depend on both the spatial correlations between $Z(s_0)$ and each of the $Z(s_i),\ i=1,...,n$ and the spatial correlations between all pairs $Z(s_i)$ and $Z(s_j)$, $i=1,2,...,n, \quad j=1,2,...,n$.

4.3.2 Universal Kriging

Let $\underline{Z}=(Z_1,Z_2,...,Z_n)^T$ be the spatial data observed at locations s_i, with coordinates (x_i,y_i), $i=1,2,...,n$. Universal kriging assumes that $\mu(\underline{s})$ in equation (4.9) is a non-stationary linear combination of p known functions $f_0(\underline{s}),...,f_p(\underline{s})$. Thus, this kriging method includes covariates. Following Cressie (1993), model (4.9) becomes

$$Z(\underline{s})=\sum_{j=1}^{p+1}f_{j-1}(\underline{s})\beta_{j-1}+\delta(\underline{s}). \tag{4.17}$$

When $p=0$ and $f_0(\underline{s})=1$ we obtain the ordinary kriging model.

Model (4.17) can be written as

$$Z(\underline{s})=(\underline{X})(\underline{\beta})+\delta(\underline{s}) \tag{4.18}$$

where $\underline{\beta}=(\beta_0,...,\beta_p)^T$ is a vector of $p+1$ unknown parameters and $\underline{\mathrm{X}}$ is an $n\times(p+1)$ array with $(i,j)th$ element $f_{j-1}(s_i),\ i=1,...,n,\ j=1,...,p+1$.

The value at location s_0 can be estimated from the nearest sampling values by model assumption (4.17) and the linear formula

$$\hat{Z}(s_0)=\sum_{i=1}^{n}\lambda_i Z(s_i); \quad \underline{\lambda}^T\underline{X}=\underline{x}^T \ \text{ and } \ \sum_{i=1}^{n}\lambda_i=1. \tag{4.19}$$

Again, the estimation of the $\lambda_i s$ is based on using the semi-variogram $\gamma(h)$, unbiasedness $E[\hat{Z}(s_0)] = E[\underline{\lambda}^T \underline{Z}] = \underline{\lambda}^T(\underline{X})(\underline{\beta})$, which is guaranteed by the constraint $\underline{\lambda}^T \underline{X} = \underline{x}^T$, $\underline{\lambda} = (\lambda_1, ..., \lambda_n)^T$, $\underline{x} = (f_0(s_0), ..., f_p(s_0))^T$ and minimising mean squared prediction error (defined as $E(\hat{Z}(s_0) - Z(s_0))^2$).

We find $\lambda_1, ..., \lambda_n$ and Lagrange multipliers $\phi_0, ..., \phi_p$ that minimise the function

$$E[(Z(s_0) - \sum_{i=1}^{n} \lambda_i Z(s_i))^2] - 2\sum_{j=1}^{p+1} \phi_{j-1}[\sum_{i=1}^{n} \lambda_i f_{j-1}(s_i) - f_{j-1}(s_0)]. \tag{4.20}$$

Assuming that $f_o(\mathbf{s})$=1 and using $\sum_{i=1}^{n} \lambda_i = 1$, then

$$[Z(s_0) - \sum_{i=1}^{n} \lambda_i Z(s_i)]^2 = [\underline{x}^T \underline{\beta} + \delta(s_0) - \underline{\lambda}^T(\underline{X})(\underline{\beta}) - \sum_{i=1}^{n} \lambda_i \delta(s_i)]^2 = [\delta(s_0) - \sum_{i=1}^{n} \lambda_i \delta(s_i)]^2 \tag{4.21}$$

which then gives

$$[Z(s_0) - \sum_{i=1}^{n} \lambda_i Z(s_i)]^2 = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j (\delta(s_i) - \delta(s_j))^2 + \sum_{i=1}^{n} \lambda_i (\delta(s_0) - \delta(s_i))^2. \tag{4.22}$$

Taking expectations of both sides of equation (4.22), (4.20) becomes

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(s_i - s_j) + 2\sum_{i=1}^{n} \lambda_i \gamma(s_0 - s_i) - 2\sum_{j=1}^{p+1} \phi_{j-1}[\sum_{i=1}^{n} \lambda_i f_{j-1}(s_i) - f_{j-1}(s_0)]. \tag{4.23}$$

Differentiating (4.23) with respect to $\lambda_1, ..., \lambda_n$ and $\phi_0, ..., \phi_p$ and equating to zero, gives estimators of $\lambda_1, ..., \lambda_n$ and $\phi_0, ..., \phi_p$.

4.4 Comparing Measles Susceptibility Maps

Here we use susceptibility to measles data described in Section 1.5 at district and postcode sector level, for 2000-2005. Firstly, we use kernel smoothing and ordinary kriging to smooth the susceptibility ratios given by observed counts /expected counts for pre-school at district level. Also ordinary kriging is used to smooth the empirical Bayes estimates obtained by smoothing using the log-normal model without a spatial component (see Section 3.3.2). The three methods are compared, and the one that gives good results is used to smooth measles susceptibility data for primary 1 and 2 school children at district level, and for both pre-school and primary 1 and 2 school children at postcode sector level. This analysis may help in comparing maps over time using isopleth maps. It is expected that the isopleth maps will show gradual changes instead of jumps as in the case of choropleth maps.

The locations used are centroids of the districts. For kernel smoothing, the Nadaraya-Watson kernel smoother from package *JLLprod*, function *Blocc*, in R was used, kernel used here is Gaussian, and the smoothing parameter (h) is automatically chosen by the R function. The package *geoR* in R was used to perform kriging.

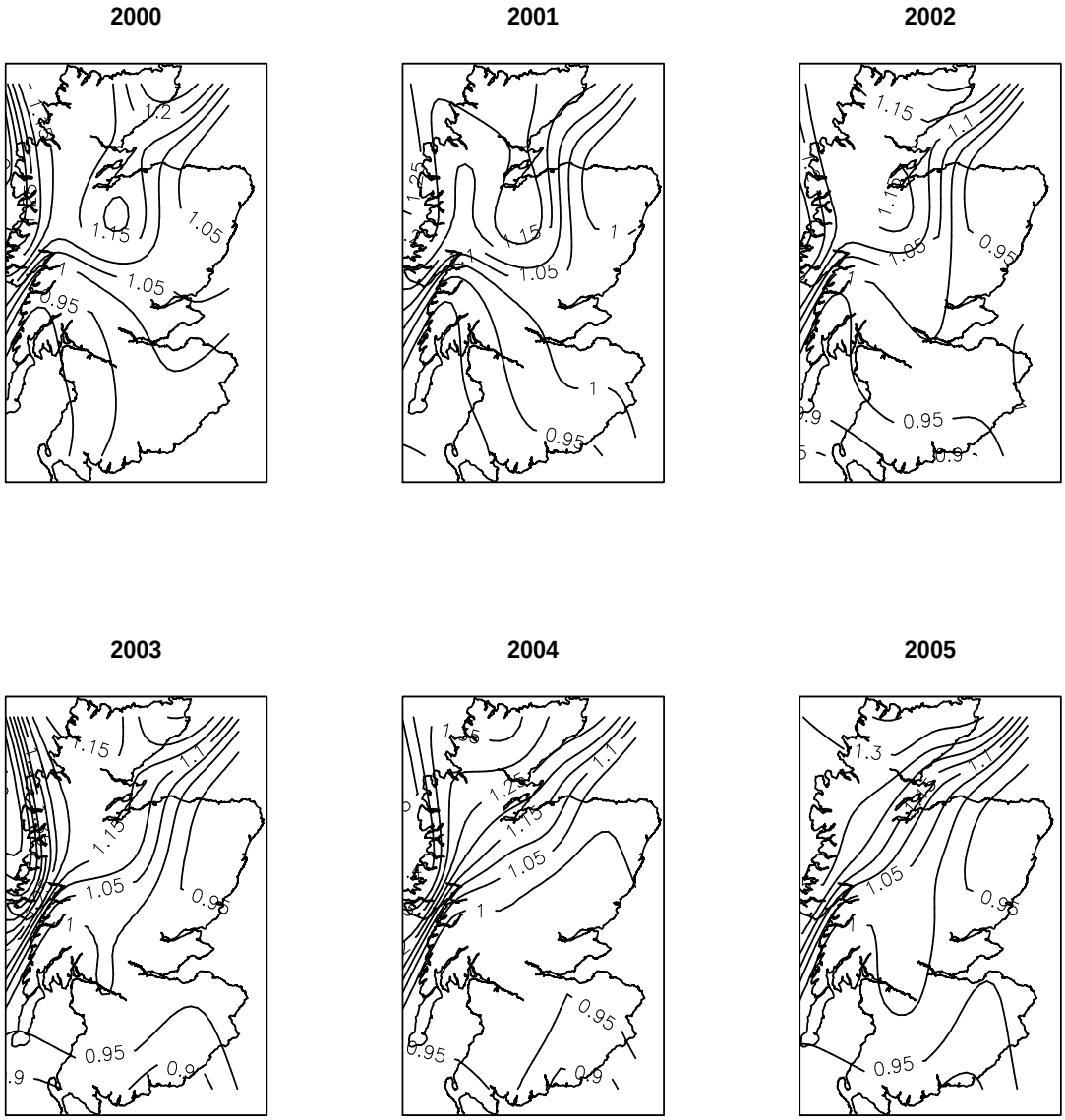

Figure 4.1: Isopleth maps from kernel smoothing of susceptibility ratios (observed/expected) for pre-school children for years 2000-2005, for districts.

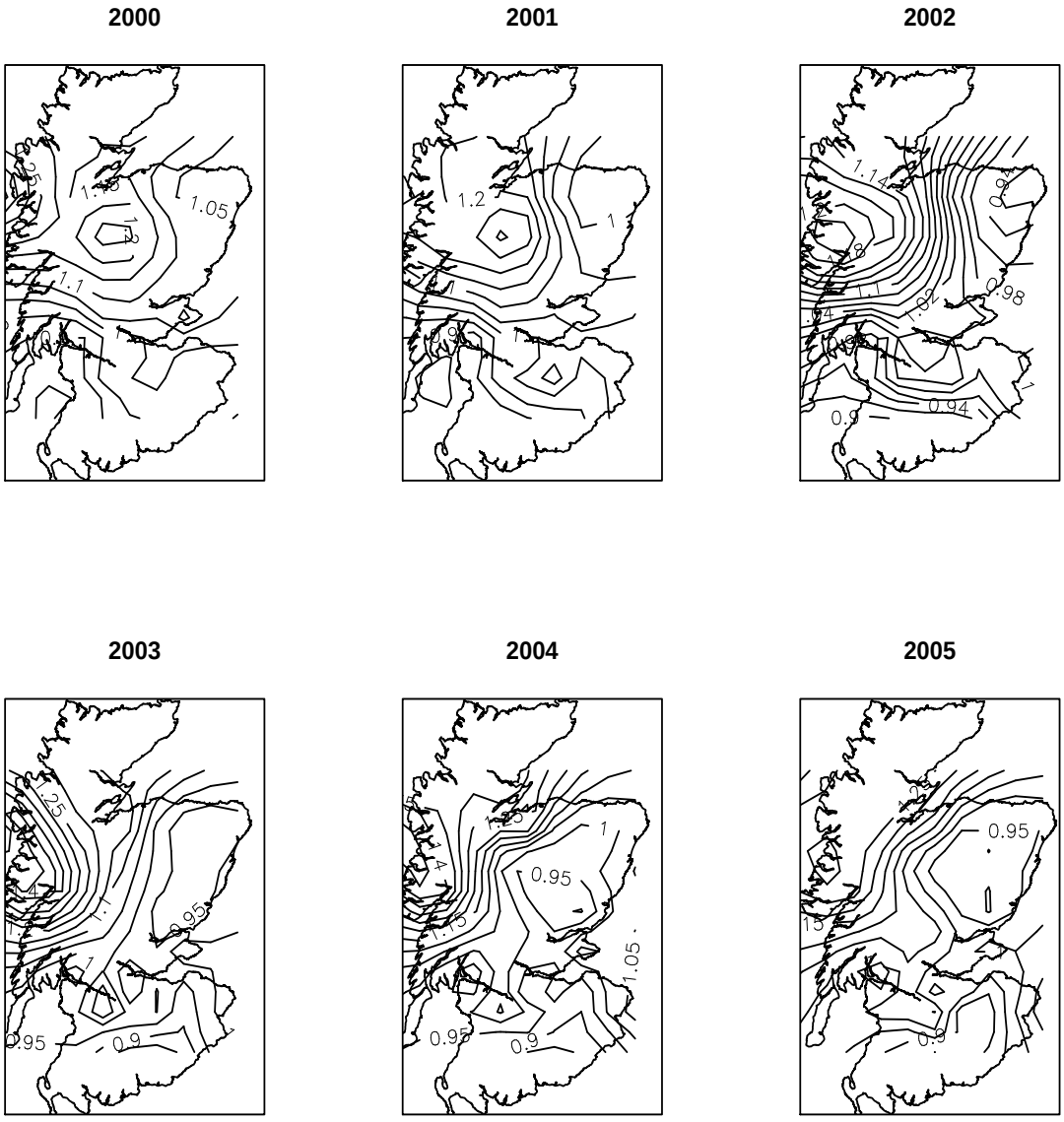

Figure 4.2: Isopleth maps from ordinary kriging of susceptibility ratios (observed/expected) for pre-school children for years 2000-2005, for districts.

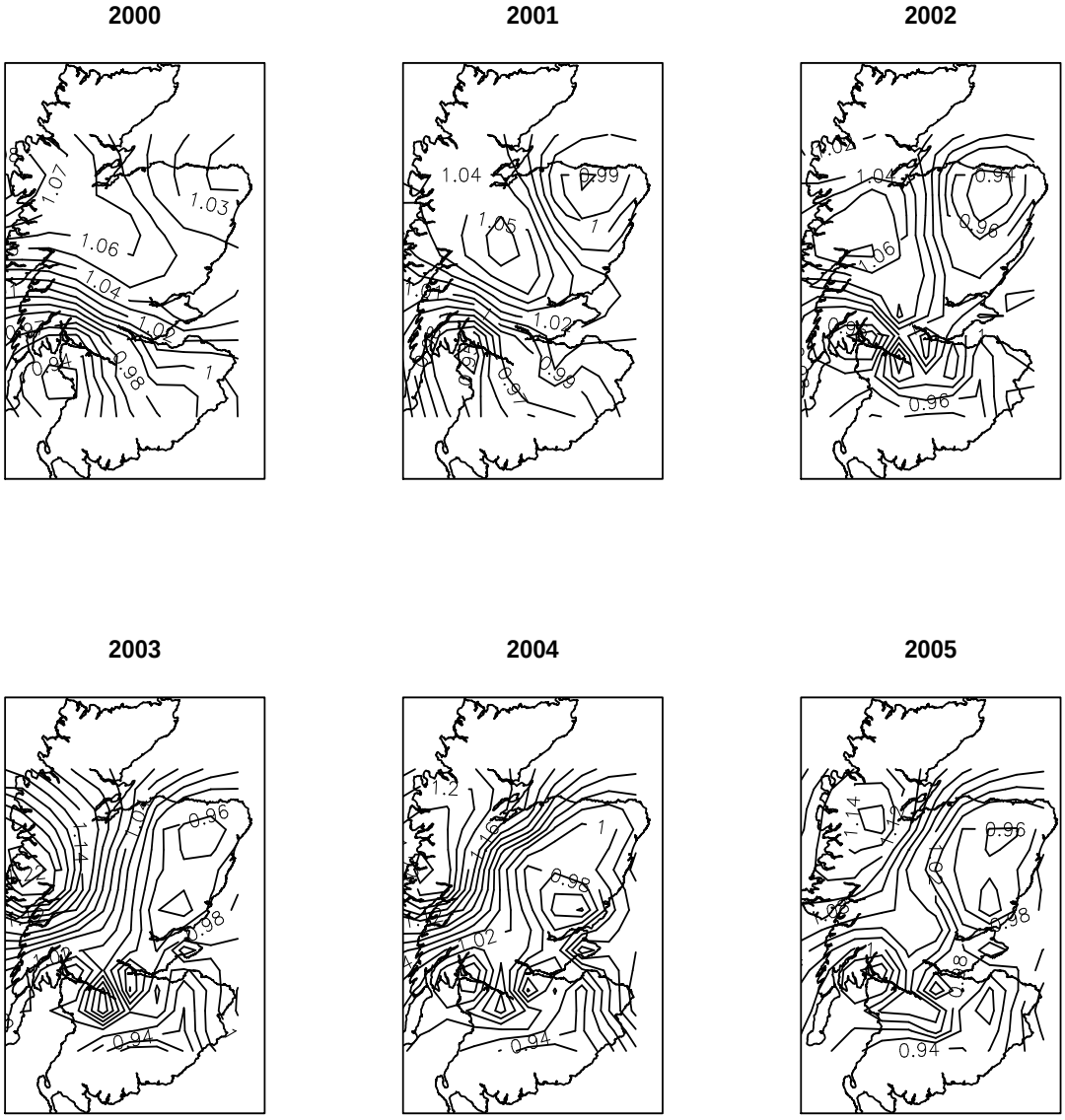

Figure 4.3: Isopleth maps from ordinary kriging of empirical Bayesian estimates for pre-school children for years 2000-2005, for districts.

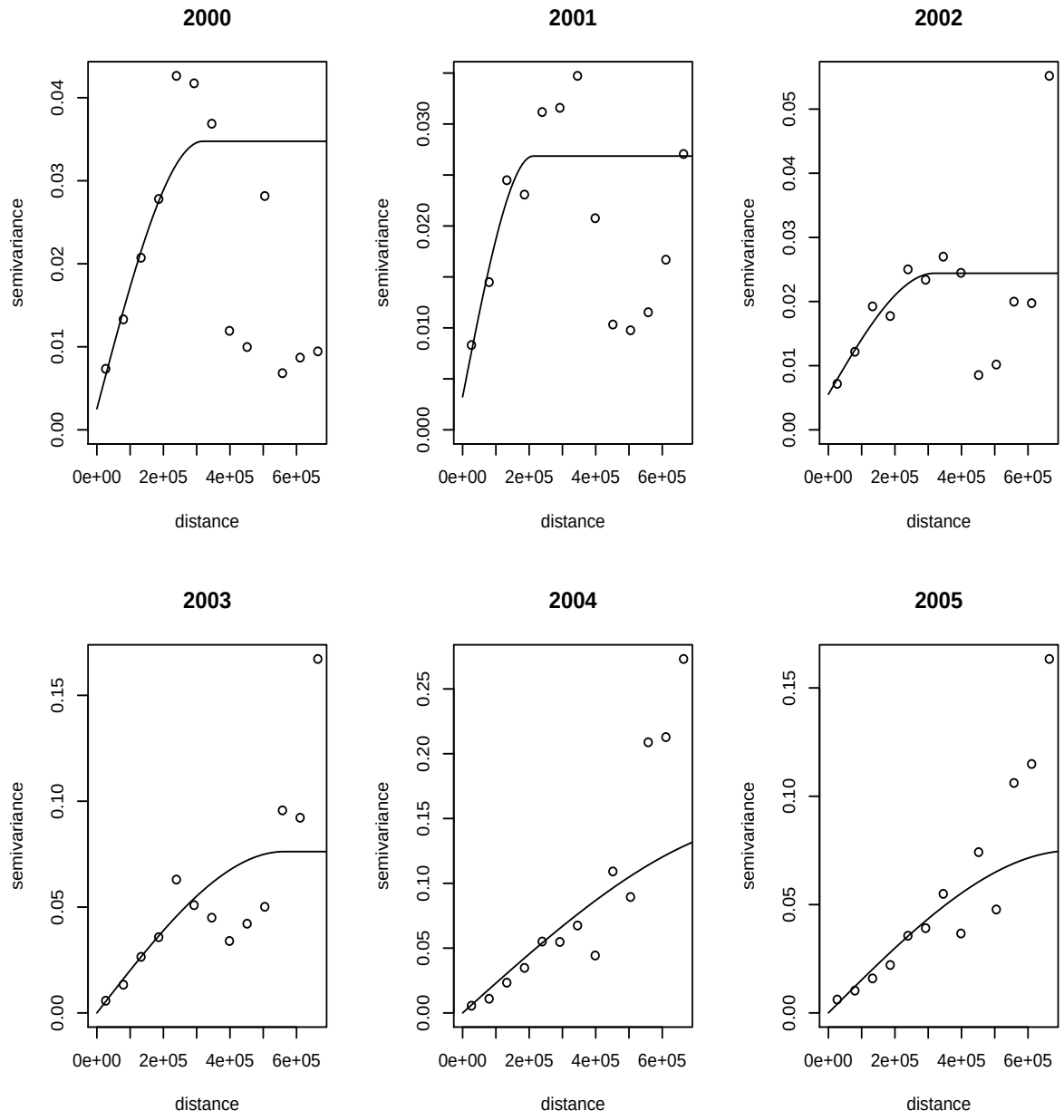

Figure 4.4: Semivariograms and fitted spherical model obtained from ordinary kriging of susceptibility ratios (observed/expected), for pre-school children for years 2000-2005 for districts. Distance is measured in metres.

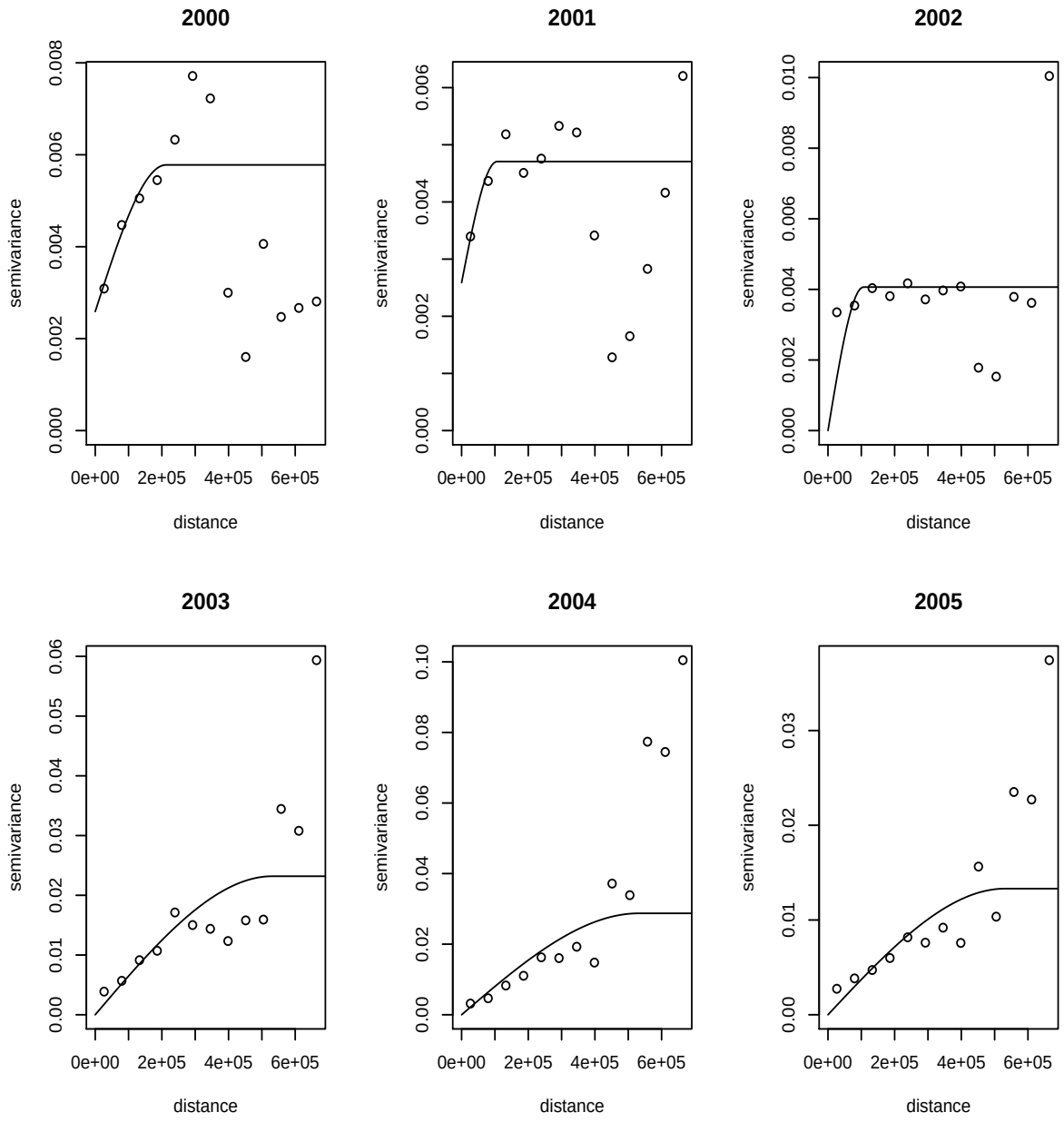

Figure 4.5: Semivariograms and fitted spherical model obtained from ordinary kriging of empirical Bayesian estimates, for pre-school children for years 2000-2005 for districts. Distance measured in metres.

The spatial structure is modelled by a spherical semi-variogram (the most widely used method; see Section 4.3, and fitted by weighted least squares estimation). The codes are given in Appendix C.

Figures 4.1-4.3 show an overlay of isopleth maps of susceptibility to measles on the map of Scotland, for kernel smoothing, ordinary kriging of susceptibility rates and ordinary kriging of empirical Bayesian estimates respectively. It can be observed that generally the three smoothing methods do not give very different results, but rates seem to vary across space more quickly with the kriging methods than kernel smoothing.

Generally for all the years, susceptibility is lower in the southern districts and higher in the north and north west regions. The results are similar to those obtained from the choropleth maps for pre-school, at district level in Chapter 3.

Figures 4.4 and 4.5 show semi-variograms and fitted spherical models for ordinary kriging of raw and smoothed susceptibility rates respectively, and the values for the nugget effect, sill and range are shown in Table 4.1.

	Raw Rates			Smoothed Rates		
Year	nugget variance	sill	range	nugget variance	sill	range
2000	0.0075	0.0322	318716	0.0026	0.0032	212477
2001	0.0032	0.0236	212477	0.0026	0.0021	106239
2002	0.0055	0.0189	318716	0.0000	0.0041	106239
2003	0.0000	0.0762	562849	0.0000	0.0232	531194
2004	0.0000	0.1474	959658	0.0000	0.0287	531671
2005	0.0000	0.0749	732287	0.0000	0.0133	531194

Table 4.1: Table of nugget variance, sill and range in metres for raw susceptibility ratios (observed/expected) and smoothed susceptibility ratios (observed/expected) obtained from ordinary kriging.

The plots indicate that the spherical model does not fit the data very well as most points are further away from the fitted model. A straight line may fit some of the years better. The exponential model was also used and the plots obtained were similar to those for the spherical model, so only the spherical model results are reported here. The semi-variogram describes the spatial correlation between the susceptibility ratios, thus should indicate how different the susceptibility ratios are at various distances

apart. Generally, the values of nugget variance, sill and range obtained from the raw data are larger than those obtained from the smoothed data as expected, as smoothing pulls rates towards the mean, so that variability between the rates decreases. The years 2000 and 2001 for the raw data and the years 2000, 2001 and 2002 for the smoothed data have a nugget variance greater than zero, though the value of the theoretical semivariogram at distance h =0 is strictly zero. These dissimilarities of susceptibility rates at extremely small distances in general may occur due to scale variation and error in measurements, though here we do not have measurements as we are using aggregated values at the centroids.

Generally, the raw data and smoothed data results give similar interpretation, i.e. the value of the sill reduces from 2001 with an increase in 2003 followed by another increase in 2004, and a decrease in 2005. Thus spatial dependence among regions was decreasing from 2000 to 2002, while from 2003 to 2004 spatial dependence among the regions increases, with a decrease in 2005. We note that for the smoothed rates a slight increase started in 2002. The range is larger for higher sill values and lower for lower values sill, i.e. spatial autocorrelation exists for large distances when most regions are clustered or have similar values, and spatial autocorrelation exists for shorter distances when fewer regions are clustered . The higher sill and largest range are in 2003 and 2004, and for these two years, both the isopleth maps and choropleth maps in Chapter 3, showed that most districts had higher susceptibility.

As kernel smoothing and kriging of raw and smoothed the data do not give very different interpretations and kernel smoothing does not require choice of model, kernel smoothing is used to smooth data for primary 1 and 2 school children at district level, and for pre-school and primary 1 and school children at postcode sector level. Isopleth maps are produced for each case and maps compared.

Figure 4.6 shows maps for primary 1 and 2 school children at district level, using kernel smoothing. As for pre-school children (Figure 4.1), lower susceptibility is present in the south and higher susceptibility in the north. The years 2002, 2003 and 2004 seem to be similar in the upper central districts.

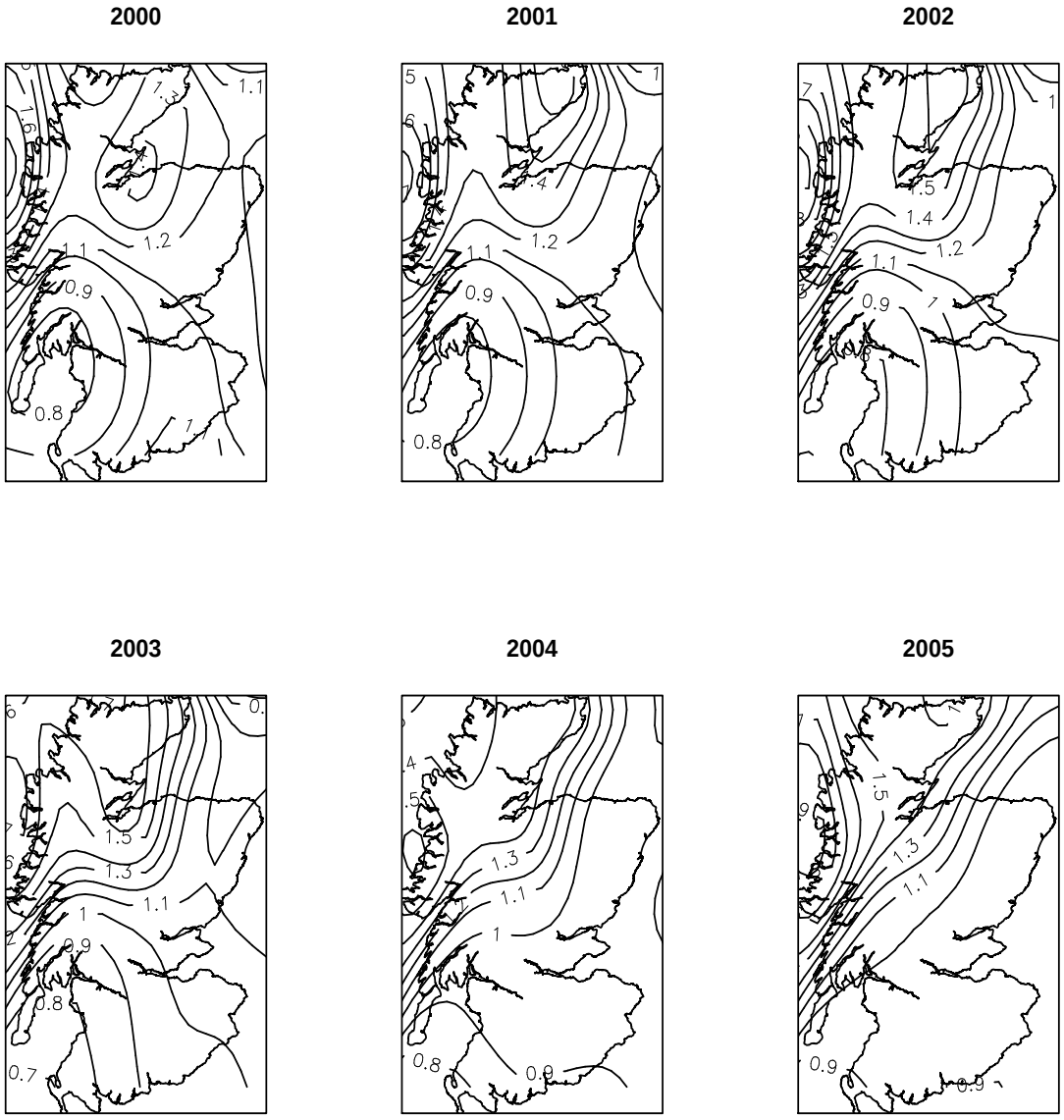

Figure 4.6: Isopleth maps from kernel smoothing for primary 1 and 2 school children for years 2000-2005 for districts.

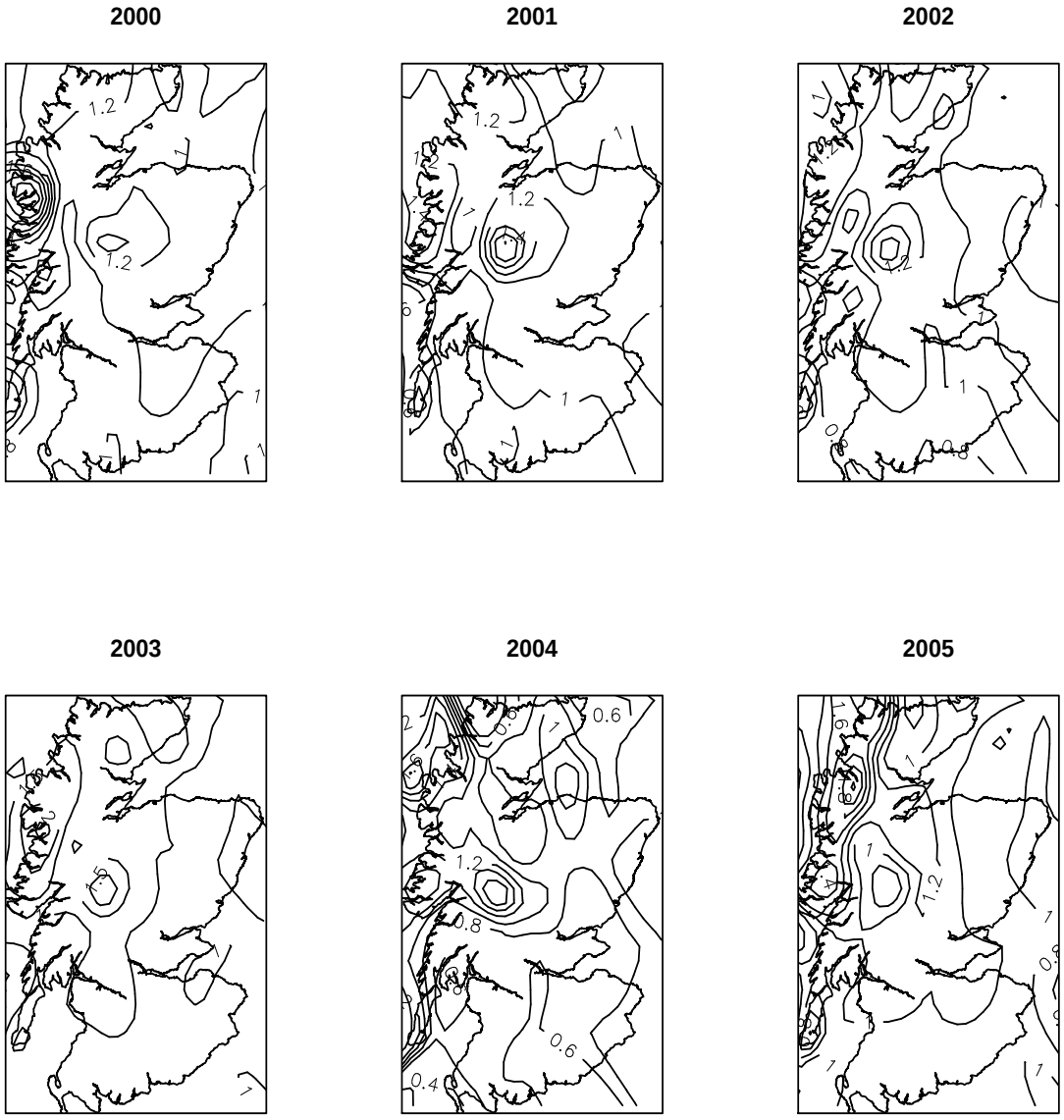

Figure 4.7: Isopleth maps from kernel smoothing for pre-school children for years 2000-2005 for postcode sectors.

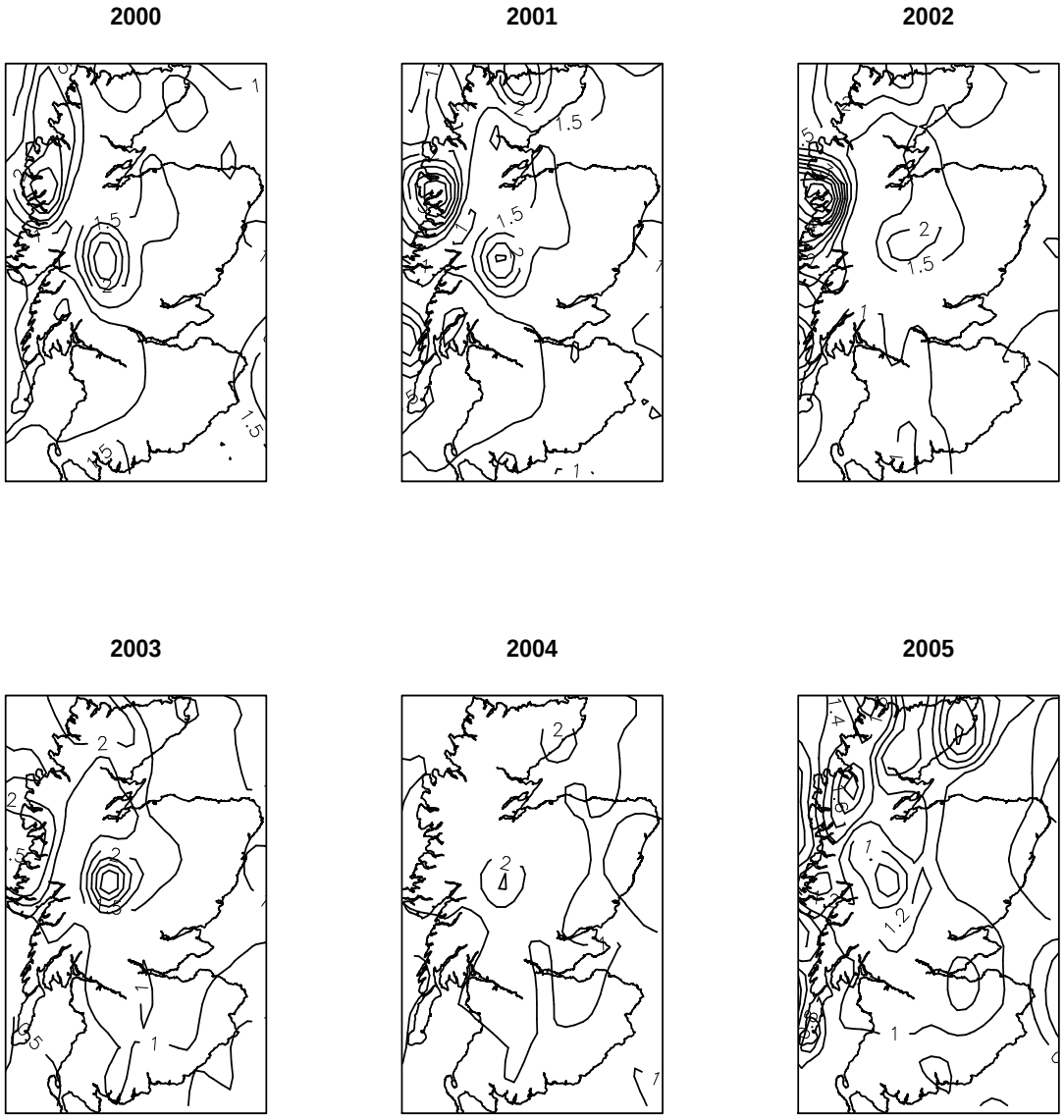

Figure 4.8: Isopleth maps from kernel smoothing for primary 1 and 2 school children for years 2000-2005 for postcode sectors.

Figures 4.7-4.8 show isopleth maps based on kernel smoothing at postcode sector level for pre-school and primary 1 and 2 school children. As for districts, lower susceptibility is found in the southern regions and higher in the northern regions. The values of the contours are higher than at district level. For pre-school, 2003 has higher values than other years. For primary 1 and 2 school children, 2003 and 2004 have highest susceptibility in the upper central region and this decreased in 2005 (Figure 4.8). For both pre-school and primary 1 and 2 and for each year, there is a peak in the west central Highlands, and these have lower population.

4.5 Discussion

In this chapter we have discussed nonparametric smoothing methods for disease mapping of count data and applied some of them to susceptibility to measles data of pre-school children in Scotland. We considered two nonparametric kernel regression methods, namely the Nadaraya-Watson kernel estimator and local linear regression estimator. Even though the Nadaraya-Watson kernel estimator is the simplest method to use, it has a disadvantage of edge effects, however the local linear estimator has edge effect correction.

We discussed the kriging methods based on ordinary and universal kriging. Universal kriging is a refinement of ordinary kriging as it allows the incorporation of covariates by means of regression modelling, so when covariates are available the relationship between the spatial distribution of a disease and possible predictors can be modelled.

Smoothing was based on smoothing of the ratios of the observed counts/expected counts for measles susceptibility, using the Nadaraya-Watson kernel estimator and ordinary kriging, and also smoothing of empirical Bayes estimates (Section 4.3) by ordinary kriging as suggested by Berke (2004), as the empirical Bayes smoothing helps to stabilise the variance by shrinking unstable estimates using neighbourhood data. For kriging, the semivariogram was used to help in interpreting the spatial dependence or structure. However, for the data used here the results were not satisfactory, as for kriging the fitted spherical model did not fit the data well as most of the values are far from the fitted curve. For both kernel and kriging methods, it was not very easy to compare the maps and see what is happening over time, even though it could be

observed that for both pre-school and primary 1 and 2 school children, and at district and postcode sector levels, the isopleth maps retain high values of susceptibility in the north regions and lower values in the south. The results obtained from the districts and the postcode sectors are quite different. The postcode sectors highlighted a peak in measles susceptibility in the west central Highlands where there are fewer people in the population, but this was not observed in the district maps. Thus maps of postcode sectors are more informative than maps of districts, and this indicates that choice of geographical unit is also important when smoothing, and in this case it may be better to use postcode sector level as it has more points than district level. Isopleth mapping has advantages over choropleth mapping, discussed in Chapter 3. The most important characteristic of isopleth mapping is that it is not restricted by the boundaries of the area of study, thus avoiding sudden jumps between two neighbouring regions (Carrat and Valleron, 1992). However, in this analysis these methods were not very useful. In practice choropleth maps are more widely used to illustrate regional patterns of disease and we focus on these in the following chapters.

In the next chapter the focus is on developing methods to compare maps. As maps can be considered as a form of image, some of the methods which we consider are based on image analysis methods, which are used to compare a distorted image to a reference image.

Chapter 5

Methods for Comparing Disease Maps

5.1 Introduction

Much of the analyses in Chapters 3 and 4 are based on comparing maps visually, and now in this chapter we look at more formal ways of comparing maps. In disease mapping, there are various possible changes that can result in maps being different. Differences can occur due to changes in the mean, changes in variability due to clustering (local variability; single area or a group of local areas with a change in rate) or changes in unstructured variability (global variability; overall mean stays the same but high areas get higher and low areas get lower), as seen in Chapter 3. In this chapter we develop descriptive methods or measures that can be used to compare maps. The aim is to develop methods that can be used to identify or detect differences/similarities between two maps, and, if possible, reveal and quantify the kind of difference that is present.

In environmental research, methods and software have been developed to compare maps of categorical data (Stehman, 1999; Hagen, 2002; Pontius, 2002; Hargrove *et al.*, 2006; Visser and Nijs, 2006; and Hagen, 2007). An example of a categorical map is a land cover or land use map. There are several possible reasons for comparing such maps. For example to find similarities or dissimilarities or to assess for which land use category both maps are similar, to detect spatial or temporal changes, for validation of land use models, analyses of sensitivity and uncertainty of models, and for

assessing map accuracy (Visser and Nijs, 2006). Some of these methods are based on pixel by pixel comparison, for example kappa statistics used by Hagen (2002). Some are pattern based comparisons which consider structural similarity, for example the fuzzy polygon based matching technique and fractal analysis approach used by White (2006). These methods do not apply directly in the area of disease mapping as the maps do not consist of pixels and regions in the maps are irregular in shape. Furthermore each region has a quantitative rate associated with it as opposed to a categorical label.

In disease mapping, space-time models have been used to compare two or more maps, and models developed for joint modelling of disease may also be useful in comparing two disease maps (see for example Dabney and Wakefield, 2005). Here we develop descriptive methods for map comparison.

The obvious starting point is the visual comparison of maps, as seen in Chapter 3 (for example, Figure 3.3). The other methods used in that chapter are based on plots against time of parameters obtained from fitted models (the overall mean, variability due to clustering and unstructured variability) (for example, Figure 3.7). Other possibilities include ratio maps, maps of differences, pseudo-colour maps produced from red, green and blue additive primary colours, spatial autocorrelation methods, and the adaption of some ideas that have been used in spatial point processes and image analysis.

In this chapter various methods are described and new ones derived. These are illustrated on two maps produced from the proportions (raw data) susceptible to measles for pre-school children in Scotland, for the years 2000 and 2001 at district level. A simulation study is done in Chapter 6 and more systematic comparisons on real data are presented in Chapter 7.

5.2 Preliminary Analysis

Different methods may be used to perform the preliminary analysis of the data. These can give better understanding of the data by showing the overall distribution of the data for each map. Here we consider histograms and boxplots. Figure 5.1 shows histograms of proportions of susceptibility to measles for 2000 and 2001 respectively.

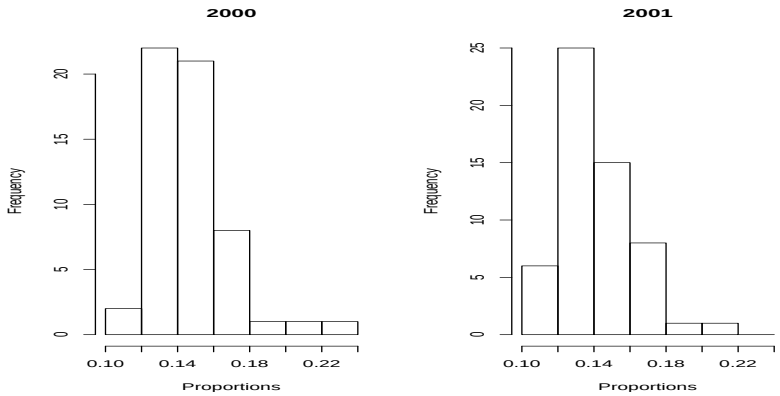

Figure 5.1: Histograms based on the proportions of pre-school susceptibility to measles data, for 2000 and 2001.

These data are slightly different. The year 2000 has more regions with higher rates than 2001. This is also shown by the boxplots in Figure 5.2, in which the median in 2000 is slightly higher than in 2001, as is the whole boxplot, indicating that there are more districts in 2000 with higher susceptibility than in 2001. These plots do not take into account the spatial distribution of the data but show how the rates in each map are distributed.

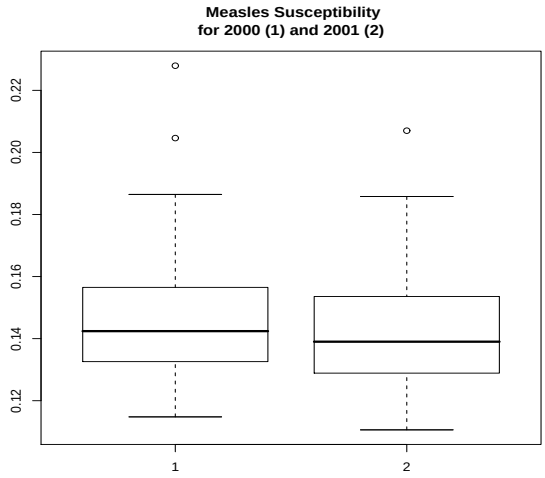

Figure 5.2: Boxplots based on the proportions of pre-school susceptibility to measles data, for 2000 and 2001.

5.3 Use of Maps

Often in disease mapping, maps are produced from the smoothed rates. These maps can be visually compared to assess whether there are any differences or similarities

in the spatial distribution over the years of the disease in question. Here we discuss the use of ratio maps, difference maps and pseudo-colour maps, which will help in comparing the rates of two maps in each region. These maps will allow us to see easily whether a disease rate in a region has changed or not, and, if the rate is different, which map has a larger/smaller rate.

Ratio and Difference Maps

To compare two maps using a ratio map, for each region the ratio of the disease rates for map 2 relative to map 1 is obtained and these ratios are mapped, i.e. we map $r_i = \frac{p_{2i}}{p_{1i}}, i = 1, ..., n$, where r_i is the ratio for region i, n is total number of regions, and p_{1i} and p_{2i} are rates for map 1 and map 2 respectively, for region i. A ratio of 1 will indicate that at that region the maps are identical, a ratio greater than 1 will indicate that the maps are dissimilar at that region and map 2 has a larger rate than map 1, while a ratio less than 1 will indicate that the maps are dissimilar at that region and map 2 has a smaller rate than map 1.

For the differences map, for each region the difference between two rates is obtained, i.e. the map 2 rate less the map 1 rate, and a map based on these differences is produced. Identical regions will have a value of zero, and dissimilar regions will have a value greater or smaller than zero, depending on whether map 2 has the larger or the smaller rate compared to map 1.

Thus the ratio and difference maps can highlight similarities and differences in individual regions, and also show which of the two maps has more or fewer regions of higher/lower rates than the other. The ratio map may not work well in the case when zero counts are present and when expected counts are small. In that case only the difference map may be used.

Pseudo-Colour Map

This map is based on the idea of a pseudo-colour image (see, for example, Glasbey and Horgan, 1995), which can be used to illustrate the combinations of different variable values available for each pixel in an image and display them simultaneously. A pseudo-colour image uses the additive primary colours, red, green and blue. To apply this idea to compare two maps, the disease rates vector for one map is used to define

one colour dimension, for example red, and the disease rates vector for the other map is used to define the other two colour dimensions, for example both green and blue colours.

In a digital greyscale image, the intensity of a pixel is usually represented as an integer in the range 0 (black) to 255 (white). In a colour image, the colour at each pixel is defined by three intensity values, one each for the amount of red, green and blue defining the colour. Each of these is usually also recorded as a value between 0 and 255 inclusive. For example, the (R,G,B) triple (255,0,0) represents pure strong red, (0,255,0) represents pure strong green and (0,0,255) represents pure strong blue. Other combinations give other colours. Equal values for R, G and B define a shade of grey. For each rate vector, the rates are allocated colour corresponding to the intensities between 0 and 255, therefore the rate vector has to be scaled first to obtain a new vector with rates between 0 and 255. In order to achieve this and produce a map, we proceed as follows:

1. For each map or rate vector find the 5th percentile and the 95th percentile.

2. Take the minimum of the two 5th percentiles and the maximum of the two 95th percentile values and obtain the difference between the two, i.e. a range.

3. For each rate vector use the minimum and the range from step 2 to scale the rate vector to obtain a new vector, as follows: $\mathbf{X_1} = (\mathbf{X} - \text{minimum}) * \frac{255}{\text{range}}$, where $\mathbf{X_1}$ and $\mathbf{X}$ are the values in the new and old rate vectors respectively. A value of less than 0 is clipped to the value 0 and a value of more than 255 is clipped to the value 255.

4. The three new scaled rate vectors together define a colour at each region. As there are three colour dimensions but only two maps, one scaled rate vector is used to define the value of one colour dimension and the other scaled rate vector is used to define the values (set equal) in each of the other two colour dimensions. For example, the value of the scaled rate in a given region of map 1 may give each of the R (red) and B (blue) values (set equal), and the value of the rescaled rate in that same region of map 2 may give the value G (green), or vice versa. There are three different ways to do this colour assignment to either map. The resulting (R,G,B) combination then gives the pseudo-colour to use to shade that

region in the pseudo-colour map.

5. Produce a pseudo-colour map based on these colours (here we use a function rgb in the software R, used for primary colour specification for displaying purposes).

Figure 5.3 gives all six sample pseudo-colour maps with colours allocated in three different ways to each map. Measles susceptibility rates for 2000 and 2001 are used. The first map (top left) results from allocating both green and blue to 2000 and red to 2001. Blue regions are those where susceptibility is higher in 2000 than 2001 and pink/brown regions are those where susceptibility is higher in 2001. The first map (bottom left) results from allocating the other way round, i.e. allocating red to 2000 and green and blue to 2001. Now pink/brown regions are those where susceptibility is higher in 2000 and blue regions are those where susceptibility is higher in 2001.

The middle map (top) is a result of allocating red and blue to 2000 and green to 2001. Pink/purple indicates regions with higher susceptibility in 2000 and a green colour indicates those with higher susceptibility in 2001. The middle map (bottom) results from allocating green to 2000 and red and blue to 2001. Green regions are those with higher susceptibility in 2000 and pink/purple are those with higher susceptibility in year 2001.

The third map (top right) results from allocating red and green to 2000 and blue to 2001, giving yellow/green regions when susceptibility is higher in 2000 than 2001 and blue/lilac regions when susceptibility is higher in 2001. The other third map (bottom right) results from allocating blue to 2000 and red and green to 2001, and now blue/lilac regions are where susceptibility is higher in 2000 and yellow regions where susceptibility is higher in 2001.

For all these maps, a grey colour indicates that the maps are similar at that region, i.e. have similar rates. For each rate vector, the lighter the colour the higher the rate, and the darker the colour the smaller the rate in relation to other rates in the map. This map allows us to identify visually and quickly regions with similar or different rates. Here we chose the option of allocating red and blue to one map and green to the other, as it is easier to know what colours to expect, i.e. red and blue gives colour not very far from red (pink/purple) if the rate in that map is higher than in the map allocated green, and the use of green gives a greenish colour if the rate in the map

defining the green is higher. The R software code for the pseudo-colour map is given in Appendix C.

5.3.1 Illustration Using Two Maps

Figure 5.4 shows maps of susceptibility to measles for 2000 and 2001, ratio maps obtained by mapping the ratio of 2001 to 2000 and a difference map (2000-2001). Three limits were chosen in the case of the difference and ratio maps, and this was done in such a way that the middle interval will be a narrow interval around 0 for the difference map and a narrow interval around 1 for the ratio map. From visual comparison, some differences can be observed in the susceptibility to measles maps (top two maps) For both years, susceptibility to measles is higher in the northern and western regions and lower in the southern regions, but in 2001 the number of regions with lower susceptibility increased to include more of the north east.

Using the suggested maps the regions with similar or different susceptibility rates can easily be identified, and more easily observed than from the susceptibility maps themselves. The ratio map plots the ratio of 2001 over 2000. Susceptibility is similar (not changed in 2001) in a few central districts for both years, i.e. with a ratio of 0.99 (inclusive)-1.01 (inclusive). This includes urban areas of Glasgow and Edinburgh. In 2001 susceptibility increased for a few districts (darker districts with ratio over 1.01). These are some districts in the south, and some in the north, including districts such as Sutherland, Lochaber and Berwickshire. In most districts all over Scotland measles susceptibility decreased in 2001 (lighter regions, with a ratio under 0.99), but mostly those in the north east and south. This confirms what we see in the maps of susceptibility, i.e. susceptibility decreases in 2001 to include the north east. The average of the ratios is 0.97, the median is 0.98 and the range is 0.30. As the median and mean ratio are both close to 1, and the range is low most regions are similar.

The difference map plots the rates for 2000 minus the rates for 2001. This gives a similar picture to the ratio map. The darker regions are where susceptibility increased in 2000, and lighter regions are where susceptibility was higher in 2001 or was lower in 2000. The regions with high susceptibility in 2000 are in the west, north east and south.

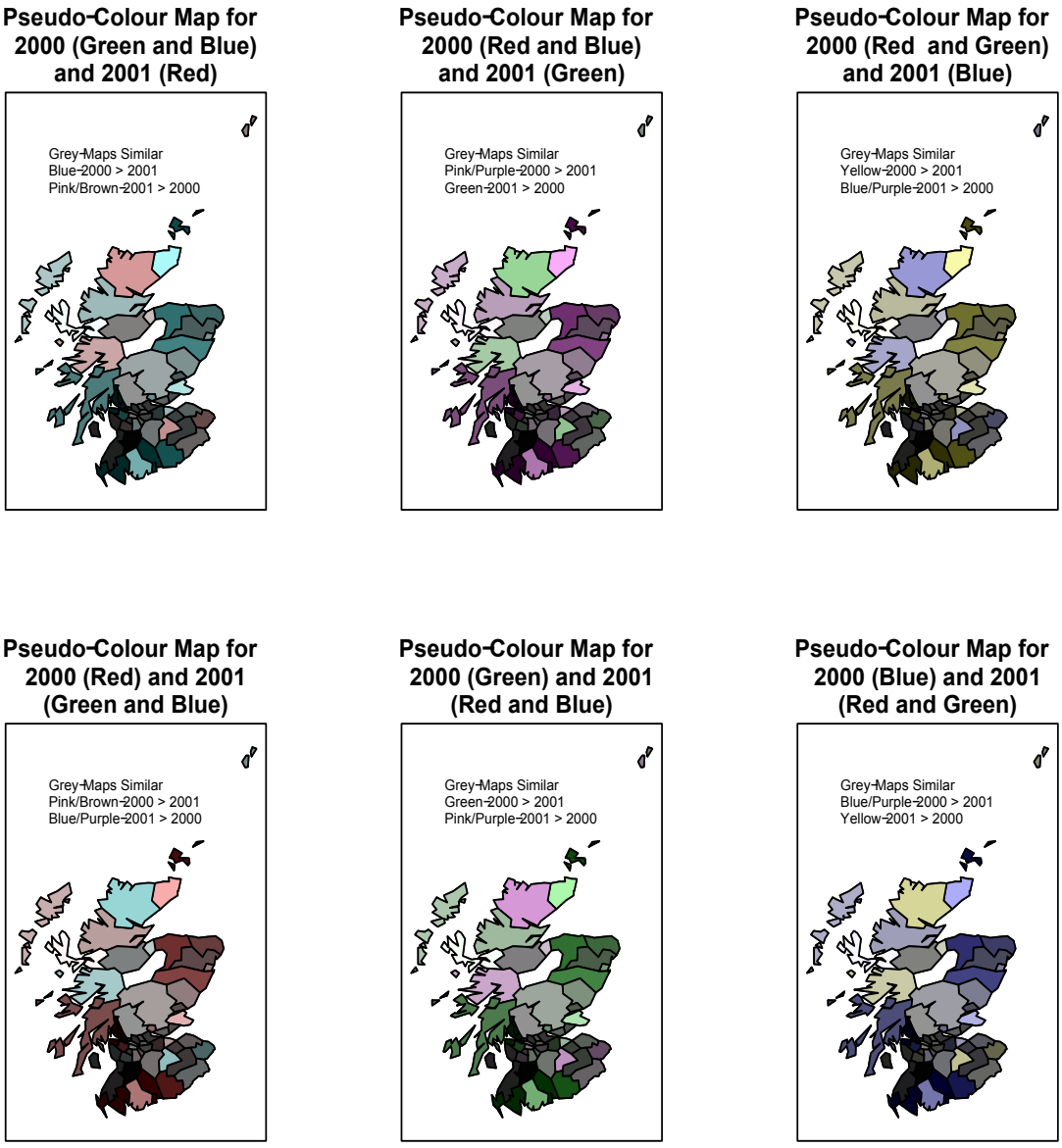

Figure 5.3: Pseudo-colour maps for proportions of pre-school children susceptible to measles, for 2000 and 2001 with colour allocated in different ways. The left maps use green and blue for 2000 and red for 2001 (top), and red for 2000 and green and blue for 2001 (bottom). The middle maps use red and blue for 2000 and green for 2001 (top), and green for 2000 and red and blue for 2001 (bottom). The right maps use red and green for 2000 and blue for 2001 (top), and blue for 2000 and red and green for 2001 (bottom).

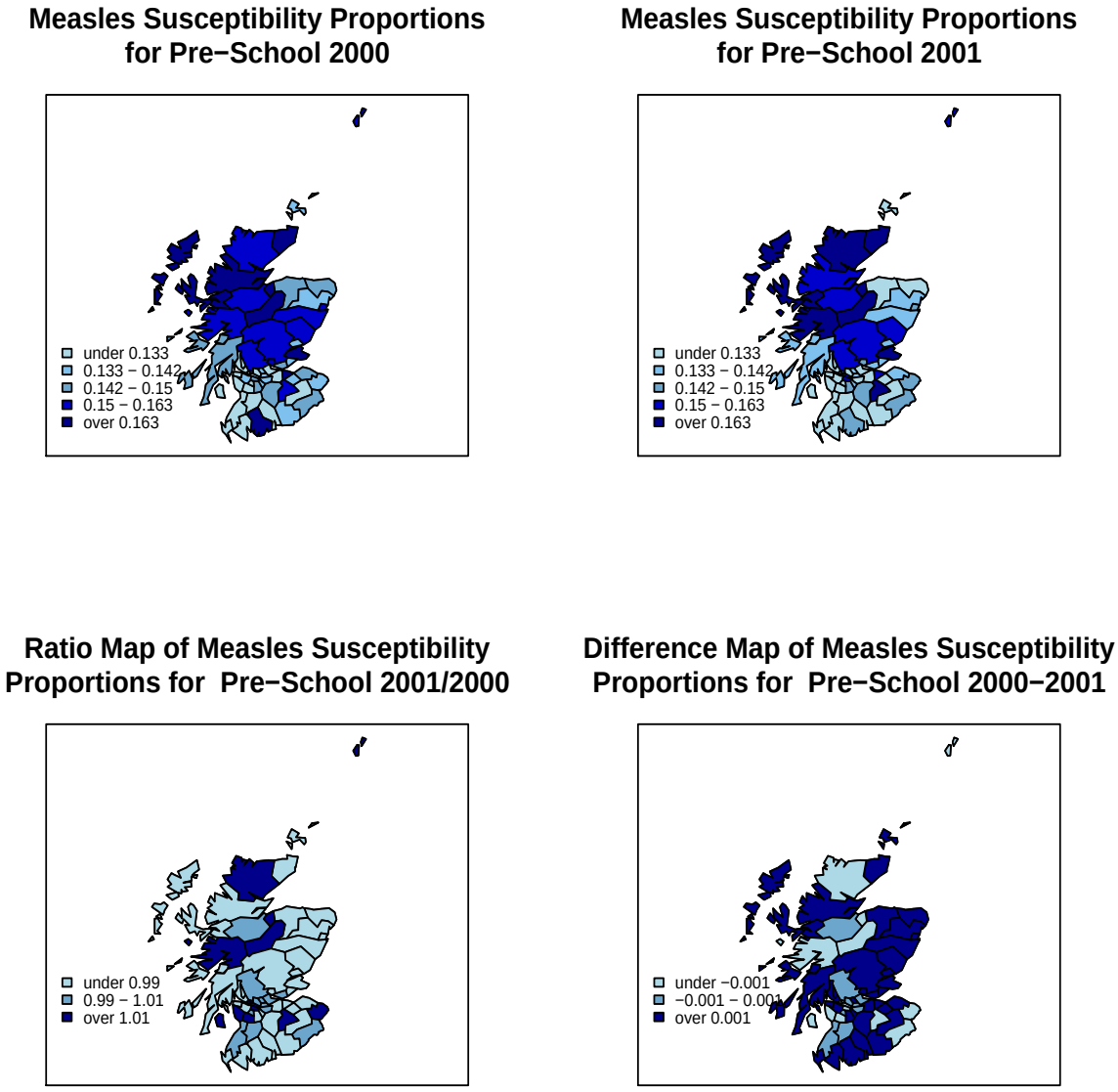

Figure 5.4: Susceptibility maps based on proportions of pre-school children susceptible to measles, for 2000 and 2001, and corresponding ratio (2001/2000) and difference maps (2000-2001).

There are pockets of districts which had high susceptibility rates in 2001 which are similar for those indicated by the ratio map. Roxburgh district (see Figure 1.4) was highlighted by the ratio map as having susceptibility rates that are similar for both 2000 and 2001, but the difference map indicates that this district had higher susceptibility in 2001.

The pseudo-colour map in Figure 5.3 (top middle plot) shows that there are districts with greyish colour, i.e. where measles susceptibility is similar in the two years, very few regions where susceptibility is higher or increased in 2001 (green) and more districts where susceptibility is higher in 2000 than 2001 (pink/purple). Regions in the north east, southern part and some in the west had higher susceptibility rates in 2000. Sunderland, Lochaber and Tweeddale are shown to have higher susceptibility in 2001 and this is consistently shown by all the maps. Inverness and Stirling are among those in which susceptibility remained similar for both years. Even though the district of Badenoch and Strathspey is indicated by the ratio map and difference map to have higher susceptibility in 2001, the pseudo-colour map shows this region as having the same higher of susceptibility rate (very pale white colour) in both years, and this higher susceptibility rate can be visually seen on the susceptibility maps.

Use of the ratio map, difference map and pseudo-colour map together with the maps based on rates may help in making better conclusions about the changes that have taken place in the regions, by identifying those regions where the rates have increased, decreased or remained the same.

5.4 Plots of Parameters

In disease mapping, when using models (spatial or space-time) to fit or smooth rates at different time points, as in Chapter 4, we obtain fitted parameters for the overall mean and variability due to structured and unstructured heterogeneity. These parameters can be plotted against time and the trend observed to summarise the nature of differences between maps.

5.4.1 Illustration Using Two Maps

Table 5.1 shows the results obtained from fitting the logistic model and the Waller *et al.* (1997) space-time model. Since here we are comparing only two maps, it is easy to compare the parameter values from the table. Even though the values of the parameters from the two models are not the same, the interpretation of the results is similar.

Year	Pre-school					
	Logistic model				Space-time model	
	α	mean	σ_u	σ_v	σ_u	σ_v
2000	−1.83 (0.03)	0.14 (0.005)	0.08 (0.03)	0.21 (0.02)	0.05 (0.06)	0.57 (0.06)
2001	−1.86 (0.03)	0.13 (0.005)	0.08 (0.04)	0.22 (0.03)	0.05 (0.04)	0.58 (0.06)

Table 5.1: Pre-school posterior means for overall mean level (α), mean ($\frac{e^{\alpha}}{1+e^{\alpha}}$), standard deviations due to correlated heterogeneity (σ_u) and uncorrelated heterogeneity (σ_v) for the logistic and space-time (Waller *et al.*, 1997) models, with standard deviations in brackets.

The logistic model indicates that the overall mean has changed very slightly between years (0.14 (2000), 0.13 (2001)). Both models indicate that standard deviation due to clustering is the same for both years (σ_u=0.08 (logistic model), σ_u=0.05 (space-time model)) and there is a very slight change in standard deviation due to unstructured heterogeneity (σ_v=0.21 (2000) and σ_v=0.22 (2001) for the logistic model; σ_v=0.57 (2000)) and σ_v=0.58 (2001) for the space-time model).

Taking into account standard errors, there is no difference between the years in any of the parameters, so the two models say the two maps are virtually identical. This suggests that the maps are over interpreted and that statistical tests are needed, and that is why we want to investigate the use of other methods.

5.5 Spatial Autocorrelation Methods

Spatial autocorrelation is present when the value of a spatial variable for a region is associated with the values of that same variable at neighbouring regions. For example, spatial autocorrelation exists when similar high/low values cluster together on a

map (Sokal and Uyterschaut, 1987; Odland, 1988; Gebhardt, 1998; and Rosenberg *et al.*, 1999). When data are mapped, the map contains the information about how the values of the variable are arranged in space, thus here we look at some descriptive statistical methods that are used to analyse spatial distribution by measuring spatial autocorrelation. These methods may not necessarily help in detecting differences between maps but may aid in comparing the spatial distribution/structure of two or more disease maps by giving a measure of strength of spatial dependency for each map, thus we will be able to know which map has the strongest spatial dependency.

The statistical methods described here are used for real-valued variables, namely *Moran's I*, *Geary's c* and the *spatial correlogram* based on each of these two measures.

5.5.1 Moran's I, Geary's c and Spatial Correlogram

Moran's I (Moran, 1950) and Geary's c (Geary, 1954), are the standard statistics for testing the independence of real-valued area data (Gebhardt, 1998). The null hypothesis is that all the values are independent. Moran's I is a product-moment correlation and Geary's c is a distance-like coefficient. Moran's I is inversely related to Geary's c. It is a measure of global spatial autocorrelation, while Geary's c is more sensitive to local spatial autocorrelation.

Let $\mathbf{x} = \{x_i, i = 1, 2, ..., n\}$ be real-valued data for n regions. Then Moran's I is given by

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2} \tag{5.1}$$

where w_{ij} is a measure of adjacency between regions i and j, with $w_{ii} = 0$.

This is a spatial autocovariance (numerator) standardised by two terms. These are $\sum_i (x_i - \overline{x})^2$, which measures variation in the values of $\mathbf{x}$ but not on their spatial arrangement, and $\frac{n}{\sum_i \sum_j w_{ij}}$, which measures connectivity of a set of regions and the value of which can change with the rearrangement of regions but not with changes in $\mathbf{x}$. The values of I range from -1 to +1 (Odland, 1988).

Various forms of w_{ij} have been used, for example,

- $w_{ij} = 1$ when regions are neighbours, i.e. share a boundary, and $w_{ij} = 0$ otherwise,

- $w_{ij} = \frac{1}{v_i}$ where v_i is the number of neighbours of a region, and

- $w_{ij} = 1$ if the distance (i.e. distance between regions or distance between their centroids) between i and j is less than a threshold, and $w_{ij} = 0$ otherwise.

The expected value of Moran's I is $-\frac{1}{n-1}$. When values of a spatial variable are independent of neighbouring values I should be close to this expected value. Values of I greater than this expected value indicate *positive spatial autocorrelation*, i.e. values of a spatial variable tend to be similar to neighbouring values (large or small values are spatially clustered), and values of I less than the expected value indicate *negative spatial autocorrelation*, i.e. neighbouring values are not independent but tend to be dissimilar. As the number of regions (n) becomes large the expected value approaches zero.

Geary's c is an alternative statistic to Moran's I. Let $\mathbf{x} = \{x_i, i = 1, 2, ..., n\}$ be a set of real-valued data for n regions. Then Geary's c is given by

$$c = \frac{n-1}{2\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - x_j)^2}{\sum_i (x_i - \overline{x})^2}, \tag{5.2}$$

again as for Moran's I w_{ij} is a measure of adjacency between regions i and j, with $w_{ii} = 0$. This statistic lies between 0 and 2.

The expected value of Geary's c is 1 for independent neighbouring values, smaller than 1 for positive spatial autocorrelation, and larger than 1 for negative spatial autocorrelation. Thus large values of I correspond to small values of c and vice versa (see also Figure 5.5).

A *spatial correlogram* is a graph of a spatial autocorrelation coefficient against distance between regions. For a given variable, this is computed by evaluating the spatial autocorrelation coefficient for sets of pairs of regions at specified distance classes. In general, in a correlogram positive autocorrelation coefficients will indicate that regions at a given distance apart are similar, and negative autocorrelation coefficients will indicate that regions at a given distance apart are dissimilar. When assessing statistical significance of the spatial correlogram, it must be considered that the individual coefficients are not independent, thus a *Bonferroni method* which approximates the adjusted significance probability for multiple testing is used (Fortin *et al.*, 2002;

Sokal and Uyttterschaut, 1987). Therefore, the significance of a spatial correlogram is tested by using the *Bonferroni criterion*, i.e. the spatial correlogram is significant if at least one of its Moran's I or Geary's c values is significant at $P < \frac{\alpha}{k}$, where k is the number of distance classes, α is the standard type I error and P is the p-value of the coefficient.

5.5.2 Illustration Using Two Maps

For both Moran's I and Geary's c (including spatial correlograms), the results obtained have a similar interpretation. The overall Moran's I and Geary's c statistics for 2000 are 0.449 and 0.421 respectively, and for 2001 the values are 0.459 and 0.456 respectively with $n = 56$ districts. These results indicate that for both years there is a positive spatial autocorrelation, i.e. neighbouring districts are similar and spatial autocorrelation is not very different. For both statistics, the value is a little lower for 2001 than for 2000.

The spatial correlograms in Figure 5.5 are all significant using the Bonferroni criterion with $\alpha = 0.05$ and $k = 12$ (number of distance classes), so $\frac{\alpha}{k} = 0.004$. The coefficients and their distance classes and p-values are shown in Tables 5.2 (Moran's I) and 5.3 (Geary's c). These calculations were done in R, and the distances used were chosen automatically by the function used (in this case as distances which are about 55.33 kilometres apart). The distances are 32.39, 87.72, 143.06, 198.39, 253.72, 309.05, 364.39, 419.72, 475.05, 530.39, 585.72 and 641.05. Here we use w_{ij} such that $w_{ij} = 1$ when regions are neighbours, i.e. share a boundary, and $w_{ij} = 0$ otherwise.

The spatial correlograms for the two years look similar. There is a positive autocorrelation at the first two short distances (32.39 and 87.72) and negative autocorrelation at long distances for both 2000 and 2001, indicating clustering. These plots confirm the existence of similar positive spatial autocorrelation, thus the spatial distributions are similar for the two maps. The conclusion is the same for both spatial measures. However, these measures do not give a comparison of the pattern in the two maps.

	2000			2001	
class	distance in kilometres	coefficient	p-value	coefficient	p-value
1	32.39	0.324	5.97e-06	0.330	5.88e-06
2	87.72	0.400	1.04e-12	0.292	1.10e-07
3	143.06	0.005	3.42e-01	0.048	1.26e-01
4	198.39	-0.074	8.21e-01	-0.051	7.02e-01
5	253.72	-0.087	7.84e-01	-0.107	8.45e-01
6	309.05	-0.460	9.99e-01	-0.277	9.94e-01
7	364.39	-0.522	9.99e-01	-0.410	9.98e-01
8	419.72	-0.245	8.53e-01	0.051	2.99e-01
9	475.05	-0.090	4.98e-01	0.109	2.20e-01
10	530.39	-0.368	9.99e-01	-0.422	9.99e-01
11	585.72	-0.580	9.99e-01	-0.492	9.99e-01
12	641.05	-0.719	9.86e-01	-0.843	9.87e-01

Table 5.2: Coefficients of Moran's I with distance classes and p-values used to produce spatial correlograms for 2000 and 2001. The p-values are compared to $\frac{\alpha}{12}$.

	2000			2001	
class	distance in kilometres	coefficient	p-value	coefficient	p-value
1	32.39	0.573	7.65e-07	0.605	1.90e-06
2	87.72	0.643	4.19e-06	0.706	5.53e-05
3	143.06	0.985	4.53e-01	1.098	8.08-01
4	198.39	1.110	7.89e-01	1.034	6.05e-01
5	253.72	1.521	9.95e-01	1.465	9.93e-01
6	309.05	2.426	9.99e-01	2.034	9.99e-01
7	364.39	2.676	9.99e-01	2.247	9.98e-01
8	419.72	1.421	8.76e-01	0.978	4.77e-01
9	475.05	0.970	4.67e-01	1.077	5.87e-01
10	530.39	3.814	9.99e-01	4.296	9.99e-01
11	585.72	4.559	9.99e-01	3.947	9.99e-01
12	641.05	2.298	9.86e-01	2.608	9.87e-01

Table 5.3: Coefficients of Geary's c with distance classes and p-values used to produce spatial correlograms for 2000 and 2001. The p-values are compared to $\frac{\alpha}{12}$.

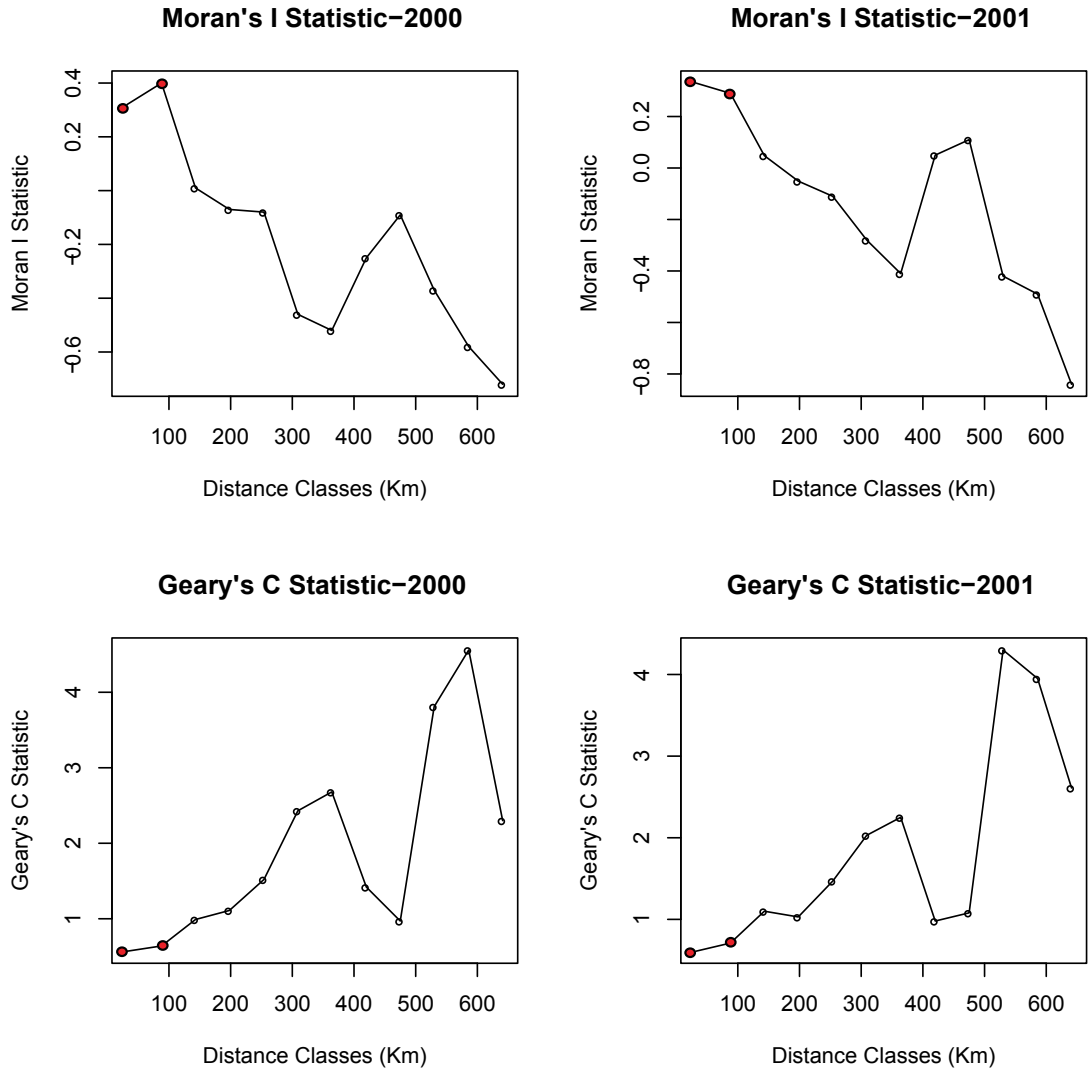

Figure 5.5: Spatial correlograms for Moran's I and Geary's c for pre-school susceptibility to measles in years 2000 and 2001 at district level, with the red points (the first two) indicating a distance in kilometres at which positive autocorrelation is significant.

5.6 Analogues of Some Point Process Methods

In a spatial point process, a spatial point pattern is a data set of points referred to as *events*, distributed within a region at different locations (Diggle, 2003). To test for complete spatial randomness (CSR) of a spatial point pattern (Diggle, 2003) several distance based methods are used. One of these is the *inter-event distances* (geographical distance between two events) method, for which a summary description of a pattern of n events in a region is the empirical cumulative distribution function (ecdf) of the $\frac{1}{2}n(n-1)$ distinct inter-event distances. Another is the *nearest neighbour distances* method for which x_i is the distance from the ith event to the geographically nearest event, $i = 1, 2, ..., n$, thus the empirical cumulative distribution is based on the n nearest neighbour distances. The ecdf of either of these sets of differences is then compared to the corresponding cumulative distribution function expected under CSR, or two ecdfs from separate samples may be compared in a two sample test. The two ecdfs can be tested for a significant difference using the Kolmogorov-Smirnoff (KS) test (Durbin, 1973), for example. In our case we compare two cumulative distribution functions (cdfs). The two-sample KS test tests whether two data sets arise from distributions with the same cdf or not. It compares the cumulative distributions by using the maximum vertical deviation (distance (D)) between the two curves as the test statistic.

Another method used in point processes is the *point to nearest event distances* method, in which x_i is the distance from the ith of m sample points in space to the nearest of the n events. This method will not be useful in comparing disease maps, as the rates corresponding to the areas in the maps typically are fairly limited in number and to sample values from these to use would be to use the values more than once (as sample points and "events").

Here in the case of disease mapping, we adapt the idea of obtaining this kind of empirical cumulative distribution to describe the maps. Instead of distances, the absolute differences between disease rates are used, for each of two disease maps the cdf of the differences is obtained, and the two cdfs are compared to test if there is any difference between the maps. Below we outline different implementations of this approach.

5.6.1 Methods

1. **Inter Region Differences (IRD)**

   This method is an analogy of the inter-event distance method. For each map, the cdf is calculated from the numerical differences between all possible distinct pairs of rates. Let x_i now be be the ith smallest inter-region difference, $i = 1, ..., l$. Then the cdf is given by

   $$\hat{F}_n(x) = \frac{i}{\frac{1}{2}n(n-1)}, \quad x_i < x < x_{i+1}, \quad i = 0, ...l, \tag{5.3}$$

   where n is the number of regions, $l = \frac{1}{2}n(n-1)$, and $x_0 = -\infty$ and $x_{i+1} = +\infty$.

2. **Most Similar Differences (MSDI) and Most Dissimilar Differences (MDD)**

   These are two methods analogous to the method of nearest neighbour distances. For each map, let x_i be the numerical difference between the rate of the ith region, $i = 1, ..., n$ and the most similar or most dissimilar rate to this amongst all other regions in the map. Then we refer to x_i as a most similar or most dissimilar difference. This method includes duplicate differences.

3. **Average Neighbour Differences (AVND)**

   This method takes the average of the differences between the disease rate of a region and the rates of its neighbouring regions. Let x_i be the average difference of the rate of the ith region and the rates of its neighbouring regions, $i = 1, ..., n$. Then we refer to this as the ith average neighbour difference.

4. **Most Similar Neighbour Differences (MSND) and Most Dissimilar Neighbour Differences (MDND)**

   These are two methods which we use based on the differences between the rate of a region and the rates of its neighbouring regions. Let x_i be the difference between the rate of ith region and the most similar or most dissimilar rate among its neighbouring regions only, $i = 1, .., n$. Then we refer to the x_i as most simil ar neighbour differences or most dissimilar neighbour differences. (These could include duplicated difference values between reciprocal nearest or farthest neighbouring values).

For approaches 2-4 there are n such differences used to construct the cdf. The empirical cumulative distribution function is given by

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{x_i \leq x} \tag{5.4}$$

where n is the number of regions and I is the 0/1 indicator function.

All of these methods are invariant to location, i.e. if $r_1, .., r_n$ are rates for map 1 and $s_1 = r_1 + c, .., s_n = r_n + c$, are rates for map 2, for regions 1,...,n, where c is a constant, then all IRD, MSDI, MDD, AVND, MSND and MDND will be the same for the two maps.

When comparing two maps, in the case when the rates in map 2 are different from those in map 1, it is expected that the differences between the rate of the ith region and the rates of other regions in one map will be different from those in the other map, and so it is expected that the IRD, MSDI and MDD will detect this change. The AVND, MSND and MDND methods use the neighbourhood structure. These methods are expected to detect differences between the two maps when the rates of a region and the rates of its neighbouring regions are different in the two maps. Thus these methods should indicate whether a change has occurred or not in the neighbourhood rate patterns. The structure of the two maps has to be the same to use these methods.

For all these methods, the cumulative distributions of the differences for each map are drawn on one plot and compared using the two-sided Kolmogorov-Smirnov test to test for differences in the cdfs. The KS test is used here as it makes no assumption about the distribution of the data. We use the p-value of the KS test to assess the significance of the difference in the cdfs.

5.6.2 Illustration Using Two Maps

Figure 5.6 gives illustrative graphs of cdfs based on the methods above. For AVND, MSND and MDND n =53 districts as the islands have been excluded since these methods uses the neighbourhood structure. The function used to implement these methods was written in R and is given in Appendix C. For the IRD method we observe that for most x values, the proportion of values less than x in 2000 is very similar to that in 2001, thus most of the differences for these two years are similar.

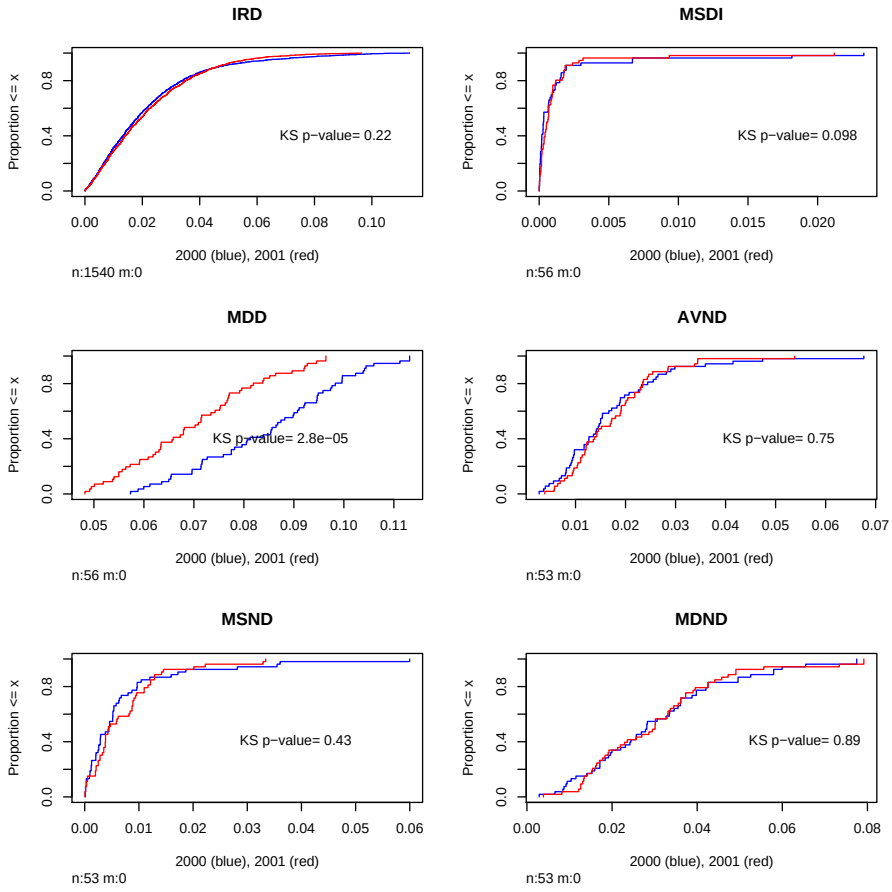

Figure 5.6: Empirical cumulative distribution function graphs (n= non-missing observations and m= missing observations) and tests based on the proposed methods IRD, MSDI, MDD, AVND, MSND and MDND for pre-school susceptibility to measles for years 2000 (shown in blue) and 2001 (shown in red).

The value of the test statistic is D=0.038 (p-value = 0.22), thus there is no significant difference between the two cdfs for the two years, indicating similarities in their inter-region differences. For MSDI (D=0.23, p-value=0.098), so generally for most x values, the proportion of values less than x in 2000 is close to that in 2001. This also applies to AVND (D=0.13, p-value= 0.75), MSND (D=0.17, p-value=0.43) and MDND (D=0.11, p-value=0.89). These test statistics all indicate similarity of the cdfs, although some differences can be seen. For MDD (D=0.45, p-value = $2 * 10^{-5}$), for all x values, the proportion of values less than x in 2000 is clearly different from 2001, and the KS test indicates that there are significant differences between the two. Therefore, except for MDD all methods indicate that the maps are similar. From Chapter 1 (Figure 1.5) we observe that about 4 districts moved to the lowest category, i.e. susceptibility of

less than 14.5% in 2001, thus as MDD depends on how different the rates are, this measure may have detected this change.

5.7 Image Analysis Methods

In this section, we discuss and adapt for disease maps some methods used to compare digital images. In the field of image analysis, degradation of visual quality of digital images may occur as a result of image distortions during acquisition, processing, compression, storage, transmission and/or reproduction (Wang *et al.*, 2004). Image distortion can be assessed subjectively by eye, where a group of people is asked to compare a distorted image with the reference image, and to provide a score on a distortion scale. The mean score may be taken as an index of image quality, thus this method is known as the *mean opinion score* (MOS) (Bouzerdoum *et al.*, 2004). However, although this method reflects well human visual perception, it is inconvenient, time-consuming and costly, and as a result researchers in this field have been developing objective quality measures.

Objective image quality measures can be classified into three approaches. These are *full-reference* measures, where the original image is known, *no-reference* measures, where no original image is available, and *reduced-reference* measures where the original is partially available (Wang *et al.*, 2004; Bouzerdoum *et al.*, 2004). Here we will only focus on full-reference image quality measures, as some of these are applicable to comparing disease maps.

There are two groups of objective image quality measures. The first are mathematically defined, such as mean square error (MSE) and peak-signal-to-noise ratio (PSNR) (Mulopulos *et al.*, 2003). These measures are simple and easy to calculate and results based on these are not influenced by the the individual observer (Wang *et al.*, 2002). However, these do not correlate well with the human visual system (HVS) (Mulopulos *et al.* 2003; Wang *et al.*, 2004). The second group of measures incoporates HVS characteristics, and includes image quality measures such as the universal image quality measure (UIQI) (Wang *et al.*, 2004), the structural similarity index (SSIM) (Wang *et al.*, 2004), the new weighted mean square error (NwMSE) (Samet *et al.*, 2005), neural network based image quality measures (Bouzerdoum *et al.*, 2004) and singular value

decomposition (SVD) based image quality measures (Shnayderman *et al.*, 2006).

The SVD based measure is not useful for maps, as here it will not be easy to use blocks and we cannot obtain singular values unless a map is rearranged in a rectangular grid. The neural network based image quality measure is not considered either, as human subjects would be needed to generate data to train the network.

The MSE and PSNR quantify size of the image error, therefore we adapt these measures for comparing disease maps as they may help in quantifying the difference between the maps. Among the objective measures that incorporate HVS characteristics, only the SSIM, which is an improved version of UIQI, will be adapted. This measures structural distortion, and in our case may help in measuring the overall structural differences. These three measures will be discussed in detail in the next sections.

5.7.1 Mean Square Error and Peak-to-signal-noise Ratio

The equations for mean square error and the peak-to-signal-noise ratio measure, as used in image analysis and given by Mulopulos *et al.* (2003), are given below. The mean square error is is not invariant to location and is calculated as

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I(x,y) - I'(x,y)]^2 \tag{5.5}$$

for an M $\times$N image, where I(x, y) is the intensity in pixel (x, y) of the original image and I'(x, y) is the intensity of the corresponding pixel in a distorted image. When MSE is equal to zero this indicates that the image to be assessed for distortion is a perfect reproduction of the original image, and the larger is MSE the more the image differs from the original.

The peak-to-signal-noise ratio (PSNR) is inversely related to MSE on a logarithmic scale, and again measures the pixel difference between the original and distorted images. This is given by

$$PSNR = 10 * log_{10}(\frac{P^2}{MSE}) \tag{5.6}$$

where P is the range of the pixel values (maximum peak-to-peak signal swing) (for 8-bit images P could be taken as 255). A PSNR of zero indicates that the images differ greatly (as MSE is large) and infinite PSNR indicates identical images. This measure

takes account of variation within the images when computing the difference between images.

Images are defined on a rectangular grid of pixels, while maps are not. However, these measures can still be used to compare corresponding regions in two maps with the same structure. Here MSE is referred to as Mean square difference (MSD) and is given by

$$MSD = \sum_{i=1}^{n}(x_i - y_i)^2/n, \quad \geq 0 \tag{5.7}$$

where x_i and y_i are rates in region i in map 1 and map 2 respectively, $i = 1, ..., n$, and n is the number of regions in each of the maps. A zero value will indicate that there are no differences in the maps examined, while large MSD indicates large difference(s).

The PSNR is taken as

$$PSNR = \frac{P^2}{MSD}, \tag{5.8}$$

where P can be taken to be the range of rates across both the two maps (the resulting measure being referred to here as PSNRR), or P can be the maximum absolute difference between rates in the two maps i.e. $\max|x_i - y_i|$, $i = 1, ..., n$ (referred to here as PSNRM). When maps are very different PSNRR or PSNRM will be near zero, while identical maps will give an infinite value. The log scale is not used here as we do not have very large values.

5.7.2 Structural Similarity Index

The SSIM is a modified version of UIQI, and was designed to correlate closely with the human visual system. It was developed by Wang *et al.* (2004) and is a more complicated measure of structural information change. This measures distortion as a product of three different factors, namely luminance, contrast and correlation (structure). Let $\mathbf{x} = \{x_i, i = 1, 2, ..., N\}$ and $\mathbf{y} = \{y_i, i = 1, 2, ..., M\}$ be the original and the distorted images, where N and M are now the total number of pixels for the reference and distorted images respectively. Then the mean luminance term (term measuring closeness of the means) is given by

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1},$$

where μ_x and μ_y are the mean pixel intensities for the reference and distorted images respectively and C_1 is a constant. The contrast term (term measuring the difference

in variability) is given by

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y+C_2}{{\sigma_x}^2+{\sigma_y}^2+C_2}.$$

where σ_x and σ_y are the standard deviations for the reference and distorted image respectively and C_2 is another constant, and the correlation structure term is given by

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3},$$

where C_3 is another constant. Wang *et al.* (2004) took $C_3 = \frac{C_2}{2}$, $C_i = (K_iL)^2, \quad i = 1,2$, where $K_i \ll 1$ is a small arbitrary constant and L is the range of pixel values (255). The constants $C_i,\ i = 1,2,3$ are included to avoid unstable results when $({\mu_x}^2 + {\mu_y}^2)$ or $({\sigma_x}^2 + {\sigma_y}^2)$ or $(\sigma_x\sigma_y)$ is close to zero (Wang *et al.*, 2004). The structural similarity index is then given by

$$\begin{aligned} SSIM(\mathbf{x},\mathbf{y}) &= l(\mathbf{x},\mathbf{y}) * c(\mathbf{x},\mathbf{y}) * s(\mathbf{x},\mathbf{y}) \\ &= \frac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1} * \frac{2\sigma_x\sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2} * \frac{\sigma_{xy} + \frac{C_2}{2}}{\sigma_x\sigma_y + \frac{C_2}{2}} \\ &= \frac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1} * \frac{2\sigma_x\sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2} * \frac{2\sigma_{xy} + C_2}{2\sigma_x\sigma_y + C_2} \\ &= \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{({\mu_x}^2 + {\mu_y}^2 + C_1)({\sigma_x}^2 + {\sigma_y}^2 + C_2)}. \end{aligned} \tag{5.9}$$

This measure satisfies the following conditions

1. Symmetry: $SSIM(\mathbf{x},\mathbf{y}) = SSIM(\mathbf{y},\mathbf{x})$,
2. Boundedness: $SSIM(\mathbf{x},\mathbf{y}) \leq 1$, and
3. Unique maximum: $SSIM(\mathbf{x},\mathbf{y}) = 1$ if and only if $\mathbf{x}$=$\mathbf{y}$ (Wang *et al.*, 2004).

This is a single overall quality measure of the entire image. Without the constants we have the UIQI.

A mean adaptive SSIM index (localised quality measure) where the local statistics (mean, standard deviation and covariance) are computed within a local square window which moves pixel by pixel over the entire image can also be used, and is given by

$$MSSIM(\mathbf{x},\mathbf{y}) = \frac{1}{Q}\sum_{i=1}^{Q} SSIM(x_i, y_i) \tag{5.10}$$

where x_i and y_i are now the image contents in the ith local window, $SSIM(x_i, y_i)$ is the SSIM of the ith local window, and Q is the number of local windows of the

image. A zero SSIM/MSSIM index indicates that the images differ greatly and an SSIM/MSSIM index of value 1 indicates identical images. The adaptive measure should help to quantify local differences.

For comparing disease maps, we consider SSIM. Let $\mathbf{x} = \{x_i, i = 1, 2, ..., n\}$ and $\mathbf{y} = \{y_i, i = 1, 2, ..., n\}$ be the rates of the regions for maps 1 and 2 respectively, where n is the number of regions in each map. In equation (5.9) for $C_i = (K_i L)^2,\ i = 1, 2$, L can be taken as P in the peak-to-signal-noise ratio, i.e. L can be taken to be the range of rates over the two maps, giving what we call SSIMR, or L can be the maximum absolute difference between values of the two maps i.e. $\max|x_i - y_i|$, $i = 1, ..., n$, giving the SSIMM measure. Here we will also use UIQI, i.e. SSIM without constants, to compare maps but we choose to refer to it as SSIM. Therefore for SSIM, SSIMR and SSIMM a value of 1 indicates that the maps are similar and a zero value will indicate that the maps are very different. Wang et al. (2004) used $K_1 = 0.01$ and $K_2 = 0.03$.

MSSIMM will not work here as this will require us to define local windows, but because of irregular regions this becomes difficult. We tried to use a region and its immediate neighbours to define a local window, however this did not work very well.

As the structural similarity index measures the structural information change between two images, in the case of comparing disease maps it may be able to give a good approximation of how much global structural change has occurred between two maps. The function used to implement these methods was written in R and is given in Appendix C.

5.7.3 Illustration Using Two Maps

Table 5.4 shows the values of the image based measures used to compare the pre-school maps of measles susceptibility in 2000 and 2001 at district level.

In Table 5.4 we observe that the value of MSD is close to zero, indicating that the maps are very similar. The values of PSNRR and PSNRM are very different but both far from zero, and, according to the interpretation of these measures, they also indicate that the maps are similar.

Year	MSD	PSNRR	PSNRM	SSIM	SSIMR	SSIMM
2000-2001	0.0001	118.6	15.3	0.88	0.88	0.88

Table 5.4: Values of MSD, PSNRR (PSNR with P=Range), PSNRM (PSNR with P= Maximum absolute difference), SSIM, SSIMR (SSIM with L= Range), SSIMM (SSIM with L= Maximum absolute difference), to compare proportions of pre-school children susceptible to measles in years 2000 and 2001 at district level.

The values of SSIM, SSIMR and SSIMM are identical to two decimal places. The value of L used for calculating SSIMR or SSIMM does not make any difference, as we obtain similar results. With the data that is used here we found that varying the values of K_1 and K_2 does not affect the value of the structural similarity index much. We tried values in the range 0.0001-0.05. Here we use $K_1 = 0.002$ and $K_2 = 0.003$. The SSIM, SSIMR and SSIMM values of 0.88 are close to 1, and are interpreted as the maps being more similar than different.

With the difference based measures in Section 5.6, we have p-values to assess the differences, but with the image analysis measures there is no frame of reference to decide if two maps are sufficiently different, therefore we will develop simulation tests.

5.8 Summary

In this chapter we have developed some methods that can be used to compare two or more disease maps. Although visual comparison of maps helps in identifying the differences and similarities of the spatial distributions of the maps, use of other measures may help in reaching more informative conclusions. Some of these methods have the advantage of comparing maps objectively without the individual observer's influence.

The methods considered were ratio maps, difference maps, pseudo-colour maps, and plots of model parameters (overall mean and structured and unstructured variability) against time. Methods used to measure the spatial autocorrelation, i.e. Moran's I and Geary's c with their spatial correlograms, were also considered. We also adapted some methods used in image analysis to compare a distorted image to a reference image, and some empirical distribution function methods used for point processes.

For the image analysis based methods, MSD and PSNRR/PSNRM do not take into account the spatial structure while SSIM does. For the empirical distribution function methods, IRD, MSDI and MDD do not take into account the neighbourhood structure, while AVND, MSND and MDSN make use of neighbourhood structure. Methods which do not use the neighbourhood structure may be more helpful in detecting the differences due to change in rate(s) of region(s) anywhere in the map, while methods that do use this structure may be able to detect differences due to change in the rates of the neighbouring regions. The spatial autocorrelation methods may confirm an existence or non-existence of positive spatial autocorrelation for the two maps, but these methods will not assist in detecting the differences or similarities that exist between the maps.

All these methods were applied to proportions (raw data) of susceptibility to measles for pre-school children for 2000 and 2001. In Chapter 7 the methods are applied to both raw and smoothed data to see if it makes any difference to the results.

Visual comparison of the two maps (Figure 5.3) suggests that the spatial distributions are not very different, with just a few districts which had high susceptibility in 2000 having slightly lower susceptibility in 2001. When visually comparing maps it is not very easy to know whether the rate of a region has changed or not, so use of the ratio, difference and pseudo-colour maps may help. For the example in this chapter, the ratio, difference and pseudo-colour maps were able to reveal the regions in which susceptibility to measles changed, i.e. was lower/higher in 2001 than 2000, and the regions where measles susceptibility was similar.

Maps of the same disease at different time points may differ because of a change in the mean level or change in variability due to structured or unstructured heterogeneity. When using disease mapping models, the values of these parameters are obtained from the models. The plots of these parameters against time will show how each of these parameters has changed over time. In the example used here, as we were just comparing two maps, it is easy to see what has happened to these parameters. These values suggested no change in structured variability, i.e. spatial correlation was similar. This similarity is confirmed by the spatial autocorrelation methods. There is a very slight change in the mean and the unstructured variability, i.e. susceptibility in districts became slightly more dissimilar in 2001 than it was in 2000.

The empirical distribution function methods are based on differences. These may give an informative idea as to how the spatial distributions of the values of the spatial variable differ or are similar between maps. For the example used here, we find that, except for MDD, all of these methods indicate that the two maps are not significantly different. It is hoped that the use of these methods can detect any change in the spatial distribution, including clustering, as some methods target the neighbourhood structure.

For the image analysis based methods, MSD and PSNRR/PSNRM may help in quantifying the existence of any differences between the maps. For PSNRR/PSNRM, since we obtain different values depending on what the range value P is, these two methods will have to be compared to see which one performs best in general. Both MSD and PSNRR/PSNRM indicate that the maps are not very different. SSIM/SSIMM/SSIMR compare the structures of two maps. These indicated some similarities in the maps.

In the next chapter, all of these methods will be tested more extensively using simulated data. In Chapter 7, the methods which perform well in the simulations will be applied to susceptibility to measles data for pre-school and primary 1 and 2 school children in Scotland. In Chapter 7 we also explore the use of these methods to compare maps based on two different spatial variables, i.e. with different means. For this we use the NHS24 data and compare call uptake for different health syndromes.

Moran's I and Geary's c will not be pursued further as they are not measures to detect differences or similarities but to inform us of the existence of spatial autocorrelation, although differences in the measures between two maps will suggest some difference in spatial pattern.

Summary of Suggested/Developed Measures

1. **Map Based Methods**: Ratio maps, Difference maps and Pseudo-colour maps.
2. **Plots of parameters**: Overall mean, Unstructured and Structured Variation.
3. **Spatial Autocorrelation Methods**: Moran's I, Geary's c and their spatial correlograms.
4. **Image Analysis Based Methods**: Mean Square Difference, Structural similarity Index (SSIM/SSIMM/SSIMR) and Peak-to-Signal-Noise Ratio (PSNRR/PSNRM).

5. **Point Process Based Methods**: Inter Region Differences (IRD), Most Similar Differences (MSDI), Most Dissimilar Difference (MDD), Average Neighbour Differences (AVND), Most Similar Neighbour Differences (MSND) and Most Dissimilar Neighbour Differences (MDND).

Chapter 6

Simulation Study

6.1 Introduction

In Chapter 5 we discussed the new methods or measures that we develop to compare two or more disease maps. In this chapter, a simulation study is carried out in order to assess the sensitivity and power of the methods to detect differences between two maps, in the case when the mean level or variances due to unstructured or structured heterogeneity have changed. This will help in understanding and knowing the suitability of the measures to detect differences. Also recommendations about the ones that are more useful will be given.

Before carrying out the full simulation described above, the methods are compared using data simulated from an existing map, by adding noise to the values of the map.

The measures used are Mean square difference (MSD), the two peak to signal noise ratio measures, referred to here as PSNRR and PSNRM, the Structural Similarity Index Measures (SSIM, SSIMM and SSIMR), and the point process based measures referred to as Inter Region Differences (IRD), Most Similar Differences (MSDI), Most Dissimilar Differences (MDD), Average Neighbour Differences (AVND), Most Similar Neighbour Differences (MSND) and Most Dissimilar Neighbour Differences (MDND) measures (see Chapter 5).

6.2 Generating Data from a Map

The map of proportions of susceptibility to measles (raw data) for 2000, for 56 districts, is used to simulate data and assess the ability of the methods to detect change. This is achieved by multiplying the proportions susceptible in each district by different multiples, and for each multiple, each method is assessed. Two extreme cases are explored. One is when the whole map or all districts change at the same rate (a large change in the map), i.e. all proportions are scaled by the same multiple, and the second case is when only one district changes (a small change in the map), i.e. the proportion for only one district is scaled by a multiple. The next section gives the results obtained, using the multiples 1.10, 1.20, 1.30, 1.40, 1.50, 1.60, 1.70, 1.80, 1.90 and 2.0 (1.10 is a 10% change and 2.0 is a 100% change).

6.2.1 Results

Figure 6.1 shows maps of susceptibility to measles for pre-school year 2000 when the proportions for all the districts are multiplied by each of 1.1, 1.2 and 1.3 to generate a new map. All the districts are getting darker very quickly, and for the multiple 1.3, all the districts except 1 are darker. In Figure 6.2 only Stirling district (see Figure 1.4) is multiplied by 1.1, 1.2 and 1.3. The district gets darker quickly.

Table 6.1 gives values of the image analysis based methods MSD, PSNRM, PSNRR, and SSIM and p-values of point process based methods, IRD, MSDI, MDD, AVND, MSND and MDND when the proportions are changing. For both cases, except for PSNRM which is constant throughout, all the image analysis based measures detect the changes.

PSNRM is a constant because of the way it is obtained. Let X_i and $Y_i = kX_i$, where $i = 1, ..., n$ are the rates of map 1 and map 2 respectively.

$$\text{PSNRM}=\frac{(max|X_i-kX_i|)^2}{MSD} = \frac{(max|X_i(1-k)|)^2}{\sum_{i=1}^{n}(X_i-kX_i)^2/n} = \frac{(1-k)^2 max(|X_i|)^2}{\sum_{i=1}^{n}X_i^2(1-k)^2/n} = \frac{max|X_i|^2}{\sum_{i=1}^{n}X_i^2/n} = \text{constant}$$

while $\text{PSNRR}=\frac{max(X_i,kX_i)-min(X_i,kX_i)}{MSD} = \frac{kX_{(n)}-X_{(1)}}{MSD}$, where $X_{(n)}$ and $X_{(1)}$ are the maximum and minimum of the $X_i s$. For the image analysis based methods, MSD is closer to zero when maps are the same. The MSD value clearly increases in both cases as the change in the maps increases. PSNRR is taken to be infinity when there is no change in the map. It can be observed that PSNRR changes fast in both cases, though the

PSNRR values for when one district changes are larger than when all districts change.

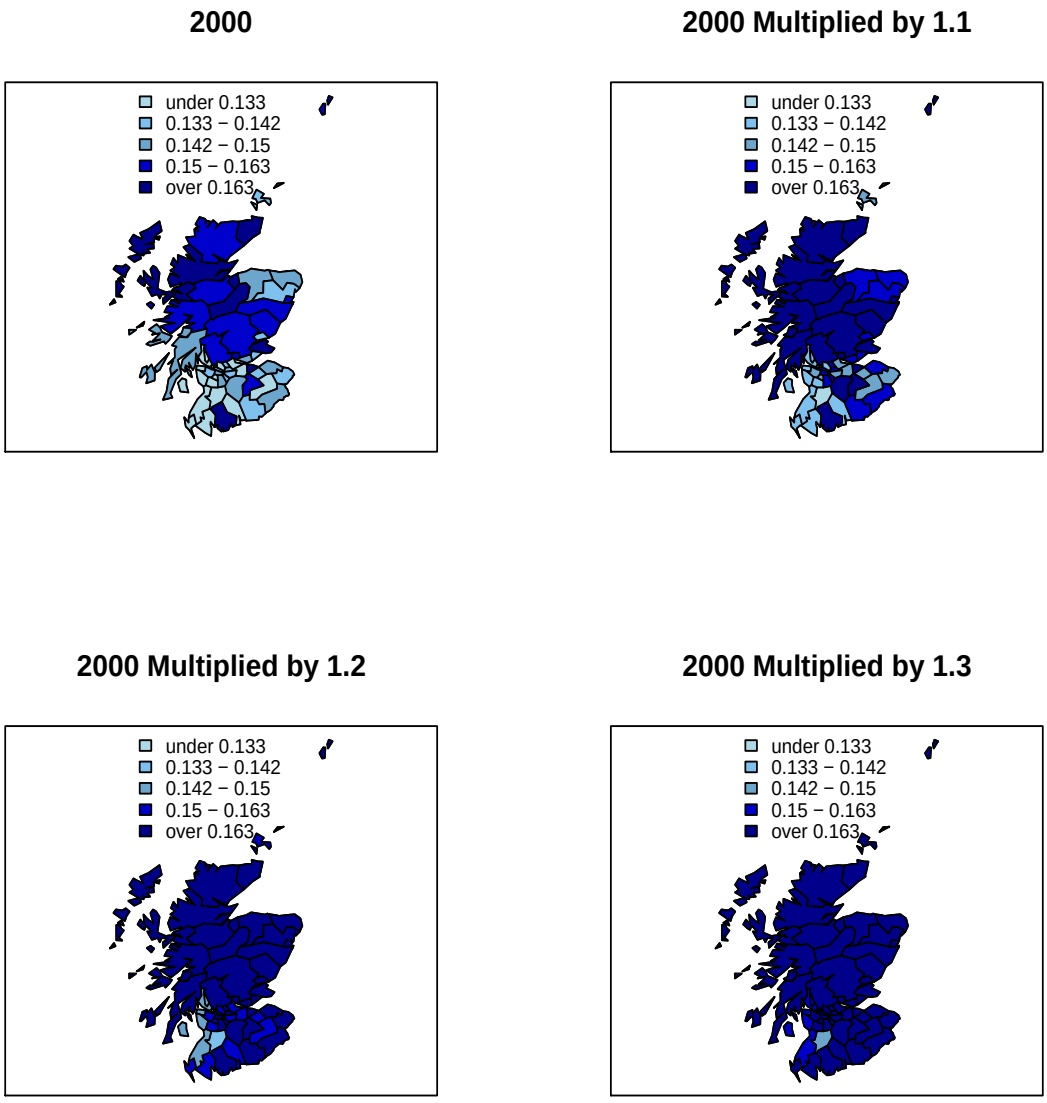

Figure 6.1: District susceptibility maps based on proportions (raw data) of pre-school children susceptible to measles for 2000, and multiplying the proportions by 1.1, 1.2 and 1.3.

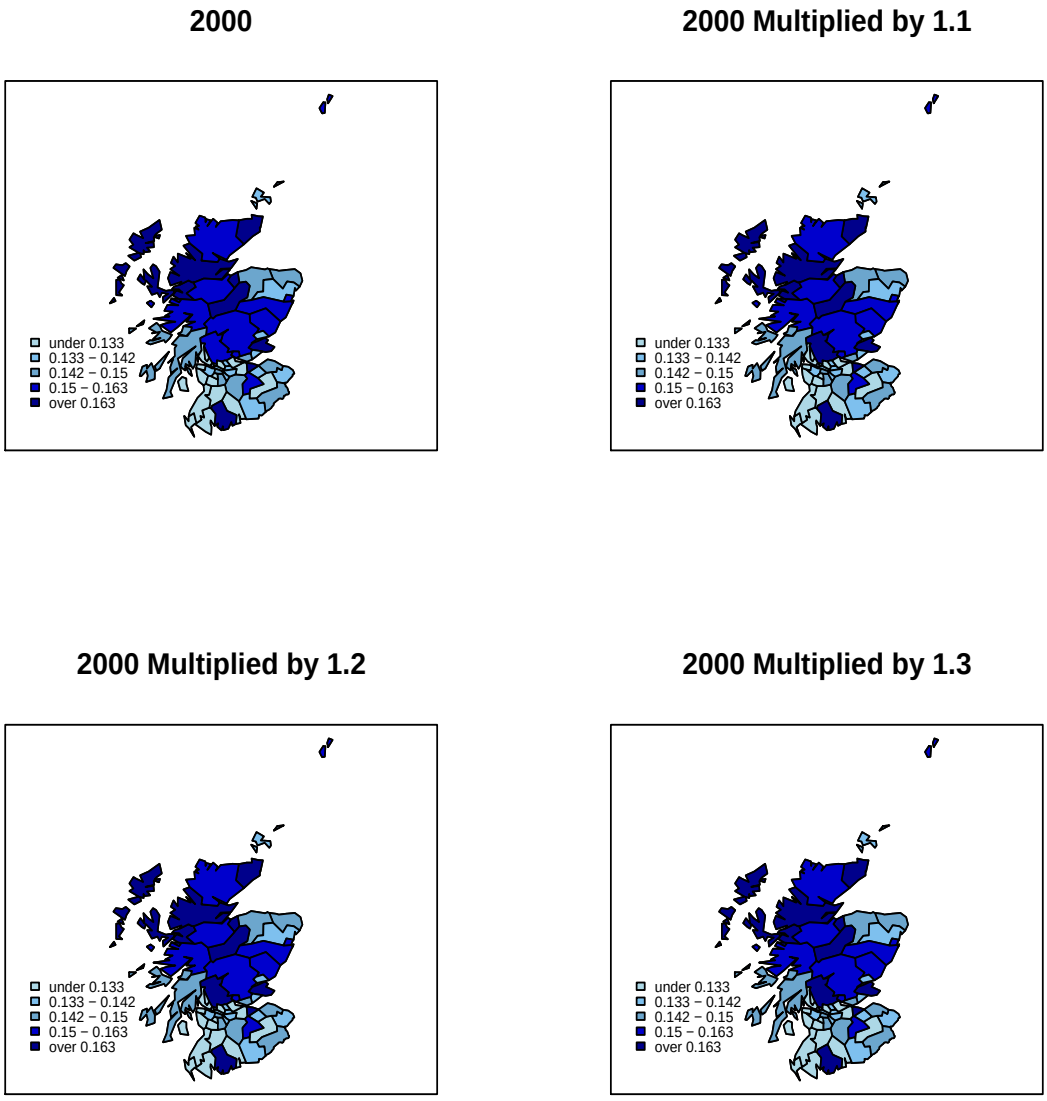

Figure 6.2: District susceptibility maps based on proportions (raw data) of pre-school children susceptible to measles for 2000, and multiplying only the proportion for Stirling district (see Figure 1.4) by 1.1, 1.2 and 1.3.

All Districts Changing										
Multiple	MSD	PSNRR	PSNRM	SSIM	IRD	MSDI	MDD	AVND	MSND	MDND
1.1	0.00022	84.3	2.4	0.99	0.038	0.98	0.036	0.59	0.89	0.74
1.2	0.00088	28.8	2.4	0.97	$1.5e-05$	0.77	$2.8e-05$	0.43	0.58	0.30
1.3	0.00197	16.7	2.4	0.93	$1.0e-10$	0.46	$2.3e-08$	0.13	0.30	0.082
1.4	0.00351	11.9	2.4	0.89	$1.1e-16$	0.46	$1.3e-11$	0.028	0.2	0.029
1.5	0.00548	9.4	2.4	0.85	0.00	0.46	$9.1e-15$	0.0082	0.082	0.0044
1.6	0.00789	7.9	2.4	0.81	0.00	0.46	$1.1e-16$	0.002	0.05	0.0011
1.7	0.01074	6.9	2.4	0.76	0.00	0.46	0.00	0.00042	0.05	0.00022
1.8	0.01402	6.2	2.4	0.72	0.00	0.23	0.00	0.00018	0.029	$3.8e-05$
1.9	0.01775	5.7	2.4	0.68	0.00	0.098	0.00	$7.3e-05$	0.029	$1.5e-05$
2.0	0.02191	5.3	2.4	0.64	0.00	0.06	000	$1.1e-05$	0.016	$5.8e-06$
Stirling District Changing										
Multiple	MSD	PSNRR	PSNRM	SSIM	IRD	MSDI	MDD	AVND	MSND	MDND
1.1	$4.4e-06$	2926	56	1.00	1.00	1.00	1.00	0.89	1.00	0.97
1.2	$1.8e-05$	731	56	0.98	0.79	1.00	1.00	0.74	1.00	0.74
1.3	$3.9e-05$	325	56	0.96	0.53	1.00	1.00	0.58	1.00	0.43
1.4	$7.0e-05$	183	56	0.93	0.42	1.00	1.00	0.43	1.00	0.30
1.5	$1.1e-04$	132	56	0.89	0.34	1.00	0.098	0.43	1.00	0.30
1.6	$1.6e-04$	117	56	0.85	0.34	1.00	$2.3e-08$	0.43	1.00	0.30
1.7	$2.2e-04$	107	56	0.81	0.32	1.00	$9.1e-15$	0.30	1.00	0.30
1.8	$2.8e-04$	100	56	0.77	0.30	1.00	0.00	0.30	1.00	0.30
1.9	$3.5e-04$	94	56	0.72	0.30	1.00	0.00	0.30	1.00	0.30
2.0	$4.4e-04$	90	56	0.68	0.28	1.00	0.00	0.30	1.00	0.30

Table 6.1: Values of MSD,PSNRR, PSNRM, SSIM and p-values for IRD, MSDI, MDD, AVND, MSND and MDND, for changing susceptibility for all districts (top) and for changing only Stirling district (bottom).

SSIM/SSIMM/SSIMR are taken to be 1 when maps are the same. The values of these three methods are all identical to two decimal places, indicating that the constants used in SSIMM and SSIMR are not useful here as they are only needed when the mean and variances of the maps are very close to zero, thus we only refer to SSIM. For a 10% change in all districts the value of 0.99 is obtained, indicating a slight difference in structure of the maps, and the value decreases steadily for larger changes. For a 10% change in one district the value of SSIM is 1, so change has not been detected here, but from 20% SSIM/SSIMM/SSIMR decreases steadily.

For the point process based methods, when changing the rates for all districts, the p-values of the measures show that IRD and MDD detect the change at 10% change, MDND and AVND at 40% change and MSND at 60% change. MSDI does not detect the change before 100% change. When only one district is changed, none of these

measures are sensitive to this change except MDD at 60% change. Thus, unlike the image analysis based methods, at least in this example, the point process measures are not sensitive when one district changes, and IRD and MDD are the most sensitive to change in rates when all the districts are changing

In this section the methods have been assessed using a specific data set, and it can been seen that some methods have the ability to detect change. When changing the proportions the spatial structure is not taken into account. These methods may behave differently under different circumstances. Thus, in the next section, we simulate data from a spatial model and assess these methods systematically, under different scenarios of changing mean, unstructured and structured variations. This analysis may help in determining both the sensitivity and power of each method.

6.3 Generating Data from a Model

In this simulation study we assess the ability of these methods to detect differences in disease maps. We change the mean level, the standard deviation due to unstructured heterogeneity or structured heterogeneity, one at a time. In order to achieve this, we chose to simulate the data using the log-normal model (Besag *et al.*, 1991), and based the simulation on the neighbourhood structure of the 53 (out of 56) districts of Scotland (mainland Scotland, i.e. excluding the islands). The log-normal model is chosen for the simulation as it allows the incorporation of the spatial structure.

6.3.1 Simulation Model

The observed counts are generated from the Poisson distribution

$$O_i \sim Pois(N_i\theta_i), \tag{6.1}$$

where N_i is the population size in the ith region, $i = 1, 2, ..., n$ $(n = 53)$. Now, the population of Scotland is approximately 5 million, and if divided equally among the 56 regions each will have approximately a population size of 100 000. To try to be realistic, we sample population size of region i from a Uniform distribution with a minimum of 50 000 below 100 000 and a maximum of 50 000 above 100 000, i.e $N_i \sim U(50000, 150000)$.

The $\theta_i s$ are the disease risks and are given by

$$\theta_i = exp(\alpha + U_i + V_i),\ i = 1, ..., 53, \tag{6.2}$$

where $\alpha = log(\mu)$ is the overall mean level, μ is the disease rate, V_i is the unstructured heterogeneity distributed as

$$V_i \sim N(0, \sigma_v^2), \tag{6.3}$$

σ_v^2 is the unstructured or global variability, and U_i is the structured heterogeneity and depends on the adjacent neighbours.

To simulate the $U_i s$, we follow Hsiao *et al.* (2000). U is written as a vector, i.e. $\underline{U} = (U_1, ..., U_{n-1}, U_n)$, with precision matrix $\underline{Q} = \underline{D} \times (\underline{I} - \underline{C})$, where $\underline{D}$ is an $n \times n$ diagonal matrix with $d_{ii} = n_i\lambda$, n_i is the number of neighbours for region i, and λ is the inverse variance of U ($\lambda = \frac{1}{\sigma_u^2}$), representing the local spatial effect, and σ_u^2 is the structured or local variability. $\underline{I}$ is the $n \times n$ identity matrix and $\underline{C}$ is the $n \times n$ matrix with (i, j)th entry $c_{ij} = \frac{w_{ij}}{n_i}$, where $w_{ij} = 1$ for neighbouring regions and 0 otherwise, $i, j = 1, 2, ..., n$. Now, based on results of Besag and Kooperberg (1995), as outlined by Hsiao *et al.* (2000), from the multivariate normal distribution we generate values $x_1, ..., x_{n-1}$, i.e.

$$x_i \sim MVN(\underline{0}, (\underline{Q}^*)^{-1}) \tag{6.4}$$

where $\underline{Q}^*$ is the upper left $(n-1) \times (n-1)$ matrix of $\underline{Q}$. To obtain $\underline{U}$,

$$U_n = \frac{\sum x_i}{n},\ U_i = x_i - U_n,\ 1 \leq i \leq n-1. \tag{6.5}$$

When comparing maps, it will be helpful to understand what change has taken place, i.e. is the change due to a change in mean, unstructured or structured variation between two maps. Thus, the data (map) are generated in such a way that the ability of the methods to detect differences in disease maps can be assessed when changing overall mean level, change in variance due to unstructured heterogeneity or change in variance due to structured heterogeneity. Thus if θ_{i1} and θ_{i2} are the disease risks in region i for maps 1 and 2 respectively, for region i, $i = 1, ..., n$,

$$\theta_{i1} = exp(log(\mu) + U_i + V_i), \quad \theta_{i2} = exp(log(k\mu) + U_i + V_i) \tag{6.6}$$

in the case of changing the mean level, where the overall mean $\alpha = log(\mu)$. A change in the mean level takes place when we change the disease rate μ, thus we take the

logarithm of the changed disease rate, i.e. $log(k\mu)$. Also we take

$$\theta_{i1} = exp(\alpha + U_i + V_i), \quad \theta_{i2} = exp(\alpha + kU_i + V_i) \tag{6.7}$$

in the case of changing structured variation, and

$$\theta_{i1} = exp(\alpha + U_i + V_i), \quad \theta_{i2} = exp(\alpha + U_i + kV_i) \tag{6.8}$$

in the case of changing the unstructured variation, where in all cases the multiple is taken as $k = 1, 1.10..., 4$, in steps of 0.10, and $k = 1$ under the null hypothesis of no change between maps 1 and 2. For $k > 1$, $k = 1.10$ is a 10% change and $k = 4$ represents a 300% percentage change. Altogether there are 31 of these multiples at intervals of 0.10. The observed data are then obtained as

$$O_{i1} \sim Pois(N_i\theta_{i1}), \;\; O_{i2} \sim Pois(N_i\theta_{i2}), \tag{6.9}$$

where N_i, U_i and V_i are identical in both sets of data. For each map and for each region we obtain a rate based on the observed count generated and the total number in the region i. The two maps are then compared based on these rates and the ability of each of the methods to detect the differences is assessed at each kth multiple.

6.3.2 Method of Generating Data

The function for simulation of data from the above model was written in R and the simulation code is given in Appendix C. In order to assess each of the developed methods, we generated data sets with different parameter values for μ, σ_v^2, λ. We considerered the case when the disease is common, rare, and when there are lots of zero counts (as for a very rare disease), i.e. with $\mu = \frac{1}{10}$ ($\alpha = -2.30$), $\mu = \frac{1}{1000}$ ($\alpha = -6.91$), $\mu = \frac{1}{100000}$ ($\alpha = -11.51$) respectively. As this simulation is not based on the results of any real data set, it is not easy to assign the values of the variances of U and V, but the important aspect here is to be able to assess the behaviour of the developed methods when these parameters change. We consider two possibilities for the variances. In the first instance, for each case of μ we take the standard deviation of V and U to be the same magnitude as $\mu = 0.1$, i.e. we have three scenarios:

- for a common disease: $\mu = \frac{1}{10}$, $\sigma_v = 0.1$, $\lambda = 100$ ($\sigma_u = 0.1$),
- for a rare disease: $\mu = \frac{1}{1000}$, $\sigma_v = 0.1$, $\lambda = 100$ ($\sigma_u = 0.1$),

- and for a very rare disease: $\mu = \frac{1}{100000},\ \sigma_v = 0.1,\ \lambda = 100\ (\sigma_u = 0.1)$.

In the second instance, since in the case of a rare disease the absolute value of α is approximately 3 times the value of α for a common disease, we take the standard deviation parameters to be 3 times the standard deviation parameters of the common disease case. For very rare disease, the value of α is approximately 5 times the value of α in the common disease case, thus the parameter values for the variabilities in this case were taken to be 5 times those in the common disease case. Thus we have

- for a common disease: $\mu = \frac{1}{10},\ \sigma_v = 0.1,\ \lambda = 100\ (\sigma_u = 0.1)$,
- for a rare disease: $\mu = \frac{1}{1000},\ \sigma_v = 0.3,\ \lambda = 100/9\ (\sigma_u = 0.3)$,
- for a very rare disease: $\mu = \frac{1}{100000},\ \sigma_v = 0.5,\ \lambda = 4\ (\sigma_u = 0.5)$.

It is expected that the size of the population in the regions will be between 50 000 and 100 000. Therefore, the numbers of observations are expected to be between 5000 and 10 000, 50 and 100, 0.5 (0) and 1 for the case of common, rare and very rare disease respectively.

In each of the three scenarios that we consider, i.e. rare disease, common disease and very rare disease, and for each multiple k, we generated different data sets, based on whether we are assessing methods when there is a change in mean level, or change in variability due to unstructured heterogeneity, or change in variability due to structured heterogeneity.

For each multiple k, we generate N_i, V_i and $U_i,\ i = 1, ..., n$, where n is number of regions, and use them to obtain θ_{i1} and θ_{i2} which are then used to obtain the observed values O_{i1} and O_{i2} for maps 1 and 2. Firstly the generation of data was done with $N = (\underline{N}_1, ..., \underline{N}_n)$ fixed for all multiples, i.e. generating $\underline{N}$ once, and it was observed that the results obtained were similar to when $\underline{\mathrm{N}}$ is not fixed. We decided to generate $\underline{N}$ separately for each multiple. This is done as follows (see detailed code in Appendix C):

- specify μ, σ_v and σ_u,
- for each of simulations 1:1000,
- for k in 1,1.1,...,4,

- generate V_i, U_i and N_i, $i = 1, ..., n$,
- generate O_i, $i = 1, ..., n$.

The number of simulations was chosen to be $n = 1000$, as this was thought to be large enough to help us obtain good estimates of the mean for each measure and simulations where quick to run. In practice, to determine the number of simulations needed to get a good estimate of the mean the following sample size formula can be used,

$$n = \frac{(z_{1-\alpha/2}s)^2}{\varepsilon^2} \text{ (Lyman, 1998)},$$

where n is now the number of simulations, $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution and α is the significance level, ε is the desired level of accuracy of the estimate (which has to be specified), and s is the estimate of the population standard deviation (obtained from a pilot study as this study is not based on a real situation).

For each measure, the desired level of accuracy ε has to be specified. Also, since the simulation is done at each kth multiple, for each measure and each kth multiple we will obtain a different value of n. Since this simulation involves generating data for all the measures at the same time, one might average the number of simulations obtained for all the measures to obtain one value of n, or choose to use the largest of all the ns depending on whether the value is sensible to use (i.e. not very large and time consuming).

To summarise the values of the different measures obtained from the simulations, we use the average of the obtained values for each measure over the 1000 simulations, together with the standard deviation and standard error. The average values are plotted against the multiple k so it can be seen how the measure behaves (showing sensitivity of the measure to deviations from $k = 1$). We also use the critical value approach to summarise results. Here, when the measure behaviour is such that when the difference in the maps increases the value of the measure increases, the 95th percentile of the null distribution of the measure is obtained (i.e. the distribution of the measure when the multiple $k = 1$), and when the behaviour is such that the value of the measure decreases when there are differences in the maps, the 5th percentile of the null distribution is obtained. For each case and each measure the proportion of obtained values that are more extreme than the percentile obtained is plotted against

the multiple k. These plots will indicate the power of the measure to detect differences.

Another approach used is a p-value approach. For each multiple k, for each new simulation a p-value was found for each observed measure by finding the proportion of the 1000 simulated values making up the null distribution ($k = 1$) that are at least as extreme as the value of the observed measure, which may be extremely large or small depending on the measure. The average of the p-values is then plotted against the multiple k, to help observe how statistically significant the value of the measure is to detect a change.

6.4 Results

In this section we discuss the results of the simulation study. Here the results reported are for the case in which the standard deviations of U and V are taken to be the same (i.e. 0.1) for each of the three scenarios considered, as the results for the other case are similar.

When there is a change in the map, it is expected that the value of MSD should increase and the values of PSNRR/PSNRM and SSIM/SSIMM/SSIMR should decrease. SSIM, SSIMM and SSIMR give similar results and therefore here we discuss SSIM as SSIMM and SSIMR are variations of SSIM and are only useful when SSIM is unstable (see Chapter 5). PSNRR and PSNRM give similar results except in a few cases, therefore the results given here will be based on PSNRR and we will refer to PSNRM whenever there are differences between the two measures. For the point process measures IRD, MSDI, MDD, AVND, MSND and MDND, the p-values should decrease as the maps change more. When there is a significant difference the p-value should be less than 0.05 (for a 5% test).

6.4.1 Effects of Changing Mean Level

Figures 6.7-6.9 show maps from the simulated data when changing mean level for common, rare and very rare disease respectively. It can be observed that when changing mean level the differences in the maps can be easily seen. It is observed that as the disease rate μ increases and overall mean increases, more regions are getting darker, i.e. the number of regions with high rates increases quickly for both common and rare

disease, while for very rare disease the change is very slow. The measures may be able to help in detecting and quantifying this kind of change. Figures 6.6-6.8 give plots of average values for each measure against the multiples of the mean when changing the mean level for the three scenarios, i.e. $\mu = \frac{1}{10},\ \sigma_v = 0.1,\ \lambda = 100$, for common disease, $\mu = \frac{1}{1000},\ \sigma_v = 0.1,\ \lambda = 100$ for rare diseases and for very rare diseases $\mu = \frac{1}{100000},\ \sigma_v = 0.1,\ \lambda = 100$ respectively.

Mean Square Difference (MSD)

Among all the measures, MSD seems to be the best measure and performs well (it is sensitive and powerful) for all three scenarios (common, rare and very rare disease), but is less sensitive and powerful for a very rare disease. Change is detected at 10% for all scenarios. A multiple close to $k = 1.75$ is required before the power becomes high or the p-value low in the case of very rare disease (Figures 6.9 and 6.10). Absolute values of MSD increase as the differences in the maps increase but there is no reference point for the method, so the Monte Carlo p-value approach will be useful to decide when a significant difference is present (this is done in Chapter 7).

Peak-to-Signal Noise Ratio (PSNRR)

PSNRR performs well only for common and rare disease. For very rare disease, PSNRR decreases very slowly and fluctuates and the power is low (Figure 6.9). The p-value plots (Figure 6.10) show that for PSNRR when the disease is very rare the p-value does not go below 0.05, so it will not be useful in this case. Change is detected at 10% for both common and rare disease.

Structural Similarity Index Measure (SSIM)

SSIM performs well for a common disease and moderately well for a rare disease. When the disease is very rare, SSIM fluctuates and is unstable. This measure does detect differences due to change in mean level for common and rare disease at 10% and 40% change in disease rate respectively, but the value of the measure does not change very fast. This may be because SSIM depends on the mean, variances and correlation of the two maps, and since here the unstructured and structured variability is the same for both maps and only the mean is changing, the value of the measure may not change very quickly.

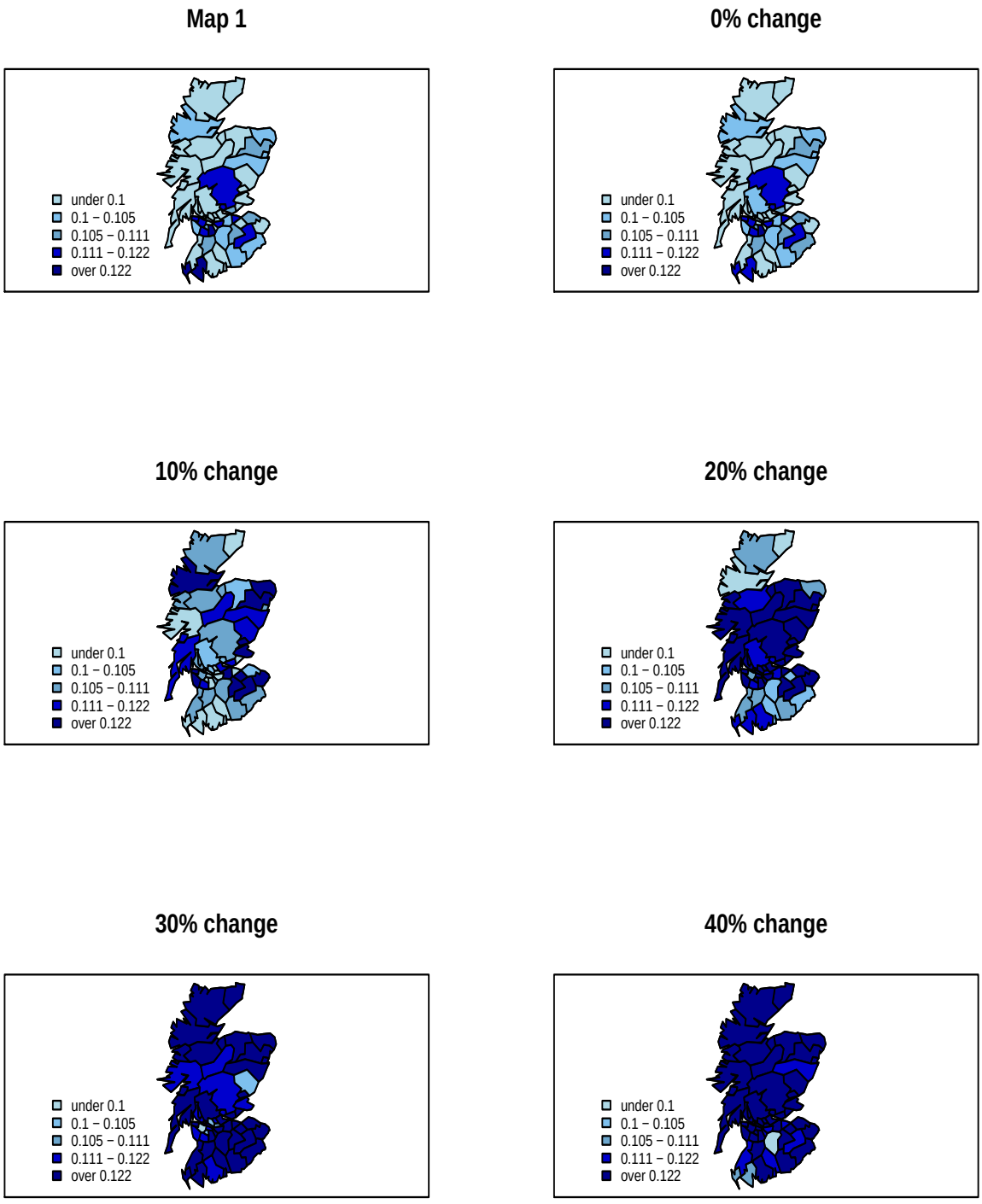

Figure 6.3: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in disease rate μ for a common disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{10}$.

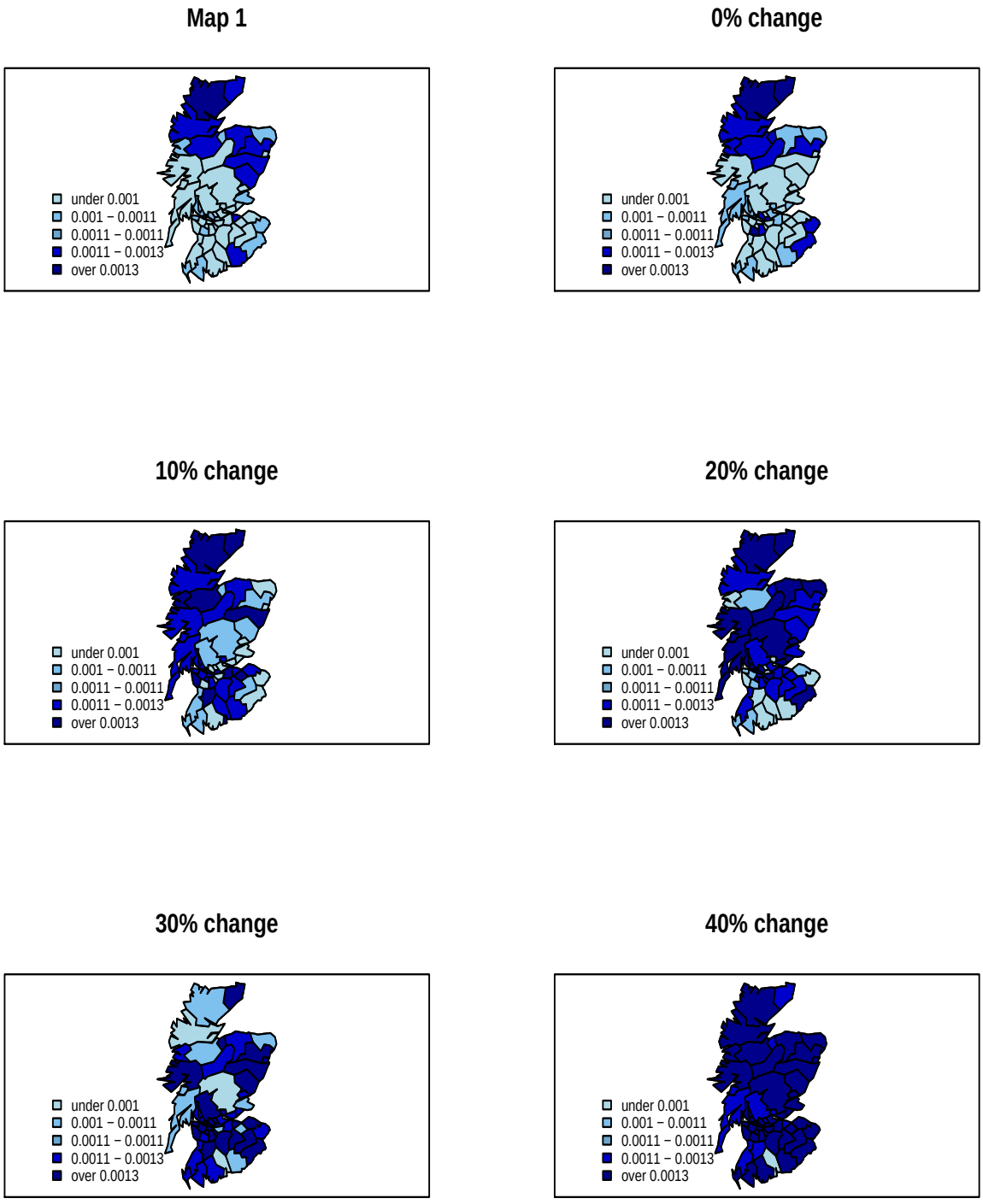

Figure 6.4: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in disease rate μ for a rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{1000}$.

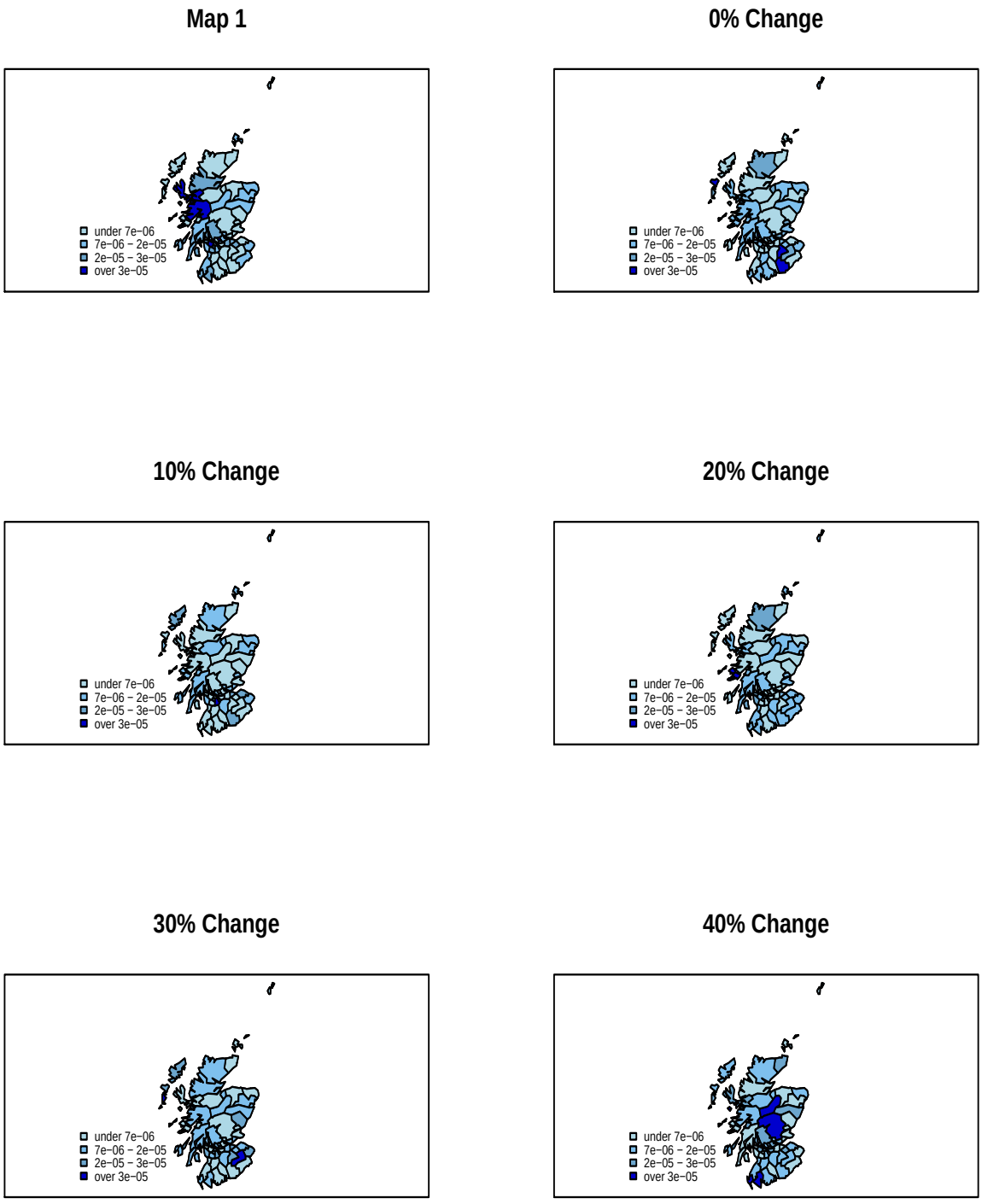

Figure 6.5: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in disease rate μ for a very rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{100000}$.

Inter Region Difference (IRD) and Most Dissimilar Difference (MDD)

Generally, IRD and MDD detect the differences in a similar manner. Among the point process based methods, these are the most sensitive and powerful. For a very rare disease, IRD and MDD do not perform well and give misleading results, i.e. we obtain a value less than 0.05 when the means are the same for two maps, indicating that there are differences. The plots of average values against multiples of the mean (Figure 6.13) show that these values fluctuate as the mean level changes. The p-value plots (Figure 6.14) show that both measures do not go below 0.05 when the disease is very rare. The power to detect changes is close to zero. Overall, IRD (at 10% and 30% for common and rare disease respectively) detects changes slightly faster than MDD (at 20% and 40% for common and rare disease respectively).

Most Similar Neighbour Difference (MSND) and Most Dissimilar Neighbour Difference (MDND)

MSND and MDND behave very similarly. They perform better for a common disease than other scenarios in terms of power and sensitivity. MSND and MDND both detect a change at 30% for a common disease and at 60% for a rare disease. For a very rare disease MSND does not detect a difference until 100% change and MDND at 110% change.

Most Similar Difference (MSDI) and Average Neighbour Difference (AVND)

AVND and MSDI perform poorly for all disease scenarios and they detect the changes very late. MSDI detects change at 110%, 140% and 120% for common, rare and very rare disease. AVND detect change at 90%, 140% and 200% for common, rare and very rare disease. Unlike other measures, the p-value plots (Figure 6.10) for MSDI and AVND show that when a disease is very rare the p-value goes below 0.05 faster than other disease cases. The power plots (Figure 6.9) also show that MSDI is more powerful in the case when the disease is very rare than other disease cases, and AVND is more powerful when the disease is very rare than when it is rare.

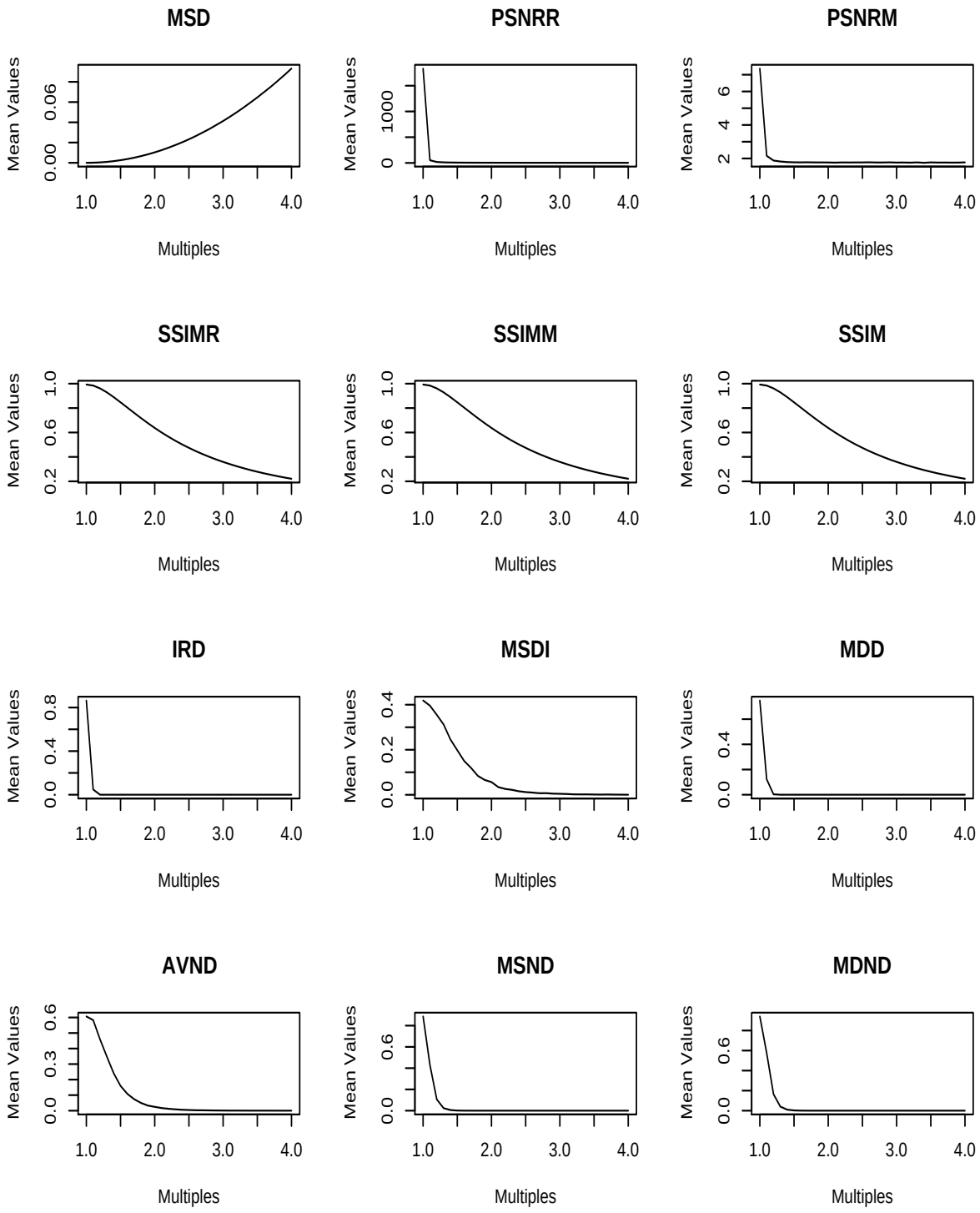

Figure 6.6: Mean of simulated values of each measure versus k for a common disease ($\mu = \sigma_u = \sigma_v = \frac{1}{10}$), when changing mean level α, with $\mu_2 = k\mu_1$, where $\mu_1 = \mu$.

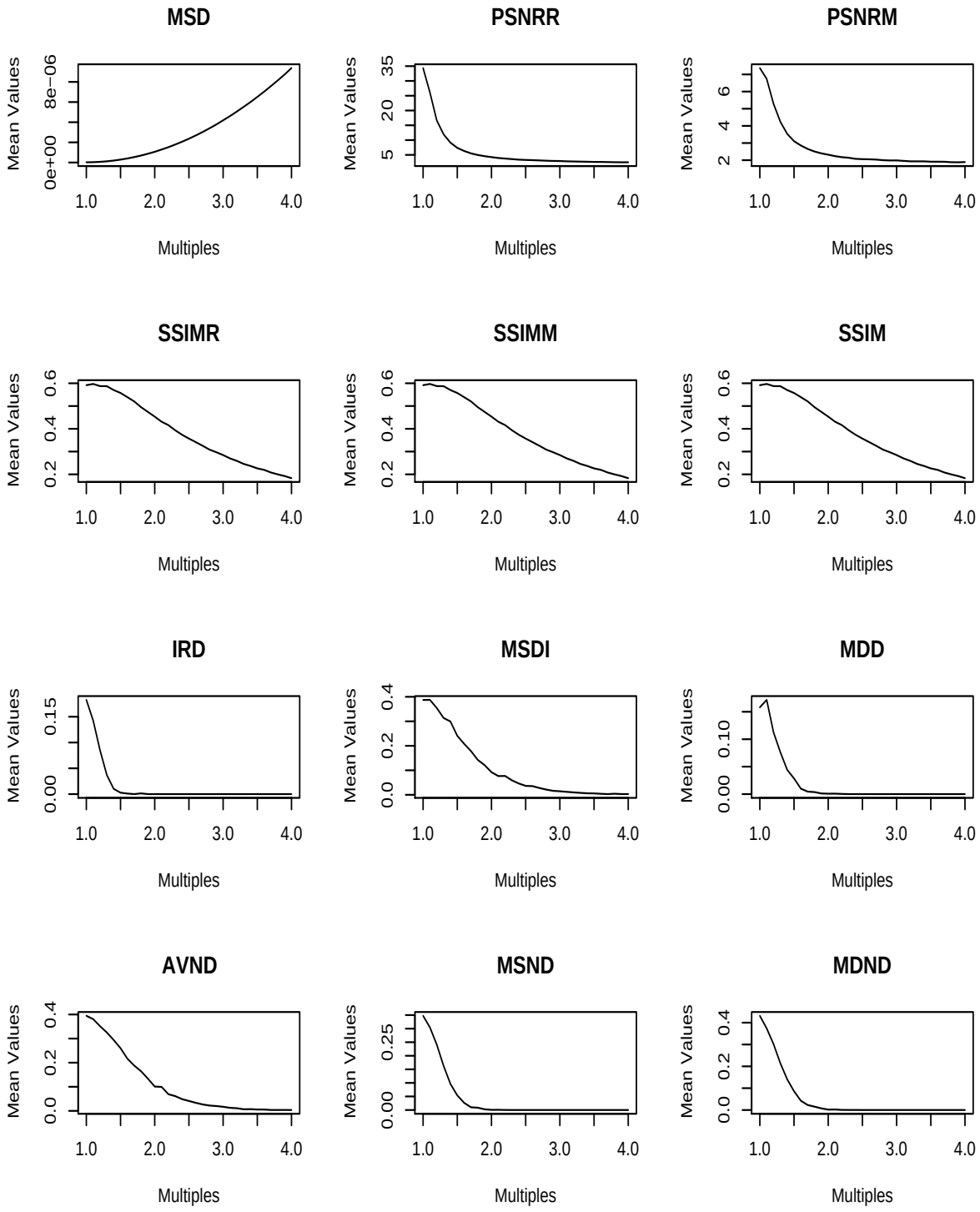

Figure 6.7: Mean of simulated values versus k for a rare disease ($\mu = \sigma_u = \sigma_v = \frac{1}{1000}$), when changing mean level α, with $\mu_2 = k\mu_1$, where $\mu_1 = \mu$.

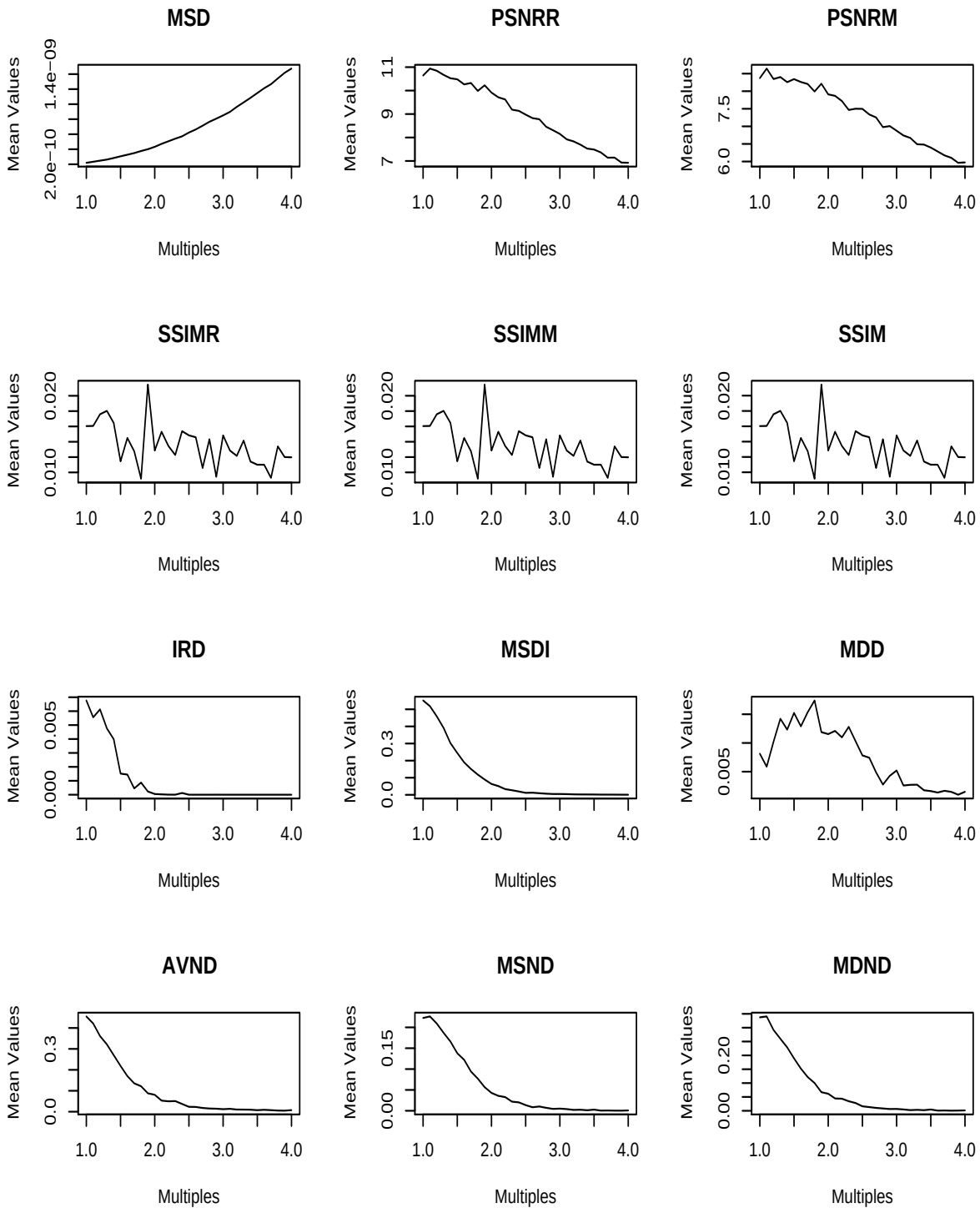

Figure 6.8: Mean of simulated values versus k for a very rare disease ($\mu = \sigma_u^2 = \sigma_v^2 = \frac{1}{100000}$), when changing mean level α, with $\mu_2 = k\mu_1$, where $\mu_1 = \mu$.

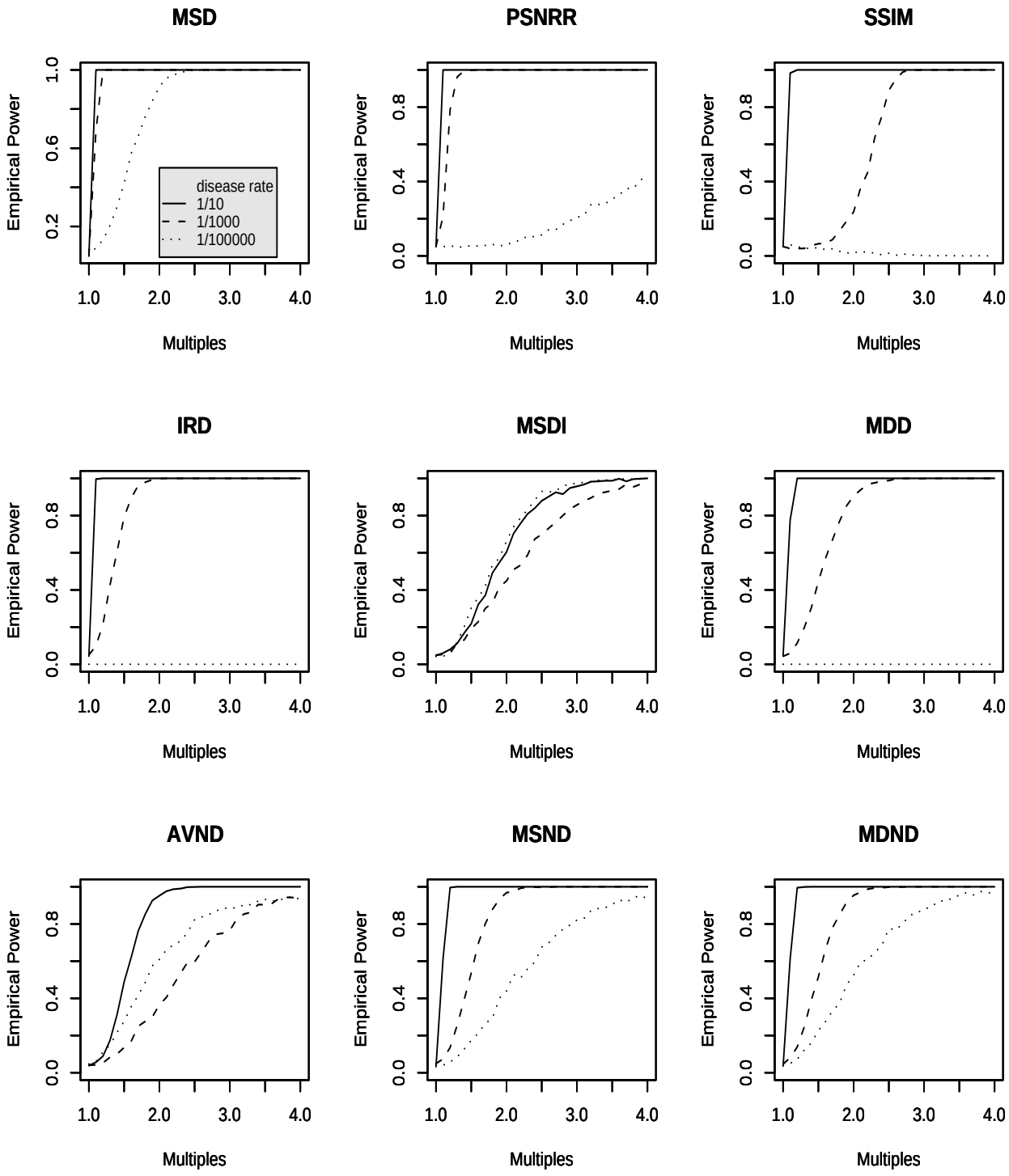

Figure 6.9: Plots of empirical power when changing mean level α, with $\mu_2 = k\mu_1$, where $\mu_1 = \mu$, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

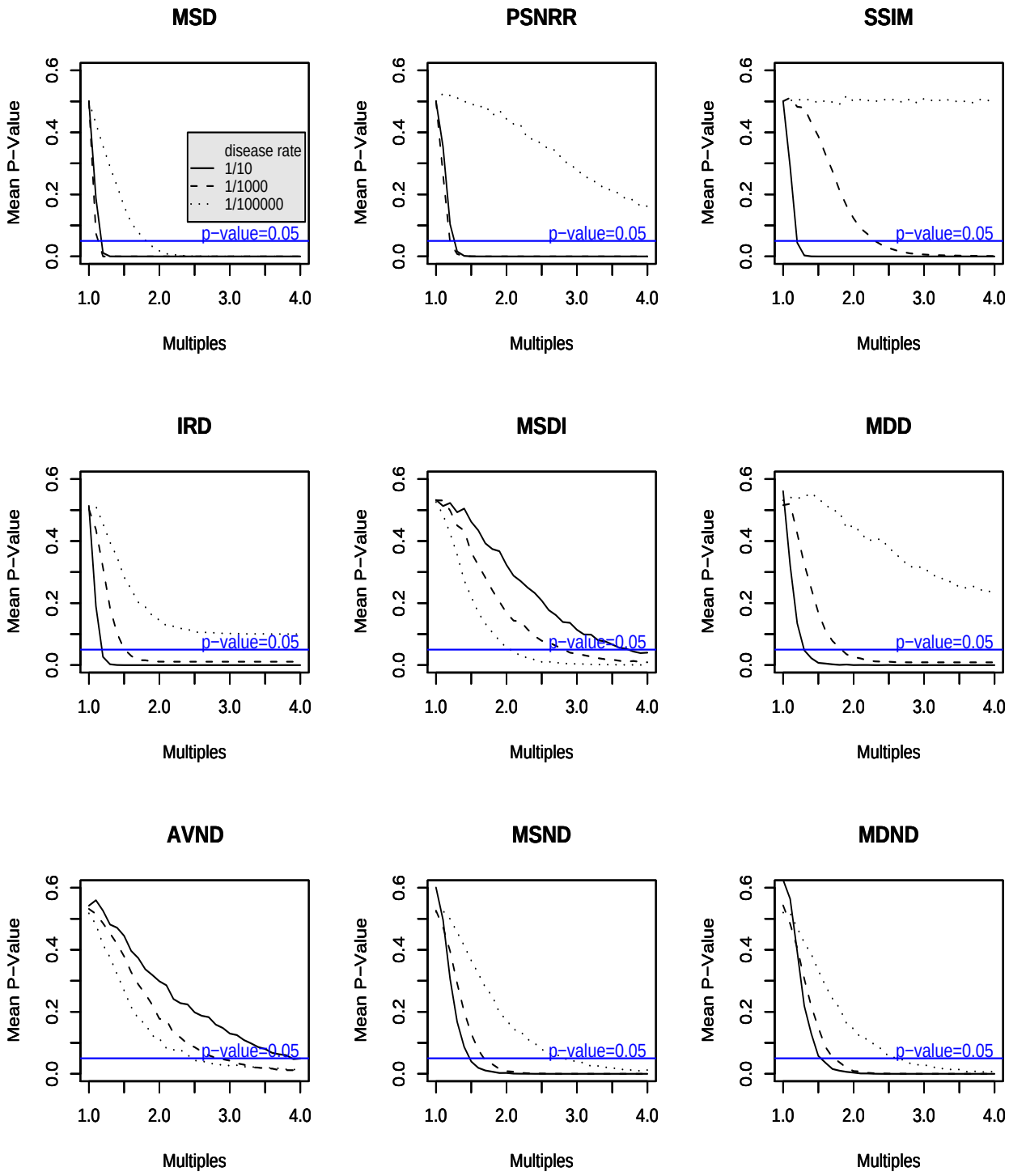

Figure 6.10: Plots of average p-values against multiples of the disease rate, with horizontal line at p=0.05, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

For common and rare disease, MSD, PSNRR/PSNRM, IRD, MDD, MSND/MDND detect differences quickly, and SSIM, MSDI and AVND change but not very fast. In the case of a rare disease, these measures do not change as fast as in the case of a common disease. For a very rare disease, only MSD detects changes at all fast, the rest of the measures are very slow in detecting the differences and MDD is also now rather unstable. Tables B.1-B.5 (Appendix B) give the average values of these measures with standard deviations and standard errors for MSD, PSNRR, SSIM, IRD, MSDI, MDD, AVND, MSND and MDND respectively.

The power plots (Figure 6.9) show that the power rises very rapidly to 1 for common disease for MSD, PSNRR and IRD, followed by SSIM, MDD, MSND/MDND. MSDI and AVND are not as powerful as the other measures. For rare disease, MSD and PSNRR are the most powerful, followed by IRD, MDD, MSND/MDND then SSIM, with MSDI and AVND being less powerful. For very rare disease, MSD is best, followed by MSDI, AVND, MSND/MDND, while the other measures have very low or near zero power.

The p-value plots (Figure 6.10), show that for common disease, the p-value of MSD, PSNRR, SSIM, IRD, MSDI, MDND/MSND changes to near zero extremely fast, while MDD and AVND change to less than 0.05 very slowly. For rare disease, MSD and PSNRR change to zero very fast, followed by IRD, MDD, MSND/MDND then SSIM, MSDI and AVND. For very rare disease, MSD changes to less than 0.05 fastest but not until k is near to about 1.75, followed by MSDI, AVND, MSND/MDND, while PSNRR, SSIM, IRD and MDD do not ever change to less than 0.05.

6.4.2 Effects of Changing Unstructured and Structured Variation

Figures 6.11-6.16 show some maps produced from the simulated data when changing structured and unstructured variability separately for common, rare and very rare disease. Unlike in the case of a change in the mean (Figures 6.3-6.5), when a change in the unstructured or structured variability has taken place the maps do not necessarily show very big differences as the variance is not changing very much. It is expected that as global variation changes regions will become less similar. For structured variation, it is expected that clusters of regions with similar rates will increase.

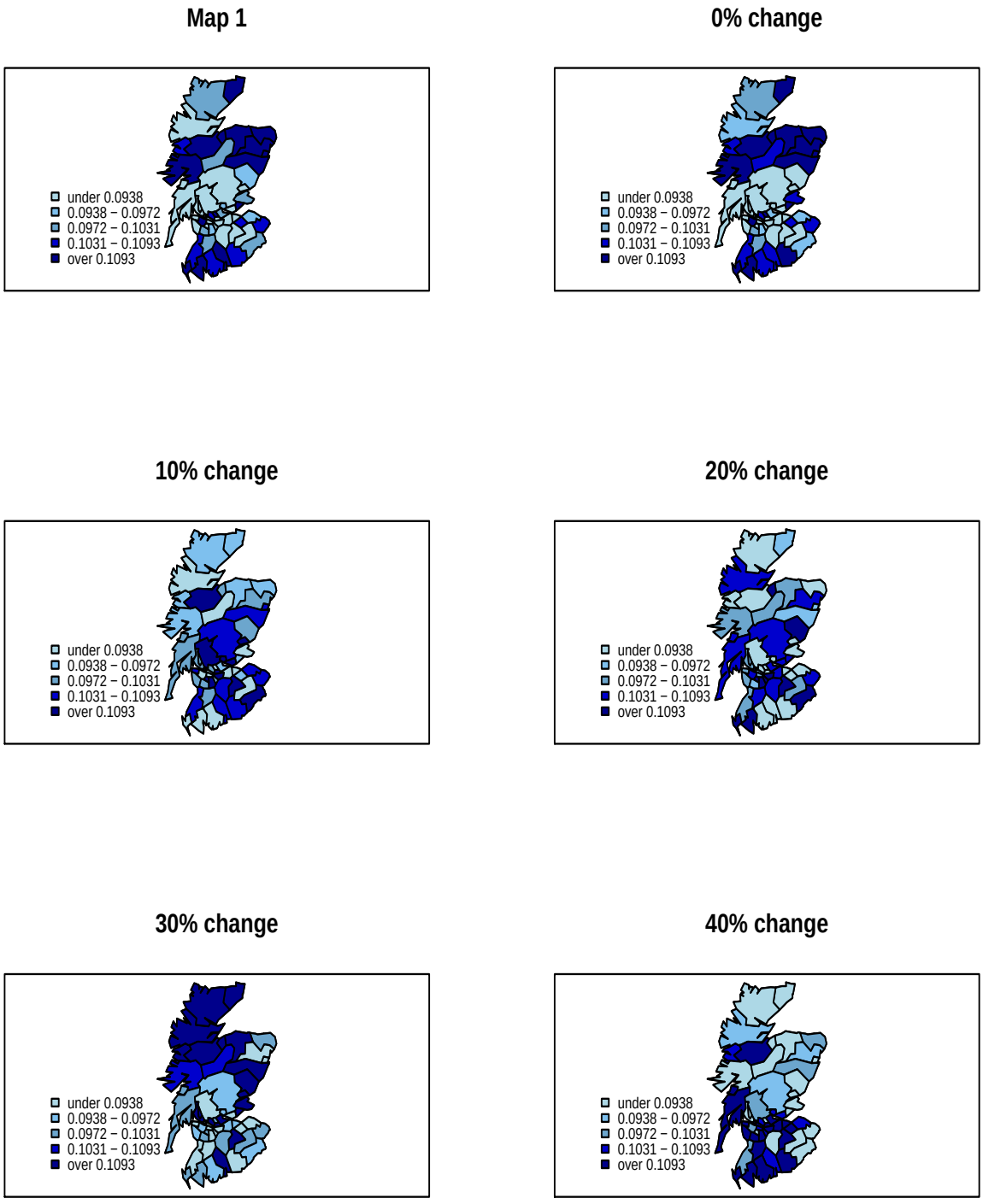

Figure 6.11: sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in variability due to unstructured heterogeneity for a common disease . Here $\mu = \sigma_u = \sigma_v = \frac{1}{10}$.

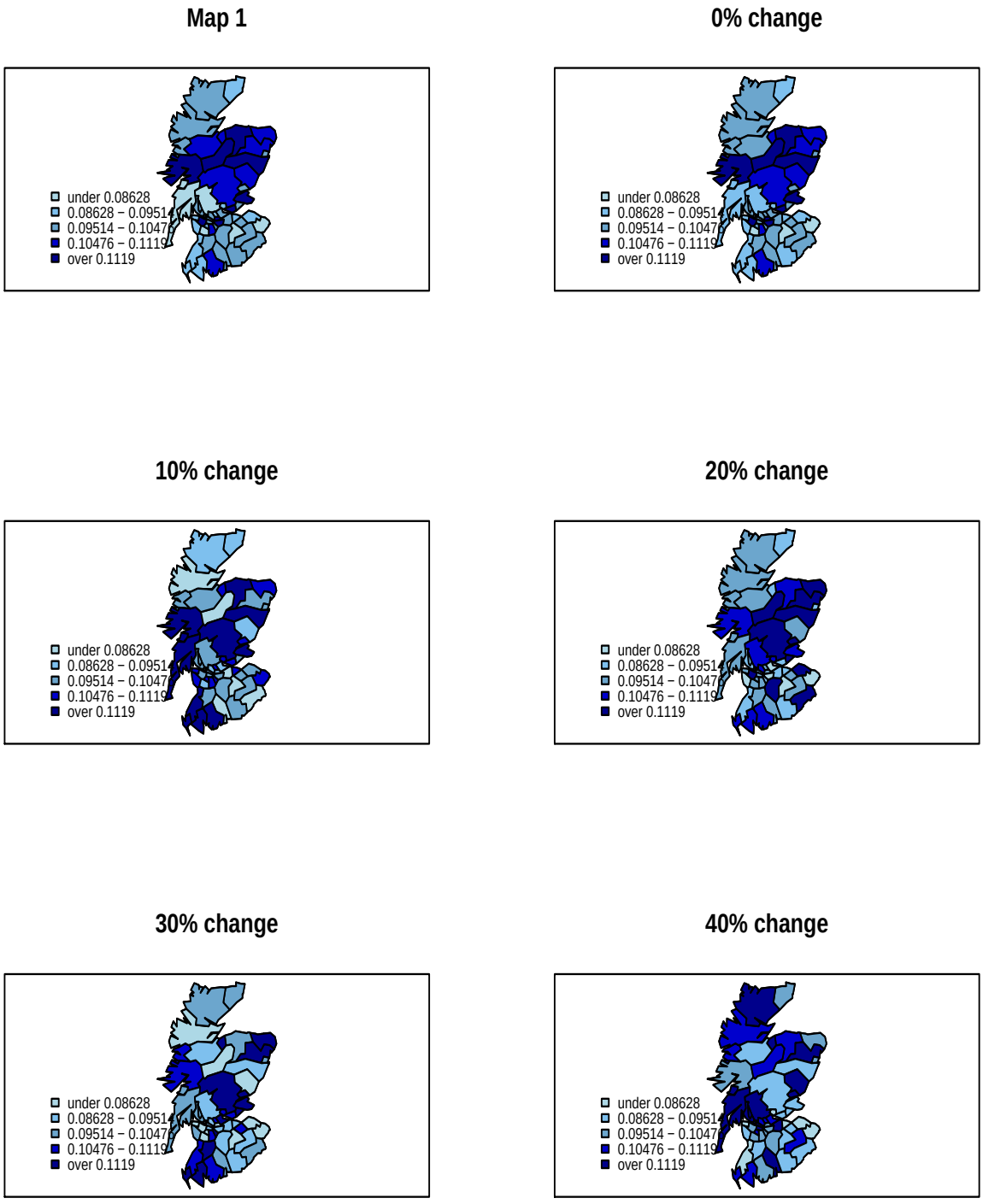

Figure 6.12: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in variability due to structured heterogeneity for a common disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{10}$.

Figure 6.13: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in unstructured variability for a rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{1000}$.

Figure 6.14: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in structured variability for a rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{1000}$.

Figure 6.15: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in unstructured variability for a very rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{100000}$.

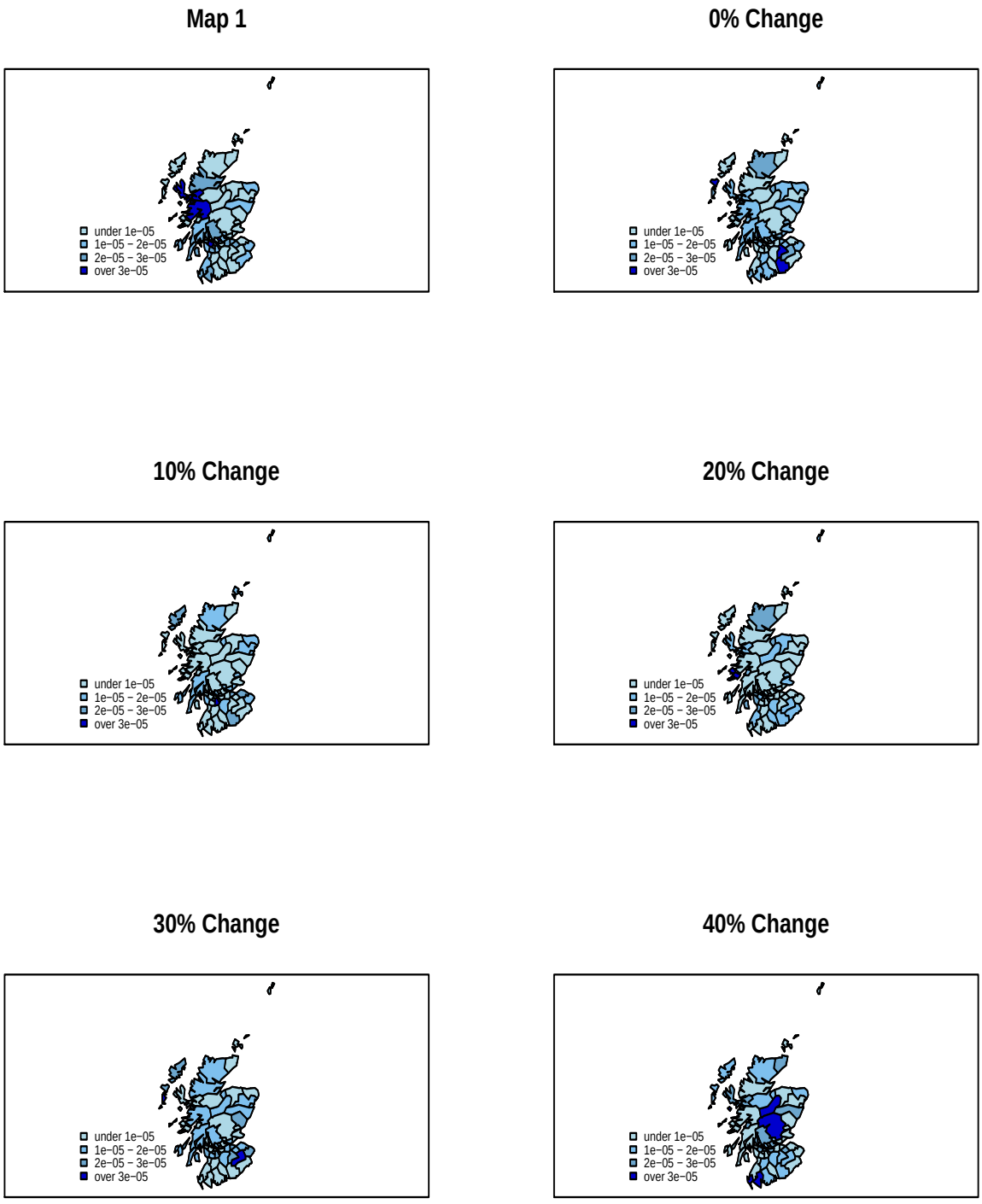

Figure 6.16: Sample map 1 and map 2 produced from the simulated data, at 0%, 10%, 20%, 30% and 40% change in structured variability for a very rare disease. Here $\mu = \sigma_u = \sigma_v = \frac{1}{100000}$.

For both unstructured and structured variation, maps of 40% change show that variation has changed to some extent.

Tables B.6-B.10 (Appendix B) give the average values, standard deviations and standard errors for the change in standard deviation due to unstructured heterogeneity for MSD, PSNRR, SSIM, IRD, MSDI, MDD, AVND, MSND and MDND respectively. Tables B.11-B.15 (Appendix B) give the corresponding results for structured heterogeneity.

Figures 6.17-6.18 give plots of average values against k for unstructured and structured heterogeneity respectively for a common disease, Figures 6.19-6.20 for a rare disease and Figures 6.21 and 6.22 for a very rare disease. Figures 6.19 and 6.20 give the power plots for unstructured and structured heterogeneity respectively. Figures 6.21 and 6.22 give the p-value plots for unstructured and structured heterogeneity respectively.

The results of structured and unstructured variation are discussed together here as they are similar. In general, for a common disease all measures perform as expected, i.e. detect change when it has taken place, except for PSNRM which increases instead of decreasing (Figures 6.17 and 6.18). For rare disease (Figures 6.19 and 6.20), most measures perform as expected, but PSNRM/PSNRR both go in the wrong direction, as do the SSIM measures initially, and these measures are also becoming unstable. These measures may be going in the wrong direction because these are ratios of two values, thus as the data are few denominators/numerators become very unstable hence unstable measures. For very rare disease (Figure 6.21 and 6.22), there is a great deal of variability and all measures are unstable. Some measures go in the wrong direction (PSNRM and SSIM measures) as the maps change and all measures have low power (Figure 6.23 and 6.24). For all measures, the ability to detect change is slower for changes in structured heterogeneity than in unstructured heterogeneity. The measures will be more useful in detecting change in variability due to unstructured variability for common diseases than other cases. No measure is powerful for very rare disease.

Mean Square Difference (MSD)

For both variabilities, MSD is sensitive for a common and a rare disease. In both cases change is detected at 10% and 30% for common and rare disease, and for a very rare disease at 80% (unstructured variation) and 110% (structured variation). The power

plots (Figures 6.23 and 6.24) indicate that MSD is powerful in detecting changes for a common disease. The p-value plots (Figures 6.25 and 6.26) show a similar picture.

Peak-to-Signal Noise Ratio (PSNRR/PSNRM)

Average value plots (Figures 6.17-6.22) indicate that for both unstructured and structured variation PSNRR is sensitive in detecting differences in the case when the disease is common. PSNRR detects change at 10% for both unstructured and structured variability. The power plots (Figures 6.23 and 6.24) indicate that this measure is fairly powerful in detecting change when disease is common, but is not at all powerful for rare and very rare disease. This is also true for the sensitivity (see the p-value plots in Figures 6.25 and 6.26). PSNRR/PSNRM are not useful for detecting changes in the case of rare and very rare disease, as they go the wrong way. PSNRM in all cases increases rather than decreases.

Structural Similarity Index Measure (SSIM/SSIMR/SSIMM)

This measure detects differences in the same way for both variabilities. The plots (Figures 6.17-6.22) of average values show that SSIM decreases slowly when the disease is common while for a rare disease it is less sensitive. SSIM detects change at 20% and 30% for unstructured and structured variability, for a common disease. For rare disease, it increases then decreases. When disease is very rare, SSIM increases rather than decreases, thus SSIM is not good in detecting differences in these cases. The power plots (Figures 6.23 and 6.24) and p-value plots (Figures 6.25 and 6.26) show that SSIM is sensitive and powerful in detecting changes only in the case of a common disease, but not for rare and very rare disease.

Inter Region Difference (IRD) and Most Dissimilar Difference (MDD)

IRD and MDD behave similarly, but IRD is slightly more sensitive than MDD. For unstructured variation, IRD detects change at 20% and 50% and MDD for 30% and 70% for common and rare disease respectively. For structured variation IRD detects changes at 30% and 80%, MDD at 40% and 100% for common and rare disease. The plots (Figures 6.17-6.22) show that IRD and MDD detect changes faster in the case of unstructured than structured variability, and are more sensitive in a case of common disease. As in the case of a change in the mean, when the disease is very rare IRD

and MDD will not work very well, as they detect differences when the variability is the same for both maps, i.e. the p-value is less than 0.05. The power plots (Figures 6.23 and 6.24) indicate that IRD and MDD are very powerful when the disease is common, and less so for a rare disease, and more powerful when changing unstructured than structured variability, but not at all powerful when the disease is very rare. The picture is the same in the p-value plots (Figures 6.25 and 6.26) for sensitivity. These measures will be more useful for detecting change in the case of common disease.

Most Similar Neighbour Difference (MSND) and Most Dissimilar Neighbour Difference (MDND)

For a common disease MSND and MDND detect differences earlier for unstructured variation and very late for structured variation (at 40% and 110% respectively). For a rare disease, MSND (at 80% and 200%) and MDND (at 90% and 230%) detect differences very late for both unstructured and structured variation. When the disease is very rare, both measures are not sensitive to change for both unstructured and structured variabilities. This can be observed from the plots (Figures 6.17-6.22) of average values against the multiples k. The power plots (Figures 6.23 and 6.24) for these measures look the same and indicate that the measures are powerful to detect changes when disease is common, and not at all powerful for a very rare disease. The p-value plots (Figures 6.25 and 6.26) show a similar picture.

Most Similar Difference (MSDI) and Average Neighbour Difference (AVND)

MSDI and AVND behave similarly. As for a change in the mean, these measures are among the least sensitive of all the measures. For common and rare disease, changes are detected very late, while for very rare disease the measures do not detect change at all. For unstructured variation, AVND detects change at 110% and 200% for a common and rare disease, and MSDI at 160% and 220% for a common and rare disease. For structured variation, except for MSDI detecting change at 230% for a common disease, these methods do not seem to detect change for any scenario. The power plots (Figures 6.23 and 6.24) indicate that for a common and a rare disease, the power rises slowly as the variability changes. The p-value plots (Figures 6.25 and 6.26) show that MSDI and AVND change very slowly for a common and a rare disease, and for a very rare disease there is no significant change in either one.

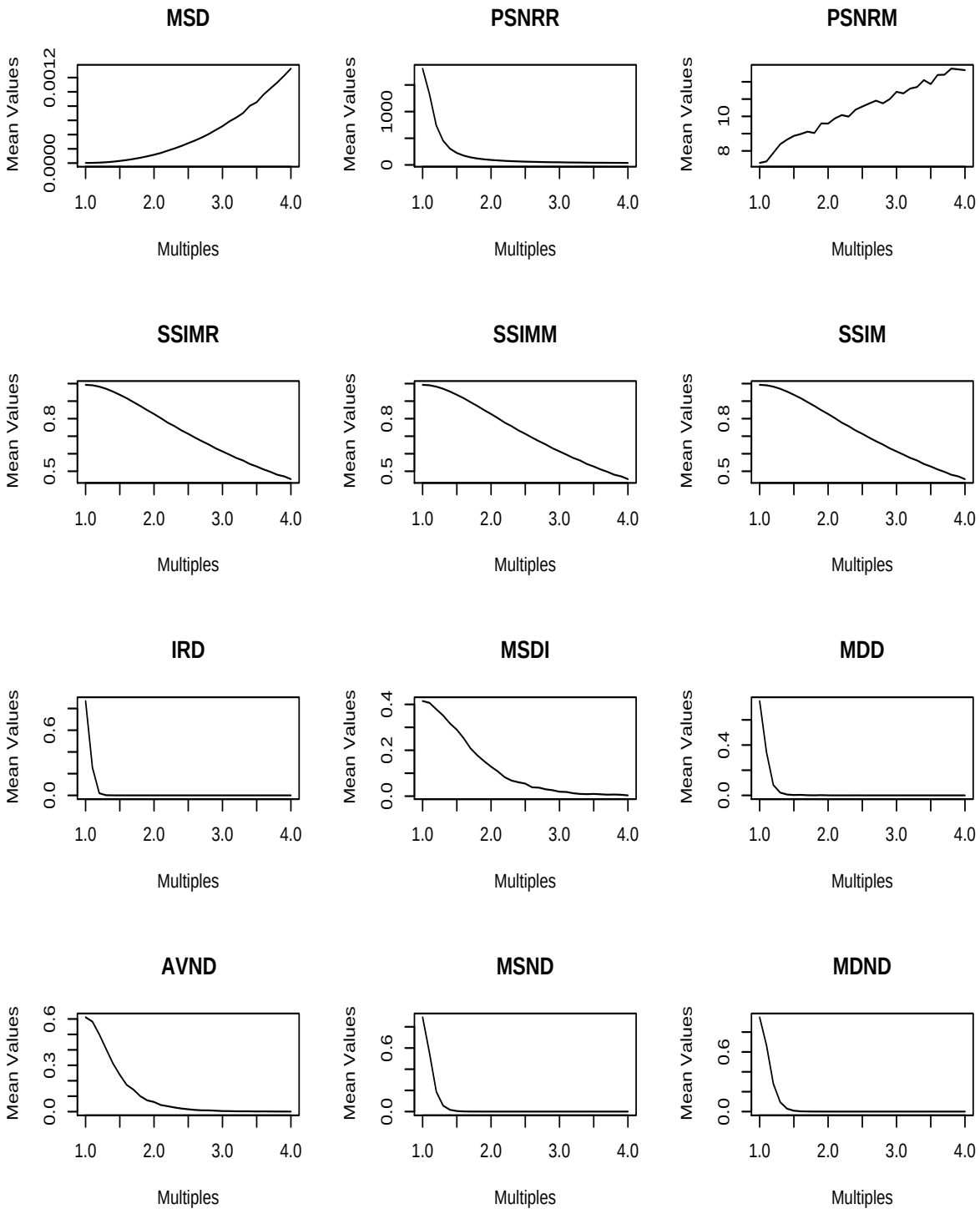

Figure 6.17: Mean of simulated values versus k for a common disease ($\mu = \sigma_u = \sigma_v = \frac{1}{10}$), when changing unstructured variation, $V_{i2} = kV_{i1}$, where $V_{i1} = V_i$.

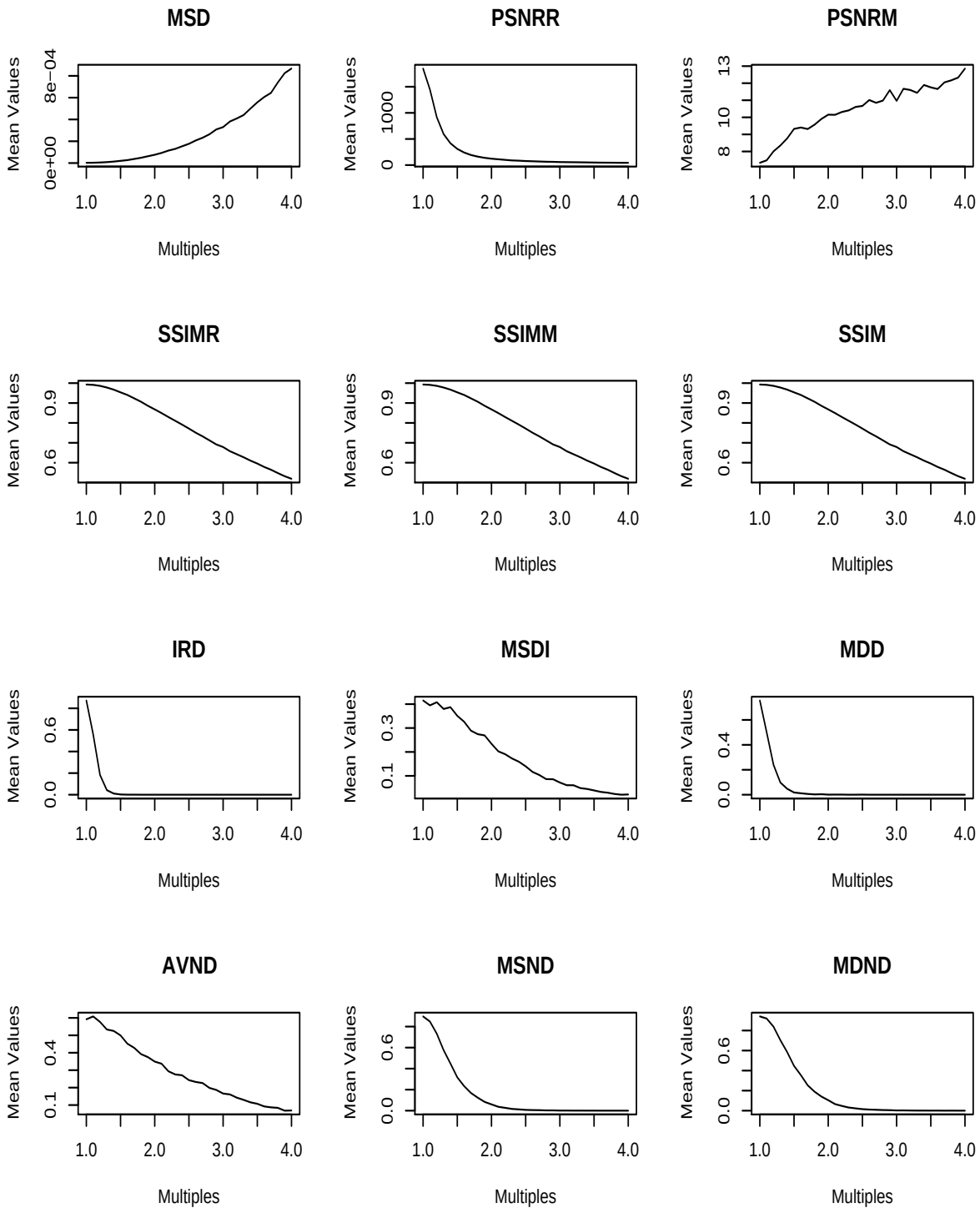

Figure 6.18: Mean of simulated values versus k for a common disease ($\mu = \sigma_u = \sigma_v = \frac{1}{10}$), when changing structured variation, $U_{i2} = kU_{i1}$, where $U_{i1} = U_i$.

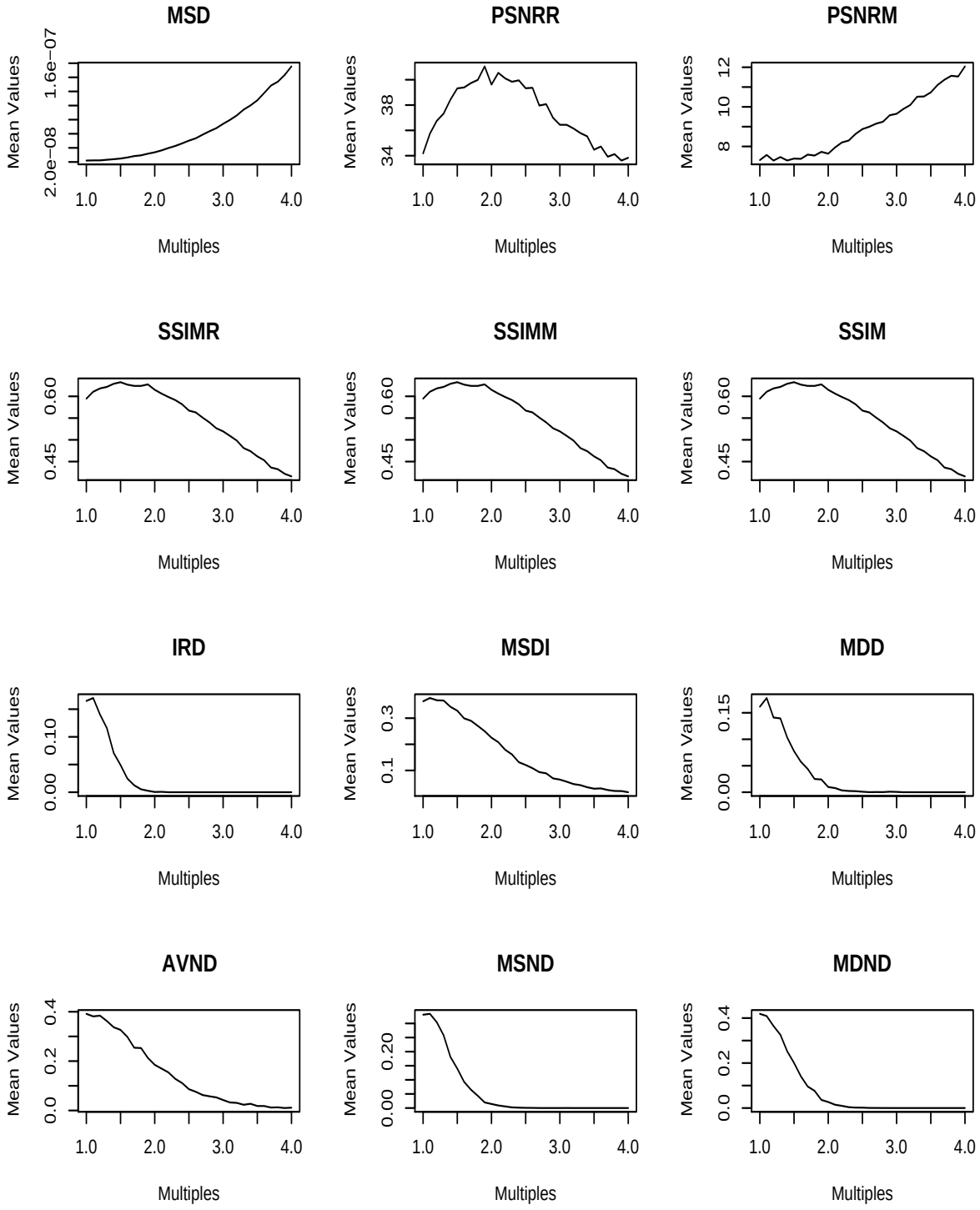

Figure 6.19: Mean of simulated values versus k for a rare disease ($\mu = \sigma_u = \sigma_v = \frac{1}{1000}$), when changing unstructured variation, $V_{i2} = kV_{i1}$, where $V_{i1} = V_i$.

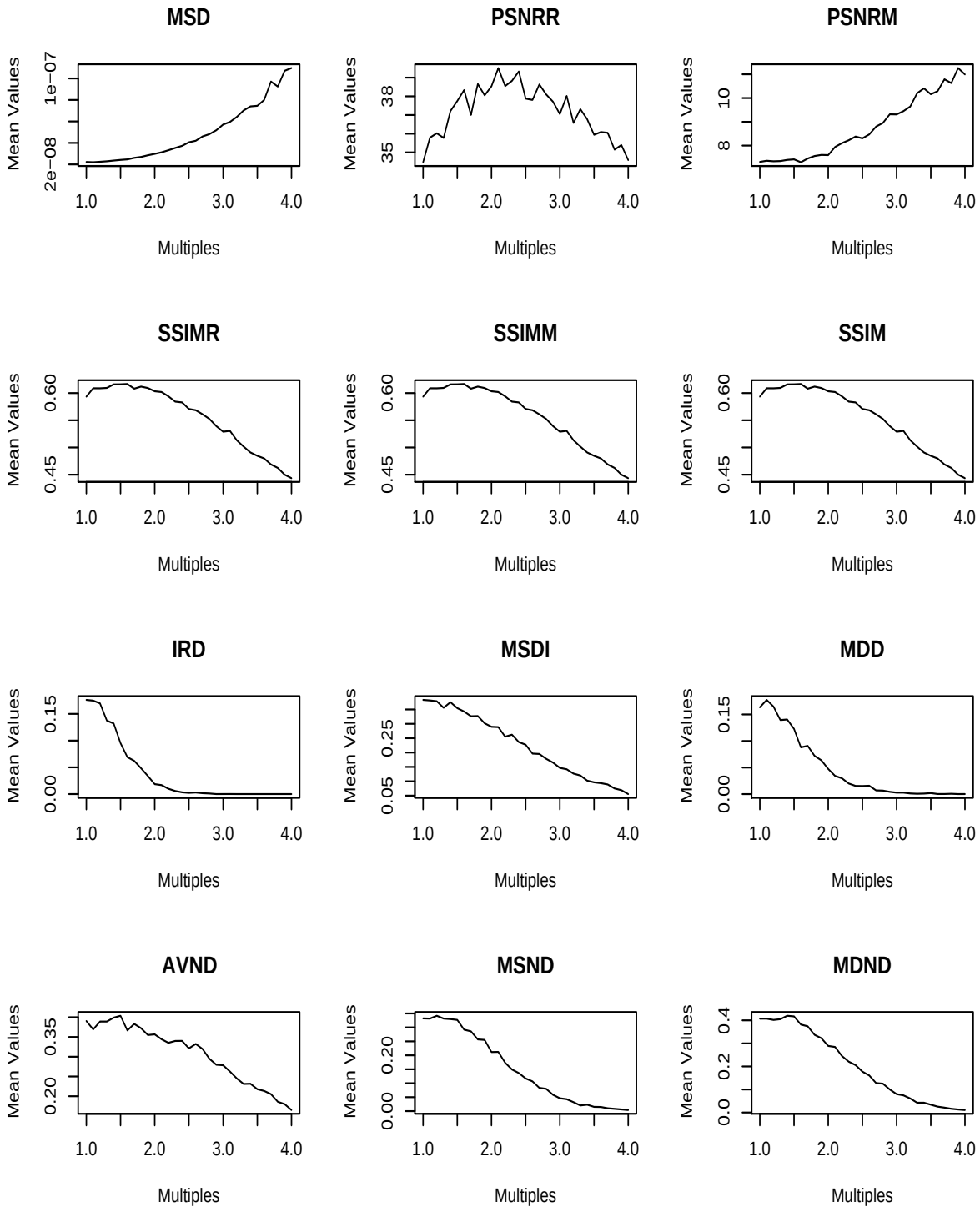

Figure 6.20: Mean of simulated values versus k for a rare disease ($\mu = \sigma_u = \sigma_v = \frac{1}{1000}$), when changing structured variation, $U_{i2} = kU_{i1}$, where $U_{i1} = U_i$.

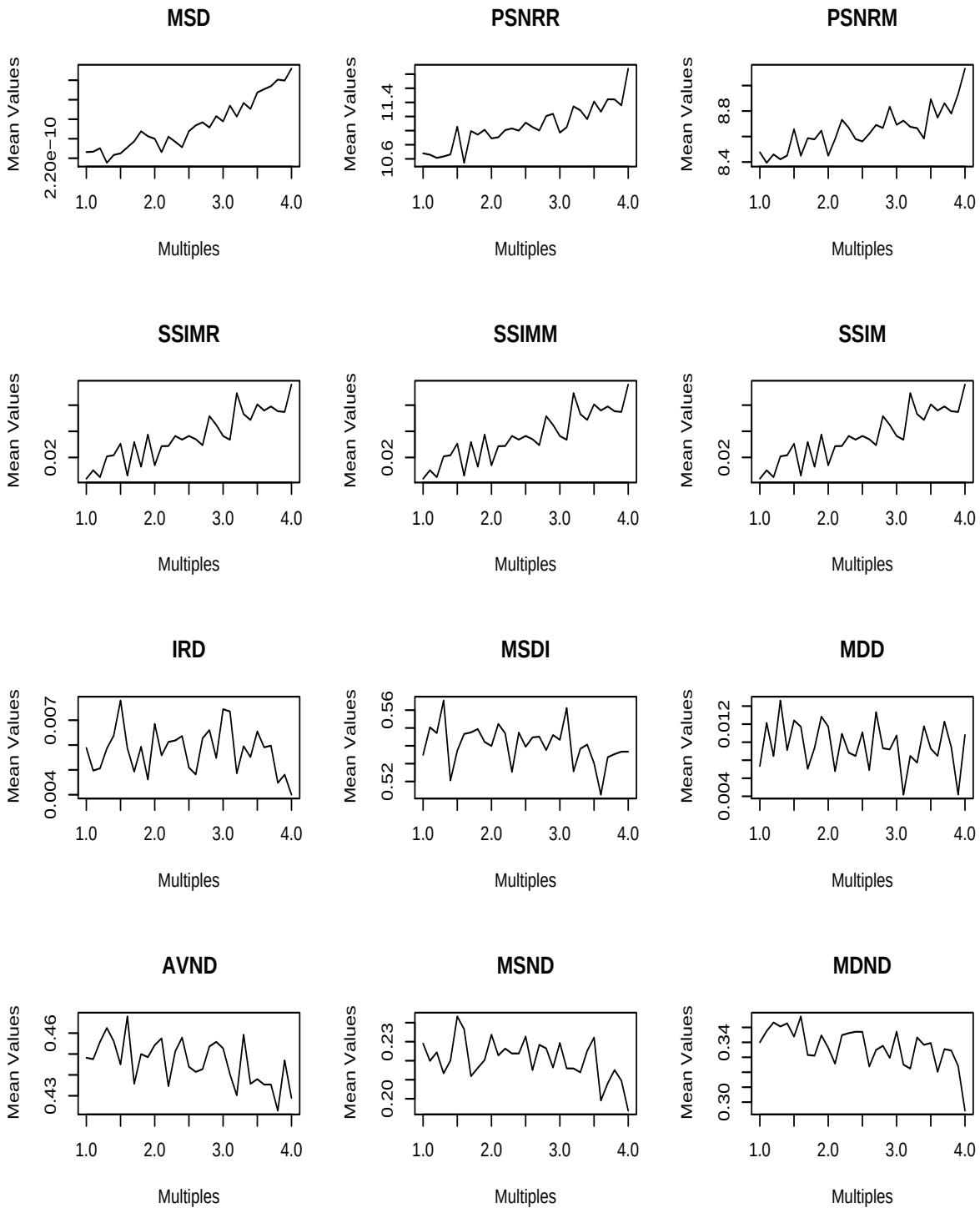

Figure 6.21: Mean of simulated values versus k for a very rare disease ($\mu = \sigma_u = \sigma_v = \frac{1}{100000}$), when changing unstructured variation, $V_{i2} = kV_{i1}$, where $V_{i1} = V_i$.

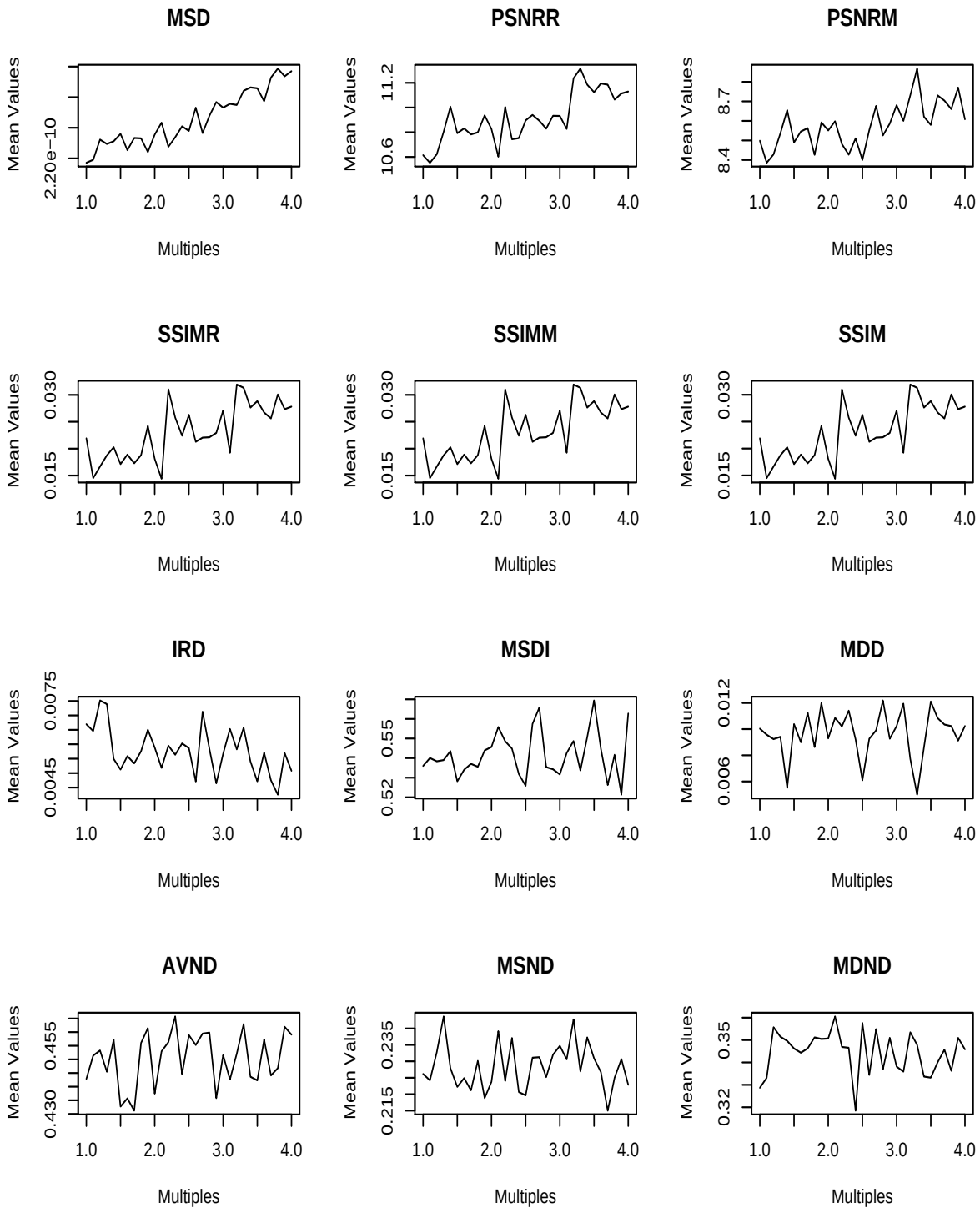

Figure 6.22: Mean of simulated values versus k for a very rare disease ($\mu = \sigma_u = \sigma_v = \frac{1}{100000}$), when changing structured variation, $U_{i2} = kU_{i1}$, where $U_{i1} = U_i$.

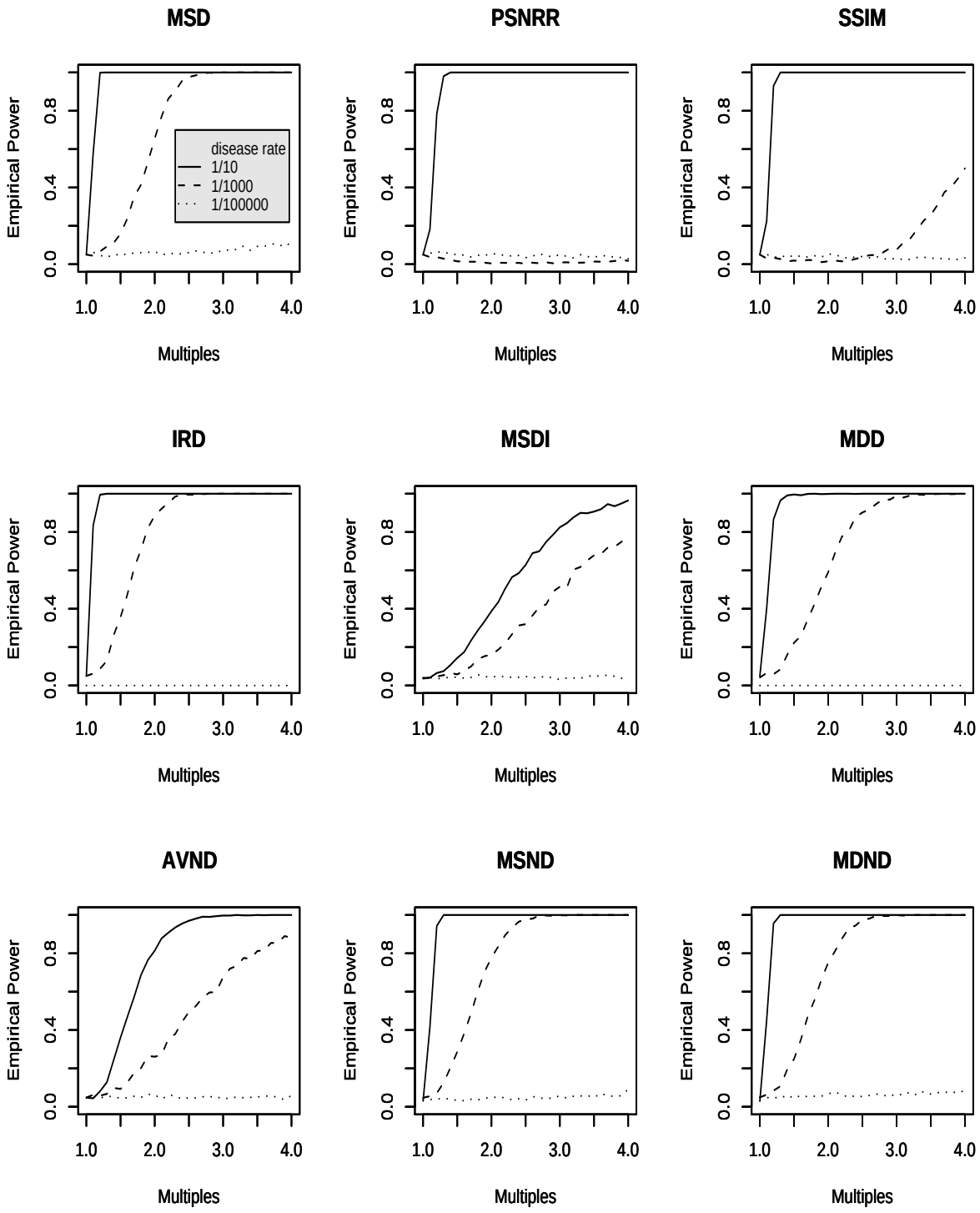

Figure 6.23: Plots of empirical power when changing unstructured variability, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

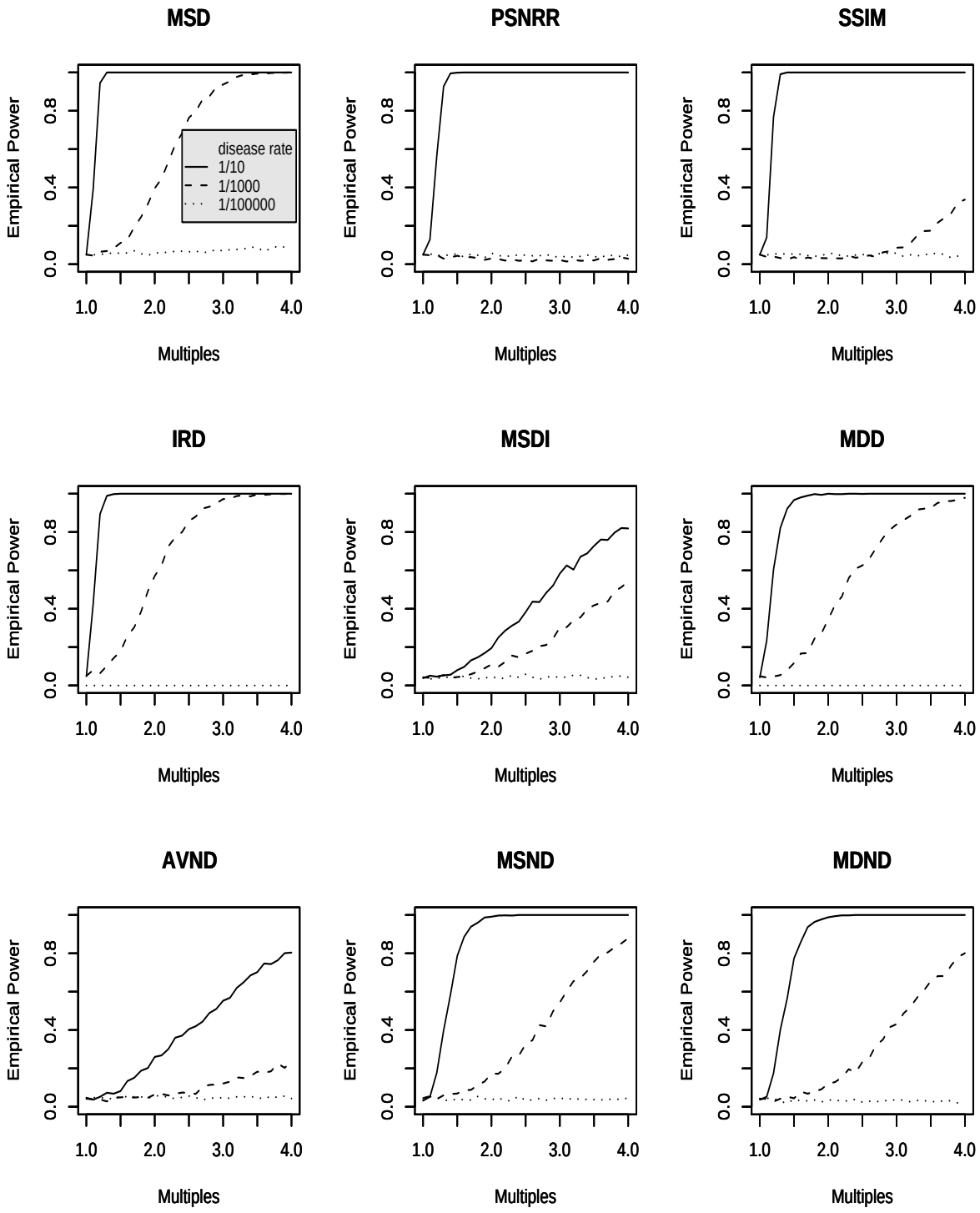

Figure 6.24: Plots of empirical power when changing structured variability, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

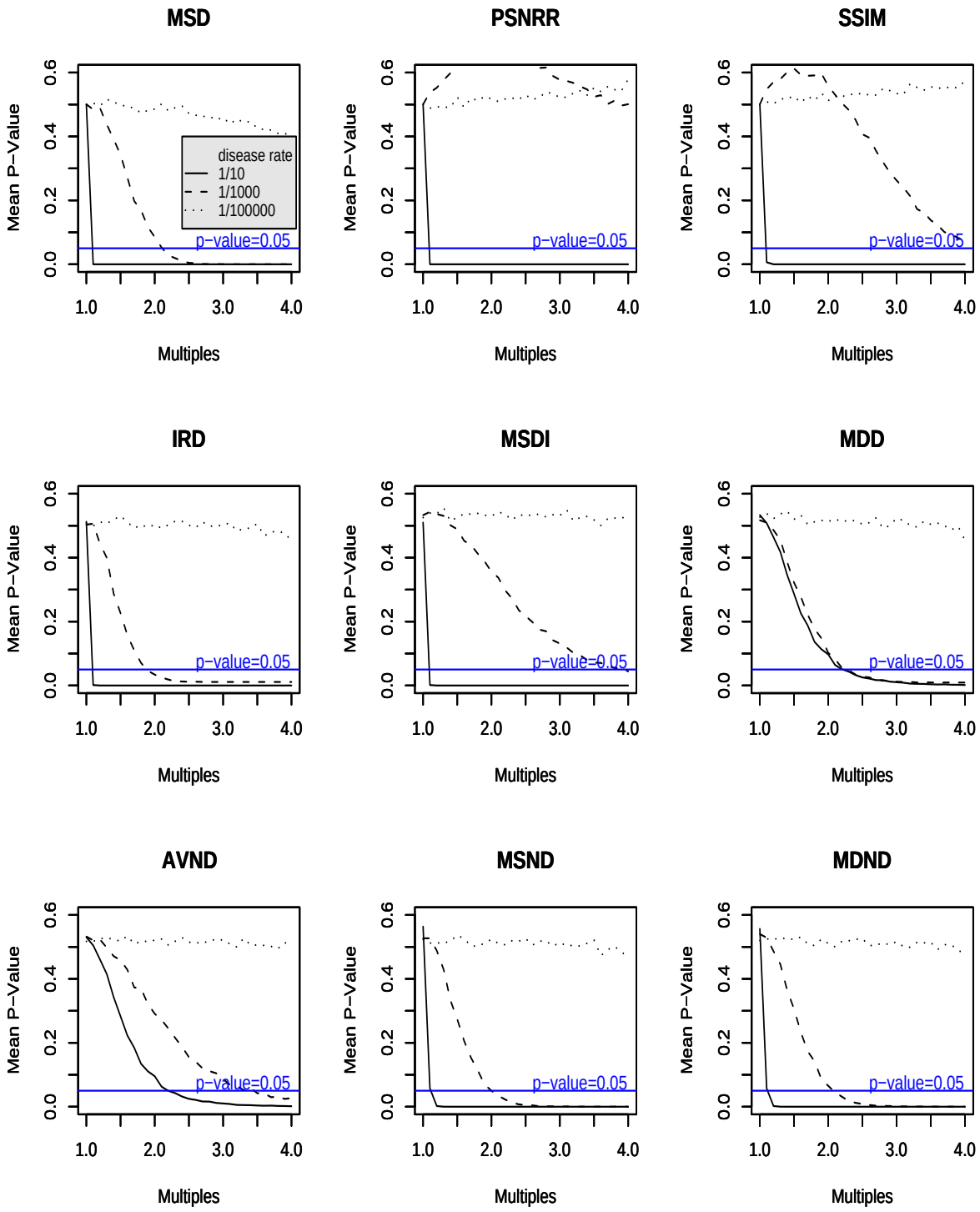

Figure 6.25: Plots of average p-values against k of the standard deviation of the unstructured heterogeneity, with horizontal line at p=0.05, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

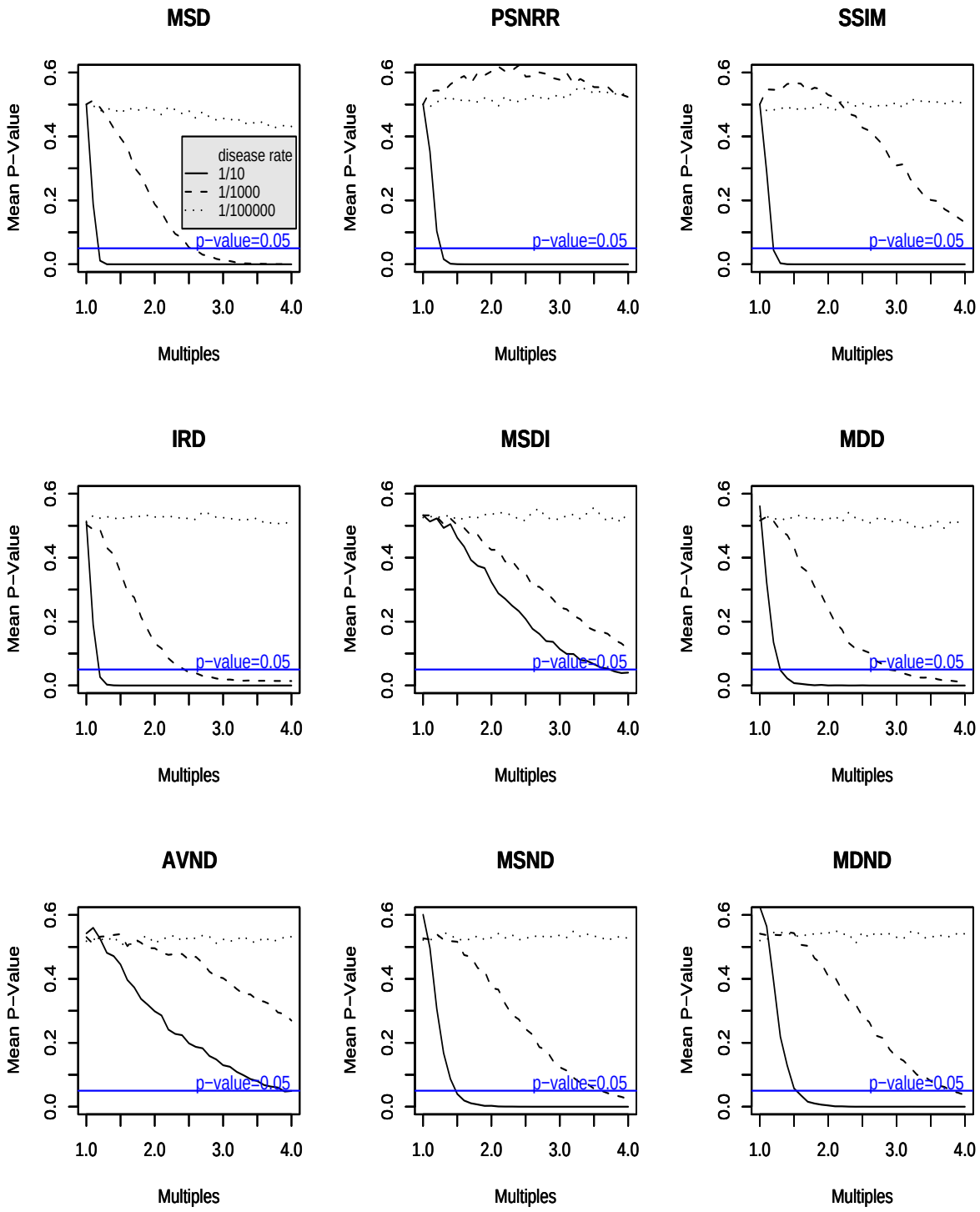

Figure 6.26: Plots of average p-values versus k of the standard deviation of the structured heterogeneity, with horizontal line at p=0.05, for common ($\mu = \frac{1}{10}$), rare ($\mu = \frac{1}{1000}$) and very rare ($\mu = \frac{1}{100000}$) diseases.

6.5 Conclusions and Recommendations

In this chapter, the developed methods have been assessed by generating data from an existing map for map comparison and, and more formally, by generating map data from a spatial model. Assessing methods by generating data from an existing map showed that when the whole map is changing most methods detect differences quickly especially the image analysis based methods MSD, PSNRM and SSIM, although PSNRR remained constant throughout. The point process based methods IRD and MDD also detected the differences. When changing the data for one region, MSD, PSNRM and SSIM do detect these changes but for the point process based methods only MDD detected the change late (80% change). This shows that some of these developed methods will be helpful when detecting change when it has occurred.

For the simulation study based on the model, it is observed that for each of the scenarios, i.e. common, rare and very rare diseases, most measures perform well for a common disease, not quite so well for rare disease, and no measure performs very well when the disease is very rare. Performance also depends on whether we are assessing change in the mean level or change in the variance due to unstructured or structured heterogeneity. The measures detect differences earlier and perform better when changing mean level than when changing either variability. Regarding variabilities, the measures detect changes earlier for changes in unstructured variability than structured variability. The measures cannot obviously distinguish between the different types of change. Tables 6.2 and 6.3 give a qualitative summary of the performance of each measure according to their sensitivity and power respectively. The measure is classified as sensitive if it detects a change at about 40% change in disease rate or variability (see Table 6.4), and classified as powerful if it reaches power of at least 70% at about $k = 1.5$. No measure works well for a very rare disease, except MSD to detect a change in mean.

6.5.1 Performance of Measures

MSD performs well for the three scenarios in the case of a change in mean level but is mainly more sensitive and more powerful for a common or a rare disease. For change in the variance of unstructured and structured heterogeneity, in terms of sensitivity and power, this measure will work well for common diseases, and in terms of sensitivity

only for rare diseases. For very rare diseases, sensitivity to changes is very small and the power to detect changes is close to zero.

Measure	Disease	Change in Mean	Change in σ_u	Change in σ_v
MSD	Common	sensitive	sensitive	sensitive
	Rare	sensitive	sensitive	sensitive
	Very rare	sensitive	not sensitive	not sensitive
PSNRR	Common	sensitive	sensitive	sensitive
	rare	sensitive	not sensitive	not sensitive
	Very rare	not sensitive	not sensitive	not sensitive
SSIM	Common	sensitive	sensitive	sensitive
	Rare	sensitive	not sensitive	not sensitive
	Very rare	not sensitive	not sensitive	not sensitive
IRD/MDD	Common	sensitive	sensitive	sensitive
	Rare	sensitive	not sensitive	not sensitive
	Very rare	not sensitive	not sensitive	not sensitive
MSDI/AVND	Common	not sensitive	not sensitive	not sensitive
	Rare	not sensitive	not sensitive	not sensitive
	Very rare	not sensitive	not sensitive	not sensitive
MSND/MDND	Common	sensitive	sensitive	sensitive
	Rare	not sensitive	not sensitive	not sensitive
	Very rare	not sensitive	not sensitive	not sensitive

Table 6.2: Performance of measures according to sensitivity. The measure is classified as sensitive if it detects a change at about 40% change in disease rate or variability.

When the mean level is changing, PSNRR is more sensitive and has greater power in detecting differences when the disease is common and rare than for very rare diseases. When the disease is very rare, this measure detects differences very slowly and is not so powerful. For a change in variability (both cases), this measure performs better for common disease in terms of sensitivity and power. In fact, for both variabilities, for both rare and very rare disease it does not seem to be sensitive to differences and power to detect change is close to zero. We note here that in relation to the Peak to Signal noise ratio, two measures were proposed, namely PSNRR and PSNRM (see Chapter 5). For a change in the mean these two measures behave similarly (Figures 6.6-6.8). Neither is useful when changing global variability or local variability (Figures

Disease	Measure	Change in Mean	Change in σ_u	Change in σ_v
MSD	Common	powerful	powerful	powerful
	Rare	powerful	not powerful	not powerful
	Very rare	powerful	not powerful	not powerful
PSNRR	Common	powerful	powerful	powerful
	Rare	powerful	not powerful	not powerful
	Very rare	not powerful	not powerful	not powerful
SSIM	Common	powerful	powerful	powerful
	Rare	not powerful	not powerful	not powerful
	Very rare	not powerful	not powerful	not powerful
IRD/MDD	Common	powerful	powerful	powerful
	Rare	powerful	not powerful	not powerful
	Very rare	not powerful	not powerful	not powerful
MSDI/AVND	Common	powerful	not powerful	not powerful
	Rare	not powerful	not powerful	not powerful
	Very rare	not powerful	not powerful	not powerful
MSND/MDND	Common	powerful	powerful	powerful
	Rare	powerful	not powerful	not powerful
	Very rare	not powerful	not powerful	not powerful

Table 6.3: Performance of measures according to power. The measure is classified as powerful if it reaches power of at least 70% at about $k = 1.5$.

6.17-6.22), except PSNRR which may be useful for common disease.

SSIM detects differences when changing the mean in the case of a common and a rare disease and when changing either of the variabilities for common disease only. It performs very poorly for a very rare disease when changing the mean level, and for a rare and a very rare disease when changing the variability (in both cases). The power to detect change is higher when the disease is common in the case of a change in the mean or a change of the variability. (We note that for SSIM two other measures were also proposed, namely SSIMM and SSIMR (see Chapter 5), and these measures perform the same as SSIM).

As for the measures involving differences, IRD and MDD perform similarly, MSDI and AVND perform similarly, and MSND and MDND perform similarly. For IRD and MDD, IRD detects differences slightly earlier than MDD. When changing the mean

level and variability due to unstructured variability, these measures are sensitive to differences when the disease is common and rare, and in the case of changing structured variability when the disease is common only. They do not work very well for a very rare disease, as they detect very small differences, i.e when there is no mean or variability change. In terms of the power, the measures are powerful when the disease is common or rare for a change in the mean and unstructured variability for IRD only. They are also powerful for a common disease only in the case of structured variability.

For MSDI and AVND, for both changes in mean and variance, these measures detect the differences very late, and are not sensitive to changes.

MSND and MDND are sensitive in detecting changes for common diseases when changing the mean and variability due to unstructured heterogeneity and not sensitive to change in structured heterogeneity. The measures are powerful in detecting change in mean for a common and a rare disease, and detecting change due to unstructured and structured variability for a common disease only.

6.5.2 Performance According to Disease Scenario

Generally the measures perform well for a common disease, with the exception of AVND and MSDI. For a rare disease, most measures perform well, except for AVND, MSDI, MSND and MDND, for change in mean. For change in unstructured and structured variabilities, only MSD performs well. Except for MSD for a change in the mean, no measure performs well for a very rare disease.

Common disease: MSD, PSNRR, SSIM, IRD, MDD, MSND and MDND can be used to help detect differences when the mean level or variability due to unstructured and structured heterogeneity has changed.

Rare disease: When the mean level has changed, MSD, PSNRR, SSIM, IRD and MDD may be useful for detecting differences. (MSND and MDND may be used when a large difference has occurred). MSD may be useful for detecting differences due to unstructured and structured variability. (IRD, MDD, MSND and MDND may be helpful only when large differences have occurred).

Very rare disease: Except for MSD, in the case of changing the mean level and also

in the case of a large change in unstructured and structured variability, no measure is good at detecting any kind of change.

6.5.3 Performance According to Type of Change

Table 6.4 gives the percentage change (roughly) in disease rate (changing mean level), unstructured variability and structured variability, at which each measure starts to detect change, for each scenario.

Change in Mean Level: MSD, PSNRR , SSIM, IRD, MDD, MSND and MDND are the most sensitive and most powerful for detecting change when the disease is common. AVND and MSDI have very poor sensitivity. For a rare disease, MSD, PSNRR and IRD are the most sensitive and most powerful, followed by MDD and SSIM. MSND and MDND detect changes late, followed by AVND and MSDI.

For a very rare disease, only MSD may be helpful in detecting changes, other measure detect change very late (MSDI , AVND, MSND and MDND). PSNRR and SSIM do not detect the changes, and IRD and MDD indicate differences when no change has occurred (which can be misleading).

Change in Unstructured Variability: For a common disease, MSD, PSNRR, IRD, SSIM and MDD are the most sensitive and powerful, followed by MSND and MDND. AVND and MSDI have very poor sensitivity.

When the disease is rare, MSD is the most sensitive and powerful, and all other measures are not sensitive as change is detected late. IRD, MDD, MSND and MDND are all sensitive and powerful only when large differences have occurred. The worst are AVND and MSDI which detect the differences far later. SSIM and PSNRR do not detect changes when disease is rare, and for a very rare disease only MSD detects changes (very late), while other measures do not detect changes (IRD and MDD give the same results as for change in mean).

Change in Structured Variability: For a common disease, MSD, PSNRR, SSIM and IRD are the most sensitive and powerful in detecting changes, followed by MDD. MSND and MDND are powerful but not sensitive, as they detect the change very late. AVND and MSDI have poor sensitivity and power.

When the disease is rare, MSD is the most sensitive and powerful, and all other measures only detect change very late or not at all. For a very rare disease, MSD detects changes very late while other measures do not detect changes. IRD and MDD give the same results as for a change in the mean and unstructured variability.

Disease	Measure	% Change in disease rate	% Change in σ_v	% Change in σ_u
MSD	Common	10	10	10
	Rare	10	30	30
	Very rare	10	80	110
PSNRR	Common	10	10	10
	Rare	10	not at all	not at all
	Very rare	not at all	not at all	not at all
SSIM	Common	10	20	30
	Rare	40	not at all	not at all
	Very rare	not at all	not at all	not at all
IRD	Common	10	20	30
	Rare	30	50	80
	Very rare	0 (misleading)	0 (misleading)	0 (misleading)
MDD	Common	20	30	40
	Rare	40	70	100
	Very rare	0 (misleading)	0 (misleading)	0 (misleading)
AVND	Common	90	110	not at all
	Rare	140	200	not at all
	Very rare	200	not at all	not at all
MSDI	Common	110	160	230
	Rare	140	220	not at all
	Very rare	120	not at all	not at all
MSND	Common	30	40	110
	Rare	60	80	200
	Very rare	100	not at all	not at all
MDND	Common	30	40	110
	Rare	60	90	230
	Very rare	110	not at all	not at all

Table 6.4: Percentage change in disease rate and standard deviations of unstructured and structured heterogeneity at which each measure starts to detect change, for each disease scenario.

In a real situation, when comparing two maps, the difference may be due to change in more than one of the three parameters (mean, unstructured and structured variability) or all three parameters together. It is not possible to consider the value of one measure and conclude whether the difference is due to a change in the mean or unstructured or

structured variability. Since the simulation study has revealed the performance of the measures under different scenarios, it may be helpful to use a combination of different measures to assess the differences. When a change is thought to have taken place, model fitting will also help to establish the nature of the change.

In the next chapter, some of the measures are used to compare susceptibility to measles maps and also to compare NHS24 call uptake maps for different types of syndromes.

Chapter 7

Comparing Maps Using Descriptive Methods

7.1 Introduction

This chapter focuses on using some of the methods developed in Chapter 5 and tested in Chapter 6, to assess the differences between real disease maps. The methods are applied to two groups of data sets. The first group consists of the susceptibility to measles data (described in Chapter 1) for Scotland for 1999-2005, both at district and postcode sector level. In the second group NHS24 call uptake data are also used, to compare call uptake for different syndromes in Scotland at postcode district level. The NHS24 data will be described later in this chapter.

The main aim is to investigate whether the methods are able to assist in the interpretation of map comparisons, to detect differences, and if so, can they indicate what kind of differences (mean, unstructured or structured variation) are there, and can the methods help quantify these differences. The susceptibility to measles results are compared to the results of the modelling of susceptibility to measles in Chapter 4.

In Chapter 6, we assessed the methods for comparing maps in the case of common disease (large mean rate), rare disease (small mean rate) and very rare disease (very small mean rate). The data sets we use here both fall under the common disease case. Therefore here we will only use the methods which were found to be most sensitive and powerful to assess the differences for the case of common disease. The methods used are MSD, SSIM, IRD and MSND. MSD and IRD just depend on the values of a

variable but do not use neighbourhood structure, whereas SSIM and MSND do take account of the neighbourhood structure, so using these methods together may give a better idea of what, if any, difference there is between maps. Pseudo-colour maps (see Chapter 5) were also used to help visualise the map regions that are different and those that are similar.

Since MSD is just a value without a reference point, we use a simulation method to find a p-value of the observed MSD value obtained to compare two maps. We use 1000 simulations, and for each simulation we obtain a MSD value. Each simulation was carried out by simulating n observed counts (n=number of districts or postcode sectors) from the Poisson distribution with means set to the observed rates of the first year, and simulated proportions are then obtained. The MSD value of this chosen first year and the new proportions is obtained, thus giving 1000 MSD values to make up a null distribution (where the null hypothesis is that the second year's map is not different to the first). The simulated proportions therefore represent the second map under the null hypothesis. A p-value of the observed MSD obtained to compare the two years is found by finding the proportion of the 1000 simulated MSD values making up the null distribution that are at least as extreme as the value of the observed MSD, i.e. as MSD is expected to increase when differences are present, the proportion of the 1000 values that are greater than or equal to MSD is found. SSIM does have a reference point, however we also used the same p-value approach for SSIM to decide at what point the value of SSIM becomes significant.

For susceptibility to measles data, the methods are applied both to the raw data (proportions) and smoothed data to see how they compare. The data were smoothed using the empirical Bayesian smoothing with the Poisson-Gamma model (see Chapter 4 for the description of the model). The function *empbaysmooth* in the package *DCluster* in R was used to obtain the estimates. The discussion of the results is based mainly on the proportion results as this will allow comparison with the modelling results.

In the next section susceptibility to measles maps are compared at district and postcode sector. This will then be followed by comparing maps based on NHS24 call uptake for health related syndromes.

7.2 Comparing Measles Maps of Scotland

In this section the measures are applied to the measles data at both district and postcode sector level for 1999-2005. This will help in comparing measles susceptibility over time at each level. The results are compared with the modelling analysis in Chapter 3. Since districts are larger and postcode sectors are smaller, this analysis may help in revealing whether the developed methods perform the same or differently for both large and small areas.

7.2.1 District Level

Table 7.1 gives values for the measures for comparing two successive time periods at a time, for pre-school children and primary 1 and 2 school children, from 1999-2005 at district level using both raw data and smoothed data.

PRE-SCHOOL CHILDREN								
	Proportions				Smoothed Rates			
Years	MSD	SSIM	IRD	MSND	MSD	SSIM	IRD	MSND
1999-2000	$9.6e-05$ (0.94)	0.88(0.91)	$8.7e-01$	0.98	0.0014	0.87	$6.8e-01$	0.979
2000-2001	$1.2e-04$ (0.80)	0.88(0.81)	$2.2e-01$	0.90	0.0012	0.90	$7.9e-01$	0.617
2001-2002	$2.0e-04$ (0.15)	0.88(0.86)	$7.4e-01$	0.46	0.0016	0.84	$3.9e-03$	0.617
2002-2003	$8.4e-04$ (0.00)	0.78(0.19)	$6.4e-08$	0.62	0.0043	0.74	$0.0e+00$	0.153
2003-2004	$2.1e-04$ (0.46)	0.90(0.58)	$5.2e-02$	0.46	0.0021	0.92	$3.9e-01$	0.098
2004-2005	$3.8e-04$ (0.04)	0.90(0.58)	$9.2e-05$	0.90	0.0030	0.85	$4.5e-10$	0.905
PRIMARY SCHOOL CHILDREN								
	Proportions				Smoothed Rates			
Years	MSD	SSIM	IRD	MSND	MSD	SSIM	IRD	MSND
1999-2000	$8.6e-05$ (0.91)	0.94(0.94)	$2.2e-07$	0.33	0.0092	0.95	$2.0e-02$	0.774
2000-2001	$7.7e-05$ (0.39)	0.92(0.30)	$5.6e-01$	0.77	0.0125	0.92	$7.1e-01$	0.905
2001-2002	$5.1e-05$ (0.86)	0.95(0.92)	$3.2e-01$	0.90	0.0095	0.94	$9.9e-02$	0.979
2002-2003	$6.4e-05$ (0.77)	0.94(0.65)	$5.6e-01$	0.77	0.0093	0.95	$4.4e-01$	0.774
2003-2004	$1.5e-04$ (0.01)	0.88(0.03)	$2.8e-02$	0.77	0.0234	0.82	$0.0e+00$	0.021
2004-2005	$3.5e-04$ (0.00)	0.84(0.03)	$4.8e-05$	0.62	0.0188	0.86	$1.0e-05$	0.774

Table 7.1: Values of measures MSD (with p-values in brackets), SSIM (with p-values in brackets), IRD (these are p-values) and MSND (these are p-values) obtained for the susceptibility to measles susceptibility raw data and smoothed rates of pre-school children(top) and primary 1 and 2 school children (bottom), at district level, comparing two successive years at a time.

The MSD values obtained after smoothing are larger than those obtained for the raw data, while the SSIM values after smoothing are closer to those obtained before smoothing. IRD results are similar for smoothed and raw data, except for pre-school 2001-2002 which gives different results. MSND results are generally similar though 2003-2004 is different. The interpretation here is based on the raw data as this will allow us to compare the use of the descriptive methods with the modelling results.

MSD

Comparing the maps using the results in Table 7.1, for the raw data, in general for pre-school children MSD is close to zero and increases over time, with a decrease in 2003-2004. This indicates that the maps are becoming less similar year on year, with 2002 and 2003 being more different. The years 2002-2003 have the highest value, followed by 2004-2005, and these are the only two cases where the p-value is significant, i.e. MSD has detected the differences in these two cases. For primary 1 and 2, MSD is very close to zero but decreases from 1999 to 2002, and increases from 2002-2005. Similarly as for pre-school, it indicates that maps are becoming less similar year on year. The highest value of MSD is in 2004-2005 followed by 2003-2004. These are the only cases where the p-value is significant.

SSIM

For pre-school, the SSIM values for 1999-2000, 2000-2001 and 2001-2002 are similar and not very far below 1, indicating that there are no major differences between each pair of years. For 2002 and 2003 the SSIM value is the smallest, showing that there were larger differences between these two years compared to other years. All the p-values indicate that if there are any differences indicated by SSIM these differences are not significant, i.e there is no difference between any consecutive pair of years.

For primary 1 and 2, SSIM is smaller than 1 but not very far from 1 and generally decreases across time. It is higher (0.95) for 2001-2002 and lowest (0.84) for 2004-2005. This indicates that 2001 and 2002 are the most similar years, and 2004 and 2005 are the most different years. The years 2003 and 2004 and the years 2004 and 2005, are the only pairs with significant p-values, thus there are differences between these pairs according to this measure. The SSIM values of 0.88 and 0.84 for primary 1 and 2 are significant (p-values of 0.03), while the SSIM values of 0.88 (p-values 0.91, 0.81, 0.86)

and 0.78 (p-values of 0.19) for pre-school are not significant. This may be because of change in disease frequency as susceptibility for pre-school is higher than at primary 1 and 2.

IRD and MSND

For pre-school, the Kolmogorov-Smirnov (KS) p-value for IRD is only significant (less than 0.05) for 2002-2003 and for 2004-2005, indicating differences between these pairs, with 2002-2003 having the smallest p-value. For other pairs of years no differences were suggested. For primary 1 and 2, the KS p-value for IRD is significant (detected differences) for 1999-2000, 2003-2004, and for 2004-2005. MSND did not indicate any differences for either pre-school or primary 1 and 2 school children. This measure is restricted to detecting changes that take place only when the difference between a rate of region and the most similar rate to it in its neighbouring regions has changed. Therefore, it may not detect changes when a different kind of change has taken place anywhere in the map.

We note that SSIM takes into account the differences in the spatial structure, and MSND takes into account the neighbourhood structure, while MSD and IRD depend on whether there has been a change anywhere in the map, so they are sensitive to any change in the disease rates, which may not be detected by SSIM and MSND. Overall, MSD and IRD indicated that there are differences in 2002-2003 and 2004-2005 for pre-school, and MSD, IRD and SSIM indicate differences in 2003-2004 and 2004-2005 for primary school. We note that for primary 1 and 2 IRD indicated a difference between 1999-2000, even though 1999-2000 has the smallest MSD and one of the highest SSIM values, which are not significant values. This was investigated further by a pseudo-colour map to allow us to visualise the differences. Pseudo-colour maps were also produced for 2002-2003 and 2004-2005 for pre-school, and 2004-2005 for primary 1 and 2. These are presented in Figure 7.1.

Pseudo-Colour Maps

The pseudo-colour map for pre-school for 2002-2003 shows that there are a few regions (grey) where the rates in these two years are similar, but most regions are green, indicating that most regions in 2003 had higher susceptibility to measles rates than in 2002, i.e. there was a global increase in susceptibility in 2003, which is greater in the

western districts. This was also observed in Chapter 4 when comparing these maps.

There are no regions where susceptibility was higher in 2002 (i.e. pink/purple regions). Therefore MSD, SSIM and IRD were able to detect these differences. This map clearly shows where the differences are. For 2004-2005, there are regions (grey) where susceptibility is the same for the two years, and there are more regions (pink/purple) where susceptibility to measles is higher for 2004 than 2005, i.e. susceptibility decreased in 2005.

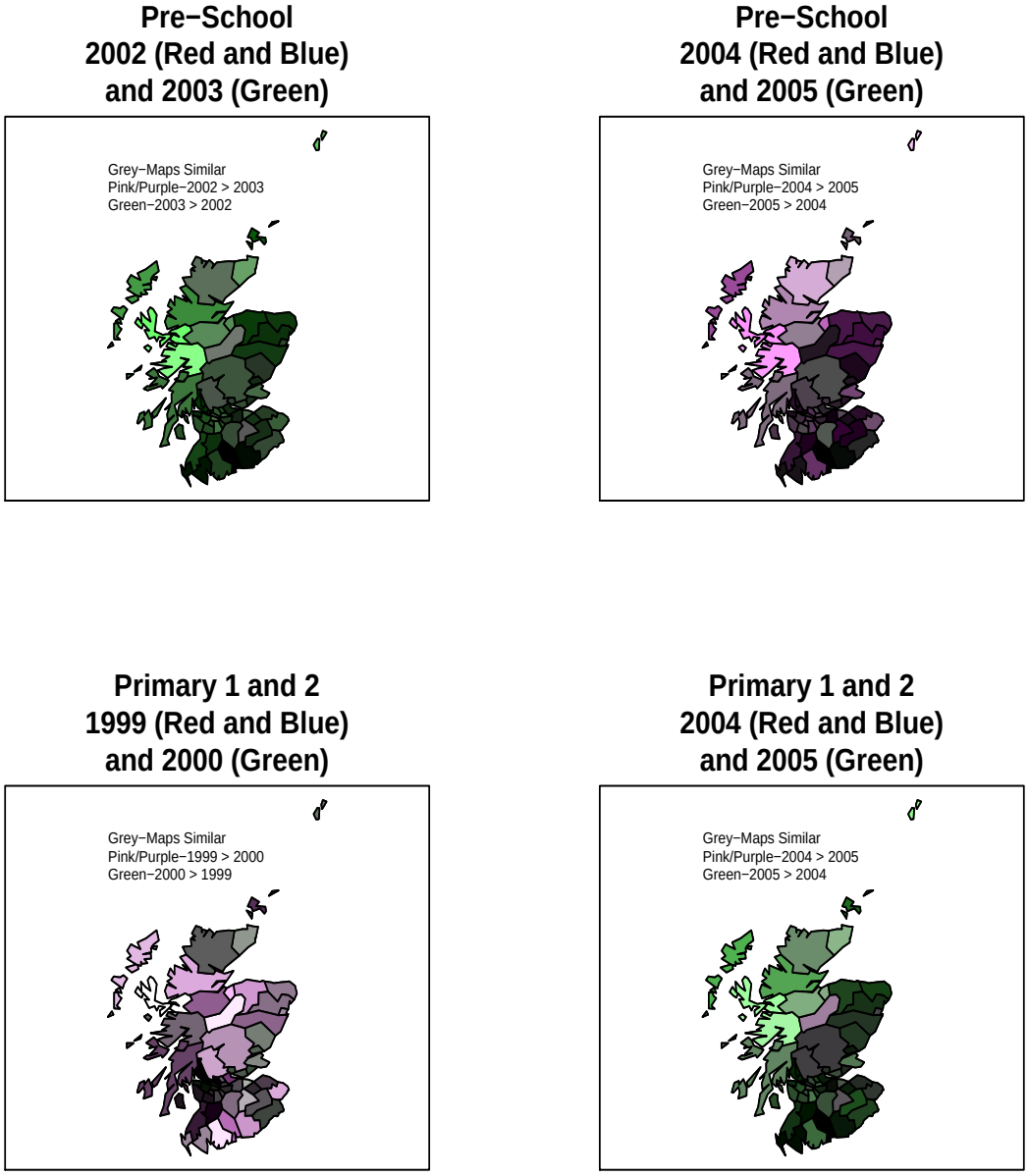

Figure 7.1: Pseudo-colour maps using raw proportions at district level for pre-school children for 2002 and 2003, 2004 and 2005, and for primary 1 and 2 for 1999 and 2000, 2004 and 2005.

There are no green regions (indicating susceptibility is higher in 2005).

For primary 1 and 2 school children, in the pseudo-colour map for 1999-2000, there are regions (grey) where both years are similar and the remaining regions (pink/purple) are where susceptibility to measles is higher in 1999. (There are no green regions where susceptibility is higher in 2000). So the differences are due mainly to the decrease of susceptibility in most regions in 2000. For 2004-2005, the pseudo-colour map shows that there are regions (grey) where these two years are similar, but most regions are green (only one region is pink/purple) indicating that susceptibility to measles increased in those regions in 2005.

So the descriptive statistics and pseudo colour-maps suggest a global increase in susceptibility in 2003 relative to 2002 and global decrease in susceptibility in 2005 relative to 2004, thus there are no pockets of local increase or decrease in measles susceptibility.

Comparison with Modelling Results

The results based on these measures compare well with the results obtained from modelling in Chapter 3 (Section 3.4), comparing 2000-2005. For pre-school, it was observed that susceptibility to measles was increasing over the years, with a decrease in susceptibility in 2005.

The year 2003 is detected by MSD, SSIM and IRD to be different from 2002. Using the logistic model results, in 2003 there was a global increase in susceptibility as shown by the maps. The overall proportion changed (from 0.17 with CI (0.16,0.18) to 0.21 with CI (0.20,0.23)). Also the clustering indicated by the structured standard deviation changed from 0.07 with CI (0.02, 0.16) to 0.15 with CI (0.07,0.24)) and the regional variation (unstructured standard deviation) changed, from 0.22 with CI (0.17,0.28) to 0.09 with CI (0.03,0.14)). This year (2003) had the highest overall proportion and unstructured standard deviation and lowest structured deviation among all the years. The credible intervals suggest an increase in overall proportion and a decrease in unstructured standard deviation in relation to 2002. The credible intervals for the structured standard deviations are wide and do not suggest any major changes. In 2005 there was a decrease in susceptibility relative to 2004. The overall proportion changed (from 0.21 with CI (0.20,0.22) to 0.19 with CI (0.18,0.20)), structured standard deviation changed (from 0.10 (0.04,0.18) to 0.09 with CI (0.03,0.18))

and unstructured standard deviation changed (from 0.17 with CI (0.12,0.22) to 0.18 with CI (0.14,0.24)). The credible intervals suggest a decrease in overall proportion in 2005 relative to 2004. The credible intervals for unstructured and structured standard deviations are wide and do not suggest any major changes. MSD, IRD and SSIM indicated differences between 2004 and 2005.

For primary 1 and 2, there was an increase from 2004 to 2005 in regions with high susceptibility. From 2004 to 2005, from Table 3.3, the logistic model shows a change in overall proportion from 0.07 with CI (0.06,0.08) to 0.08 with CI (0.08,0.09), a change in structured variation (from 0.18 with CI (0.04,0.32) to 0.21 with CI (0.10,0.35)) and a change in unstructured standard deviations (from 0.25 with CI (0.18,0.32) to 0.21 with CI (0.14,0.28)). The credible intervals suggest an increase in overall proportion in 2005 relative to 2004. The credible intervals for unstructured and structured standard deviations are wide and do not suggest any major changes. MSD, IRD and SSIM again suggested differences between these two years.

Susceptibility measured in 2003 for pre-school and 2005 for primary 1 and 2 is for the same birth cohort. The results of modelling highlighted this birth cohort as having the largest susceptibility to measles. MSD, SSIM and IRD were able to detect these differences. Pre-school 2002-2003 and primary 1 and 2 2004-2005 have the highest MSD with a significant p-value, lowest SSIM (but not significant for pre-school but significant for primary 1 and 2 school children), and a significant p-value for IRD. All of these indicate that the largest change was between 2002 and 2003 for pre-school, and between 2004 and 2005 for primary 1 and 2 school children. The pseudo-colour maps showed that most regions in this cohort had higher susceptibility in 2003 than 2002 (pre-school) and in 2005 than 2004 (primary 1 and 2).

7.2.2 Postcode Sector Level

Table 7.2 gives values for the measures for comparing two successive periods at a time at postcode sector level, for pre-school and primary 1 and 2 school children, for 1999-2005. As at district level, the MSD values for the smoothed data are larger than for the raw data. For pre-school the SSIM values are mostly smaller for the smoothed data than the raw data, while for primary school the SSIM values are larger for the smoothed data than the raw data. IRD gives a similar interpretation for both

smoothed and raw data, while MSDN gives similar interpretation in some cases and in other cases it gives different results for smoothed and raw data.

PRE-SCHOOL CHILDREN								
Proportions					Smoothed Rates			
Years	MSD	SSIM	IRD	MSND	MSD	SSIM	IRD	MSND
1999-2000	$1.7e-05$ (1.00)	0.82(1.00)	$0.0e+00$	$2.5e-07$	0.0048	0.53	$0.0e+00$	$1.5e-01$
2000-2001	$1.4e-05$ (1.00)	0.81(1.00)	$0.0e+00$	$1.2e-13$	0.0059	0.44	$0.0e+00$	$9.8e-01$
2001-2002	$1.4e-05$ (1.00)	0.79(1.00)	$0.0e+00$	$2.8e-01$	0.0054	0.59	$0.0e+00$	$2.2e-03$
2002-2003	$2.7e-04$ (1.00)	0.36(0.00)	$0.0e+00$	$0.0e+00$	0.0084	0.52	$0.0e+00$	$1.2e-06$
2003-2004	$3.9e-04$ (1.00)	0.10(0.00)	$0.0e+00$	$0.0e+00$	0.0274	0.13	$0.0e+00$	$0.0e+00$
2004-2005	$1.1e-04$ (1.00)	0.16(0.00)	$0.0e+00$	$0.0e+00$	0.0259	0.05	$0.0e+00$	$0.0e+00$
PRIMARY SCHOOL CHILDREN								
Proportions					Smoothed Rates			
Years	MSD	SSIM	IRD	MSND	MSD	SSIM	IRD	MSND
1999-2000	0.0028 (0.16)	0.64(0.01)	$0.0e+00$	$4.0e-09$	0.034	0.83	$0.0e+00$	$6.2e-06$
2000-2001	0.0024 (0.66)	0.74(0.47)	$0.0e+00$	$8.3e-01$	0.045	0.81	$0.0e+00$	$4.2e-03$
2001-2002	0.0044 (0.26)	0.59(0.06)	$0.0e+00$	$1.9e-01$	0.036	0.82	$0.0e+00$	$1.7e-02$
2002-2003	0.0057 (0.40)	0.53(0.03)	$0.0e+00$	$4.9e-01$	0.029	0.83	$0.0e+00$	$7.0e-02$
2003-2004	0.0052 (0.40)	0.61(0.17)	$0.0e+00$	$1.0e-01$	0.037	0.73	$0.0e+00$	$3.8e-02$
2004-2005	0.0080 (0.23)	0.43(0.001)	$0.0e+00$	$1.1e-04$	0.031	0.70	$0.0e+00$	$4.3e-04$

Table 7.2: Values of measures MSD (with p-values in brackets), SSIM (with p-values in brackets), IRD (these are p-values) and MSND (these are p-values) obtained for the susceptibility to measles susceptibility raw data and smoothed rates of pre-school children (top) and primary 1 and 2 school children (bottom), at postcode sector level, comparing two successive years at a time.

MSD

For pre-school, the value of MSD initially decreases but increases for 2002-2003 (the second highest value) and for 2003 and 2004 (highest value). This indicates that the maps become more dissimilar year on year, with 2003 and 2004 being the least similar pair of successive years than all other years. For all the pairs the obtained p-value is 1, so that this measure indicates that there are no differences. For primary 1 and 2, generally the MSD values increase over the years, indicating differences are increasing over the years, with 2004 and 2005 having the highest MSD. Again, the p-values are not significant.

SSIM

For both pre-school and primary 1 and 2 school children, SSIM decreases over the years. For pre-school, 2003-2004 have the smallest SSIM, and for primary 1 and 2, 2004-2005 have the smallest SSIM, showing these pairs are more different. For pre-school, the p-value is significant for 2002-2003, 2003-2004, and 2004-2005, and for primary 1 and 2, for 1999-2000, 2002-2003, and 2004-2005, indicating differences between these pairs.

IRD and MSND

IRD is significant for all pairs, for both pre-school and primary 1 and 2 children, indicating differences between the years. For pre-school, MSND is significant for all pairs except 2001-2002. For primary school, MSND is only significant for the pairs 1999-2000 and 2004-2005.

Overall, even though MSD p-values are not significant, the other measures still highlight that there are differences year on year. Similarly as at district level, the birth cohort for which susceptibility was measured at 2003 for pre-school and at 2005 for primary 1 and 2 school children shows greater differences when compared to other years.

The pseudo-colour maps for pre-school, 2002-2003 and 2003-2004, and primary school, 1999-2000, 2004-2005 are shown in Figure 7.2. These are the pairs of years which the measures detected to be very different compared to others. For pre-school, even though 2004-2005 has the smallest SSIM, 2002-2003 was chosen because of its larger MSD. For primary school, 1999-2000 has the second highest SSIM and second smallest MSD, but it was chosen because MSND was significant.

Pseudo-Colour Map

For pre-school, the map for 2002-2003 shows that there are postcode sectors (grey) where these two years are similar, but there are more green postcode sectors where in 2003 susceptibility to measles was higher than in 2002, and very few pink/purple (mainly on the eastern side), where susceptibility to measles was higher in 2002 than 2003. This indicates that susceptibility to measles increased in 2003, but there may also be an increased variance as the trends are different in the east and west.

Figure 7.2: Pseudo-colour maps using raw proportions at postcode sector level for pre-school children for 2002 and 2003, and 2003 and 2004, and for primary 1 and 2 children for 1999 and 2000, and 2004 and 2005.

For 2003-2004, there are postcode sectors (grey) where susceptibility to measles is similar for the two years, but more pink/purple postcode sectors indicate that in 2004 susceptibility generally increased.

For primary school children, the pseudo-colour map for 1999-2000 indicates that there are grey regions where these years are similar, but there are also more pink/purple regions where susceptibility is higher in 1999. For 2004 and 2005, there are also grey regions where susceptibility is similar for both years, but there are more green regions where susceptibility was higher in 2005 than 2004.

Comparing with Modelling Results

These results compare fairly well with the modelling results. For pre-school, the modelling results for 2002 to 2004 for susceptibility to measles were as follows. From 2002 to 2003 there was a change in overall proportion, from 0.18 (0.176, 0.184) to 0.21 with CI (0.20,0.213). Structured standard deviation changed from 0.16 with CI (0.12,0.19) to 0.17 with CI (0.12,0.20) and unstructured standard deviation changed from 0.03 with CI (0.01,0.05) to 0.04 with CI (0.03,0.06). The credible intervals suggest an increase in overall proportion, while for unstructured and structured standard deviations the credible intervals do not suggest any real differences. From 2003 to 2004 there was a slight change in overall proportion from 0.21 with CI (0.20,0.213) to 0.20 with CI (0.19,0.21), a change in structured standard deviation from 0.17 with CI (0.12,0.20) to 0.15 with CI (0.11,0.19) and change in unstructured standard deviation from 0.04 with CI (0.03,0.06) to 0.06 with CI (0.04,0.08). The credible intervals for overall proportion, unstructured and structured deviations do not suggest any real change from 2003 to 2004. However, SSIM, IRD and MSND detected differences between 2002 and 2003 and also between 2003 and 2004, confirming visual impression. This suggests that the modelling results may not always detect changes as the map comparison measures.

For primary school, from 2004 to 2005, there was a change in overall proportion from 0.071 with CI (0.068,0.073) to 0.084 with CI (0.081,0.086), a similar structured standard deviation 2004 (0.40 with CI (0.32,0.48)) and 2005 (0.40 with CI (0.34,0.36)), and unstructured standard deviation changed from 0.11 with CI (0.02,0.17) to 0.05 with CI (0.01,0.12). The credible interval suggests an increase in overall proportion. SSIM, IRD and MSND detected differences between 2004 and 2005.

7.2.3 Comparing District and Postcode Sector Level

Generally for districts the results given by all measures more or less agree, while for postcode sectors there are different results for MSD, SSIM, IRD and MSND. We note that for postcode sectors regions are smaller than the districts, therefore the sample size is smaller. So postcode sectors have sampling variation which is greater than for districts. Thus some of the measures may not be able to detect differences for postcode sectors but detect differences at district level. MSD p-values at postcode sector level are not significant for any pair of years, while for some years at district level MSD has detected some differences. SSIM detected differences in some of the pairs at postcode level for pre-school, but did not detect difference for the same pairs at district level, and the measure detected differences for the pair 2004-2005 at district level but not at postcode sector level. MSND detected differences in most pairs at postcode sector level, while at district level the p-values were not significant for any pair of years. This may be because MSND depends on the neigbourhood structure and a postcode sector has more neighbouring regions than a district.

7.3 Comparing Maps of Proportions of NHS24 Call Uptake

NHS24 is a Scotland wide 24 hour emergency telephone and health information service. In this section we use NHS24 data and compare maps of different health syndromes. These data represents NHS24 call rates for different health problems for each postcode district in Scotland. We assess the spatial differences between maps of proportions of NHS24 calls attributed to four syndromes, namely colds/flu and fever (CFF), rash (RASH), difficulty in breathing and cough (DBC), and for diarrhoea and vomiting (DV). Kavanagh and Robertson (2008) investigated the spatial clustering of the call rates between regions and the effect of deprivation on call rates.

7.3.1 Data

The raw data used here are a 6 months (September 2007-March 2008) data at postcode district level. The raw data contains counts for the number of individuals calling NHS24 with different syndromes, namely colds and flu, coughs, diarrhoea, difficulty

in breathing, double vision, fever, eye problem, lumps, rash, vomiting and other.

Here we chose to put together some of these syndromes to form four new groups of syndromes, to compare their call uptake. The colds/flu and fever (CFF) syndrome combines colds/flu syndrome and the fever syndrome; difficulty in breathing and cough (DBC) combines difficulty in breathing and cough syndromes; diarrhoea and vomiting (DV) syndrome combines diarrhoea and vomiting syndromes; and RASH is the rash syndrome alone. For each of the new four syndromes, the total counts were converted into call proportions by dividing by the total number of calls for each individual postcode district.

Descriptive Statistics

Table 7.3 shows descriptive statistics for CFF, RASH, DBC and DV. It can be observed that DBC and DV have the same mean, with DBC having a higher median, followed by CFF then RASH. Therefore overall DBC has highest proportion of call uptake while RASH has the least proportion of call uptake. The maximum observed proportion of calls for DV is 1, as some postcode districts have call uptake for DV only.

Syndrome	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
CFF	0.000	0.072	0.088	0.082	0.100	0.235
RASH	0.000	0.029	0.041	0.039	0.048	0.200
DBC	0.000	0.100	0.118	0.109	0.131	0.500
DV	0.000	0.086	0.104	0.109	0.117	1.000

Table 7.3: Summary statistics for NHS24 syndromes.

Figure 7.3 shows maps of proportions of calls due to the four syndromes, at postcode district level. For colds/flu and fever (CFF) syndrome, there is a higher number of calls made in the central and north west regions, but most postcode districts have lower call uptake. Call uptake for difficulty in breathing and cough (DBC) syndrome is higher in the north west, eastern, central and south regions, with pockets of lower call uptake in the north. For RASH syndrome, most postcode districts have a lower call uptake with pockets of higher call uptake in the western regions.

There is higher call uptake for diarrhoea and vomiting (DV) syndrome mainly in the north west, south, central and eastern regions and lower call uptake in the north.

Figure 7.3: Maps of proportions of calls to NHS24 for Colds/Flu and Fever (CFF), Difficulty in Breathing and Cough (DBC), RASH, and Diarrhoea and Vomiting (DV), at postcode district level.

This seems similar to the spatial pattern of calls due to difficulty in breathing and cough syndrome (DBC). DBC has more regions with higher call uptake than other syndromes, while RASH has lower call uptake in most regions than other syndromes.

Looking at the correlations between the proportions of calls due to the different syndromes, it can be observed that the correlations are small: CFF and DBC (0.47), CFF and RASH (0.26), CFF and DV (-0.17), DBC and RASH (0.19), DBC and DV (-0.21), RASH and DV (-0.13). Thus it is expected that the descriptive methods may indicate big differences between these syndromes in terms of proportions of calls made.

7.3.2 Analysis

The descriptive methods proposed can certainly be used if comparing, for example, call proportions for DV in one time period with call proportions for DV in another time period, but not necessarily when comparing for example call proportions for DV and call proportions for RASH as mean levels are different. It is expected that for the latter all the methods would indicate a difference in the maps owing to a different mean level rather than from a different spatial distribution. Standardisation of data in some way or use of spatial correlation measures may be helpful in this case.

Here the means are standardised before applying the methods. To compare two syndromes, we find the mean of the proportions for each syndrome and multiply one group of proportions by the ratio of the two means so that the two groups end up with the same mean.

Results

Table 7.4 gives the results of the measures applied to the standardised proportions. MSD p-values indicate that all the pairs of syndromes are different except for CFF and RASH. The SSIM and IRD p-values are all significant for all the pairs of syndromes, thus according to all these measures there are significant differences between each of the syndrome pairs. MSND indicates that there are differences between all pairs of syndromes except CFF and DBC, and DBC and DV.

The pseudo colour maps are shown in Figures 7.4-7.5. For CFF and DBC, there are pockets of regions in the north east, south and west where call uptake is higher for

DBC than CFF (green), and similarly, there are pockets of regions in the west and north east where call uptake is higher for CFF than DBC (pink/purple). For CFF and RASH there are more regions where call uptakes for these syndromes are similar, but call uptake attributed to CFF is higher in most regions than for RASH. CFF and DV show a few more pink/purple regions, indicating that CCF has a higher number of regions with higher call uptake than DV, mostly in the north eastern and central regions. The spatial distributions of the CFF and DV map, and the CFF and DBC map look similar, as in both maps DV has few regions with higher call uptake (these are in the eastern part of Scotland).

For DV and RASH, and DBC and RASH, there are a few regions where RASH call uptake is higher than for other syndromes, mainly in the eastern part of Scotland, but these maps mostly indicate a higher DV or DBC call uptake.

Generally, RASH call uptake is lower in most regions compared to the other three syndromes, CFF call uptake is higher compared to RASH and DV but it is similar to the call uptake of DBC.

Standardised Proportions				
Syndromes	MSD	SSIM	IRD	MSND
CFF and DBC	0.0013 (0.063)	0.46(0.00)	$0.0e+00$	$6.3e-02$
CFF and RASH	0.0029 (0.002)	0.24(0.00)	$0.0e+00$	$4.0e-05$
CFF and DV	0.0107 (0.000)	-0.12(0.00)	$0.0e+00$	$2.6e-03$
DBC and RASH	0.0054 (0.002)	0.17(0.00)	$0.0e+00$	$5.8e-09$
DBC and DV	0.0193 (0.000)	-0.15(0.00)	$0.0e+00$	$2.6e-01$
RASH and DV	0.0027 (0.000)	-0.12(0.00)	$0.0e+00$	$4.6e-07$

Table 7.4: Values of measures MSD (with p-values in brackets), SSIM (with p-values), IRD and MSND for comparison of maps based on call uptake for NHS24 for Colds/Flu and Fever (CFF), RASH, Difficulty in Breathing and Cough (DBC) and Diarrhoea and Vomiting (DV), for standardised proportions at postcode district level.

Figure 7.4: Pseudo-colour maps at postcode district level for the standardised proportions of calls to NHS24 for Colds/Flu and Fever (CFF) versus each of Difficulty in Breathing and Cough (DBC), RASH, Diarrhoea and Vomiting (DV), and DBC versus DV.

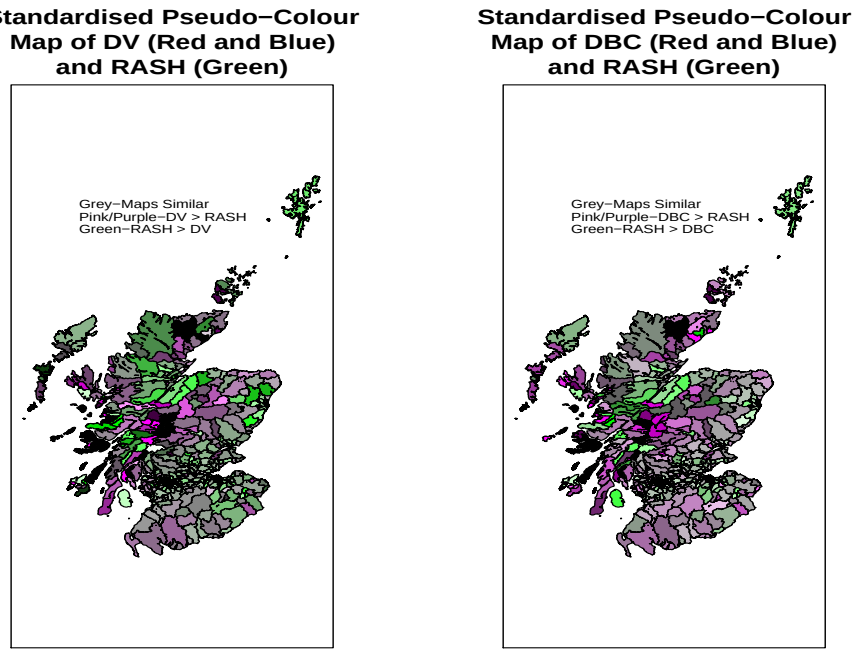

Figure 7.5: Pseudo-colour maps at postcode district level for standardised proportions of calls to NHS24 for Diarrhoea and Vomiting (DV) versus RASH, and Difficulty in Breathing and cough (DBC) versus RASH.

7.4 Conclusions

In this chapter we have applied some of the proposed methods, i.e. the pseudo-colour map, MSD, IRD, SSIM and MSND, from Chapter 5, to susceptibility to measles of pre-school and primary 1 and 2 school children from 1999 to 2005, at both district and postcode sector level. The methods were also applied to proportions of NHS24 calls in Scotland attributed to four syndromes. For all data sets, the aim was to assess if the methods could detect the differences, and if differences were detected, to assess what change has taken place and if possible quantify the differences.

These methods may be helpful in establishing whether differences exist or not between maps, but will not be able to establish what kind of change has occurred. The methods can also detect changes which are visually seen on maps but not indicated by the modelling results. Thus these methods may be used together with model fitting in order to try to establish the kind of change that has taken place. Thus these methods may be most suitable for use in exploratory analysis.

The results for susceptibility to measles were compared with the modelling results in

Chapter 3, for both districts and postcode sectors and led to similar conclusions. The pseudo-colour maps were found to be a useful way of highlighting the differences in the data between years. These maps were able to show the regions where susceptibility decreased, increased or did not change for the years that were compared. For example in Chapter 4, for pre-school at district level, susceptibility to measles increased in most districts and the pseudo-colour map has highlighted this very well. The MSD, SSIM and IRD measures also detected these differences. The measures were able to highlight that the birth cohort for 2003 pre-school and 2005 primary 1 and 2 differ from other years.

MSD p-values were not significant at postcode sector level for any pair of years, suggesting this measure may not work well for comparing maps based on very small regions. SSIM detected changes mostly at postcode sector level. We note that there is a limitation to the way the p-values were obtained for MSD and SSIM. The method of obtaining monte carlo p-values assumes that map 1 is fixed and that all the sampling variation is contained in map 2. An alternative strategy would be to treat both maps as subject to sampling variation, i.e. get rates from map 1 or average of map 1 and 2, simulate map 1 and map 2 using same rates, calculated MSD/SSIM, and repeat the process to get p-values.

MSND was able to detect differences only at postcode sector level, which suggests that these measures may not work well when comparing maps based on large regions but may do well for smaller regions. Since postcode sectors are smaller than districts, each postcode sector has a larger number of neighbouring regions than each district. Furthermore there is more variation in susceptibility rates at postcode sector level than districts as postcode sectors are smaller than districts, and have smaller sample size. Thus there is more potential for MSND to detect differences at postcode sector level as this measure depends on this neighbourhood structure.

When smoothing, the rates are pulled towards their means, thus they become more similar and hence there is less variation between them. For susceptibility to measles data, for most methods, similar interpretations were obtained from the results based on smoothed data and those from the raw proportions, except that the MSD values become larger when the data are smoothed.

Comparing measles susceptibility over time involves a relatively small change between maps as the mean levels are similar. These descriptive measures have shown to be useful when comparing maps where small changes have occurred, as in the case of measles susceptibility. When comparing maps with different mean level, as in the case of the NHS24 data, mean standardisation is needed. Here standardisation was carried out by multiplying one map by the ratio of the means of the two maps which are being compared. Other possibilities are: multiplying by the difference between the means of the two maps, or finding $Z_1 = \frac{R_1 - M_1}{Sd_1}$ and $Z_2 = \frac{R_2 - M_2}{Sd_2}$ where R_1, R_2 are the rates of map 1 and 2, M_1, M_2 are the means of map 1 and 2, Sd_1 and Sd_2 are the standard deviations of the rates of map 1 and 2, and Z_1 and Z_2 are the new standardised rates which can be mapped and compared. But for these two cases there are issues of negative values and a way of dealing with these needs to be addressed.

Thus the descriptive methods may be useful when comparing susceptibility to measles map, but probably are not so useful in the case of the NHS24 data. In this case spatial correlation methods may be more appropriate, and these were used for these data by Kavanagh and Robertson (2008).

Chapter 8

Summary, Conclusions and Further Work

The work carried out in this thesis is divided into two parts. For the first part of this thesis, we investigated and compared different approaches to the analysis of disease maps through modelling. These methods are based on the modelling of relative risk, which allows the production of choropleth maps (Chapter 3), and smoothing of relative risk using non-parametric methods, which allows production of isopleth maps (continuous surface maps) (Chapter 4). As we make use of susceptibility to measles data of Scotland, for pre-school and primary 1 and 2 school children, for the period 1999-2005, some of these methods are used to analyse the susceptibility to measles. Our aim also was to analyse spatial trends in susceptibility to measles over time, and with the use of spatial ecological models, to investigate the relationship between susceptibility in areas and markers of inequality between areas.

The second half of the thesis focuses on the comparison of disease maps and on the investigation and development of methods to compare disease maps. The methods were investigated through simulation and were again illustrated on the measles data as well as on NHS24 call uptake data.

8.1 Modelling

8.1.1 Comparing Maps Over Time

A spatial logistic model based on the model proposed by Besag *et al.* (1991) and a space-time model proposed by Waller *et al.* (1997) were used to analyse the susceptibility to measles data, and to compare maps over time at district and postcode sector levels. The results were satisfying in that we were able to visually compare the maps and identify regions where susceptibility to measles had changed or had not changed over the years. For both district and postcode sector levels, the maps were seen to become less similar over the years, as in general, susceptibility was increasing over time. Maps from both models indicated that for all the years, measles susceptibility tended to be higher in the rural regions of the north west part of Scotland (the Highlands). For pre-school children, measles susceptibility was high in 2003 and 2004, while for primary 1 and 2 school children, measles susceptibility was high in 2005. The pre-school children group whose susceptibility was assessed in 2003, and the primary 1 and 2 school children group whose susceptibility was assessed in 2005 is the same group, i.e. children born between 1/3/1999 and 28/2/2001. Therefore, for this birth cohort many regions may have children who did not receive first and second uptake of measles, mumps and rubella (MMR) vaccine. This means that if there was to be an outbreak of measles in the future, these birth cohorts are likely to be affected the most.

The proportion, and the structured (local) and unstructured (global) standard deviation values obtained from the models also helped to a certain extent to understand whether the change in susceptibility was a result of change in these parameters (see Chapter 3 for plots of these parameters). In general, both models indicate that, for each year, at district level the difference in the rates is more due to unstructured variation than structured variation, and at postcode sector level the difference is due more to structured variation than unstructured. Therefore there is a difference in analysing the data at district and postcode sector level. This is because districts are a combination of postcode sectors, so by adding them up we accumulate extra binomial dispersion giving unstructured variability. Postcode sectors are smaller than districts and we expect more clustering at postcode sector level than at district level, thus obtain higher structured variation.

Only the logistic model gives the values of the overall proportion for each year separately, while the space-time model gives one single overall mean value, but could be extended by including a trend. As these models are different, they tend not to give similar results in some cases. For example, for pre-school children, the logistic model indicates that the change in unstructured standard deviation between 2002 and 2003 is about 0.13, while the space-time model indicates that the change is about 0.02. From the logistic model, for pre-school children, for districts, the overall proportion is highest in 2003, and overall proportion is the highest for 2003 and 2004 for postcode sectors. For primary 1 and 2, the overall proportion is highest in 2005 for both district and postcode sectors.

In Bayesian modelling, choosing a hyperprior for the variance parameters is not easy, therefore a sensitivity test is normally carried out prior to the analysis of the data, to make sure that a reasonable hyperprior is chosen. In our analysis an exhaustive sensitivity analysis was not carried out but a noninformative prior, a gamma distribution, was assigned to the inverse variance parameters, i.e. Gamma(0.1, 0.001). This was assigned with an understanding that the hyperprior will give a large variance, therefore it will be relatively flat over a large range, thus will have little influence on the likelihood of the data. However, carrying out a detailed sensitivity test may have helped us assess the suitability of this hyperprior and its influence on the values of parameters obtained.

8.1.2 Ecological Analysis

Models were also fitted to assess the effect of covariates on measles susceptibility. The census variables considered were percentage of people in households with no car, percentage of people in overcrowded households, percentage of unemployed males, percentage of people in households with low social class, percentage of children aged 0-4 years, percentage of lone parent households, percentage of educational qualifications (there were five levels but here we considered percentage of people with no qualifications and high qualifications only, as other levels are correlated with these), percentage of people born in a country outside the EU (born elsewhere), percentage of people born in a country within the EU but outside the UK (born other EU), and percentage of people working in agricultural (a rurality marker). The first four markers are the

components of the Carstairs scores for deprivation, but they were put in the model separately to check if they all have similar effects. In fact the results show that they do not have similar effects. Percentage of people in overcrowded households and percentage of people in low social class households have a negative effect and percentage of people in households with no car and percentage of unemployed males showed a positive effect on measles susceptibility.

An ecological analysis was performed at district and postcode sector level. At district level, the overall results indicated that the geographical differences in susceptibility to measles cannot be explained by any of the census variables, as the parameters are very small and the credible intervals span zero. At postcode sector level, some of the geographical differences in susceptibility to measles can be explained by some of the explanatory variables. Susceptibility increases with high percentage of people born in other EU countries, working in agriculture, no car, and unemployed, and susceptibility decreases where there is a high percentage of people in overcrowded households and low social class. Among these variables, the most important variables for prediction of measles susceptibility are percentage of people in households with no car and percentage of people in households with low social class. Thus more rural areas are associated with high susceptibility (this may be due to difficulty in accessing or getting to a doctor).

8.2 Smoothing

As our interest was to compare maps over time and develop methods that can be used to compare maps over time, we investigated smoothing methods and produced isopleth maps, which can be used to compare maps over time. The smoothing or interpolation methods considered were nonparametric kernel regression methods and kriging methods (Chapter 4). The analysis was based on smoothing the ratio of the observed counts/expected counts, using the Nadaraya-Watson kernel estimator and ordinary kriging, and smoothing of empirical Bayes estimates by ordinary kriging. Overall, the isopleth maps produced from these three methods did not differ that much. The isopleth maps strengthen the ability to visualise data over time, and as they are not restricted by region boundaries of study, sudden jumps are avoided between two neighbouring regions. However, for the data and map area used here, these methods did not

work very well. For kriging, models used did not fit the data well, and for the produced maps it was not easy to see change over time in regions of high/low susceptibility.

When smoothing data, the interpolation methods "borrow" information from the neighbouring regions and estimates of regions close to the edge are more likely to be biased. Use of methods that allow for inclusion of edge effects may help this situation. The local linear estimator is a nonparametric regression method with edge effect correction. This method was discussed in Chapter 2, but it was not used to analyse the susceptibility to measles data. The use of this method may have allowed us to compare the results to those produced by the Nadaraya-Watson kernel estimator, which is a simpler method to use but does not allow for edge effect correction.

8.3 Methods for Comparing Disease Maps

Even though the use of the models provides estimates that can be mapped to visualise the spatial distribution of a disease, and provides values of parameters that can be compared over time, we sought to develop measures to compare two maps on a number of facets. These methods were explained in Chapter 5. The aim was to develop measures that can detect differences between maps when mean or variability due to unstructured or structured heterogeneity has changed.

The first methods considered are visual ones to help in identifying a region whose rate has changed or whose rate remained similar, namely the ratio maps, difference maps and pseudo-colour maps. The second group of methods were methods based on differences between rates, which are analogues of distance methods used in point processes to test for complete spatial randomness or for differences in spatial distribution of two point processes. These methods were thought to be appropriate here as they would allow us to compare the difference between maps based on the rates, and some of these methods will be able to assess the change in the neighbourhood structure. These methods included Inter Region Differences (IRD), Most Similar Differences (MSDI) and Most Dissimilar differences (MDD), which may help in detecting change when differences between the rate of a region and rates of other regions have changed. The Average Neighbour Differences (AVND), Most Similar Neighbour Differences (MSND) and Most Dissimilar Neighbour Differences (MDND) methods, consider the change in

the rates within the neighbourhood structure of the map. All of these methods give a p-value of a test of difference in two cdfs as an assessment of change. As maps are similar to images, the third group of methods were adaptations of methods used in image analysis to compare a distorted image to a reference image. These methods were mean square error, referred to here as mean square difference (MSD), two variations of peak to signal noise ratio, namely PSNRR and PSNRM, and three versions of the structural similarity index measure, namely SSIM, SSIMM and SSIMR. MSD and PSNRR may help in quantifying the difference between the maps, and SSIM/SSIMM/SSIMR compare the structure of disease maps. The last group of methods considered were methods that measure spatial autocorrelation (Moran's I and Geary's c), but these methods were not pursued as they can only indicate if spatial autocorrelation exists or not, but cannot detect the differences and similarities between maps.

The results of this work are disappointing in the sense that no measure is able to detect exactly what kind of change has taken place (change in mean, structured or unstructured variation). However, useful results have been obtained, as most measures are able to indicate that there are differences when differences exist, and measures like SSIM are able to indicate by how much the map structures differ. The pseudo-colour map is a very useful tool, as it highlights clearly where the differences are on the map.

8.3.1 Simulation Study

The simulation study in Chapter 6 helped to determine the ability of the methods to detect change. Firstly some informal investigations were carried out. Data were generated from an existing map by changing all the values in the map by a certain multiple, and also by just changing one region in a map. Different multiples were used. This analysis showed that the image based analysis methods are better in detecting change in both cases, while the point process based methods are not so good at detecting change in a single region. When all regions are changing, IRD and MDD seem to detect the difference fastest.

A more formal simulation was conducted by simulating data from a spatial model. This helped in determining the sensitivity and power of the methods (point process and image analysis methods), to detect change (in the mean, unstructured or structured variability). Some of the methods are very sensitive and powerful in detecting

changes, some are not. This also depends on what kind of change has taken place and on whether the disease is common, rare or very rare. Overall, some methods will be helpful, depending on the situation addressed, and some methods will not be helpful at all.

For the difference based methods, IRD and MDD behave similarly, MSND behaves like MDND and AVND behaves like MSDI. All the measures detect change earlier in the case of change in mean than for a change in unstructured and structured variability. For variability, the results are similar but slightly better for change in unstructured rather than structured variability. In the case of a common disease, for a change in mean, unstructured and structured variability, all the methods show ability to detect changes. The most sensitive and powerful are MSD, PSNRR, SSIM, IRD, MDD, MSND and MDND, with the exception of MSND and MDND for structured variability as these detect differences very late. For rare disease, for change in mean, the results were as for common disease with the exception of MSND and MDND, which detect differences late. For unstructured and structured variability, MSD is the most sensitive and powerful. For very rare disease, no method really works well in this case except for MSD which may be helpful in the case of change in mean.

For structured and unstructured variation, PSNRM changes in the wrong direction for all disease scenarios, and also does this for very rare disease in the case of change in mean. PSNRR changes in the wrong direction for rare and very rare disease for change in unstructured variation. SSIM, SSIMM ansd SSIMR give similar results, as the only difference between them is that SSIMM and SSIMR include constants to stabilise the denominator when the mean or variance is zero.

The weakness of the simulation study is that the performance of the methods is assessed based only on changing mean or unstructured or structured variability one at a time, and therefore the results are not known when more than one of these three has changed. This simulation could be extended to assess this, however if this was done we will still not know what has changed. The performance of methods like SSIM/SSIMM/SSIMR cannot be truly assessed, as this depends on the mean, variability and correlation of the rates. These measures should give values between -1 and 1, and about 1 if there are no differences in the maps. In generating data when the value of the variability is very low, these measures give very low values even when

the compared maps arise from the same distribution. This was observed when generating data from the Poisson distribution without the inclusion of the unstructured and structured heterogeneity term. In this situation the correlation is about zero so the SSIM becomes very unstable. However, in disease mapping this is not an issue of concern as structured spatial variability is always present leading to spatial correlation.

All of the point process based type methods based the test of the difference between cdfs on the Kolmogorov-Smirnoff test statistic. Other such test statistics could have been used as well for comparison of results, for example, Cramer-von-Mises (en.wikipedi.org/wiki/cramer-van-mises_criterion).

8.3.2 Comparing Maps

The methods which worked well in the simulation study were used to compare maps of susceptibility to measles over time, at district and postcode sector level, and on the NHS24 data to compare proportions of NHS24 calls based on different syndromes. Both data sets fall under the scenario of common disease, therefore we chose methods that were shown by the simulation study to perform well in this case. These were MSD, IRD, MDD, SSIM and MSND and MDND, but since IRD and MDD behave similarly and MSND and MDND behave similarly, only IRD and MSND were used. The pseudo-colour maps were also plotted to help in visualising the differences between maps.

For measles susceptibility data, the methods were applied to both raw and smoothed data (empirical Bayes using the Poisson-Gamma model), to see how the two compare, but interpretation of the results was based on the results of the raw data. (The results obtained from raw and smoothed data were similar in interpretation). For the susceptibility to measles data, MSND could not detect any differences at district level, but could detect change at postcode sector level, and MSD was not significant at postcode sector level. Therefore, some of the methods may be affected by whether the analysis is based on large or small areas. Further investigation is needed to assess this issue. These methods detect differences in the maps, but the problem is that it is difficult to know whether it is the mean, or the variabilities that have changed. The simulation study was based at district level. On the real data it can be observed that

MSD and IRD (Table 7.1) detect changes together at district level. MSD and IRD were found to detect changes earlier than other methods in the simulation study. At postcode sector level the methods that seem to detect differences together are SSIM, IRD and MSND (Table 7.2).

The methods can also be used on two different spatial variables, for example, in the case of comparing NHS24 syndromes in Chapter 7. In such a case when the two spatial variables have different means, these means will have to be standardised so that they are comparable before the methods can be used. If this is done, the methods proposed can be applied to pairs of maps, differences visualised by plotting the pseudo-colour map, and differences quantified by model fitting using the two different data sets to compare and interpret model parameters. Standardisation of the mean of the proportions of NHS24 syndromes helped in comparing the maps.

8.4 Suggestion for Analysing Distribution of Disease

8.4.1 Selection of Developed Methods for Comparing Disease Maps

In this thesis descriptive methods have been developed/suggested to compare two or more disease maps. These methods can only indicate that a difference does or does not exist between maps, but cannot indicate what kind of change has taken place, therefore these can be used for exploratory analysis. A suitable model from Table 8.2 can then be used to identify whether the mean level, local or global variation has changed.

The first place to start is to produce maps of the disease in question based on, for example, proportions, the ratio, difference, and pseudo-colour maps, as these will indicate where change has taken place in the map. Then Table 8.1 can be used to select suitable methods to quantify the amount of change. Table 8.1 show descriptive methods according to disease type. Of MSD and PSNRR one is advised to use MSD as it performs better than PSNRR. Of IRD and MDD, one should use IRD, and of MSND and MDND either of the two methods can be used.

Type of Disease	Suitable Methods
Common	Mean Square Difference (MSD)
	Peak-to-Signal-noise-ratio (PSNRR)
	Structural Similarity Index (SSIM)
	Inter Region Differences (IRD)
	Most Dissimilar Differences (MDD)
	Most Similar Neighbour Difference (MSND)
	Most Disimilar Neighbour Difference (MDND)
Rare	Mean Square Difference (MSD)
	Peak-to-Signal-Noise-Ratio (PSNRR)
	Structural Similarity Index (SSIM)
	Inter Region Differences (IRD)
	Most Dissimilar Differences (MDD)
Very rare	Mean Square Difference (MSD)

Table 8.1: Table showing descriptive methods that can be used to compare disease maps, according to disease type.

8.4.2 Selection of Models

In disease mapping often the ratios of the observed to the expected counts (the standardised mortality/morbidity ratios (SMRs) are obtained and mapped, but there are problems associated with mapping SMRs (see Section 1.3.2). In Chapter 2, we reviewed models that can be used to address these problems by smoothing the SMRs, to allow better interpretation of the map.

The choice of the model will depend on the aim of the analysis. For example, one may have data available at one time point and be interested to see how the disease is spatially distributed, and to identify areas with high/low disease rates. Also one may have data for a specific disease available at different time points, and the aim might be to observe and compare the spatial distribution of the disease over time, and see if the disease rates have decreased or increased over time. As disease counts of neighbouring regions tend to be similar, due to clustering or spatial autocorrelation, it will be better to consider models that take this into account. The other aim may be to see if any ecological variable can predict areas of high disease rates and explain some of the spatial differences that may exist. Including covariates in the models will help account for unobserved effects, if these exist, thus this kind of models may be

preferred.

The other issue to consider is that some models smooth the data without considering jumps in relative risk surface, and these jumps are of great importance to help in allocation of resources. Some models consider jumps in the relative risk but ignore smoothing, and some models take both of these into account. The models that allow for smoothness and discontinuities on the map may be a better choice. Table 8.2 gives a summary of some disease mapping models that one can consider to analyse the distribution of disease.

Model	Allow Spatial Autocorrelation	Allow Covariates	Allow Smoothness	Allow Discontinuities
Poisson-Gamma	No	Yes	Yes	No
Lognormal/Logistic	Yes	Yes	Yes	No
Mixture Models				
NPML	No	Yes	No	Yes
TNPMPL	Yes	Yes	Yes	No
Lawson and Clark (2002)	Yes	Yes	Yes	Yes
Linear Bayes Method	Yes	No	Yes	No
Space-time Models				
Waller *et al.* (1997)	Yes	Yes	Yes	No
Bernardinelli *et al.* (1997)	Yes	Yes	Yes	No

Table 8.2: Table showing whether a disease mapping model takes into account spatial autocorrelation, covariates, smoothness or discontinuities. The models are Poisson-Gamma, lognormal/logistic model, mixture models based on nonparametric maximum likelihood estimation (NPML), transitional nonparametric maximum pseudo-likelihood estimator (TNPMPL), and the Lawson and Clark (2002) model, linear Bayes methods and the space-time models of Waller *et al.* (1997) and Bernardinelli *et al.* (1997).

8.5 Further Work

More work still needs to be done in developing methods for comparing disease maps. In the work done here, further simulation study may be done to further assess the performance of the developed methods. This study may be based on changing mean,

unstructured and structured variability at the same time and also changing two at the same time while the other one is not changed. It would be useful to try more systematically to relate the value of the measure to the change in the value of the model parameters or type of change that has taken place, as in practice we would like to be able to compute one of the proposed measures and know from that what is different between the two maps.

Models used in disease mapping give values for the mean, and degree of unstructured and structured variation. These values can vary depending on what model has been used. It may be helpful to develop a model that can include a parameter which measures the difference between change in mean, and degree of unstructured or structured variability between two maps. In disease mapping there are models that have been developed for mapping two or more diseases, and these models help in identifying similarities and dissimilarities in the spatial distribution of the disease risk (see Chapter 3). This sort of idea may be adapted to develop a model that can be used to compare disease maps. For example, the proportional mortality model proposed by Dabney and Wakefield (2005) to model two diseases may be possibly modified to the situation of comparing two maps. This will allow us to identify regions where either the mean, or the unstructured or structured variation has changed, by mapping the values obtained for each region. These values are the differences between the mean, or unstructured or unstructured variability of map 1 and map 2, and by looking at these values we may be able to see how much change has taken place and where, and know what type of change has occurred.

For comparing maps with different mean levels as in the NHS24 data, investigating different strategies for standardising the means of rates will be generally useful.

Appendix A

Chapter 3 Tables

A.1 Relative Risks (Section 3.2.1)

District	Observed	SMR	EB-Poisson-Gamma	FB-Poisson-Gamma	EB-log-normal	FB-log-normal
1	58	2.34	1.77	1.74	1.87	1.93
2	313	1.26	1.24	1.24	1.24	1.25
3	127	1.24	1.20	1.20	1.19	1.23
4	67	1.62	1.45	1.44	1.46	1.44
5	200	1.63	1.55	1.54	1.56	1.58
6	65	0.97	1.00	1.00	0.99	1.39
7	361	1.57	1.52	1.52	1.53	1.54
8	86	1.05	1.06	1.06	1.05	1.41
9	70	1.10	1.09	1.09	1.08	1.12
10	278	1.21	1.19	1.19	1.19	1.20
11	108	1.67	1.50	1.50	1.52	1.86
12	37	0.99	1.03	1.03	1.01	1.09
13	52	1.79	1.48	1.47	1.50	1.57
14	104	1.30	1.25	1.24	1.24	1.21
15	269	1.19	1.18	1.18	1.17	1.17
16	239	1.23	1.21	1.21	1.20	1.21
17	51	1.79	1.48	1.47	1.50	1.54
18	120	0.98	1.00	1.00	0.99	0.98
19	206	1.25	1.22	1.22	1.22	1.24
20	121	0.84	0.89	0.89	0.89	0.88
21	347	1.11	1.11	1.11	1.11	1.11
22	725	1.33	1.31	1.31	1.31	1.32
23	170	1.04	1.04	1.04	1.04	1.03
24	206	1.17	1.16	1.16	1.15	1.13
25	499	0.83	0.84	0.85	0.84	0.84
26	467	0.91	0.93	0.93	0.92	0.92
27	193	1.44	1.39	1.39	1.39	1.38
28	354	0.91	0.93	0.93	0.93	0.92
29	503	1.44	1.41	1.41	1.41	1.41
30	611	0.94	0.94	0.94	0.94	0.94
31	119	0.64	0.73	0.73	0.74	0.72
32	80	1.91	1.63	1.61	1.66	1.63
33	290	0.77	0.80	0.80	0.80	0.80
34	348	1.55	1.50	1.50	1.50	1.48
35	274	0.80	0.83	0.83	0.83	0.81
36	229	0.61	0.66	0.66	0.67	0.63
37	416	0.69	0.71	0.71	0.72	0.70
38	391	0.88	0.90	0.90	0.90	0.88
39	191	0.56	0.62	0.63	0.64	0.62
40	120	0.51	0.62	0.63	0.65	0.61
41	539	0.55	0.58	0.58	0.59	0.57
42	450	1.11	1.11	1.11	1.11	1.10
43	175	0.86	0.90	0.90	0.89	0.88
44	502	0.75	0.77	0.78	0.78	0.77
45	1486	1.39	1.39	1.38	1.38	1.38
46	259	0.67	0.71	0.72	0.72	0.70
47	320	0.88	0.90	0.90	0.90	0.88
48	356	0.91	0.92	0.92	0.92	0.91
49	1857	0.67	0.68	0.68	0.68	0.68
50	527	0.95	0.96	0.96	0.96	0.96
51	224	1.05	1.05	1.05	1.05	1.03
52	131	0.77	0.78	0.78	0.75	
53	195	0.77	0.81	0.81	0.81	0.78
54	255	0.55	0.60	0.61	0.62	0.60
55	69	1.44	1.30	1.30	1.30	1.27
56	125	1.54	1.44	1.44	1.45	1.43

Table A.1: Observed values, SMR values and empirical Bayes (EB) and full Bayes (FB) estimates relative risks from Poisson-Gamma and log normal models for pre-school 1999.

A.2 Census variable selection for 56 districts (Section 3.7.1)

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.3737	0.618	0.0314	−1.9677	0.661	0.0047	−2.1744	0.660	0.0020
% no car	0.0008	0.005	0.8716	−0.0012	0.005	0.8148	0.0036	0.005	0.4836
% overcrowded	−0.0031	0.026	0.9028	−0.0067	0.027	0.8077	0.0111	0.027	0.6840
% unemployed	−0.0011	0.014	0.9374	0.0082	0.015	0.5835	0.0136	0.015	0.3642
% low social class	−0.0091	0.008	0.2697	−0.0101	0.009	0.2515	−0.0142	0.009	0.1093
% children age 0-4	−0.0687	0.051	0.1809	−0.0357	0.054	0.5094	0.0123	0.054	0.8198
% born other EU	0.1078	0.074	0.1502	0.0907	0.078	0.2532	0.0415	0.078	0.5991
% born elsewhere	0.0010	0.034	0.9769	0.0126	0.037	0.7321	−0.0107	0.036	0.7697
% no qualifications	0.0018	0.012	0.8775	0.0089	0.012	0.4773	0.0093	0.012	0.4556
% high qualifications	−0.0015	0.010	0.8843	0.0053	0.011	0.6330	0.0105	0.011	0.3446
% working in agriculture	0.0007	0.013	0.9541	−0.0058	0.013	0.6668	−0.0072	0.013	0.5929
% lone parent households	−0.0005	0.029	0.9875	0.0041	0.031	0.8952	−0.0169	0.031	0.5876
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.4950	0.780	0.0617	−1.1774	0.801	0.1487	−1.9889	0.716	0.0080
% no car	0.0053	0.006	0.3907	0.0029	0.006	0.6481	0.0065	0.006	0.2589
% overcrowded	0.0210	0.032	0.5153	0.0080	0.033	0.8088	−0.0099	0.029	0.7391
% unemployed	0.0195	0.017	0.2702	0.0222	0.018	0.2230	0.0135	0.016	0.4076
% low social class	**−0.0213**	**0.010**	**0.0438**	**−0.0224**	**0.011**	**0.0398**	**−0.0223**	**0.009**	**0.0232**
% children age 0-4	−0.0163	0.063	0.7983	−0.0140	0.065	0.8302	0.0308	0.058	0.5998
% born other EU	0.0230	0.093	0.8066	0.0085	0.096	0.9302	0.0524	0.087	0.5485
% born elsewhere	−0.0355	0.043	0.4144	−0.0121	0.045	0.7874	−0.0419	0.040	0.3012
% no qualifications	0.0073	0.015	0.6180	0.0068	0.015	0.6521	0.0129	0.013	0.3400
% high qualifications	0.0075	0.013	0.5633	−0.0001	0.013	0.9911	0.0097	0.012	0.4245
% working in agriculture	−0.0106	0.016	0.5032	−0.0117	0.016	0.4708	−0.0082	0.014	0.5741
% lone parent households	−0.0414	0.036	0.2610	−0.0518	0.037	0.1728	−0.0257	0.033	0.4468

Table A.2: Parameter estimates, standard errors and p-values of all census variables for pre-school, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.9232	0.030	0.0000	1.8508	0.087	0.000	−1.6635	0.065	0.0000
% no car	–	–	–	–	–	–	0.0082	0.002	0.0000
% unemployed	–	–	–	0.0169	0.007	0.0197	–	–	–
% low social class	–	–	–	−0.0118	0.005	0.0335	−0.0145	0.001	0.0011
% born other EU	0.1496	0.028	0.0000	0.1259	0.032	0.0003	–	–	–
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.4202	0.080	0.0000	−1.3620	0.081	0.0000	−1.5200	0.07	0.0000
% no car	–	–	–	–	–	–	0.0037	0.002	0.0507
% unemployed	0.0302	0.008	0.0008	0.0190	0.009	0.0311	–	–	–
% low social class	−0.0213	0.006	0.0015	−0.0186	0.006	0.0054	−0.0117	0.004	0.0140

Table A.3: Parameter estimates, standard errors and p-values of census variables significant at the 5% level for pre-school, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−4.2509	1.36	0.0030	−2.8189	1.36	0.0446	−1.8551	1.34	0.1736
% no car	−0.0001	0.011	0.9955	0.0035	0.011	0.7386	−0.0009	0.010	0.9311
% overcrowded	**−0.1284**	**0.057**	**0.0297**	−0.0814	0.010	0.1613	−0.0385	0.056	0.4985
% unemployed	−0.0293	0.030	0.3403	−0.0173	0.031	0.5760	0.0115	0.030	0.7050
% low social class	−0.0049	0.018	0.7869	−0.0063	0.018	0.7288	−0.0174	0.018	0.3355
% children age 0-4	−0.0114	0.112	0.9192	−0.0867	0.112	0.4465	−0.1352	0.111	0.2309
% born other EU	**0.4709**	**0.158**	**0.0047**	**0.3364**	**0.158**	**0.0394**	0.2607	0.156	0.1028
% born elsewhere	−0.0318	0.077	0.6794	0.00337	0.076	0.9652	0.0227	0.075	0.7627
% no qualifications	0.0446	0.025	0.0850	0.0236	0.025	0.3587	0.0129	0.025	0.6094
% high qualifications	0.0220	0.022	0.3259	0.0038	0.022	0.8665	−0.0053	0.022	0.8119
% working in agriculture	−0.0112	0.027	0.6784	−0.0124	0.027	0.6485	−0.0138	0.027	0.6115
% lone parent households	0.0370	0.063	0.5591	−0.0153	0.063	0.8103	−0.0558	0.063	0.3787
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.8684	1.37	0.1794	−2.0608	1.28	0.1137	−4.2593	1.44	0.0049
% no car	−0.0019	0.011	0.8601	0.0081	0.010	0.4185	−0.0076	0.011	0.5037
% overcrowded	−0.0191	0.058	0.7410	0.0242	0.052	0.6495	**−0.1288**	**0.061**	**0.0400**
% unemployed	0.0241	0.031	0.4383	0.0346	0.028	0.2342	−0.0258	0.032	0.4238
% low social class	−0.0208	0.018	0.2615	−0.0262	0.016	0.1286	−0.0153	0.019	0.4280
% children age 0-4	−0.1337	0.112	0.2394	−0.0325	0.104	0.7559	−0.0045	0.119	0.9697
% born other EU	0.1949	0.159	0.2285	0.0368	0.150	0.8076	**0.5162**	**0.168**	**0.0036**
% born elsewhere	0.0246	0.075	0.7468	−0.0218	0.070	0.7587	−0.0201	0.082	0.8083
% no qualifications	0.0111	0.025	0.6683	0.0014	0.024	0.9542	**0.0613**	**0.027**	**0.0257**
% high qualifications	−0.0029	0.022	0.8997	0.0039	0.021	0.8527	0.0196	0.024	0.4086
% working in agriculture	−0.0229	0.027	0.4131	−0.0167	0.026	0.5169	−0.0165	0.028	0.5614
% lone parent households	−0.0614	0.064	0.3438	−0.0905	0.060	0.1364	0.0073	0.066	0.9124

Table A.4: Parameter estimates, standard errors and p-values of all census variables for primary 1 and 2, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−2.6670	0.287	0.0000	−2.8369	0.110	0.0000	−1.9161	0.433	0.0001
% overcrowded	−0.0808	0.023	0.0001	−0.0560	0.018	0.0035	−0.0520	0.017	0.0043
% children age 0-4	–	–	–	–	–	–	−0.1637	0.072	0.0269
% born other EU	0.3136	0.082	0.0000	0.3386	0.060	0.0000	0.2692	0.065	0.0001
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.6802	0.497	0.0014	−2.3924	0.125	0.0000	−3.1723	0.319	0.0000
% no car	–	–	–	0.0100	0.003	0.0043	–	–	–
% overcrowded	–	–	–	–	–	–	−0.1531	0.026	0.0000
% low social class	−0.0177	0.008	0.0321	−0.0288	0.008	0.0010	–	–	–
% children age 0-4	−0.1827	0.008	0.0177	–	–	–	–	–	–
% born other EU	0.1802	0.073	0.0179	–	–	–	0.4351	0.092	0.0000
% no qualifications	–	–	–	–	–	–	0.0234	0.009	0.0145

Table A.5: Parameter estimates, standard errors and p-values of census variables significant at the 5% level for primary 1 and 2, derived from logistic regression model using Penalized-quasi likelihood.

A.3 Census variable selection for 937 postcode sectors (Section 4.5.2)

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
(Intercept)	−2.0530	0.110	0.0000	−1.9463	0.108	0.0000	−1.7627	0.105	0.0000
% children age 0-4	0.0050	009	0.5731	−0.0041	0.009	0.6386	−0.0084	0.009	0.3315
% born other EU	**0.0672**	**0.014**	**0.0000**	**0.0454**	**0.014**	**0.0009**	0.0215	0.013	0.1113
% born elsewhere	−0.0090	0.006	0.1061	−0.0076	0.006	0.1662	−0.0050	0.005	0.3560
% no qualifications	**0.0053**	**0.002**	**0.0075**	0.0024	0.002	0.2217	0.0003	0.002	0.8805
% high qualifications	**0.0052**	**0.002**	**0.0051**	**0.0038**	**0.002**	**0.0369**	0.0027	0.002	0.1330
% lone parent households	−0.0017	0.004	0.6801	0.0002	0.004	0.9579	0.0041	0.004	0.3118
% working in agriculture	**0.0074**	**0.002**	**0.0076**	**0.0084**	**0.003**	**0.0022**	**0.0084**	**0.003**	**0.0022**
% no car	**0.0075**	**0.001**	**0.0000**	**0.0076**	**0.001**	**0.0000**	**0.0067**	**0.001**	**0.0000**
% overcrowded	**−0.0168**	**0.005**	**0.0008**	**−0.0135**	**0.004**	**0.0063**	−0.0017	0.005	0.7203
% unemployed	0.0044	0.004	0.2453	**0.0102**	**0.004**	**0.0065**	**0.0158**	**0.004**	**0.0000**
% low social class	**−0.0118**	**0.002**	**0.0000**	**−0.0134**	**0.002**	**0.0000**	**−0.0170**	**0.002**	**0.0000**
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
(Intercept)	−1.6442	0.110	0.0000	−1.8007	0.160	0.0000	−1.6296	0.102	0.0000
% children age 0-4	−0.0112	0.009	0.2101	−0.0028	0.014	0.8336	−0.0037	0.008	0.6508
% born other EU	**0.0305**	**0.014**	**0.0292**	0.0196	0.022	0.3651	0.0195	0.013	0.1345
% born elsewhere	−0.0081	0.006	0.1488	−0.0033	0.008	0.6621	−0.0093	0.005	0.0741
% no qualifications	0.0027	0.002	0.1675	0.0032	0.003	0.2780	0.0012	0.002	0.5135
% high qualifications	**0.0040**	**0.002**	**0.0351**	**0.0082**	**0.003**	**0.0030**	0.00317	0.002	0.0746
% lone parent households	0.0047	0.004	0.2628	−0.0006	0.006	0.9282	−0.00097	0.004	0.8249
% working in agriculture	**0.0075**	**0.003**	**0.0116**	0.0002	0.004	0.9508	**0.0077**	**0.003**	**0.0047**
% no car	**0.0052**	**0.001**	**0.0000**	0.0001	0.002	0.9478	**0.0039**	**0.001**	**0.0009**
% overcrowded	−0.0025	0.005	0.6041	−0.0023	0.007	0.7352	−0.0007	0.004	0.8732
% unemployed	**0.0213**	**0.004**	**0.0000**	0.0056	0.005	0.2150	**0.0152**	**0.004**	**0.0000**
% low social class	**−0.0205**	**0.002**	**0.0000**	−0.0022	0.003	**0.4204**	**−0.0170**	**0.002**	**0.0000**

Table A.6: Parameter estimates of all census variables in the model with their standard errors and p-values, pre-school, 937 postcode sectors for years 2000-2005, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.7817	0.024	0.0000	−1.8222	0.024	0.0000	−1.7251	0.023	0.0000
% born other EU	0.0462	0.009	0.0000	0.0394	0.008	0.0000	0.0240	0.008	0.0038
% working in agriculture	0.0103	0.002	0.0000	0.0100	0.002	0.0001	0.0087	0.002	0.0004
% no car	0.0085	0.001	0.0000	0.0077	0.001	0.0000	0.0068	0.001	0.0000
% overcrowded	−0.0147	0.005	0.0025	−0.0135	0.005	0.0050	–	–	–
% unemployed	–	–	–	0.0112	0.004	0.0026	0.0167	0.003	0.0000
% low social class	−0.0113	0.002	0.0000	−0.0144	0.002	0.0000	−0.0185	0.002	0.0000
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−1.5334	0.024	0.0000	−1.6620	0.025	0.0000	−1.6013	0.034	0.0000
% born other EU	0.0213	0.009	0.0133	–	–	–	–	–	–
% high qualifications	–	–	–	0.0060	0.001	0.0000	0.0018	0.001	0.0345
% working in agriculture	0.0087	0.003	0.0010	–	–	–	0.0090	0.002	0.0002
% no car	0.0056	0.001	0.0000	–	–	–	0.0038	0.001	0.0004
% unemployed	0.0226	0.004	0.0000	–	–	–	0.0149	0.004	0.0000
% low social class	−0.0215	0.002	0.0000	–	–	–	−0.0169	0.002	0.0000

Table A.7: Parameter estimates of significant census variables at 5% level, with their standard errors and p-values, for pre-school, 937 postcode sectors, for 2000-2005, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−2.9859	0.227	0.0000	−3.1258	0.229	0.0000	−3.3335	0.227	0.0000
% children age 0-4	0.0186	0.018	0.3106	0.0284	0.019	0.1251	0.0273	0.018	0.1327
% born other EU	**0.1446**	**0.028**	**0.0000**	**0.1294**	**0.028**	**0.0000**	**0.1296**	**0.027**	**0.0000**
% born elsewhere	−0.0179	0.012	0.1311	−0.0113	0.012	0.3363	−0.0110	0.011	0.3374
% no qualifications	0.0063	0.004	0.1287	0.0078	0.004	0.0609	**0.0129**	**0.004**	**0.0017**
% high qualifications	0.0052	0.004	0.1668	0.0062	0.004	0.1035	**0.0097**	**0.004**	**0.0100**
% lone parent households	−0.0083	0.009	0.3572	−0.0135	0.009	0.1370	−0.0158	0.009	0.0725
% working in agriculture	**0.0330**	**0.005**	**0.0000**	**0.0232**	**0.006**	**0.0000**	**0.0158**	**0.006**	**0.0050**
% no car	**0.0171**	**0.002**	**0.0000**	**0.0184**	**0.002**	**0.0000**	**0.0150**	**0.002**	**0.0000**
% overcrowded	**−0.0578**	**0.011**	**0.0000**	**−0.0583**	**0.011**	**0.0000**	**−0.0545**	**0.011**	**0.0000**
% unemployed	−0.0069	0.008	0.3841	−0.0033	0.008	0.6805	0.0076	0.008	0.3374
% low social class	**−0.0210**	**0.004**	**0.0000**	**−0.0222**	**0.004**	**0.0000**	**−0.0228**	**0.004**	**0.0000**
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−2.8828	0.219	0.0000	−2.5016	0.199	0.0000	−2.3477	0.203	0.0000
% children age 0-4	−0.0046	0.018	0.7980	−0.0242	0.017	0.1459	−0.0208	0.016	0.2063
% born other EU	**0.1081**	**0.027**	**0.0001**	**0.0606**	**0.025**	**0.0164**	0.0394	0.025	0.1217
% born elsewhere	−0.0140	0.011	0.2179	−0.0095	0.010	0.3564	−0.0129	0.010	0.2154
% no qualifications	0.0036	0.004	0.3616	−0.0009	0.004	0.7996	0.0019	0.004	0.6119
% high qualifications	0.0047	0.004	0.2036	0.0028	0.003	0.4108	0.0035	0.003	0.3121
% lone parent households	−0.0036	0.009	0.6800	0.0079	0.008	0.3245	0.0018	0.008	0.8185
% working in agriculture	**0.0202**	**0.006**	**0.0003**	**0.0166**	**0.005**	**0.0017**	**0.0133**	**0.005**	**0.0109**
% no car	**0.0138**	**0.002**	**0.0000**	**0.0109**	**0.002**	**0.0000**	**0.0063**	**0.002**	**0.0073**
% overcrowded	**−0.0411**	**0.010**	**0.0001**	**−0.0196**	**0.009**	**0.0396**	**−0.0242**	**0.009**	**0.0103**
% unemployed	0.0146	0.008	0.0582	**0.0271**	**0.007**	**0.0002**	**0.0307**	**0.007**	**0.0000**
% low social class	**−0.0265**	**0.004**	**0.0000**	**−0.0312**	**0.004**	**0.0000**	**−0.0281**	**0.004**	**0.0000**

Table A.8: Parameter estimates of all census variables in the model with their standard errors and p-values, primary 1 and 2, 937 postcode sectors, for years 2000-2005, derived from logistic regression model using Penalized-quasi likelihood.

	2000			2001			2002		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−2.6426	0.050	0.0000	−2.6765	0.051	0.0000	−3.1436	0.177	0.0000
% born other EU	0.1056	0.017	0.0000	0.0993	0.017	0.0000	0.0967	0.020	0.0000
% no qualifications	–	–	–	–		–	− 0.0094	0.004	0.0099
% high qualifications	–	–	–	–	–	–	0.0086	0.003	0.0144
% working in agriculture	0.0368	0.005	0.0000	0.0282	0.005	0.0000	0.0198	0.005	0.0002
% no car	0.0154	0.002	0.0000	0.0173	0.002	0.0000	0.0149	0.002	0.0000
% overcrowded	−0.0590	0.010	0.0000	−0.0579	0.011	0.0000	−0.0550	0.010	0.0000
% low social class	−0.0218	0.004	0.0000	−0.0232	0.004	0.0000	−0.0201	0.004	0.0000
	2003			2004			2005		
Variable	Estimate	sd	p-value	Estimate	sd	p-value	Estimate	sd	p-value
Intercept	−2.7329	0.049	0.0000	−2.5735	0.045	0.0000	−2.3477	0.044	0.0000
% born other EU	0.0917	0.017	0.0000	0.0653	0.015	0.0000	0.0316	0.016	0.0421
% working in agriculture	0.0231	0.005	0.0000	0.0162	0.005	0.0007	0.0151	0.005	0.0014
% no car	0.0137	0.002	0.0000	0.0112	0.002	0.0000	0.0066	0.002	0.0039
% overcrowded	−0.0423	0.010	0.0000	−0.0207	0.009	0.0256	−0.0257	0.009	0.0054
% unemployed	0.0155	0.008	0.0414	0.0286	0.007	0.0001	0.0319	0.007	0.0000
% low social class	−0.0272	0.004	0.0000	−0.0328	0.004	0.0000	−0.0289	004	0.0000

Table A.9: Parameter estimates of significant census variables at 5% level, with their standard errors and p-values, for primary 1 and 2, 937 postcode sectors, for 2000-2005, derived from logistic regression model using Penalized-quasi likelihood.

Appendix B

Chapter 6 Tables

B.1 Changing Mean Level (Section 6.4.1)

	Mean Square Difference								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$2.2e-06$	$4.8e-07$	$1.5e-08$	$2.2e-08$	$4.7e-09$	$1.5e-10$	$2.2e-10$	$5.2e-11$	$1.7e-12$
1.1	$1.1e-04$	$5.3e-06$	$1.7e-07$	$3.3e-08$	$6.3e-09$	$2.0e-10$	$2.3e-10$	$5.5e-11$	$1.7e-12$
1.2	$4.2e-04$	$1.6e-05$	$5.0e-07$	$6.6e-08$	$1.1e-08$	$3.4e-10$	$2.5e-10$	$5.9e-11$	$1.9e-12$
1.3	$9.3e-04$	$2.9e-05$	$9.3e-07$	$1.2e-07$	$1.5e-08$	$4.7e-10$	$2.6e-10$	$6.0e-11$	$1.9e-12$
1.4	$1.7e-03$	$5.2e-05$	$1.7e-06$	$1.9e-07$	$2.0e-08$	$6.3e-10$	$2.8e-10$	$6.9e-11$	$2.2e-12$
1.5	$2.6e-03$	$7.9e-05$	$2.5e-06$	$2.9e-07$	$2.6e-08$	$8.1e-10$	$3.1e-10$	$7.3e-11$	$2.3e-12$
1.6	$3.7e-03$	$1.2e-04$	$3.7e-06$	$4.0e-07$	$3.1e-08$	$9.8e-10$	$3.3e-10$	$7.4e-11$	$2.3e-12$
1.7	$5.1e-03$	$1.5e-04$	$4.7e-06$	$5.4e-07$	$3.9e-08$	$1.2e-09$	$3.5e-10$	$7.9e-11$	$2.5e-12$
1.8	$6.6e-03$	$2.0e-04$	$6.4e-06$	$6.9e-07$	$4.3e-08$	$1.4e-09$	$3.8e-10$	$8.7e-11$	$2.7e-12$
1.9	$8.4e-03$	$2.5e-04$	$8.0e-06$	$8.7e-07$	$5.5e-08$	$1.7e-09$	$4.0e-10$	$9.0e-11$	$2.8e-12$
2.0	$1.0e-02$	$3.1e-04$	$9.7e-06$	$1.1e-06$	$6.0e-08$	$1.9e-09$	$4.3e-10$	$9.9e-11$	$3.1e-12$
2.1	$1.2e-02$	$3.6e-04$	$1.1e-05$	$1.3e-06$	$6.9e-08$	$2.2e-09$	$4.7e-10$	$1.1e-10$	$3.3e-12$
2.2	$1.5e-02$	$4.4e-04$	$1.4e-05$	$1.5e-06$	$7.7e-08$	$2.4e-09$	$5.1e-10$	$1.1e-10$	$3.6e-12$
2.3	$1.7e-02$	$5.3e-04$	$1.7e-05$	$1.8e-06$	$8.8e-08$	$2.8e-09$	$5.4e-10$	$1.2e-10$	$3.8e-12$
2.4	$2.0e-02$	$6.1e-04$	$1.9e-05$	$2.1e-06$	$1.0e-07$	$3.1e-09$	$5.7e-10$	$1.2e-10$	$3.9e-12$
2.5	$2.3e-02$	$6.8e-04$	$2.2e-05$	$2.4e-06$	$1.1e-07$	$3.4e-09$	$6.2e-10$	$1.3e-10$	$4.1e-12$
2.6	$2.6e-02$	$7.7e-04$	$2.4e-05$	$2.7e-06$	$1.2e-07$	$3.8e-09$	$6.6e-10$	$1.4e-10$	$4.4e-12$
2.7	$3.0e-02$	$9.2e-04$	$2.9e-05$	$3.0e-06$	$1.3e-07$	$4.3e-09$	$7.1e-10$	$1.5e-10$	$4.8e-12$
2.8	$3.3e-02$	$1.0e-03$	$3.2e-05$	$3.4e-06$	$1.4e-07$	$4.4e-09$	$7.7e-10$	$1.6e-10$	$5.0e-12$
2.9	$3.7e-02$	$1.1e-03$	$3.5e-05$	$3.8e-06$	$1.6e-07$	$5.0e-09$	$8.1e-10$	$1.6e-10$	$5.0e-12$
3.0	$4.1e-02$	$1.2e-03$	$3.9e-05$	$4.2e-06$	$1.7e-07$	$5.4e-09$	$8.5e-10$	$1.6e-10$	$5.2e-12$
3.1	$4.6e-02$	$1.3e-03$	$4.1e-05$	$4.6e-06$	$1.9e-07$	$6.0e-09$	$9.0e-10$	$1.7e-10$	$5.4e-12$
3.2	$5.0e-02$	$1.4e-03$	$4.6e-05$	$5.1e-06$	$1.9e-07$	$6.1e-09$	$9.6e-10$	$1.8e-10$	$5.7e-12$
3.3	$5.5e-02$	$1.5e-03$	$4.9e-05$	$5.5e-06$	$2.2e-07$	$6.9e-09$	$1.0e-09$	$1.9e-10$	$5.9e-12$
3.4	$6.0e-02$	$1.7e-03$	$5.4e-05$	$6.0e-06$	$2.3e-07$	$7.4e-09$	$1.1e-09$	$2.0e-10$	$6.4e-12$
3.5	$6.5e-02$	$1.9e-03$	$5.9e-05$	$6.5e-06$	$2.5e-07$	$7.8e-09$	$1.1e-09$	$2.1e-10$	$6.6e-12$
3.6	$7.0e-02$	$2.1e-03$	$6.6e-05$	$7.0e-06$	$2.7e-07$	$8.7e-09$	$1.2e-09$	$2.2e-10$	$7.1e-12$
3.7	$7.5e-02$	$2.3e-03$	$7.2e-05$	$7.6e-06$	$2.8e-07$	$8.9e-09$	$1.3e-09$	$2.3e-10$	$7.3e-12$
3.8	$8.1e-02$	$2.4e-03$	$7.7e-05$	$8.2e-06$	$3.0e-07$	$9.5e-09$	$1.3e-09$	$2.4e-10$	$7.5e-12$
3.9	$8.7e-02$	$2.6e-03$	$8.3e-05$	$8.7e-06$	$3.1e-07$	$9.8e-09$	$1.4e-09$	$2.5e-10$	$7.8e-12$
4.0	$9.3e-02$	$2.7e-03$	$8.6e-05$	$9.4e-06$	$3.3e-07$	$1.1e-08$	$1.5e-09$	$2.6e-10$	$8.1e-12$

Table B.1: MSD summary measures (mean, standard deviation (sd), standard error (se)), for change in mean.

	Peak-to Signal Noise Ratio								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	1832.94	827.260	26.1603	34.36	11.503	0.3638	10.64	3.82	0.1206
1.1	51.28	17.046	0.5390	26.09	8.565	0.2708	10.94	4.00	0.1265
1.2	18.50	5.387	0.1704	16.65	4.663	0.1475	10.85	3.86	0.1222
1.3	11.09	3.093	0.0978	11.93	3.617	0.1144	10.67	3.65	0.1155
1.4	7.86	2.041	0.0646	9.14	2.433	0.0769	10.53	3.64	0.1151
1.5	6.28	1.480	0.0468	7.38	1.788	0.0565	10.48	3.64	0.1150
1.6	5.36	1.248	0.0395	6.33	1.467	0.0464	10.26	3.34	0.1057
1.7	4.75	1.059	0.0335	5.51	1.207	0.0382	10.32	3.39	0.1071
1.8	4.28	0.942	0.0298	4.97	1.044	0.0330	9.98	3.27	0.1033
1.9	3.94	0.815	0.0258	4.57	0.930	0.0294	10.23	3.52	0.1112
2.0	3.67	0.765	0.0242	4.25	0.875	0.0277	9.91	3.48	0.1100
2.1	3.45	0.652	0.0206	3.96	0.737	0.0233	9.71	3.24	0.1024
2.2	3.33	0.626	0.0198	3.76	0.721	0.0228	9.62	3.08	0.0975
2.3	3.17	0.585	0.0185	3.58	0.653	0.0207	9.19	2.96	0.0936
2.4	3.06	0.592	0.0187	3.39	0.605	0.0191	9.14	2.87	0.0906
2.5	2.98	0.546	0.0173	3.30	0.591	0.0187	8.98	2.79	0.0883
2.6	2.90	0.505	0.0160	3.23	0.584	0.0185	8.83	2.83	0.0896
2.7	2.81	0.497	0.0157	3.14	0.568	0.0180	8.78	2.82	0.0892
2.8	2.74	0.512	0.0162	3.03	0.520	0.0164	8.46	2.62	0.0830
2.9	2.70	0.448	0.0142	2.96	0.499	0.0158	8.31	2.58	0.0816
3.0	2.62	0.480	0.0152	2.93	0.511	0.0162	8.15	2.42	0.0766
3.1	2.58	0.455	0.0144	2.83	0.472	0.0149	7.92	2.32	0.0732
3.2	2.53	0.403	0.0127	2.78	0.448	0.0142	7.84	2.43	0.0768
3.3	2.52	0.404	0.0128	2.73	0.474	0.0150	7.70	2.38	0.0752
3.4	2.44	0.383	0.0121	2.69	0.464	0.0147	7.53	2.14	0.0677
3.5	2.45	0.422	0.0134	2.63	0.440	0.0139	7.49	2.29	0.0725
3.6	2.41	0.398	0.0126	2.63	0.457	0.0144	7.36	2.12	0.0670
3.7	2.39	0.420	0.0133	2.58	0.451	0.0143	7.14	1.96	0.0619
3.8	2.35	0.361	0.0114	2.53	0.441	0.0139	7.14	2.03	0.0643
3.9	2.34	0.394	0.0125	2.51	0.409	0.0129	6.93	1.90	0.0600
4.0	2.34	0.393	0.0124	2.50	0.436	0.0138	6.92	1.96	0.0620
	Structural similarity Index Measure								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	0.99	0.0023	$7.3e-05$	0.59	0.100	0.00317	0.0160	0.142	0.0045
1.1	0.98	0.0028	$8.8e-05$	0.60	0.097	0.00307	0.0160	0.143	0.0045
1.2	0.96	0.0033	$1.1e-04$	0.59	0.094	0.00297	0.0176	0.139	0.0044
1.3	0.93	0.0044	$1.4e-04$	0.59	0.091	0.00289	0.0180	0.136	0.0043
1.4	0.89	0.0049	$1.6e-04$	0.57	0.093	0.00293	0.0164	0.129	0.0041
1.5	0.85	0.0055	$1.7e-04$	0.56	0.086	0.00272	0.0114	0.132	0.0042
1.6	0.80	0.0055	$1.7e-04$	0.54	0.081	0.00255	0.0145	0.118	0.0037
1.7	0.76	0.0057	$1.8e-04$	0.52	0.076	0.00240	0.0127	0.123	0.0039
1.8	0.72	0.0059	$1.9e-04$	0.49	0.075	0.00238	0.0092	0.114	0.0036
1.9	0.68	0.0060	$1.9e-04$	0.47	0.068	0.00214	0.0215	0.109	0.0034
2.0	0.64	0.0059	$1.9e-04$	0.45	0.062	0.00196	0.0128	0.109	0.0034
2.1	0.60	0.0056	$1.8e-04$	0.43	0.062	0.00195	0.0153	0.107	0.0034
2.2	0.57	0.0054	$1.7e-04$	0.42	0.059	0.00186	0.0134	0.100	0.0032
2.3	0.53	0.0054	$1.7e-04$	0.39	0.054	0.00171	0.0123	0.099	0.0031
2.4	0.50	0.0051	$1.6e-04$	0.37	0.054	0.00169	0.0154	0.095	0.0030
2.5	0.47	0.0049	$1.5e-04$	0.36	0.050	0.00159	0.0148	0.092	0.0029
2.6	0.45	0.0047	$1.5e-04$	0.34	0.048	0.00153	0.0146	0.087	0.0028
2.7	0.42	0.0046	$1.5e-04$	0.33	0.045	0.00141	0.0106	0.086	0.0027
2.8	0.40	0.0045	$1.4e-04$	0.31	0.043	0.00135	0.0143	0.080	0.0025
2.9	0.38	0.0041	$1.3e-04$	0.30	0.041	0.00131	0.0094	0.076	0.0024
3.0	0.36	0.0040	$1.3e-04$	0.28	0.038	0.00120	0.0148	0.075	0.0024
3.1	0.34	0.0039	$1.2e-04$	0.27	0.037	0.00116	0.0128	0.073	0.0023
3.2	0.32	0.0039	$1.2e-04$	0.26	0.035	0.00112	0.0121	0.069	0.0022
3.3	0.31	0.0036	$1.1e-04$	0.25	0.034	0.00108	0.0141	0.070	0.0022
3.4	0.29	0.0033	$1.0e-04$	0.24	0.033	0.00103	0.0114	0.066	0.0021
3.5	0.28	0.0033	$1.0e-04$	0.23	0.032	0.00100	0.0110	0.065	0.0020
3.6	0.27	0.0030	$9.6e-05$	0.22	0.029	0.00091	0.0110	0.059	0.0019
3.7	0.25	0.0030	$9.6e-05$	0.21	0.027	0.00086	0.0093	0.060	0.0019
3.8	0.24	0.0029	$9.1e-05$	0.20	0.026	0.00083	0.0134	0.060	0.0019
3.9	0.23	0.0027	$8.6e-05$	0.19	0.026	0.00082	0.0120	0.057	0.0018
4.0	0.22	0.0027	$8.5e-05$	0.18	0.025	0.00078	0.0120	0.054	0.0017

Table B.2: PSNRR (above) and SSIM (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in mean.

	Inter Region Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$8.6e-01$	$1.8e-01$	$5.5e-03$	$1.8e-01$	$2.8e-01$	$8.8e-03$	$6.8e-03$	$2.9e-02$	$9.1e-04$
1.1	$4.8e-02$	$7.2e-02$	$2.3e-03$	$1.4e-01$	$2.5e-01$	$7.8e-03$	$5.6e-03$	$2.7e-02$	$8.4e-04$
1.2	$4.6e-05$	$2.6e-04$	$8.2e-06$	$8.5e-02$	$2.1e-01$	$6.6e-03$	$6.1e-03$	$3.0e-02$	$9.5e-04$
1.3	$8.6e-09$	$6.7e-08$	$2.1e-09$	$3.7e-02$	$1.3e-01$	$4.3e-03$	$4.8e-03$	$2.9e-02$	$9.3e-04$
1.4	$2.5e-12$	$5.5e-11$	$1.7e-12$	$1.0e-02$	$6.8e-02$	$2.1e-03$	$4.0e-03$	$2.8e-02$	$8.8e-04$
1.5	$1.2e-15$	$3.7e-14$	$1.2e-15$	$2.8e-03$	$3.0e-02$	$9.4e-04$	$1.5e-03$	$1.4e-02$	$4.4e-04$
1.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.1e-03$	$2.0e-02$	$6.2e-04$	$1.4e-03$	$1.8e-02$	$5.7e-04$
1.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$9.6e-07$	$1.9e-05$	$5.9e-07$	$4.5e-04$	$7.5e-03$	$2.4e-04$
1.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.4e-03$	$3.3e-02$	$1.0e-03$	$8.7e-04$	$1.7e-02$	$5.5e-04$
1.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.2e-07$	$3.6e-06$	$1.2e-07$	$2.3e-04$	$3.7e-03$	$1.2e-04$
2.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.6e-12$	$4.9e-11$	$1.6e-12$	$5.1e-05$	$8.3e-04$	$2.6e-05$
2.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.8e-13$	$8.5e-12$	$2.7e-13$	$2.5e-05$	$4.9e-04$	$1.5e-05$
2.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$4.8e-06$	$1.1e-04$	$3.5e-06$
2.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.9e-06$	$3.1e-05$	$9.9e-07$
2.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.2e-04$	$3.6e-03$	$1.1e-04$
2.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$9.9e-09$	$2.5e-07$	$8.0e-09$
2.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$7.5e-09$	$1.7e-07$	$5.3e-09$
2.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$6.9e-09$	$1.7e-07$	$5.4e-09$
2.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.4e-09$	$5.4e-08$	$1.7e-09$
2.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.2e-06$	$3.9e-05$	$1.2e-06$
3.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$6.1e-11$	$1.7e-09$	$5.4e-11$
3.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$8.9e-11$	$2.7e-09$	$8.4e-11$
3.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.6e-09$	$5.1e-08$	$1.6e-09$
3.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.7e-13$	$8.3e-12$	$2.6e-13$
3.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$4.1e-12$	$1.3e-10$	$4.1e-12$
3.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.1e-09$	$3.4e-08$	$1.1e-09$
3.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.3e-16$	$6.6e-15$	$2.1e-16$
3.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$5.7e-13$	$1.8e-11$	$5.7e-13$
3.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$5.6e-14$	$1.8e-12$	$5.6e-14$
3.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.3e-18$	$6.7e-17$	$2.1e-18$
4.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.8e-15$	$8.7e-14$	$2.8e-15$
	Most Similar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	0.41888	0.2863	$9.1e-03$	0.3871	0.284	0.00899	0.55180	0.3024	$9.6e-03$
1.1	0.39574	0.2859	$9.0e-03$	0.3876	0.285	0.00902	0.51833	0.3009	$9.5e-03$
1.2	0.35504	0.2858	$9.0e-03$	0.3543	0.273	0.00864	0.45881	0.3071	$9.7e-03$
1.3	0.31197	0.2687	$8.5e-03$	0.3136	0.273	0.00864	0.39074	0.3018	$9.5e-03$
1.4	0.24557	0.2407	$7.6e-03$	0.2992	0.268	0.00849	0.30219	0.2810	$8.9e-03$
1.5	0.19794	0.2238	$7.1e-03$	0.2408	0.253	0.00802	0.24470	0.2746	$8.7e-03$
1.6	0.15035	0.1953	$6.2e-03$	0.2082	0.243	0.00768	0.18980	0.2374	$7.5e-03$
1.7	0.11985	0.1690	$5.3e-03$	0.1786	0.229	0.00723	0.15102	0.2098	$6.6e-03$
1.8	0.08412	0.1472	$4.7e-03$	0.1423	0.193	0.00610	0.11797	0.1957	$6.2e-03$
1.9	0.06598	0.1182	$3.7e-03$	0.1209	0.191	0.00605	0.08956	0.1575	$5.0e-03$
2.0	0.05642	0.1131	$3.6e-03$	0.0922	0.155	0.00490	0.06342	0.1227	$3.9e-03$
2.1	0.03395	0.0751	$2.4e-03$	0.0765	0.142	0.00450	0.05120	0.1106	$3.5e-03$
2.2	0.02609	0.0646	$2.0e-03$	0.0769	0.149	0.00472	0.03332	0.0738	$2.3e-03$
2.3	0.02193	0.0544	$1.7e-03$	0.0590	0.121	0.00381	0.02673	0.0716	$2.3e-03$
2.4	0.01561	0.0466	$1.5e-03$	0.0461	0.104	0.00329	0.01948	0.0570	$1.8e-03$
2.5	0.01191	0.0407	$1.3e-03$	0.0364	0.088	0.00277	0.01144	0.0368	$1.2e-03$
2.6	0.00975	0.0327	$1.0e-03$	0.0355	0.101	0.00320	0.01237	0.0417	$1.3e-03$
2.7	0.00694	0.0228	$7.2e-04$	0.0278	0.073	0.00230	0.00926	0.0319	$1.0e-03$
2.8	0.00711	0.0239	$7.6e-04$	0.0214	0.057	0.00181	0.00653	0.0251	$7.9e-04$
2.9	0.00493	0.0204	$6.5e-04$	0.0162	0.054	0.00170	0.00472	0.0181	$5.7e-04$
3.0	0.00432	0.0246	$7.8e-04$	0.0145	0.053	0.00166	0.00476	0.0197	$6.2e-04$
3.1	0.00339	0.0184	$5.8e-04$	0.0121	0.040	0.00126	0.00404	0.0160	$5.1e-04$
3.2	0.00188	0.0089	$2.8e-04$	0.0099	0.034	0.00109	0.00262	0.0101	$3.2e-04$
3.3	0.00173	0.0090	$2.8e-04$	0.0079	0.033	0.00105	0.00214	0.0083	$2.6e-04$
3.4	0.00170	0.0123	$3.9e-04$	0.0062	0.024	0.00076	0.00200	0.0150	$4.7e-04$
3.5	0.00111	0.0051	$1.6e-04$	0.0057	0.022	0.00069	0.00175	0.0144	$4.6e-04$
3.6	0.00083	0.0071	$2.3e-04$	0.0041	0.015	0.00047	0.00116	0.0048	$1.5e-04$
3.7	0.00111	0.0070	$2.2e-04$	0.0029	0.012	0.00038	0.00110	0.0058	$1.8e-04$
3.8	0.00072	0.0074	$2.3e-04$	0.0044	0.022	0.00068	0.00096	0.0049	$1.5e-04$
3.9	0.00049	0.0025	$7.8e-05$	0.0030	0.016	0.00049	0.00074	0.0048	$1.5e-04$
4.0	0.00028	0.0011	$3.5e-05$	0.0029	0.022	0.00070	0.00058	0.0030	$9.4e-05$

Table B.3: IRD (above) and MSDI (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in mean.

	Most Dissimilar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$7.5e-01$	$2.7e-01$	$8.5e-03$	$1.6e-01$	$2.6e-01$	$8.4e-03$	0.00813	0.075	0.00236
1.1	$1.2e-01$	$1.8e-01$	$5.7e-03$	$1.7e-01$	$2.8e-01$	$8.8e-03$	0.00586	0.052	0.00164
1.2	$3.0e-03$	$9.6e-03$	$3.0e-04$	$1.1e-01$	$2.3e-01$	$7.3e-03$	0.01022	0.070	0.00221
1.3	$3.7e-05$	$2.4e-04$	$7.6e-06$	$7.6e-02$	$2.0e-01$	$6.2e-03$	0.01422	0.081	0.00257
1.4	$3.1e-07$	$3.5e-06$	$1.1e-07$	$4.4e-02$	$1.5e-01$	$4.7e-03$	0.01232	0.063	0.00198
1.5	$1.9e-09$	$2.6e-08$	$8.2e-10$	$2.8e-02$	$1.2e-01$	$3.9e-03$	0.01524	0.066	0.00208
1.6	$1.9e-12$	$2.4e-11$	$7.5e-13$	$9.9e-03$	$6.6e-02$	$2.1e-03$	0.01289	0.044	0.00140
1.7	$5.8e-14$	$1.5e-12$	$4.8e-14$	$4.5e-03$	$4.7e-02$	$1.5e-03$	0.01537	0.057	0.00181
1.8	$1.1e-15$	$2.3e-14$	$7.2e-16$	$3.6e-03$	$3.6e-02$	$1.1e-03$	0.01741	0.061	0.00193
1.9	$2.2e-19$	$7.0e-18$	$2.2e-19$	$1.2e-03$	$2.0e-02$	$6.4e-04$	0.01189	0.048	0.00152
2.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$7.7e-04$	$1.5e-02$	$4.7e-04$	0.01154	0.044	0.00138
2.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$8.0e-04$	$1.9e-02$	$6.0e-04$	0.01210	0.048	0.00150
2.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.3e-04$	$9.6e-03$	$3.0e-04$	0.01097	0.046	0.00144
2.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.5e-06$	$7.0e-05$	$2.2e-06$	0.01281	0.057	0.00181
2.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$5.3e-07$	$1.5e-05$	$4.9e-07$	0.01030	0.040	0.00128
2.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$7.2e-07$	$1.7e-05$	$5.3e-07$	0.00783	0.037	0.00118
2.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.1e-09$	$6.7e-08$	$2.1e-09$	0.00743	0.037	0.00116
2.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.7e-11$	$8.4e-10$	$2.7e-11$	0.00488	0.026	0.00082
2.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.0e-15$	$9.3e-14$	$2.9e-15$	0.00275	0.022	0.00070
2.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.3e-13$	$6.0e-12$	$1.9e-13$	0.00427	0.024	0.00076
3.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$6.9e-13$	$2.1e-11$	$6.7e-13$	0.00521	0.031	0.00098
3.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$5.8e-09$	$1.8e-07$	$5.8e-09$	0.00259	0.016	0.00051
3.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00271	0.019	0.00062
3.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00273	0.021	0.00066
3.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00176	0.015	0.00047
3.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00161	0.016	0.00051
3.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00136	0.013	0.00041
3.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00167	0.015	0.00048
3.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00148	0.015	0.00046
3.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00098	0.011	0.00036
4.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00149	0.016	0.00051
	Average Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$6.1e-01$	0.26970	$8.5e-03$	0.3946	0.287	0.00907	0.4552	0.311	0.00982
1.1	$5.8e-01$	0.27304	$8.6e-03$	0.3807	0.286	0.00905	0.4214	0.311	0.00984
1.2	$4.6e-01$	0.26226	$8.3e-03$	0.3514	0.283	0.00894	0.3611	0.302	0.00957
1.3	$3.5e-01$	0.23970	$7.6e-03$	0.3252	0.284	0.00899	0.3217	0.298	0.00943
1.4	$2.4e-01$	0.19571	$6.2e-03$	0.2938	0.273	0.00863	0.2702	0.286	0.00905
1.5	$1.6e-01$	0.15420	$4.9e-03$	0.2600	0.265	0.00838	0.2188	0.263	0.00831
1.6	$1.1e-01$	0.11859	$3.8e-03$	0.2167	0.243	0.00768	0.1702	0.242	0.00766
1.7	$7.4e-02$	0.08919	$2.8e-03$	0.1879	0.237	0.00750	0.1355	0.211	0.00667
1.8	$5.0e-02$	0.06646	$2.1e-03$	0.1643	0.224	0.00710	0.1216	0.201	0.00636
1.9	$3.3e-02$	0.05033	$1.6e-03$	0.1338	0.192	0.00607	0.0880	0.170	0.00537
2.0	$2.5e-02$	0.04111	$1.3e-03$	0.1009	0.165	0.00522	0.0807	0.165	0.00521
2.1	$1.7e-02$	0.03021	$9.6e-04$	0.0995	0.164	0.00519	0.0531	0.124	0.00392
2.2	$1.2e-02$	0.02650	$8.4e-04$	0.0691	0.129	0.00409	0.0497	0.119	0.00376
2.3	$8.8e-03$	0.02022	$6.4e-04$	0.0612	0.124	0.00394	0.0509	0.125	0.00395
2.4	$5.7e-03$	0.01318	$4.2e-04$	0.0487	0.110	0.00347	0.0371	0.100	0.00317
2.5	$4.4e-03$	0.00998	$3.2e-04$	0.0410	0.094	0.00299	0.0233	0.074	0.00234
2.6	$2.8e-03$	0.00668	$2.1e-04$	0.0335	0.084	0.00265	0.0227	0.073	0.00231
2.7	$2.4e-03$	0.00661	$2.1e-04$	0.0267	0.076	0.00239	0.0178	0.061	0.00194
2.8	$1.9e-03$	0.00613	$1.9e-04$	0.0222	0.064	0.00204	0.0152	0.054	0.00172
2.9	$1.2e-03$	0.00341	$1.1e-04$	0.0200	0.059	0.00186	0.0143	0.055	0.00174
3.0	$8.5e-04$	0.00279	$8.8e-05$	0.0170	0.046	0.00147	0.0121	0.043	0.00136
3.1	$6.0e-04$	0.00174	$5.5e-05$	0.0127	0.037	0.00118	0.0138	0.055	0.00173
3.2	$5.0e-04$	0.00141	$4.5e-05$	0.0110	0.034	0.00109	0.0103	0.040	0.00127
3.3	$4.1e-04$	0.00180	$5.7e-05$	0.0067	0.021	0.00067	0.0101	0.041	0.00129
3.4	$2.8e-04$	0.00103	$3.3e-05$	0.0071	0.033	0.00103	0.0095	0.042	0.00133
3.5	$2.5e-04$	0.00106	$3.4e-05$	0.0056	0.023	0.00074	0.0073	0.038	0.00121
3.6	$2.0e-04$	0.00095	$3.0e-05$	0.0056	0.024	0.00075	0.0091	0.038	0.00120
3.7	$1.7e-04$	0.00104	$3.3e-05$	0.0037	0.021	0.00065	0.0074	0.031	0.00098
3.8	$1.3e-04$	0.00071	$2.2e-05$	0.0037	0.021	0.00066	0.0054	0.022	0.00068
3.9	$7.8e-05$	0.00035	$1.1e-05$	0.0038	0.022	0.00069	0.0050	0.025	0.00080
4.0	$1.0e-04$	0.00056	$1.8e-05$	0.0036	0.024	0.00076	0.0075	0.032	0.00100

Table B.4: MDD (above) and AVND (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in mean.

	Most Similar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$8.9e-01$	$1.4e-01$	$4.5e-03$	$3.5e-01$	$2.9e-01$	$9.2e-03$	0.22248	0.2523	$8.0e-03$
1.1	$4.3e-01$	$2.3e-01$	$7.4e-03$	$3.0e-01$	$2.9e-01$	$9.2e-03$	0.22627	0.2537	$8.0e-03$
1.2	$1.1e-01$	$1.0e-01$	$3.3e-03$	$2.4e-01$	$2.8e-01$	$8.8e-03$	0.20908	0.2501	$7.9e-03$
1.3	$2.3e-02$	$3.6e-02$	$1.1e-03$	$1.6e-01$	$2.4e-01$	$7.6e-03$	0.18683	0.2431	$7.7e-03$
1.4	$5.7e-03$	$1.2e-02$	$3.9e-04$	$9.6e-02$	$1.8e-01$	$5.8e-03$	0.16597	0.2367	$7.5e-03$
1.5	$1.2e-03$	$4.4e-03$	$1.4e-04$	$5.4e-02$	$1.4e-01$	$4.3e-03$	0.13793	0.2206	$7.0e-03$
1.6	$3.0e-04$	$1.2e-03$	$3.8e-05$	$2.6e-02$	$9.6e-02$	$3.0e-03$	0.12222	0.2226	$7.0e-03$
1.7	$1.0e-04$	$6.5e-04$	$2.1e-05$	$1.0e-02$	$4.8e-02$	$1.5e-03$	0.09408	0.1916	$6.1e-03$
1.8	$2.8e-05$	$1.9e-04$	$6.1e-06$	$8.5e-03$	$5.4e-02$	$1.7e-03$	0.07703	0.1727	$5.5e-03$
1.9	$5.3e-06$	$3.5e-05$	$1.1e-06$	$2.6e-03$	$1.7e-02$	$5.4e-04$	0.05662	0.1462	$4.6e-03$
2.0	$2.1e-06$	$2.0e-05$	$6.2e-07$	$5.5e-04$	$3.3e-03$	$1.0e-04$	0.04281	0.1204	$3.8e-03$
2.1	$7.8e-07$	$8.4e-06$	$2.7e-07$	$7.2e-04$	$6.6e-03$	$2.1e-04$	0.03577	0.1154	$3.6e-03$
2.2	$3.8e-07$	$4.6e-06$	$1.4e-07$	$2.1e-04$	$2.3e-03$	$7.3e-05$	0.03280	0.1113	$3.5e-03$
2.3	$7.9e-08$	$1.3e-06$	$4.0e-08$	$4.6e-05$	$4.3e-04$	$1.4e-05$	0.02187	0.0875	$2.8e-03$
2.4	$4.8e-08$	$5.9e-07$	$1.9e-08$	$3.5e-05$	$5.3e-04$	$1.7e-05$	0.02018	0.0924	$2.9e-03$
2.5	$2.3e-08$	$4.9e-07$	$1.5e-08$	$2.5e-05$	$3.4e-04$	$1.1e-05$	0.01351	0.0659	$2.1e-03$
2.6	$1.9e-09$	$2.6e-08$	$8.2e-10$	$4.1e-05$	$9.5e-04$	$3.0e-05$	0.00849	0.0505	$1.6e-03$
2.7	$1.4e-09$	$2.5e-08$	$8.0e-10$	$9.0e-07$	$1.1e-05$	$3.4e-07$	0.01025	0.0584	$1.8e-03$
2.8	$1.1e-09$	$2.5e-08$	$8.0e-10$	$2.4e-07$	$2.1e-06$	$6.7e-08$	0.00708	0.0427	$1.3e-03$
2.9	$2.3e-09$	$6.7e-08$	$2.1e-09$	$4.3e-07$	$7.5e-06$	$2.4e-07$	0.00446	0.0277	$8.7e-04$
3.0	$5.2e-11$	$9.2e-10$	$2.9e-11$	$5.5e-07$	$1.5e-05$	$4.9e-07$	0.00518	0.0403	$1.3e-03$
3.1	$9.5e-11$	$2.7e-09$	$8.4e-11$	$6.7e-08$	$1.3e-06$	$4.0e-08$	0.00403	0.0346	$1.1e-03$
3.2	$3.0e-10$	$8.2e-09$	$2.6e-10$	$6.7e-08$	$1.3e-06$	$4.2e-08$	0.00210	0.0243	$7.7e-04$
3.3	$4.2e-12$	$7.9e-11$	$2.5e-12$	$3.0e-09$	$6.8e-08$	$2.1e-09$	0.00248	0.0235	$7.4e-04$
3.4	$5.7e-13$	$6.8e-12$	$2.1e-13$	$1.1e-08$	$2.0e-07$	$6.3e-09$	0.00125	0.0120	$3.8e-04$
3.5	$1.1e-11$	$2.7e-10$	$8.5e-12$	$4.3e-09$	$7.6e-08$	$2.4e-09$	0.00273	0.0319	$1.0e-03$
3.6	$1.0e-13$	$9.0e-13$	$2.8e-14$	$2.4e-10$	$2.0e-09$	$6.4e-11$	0.00044	0.0048	$1.5e-04$
3.7	$8.0e-13$	$2.1e-11$	$6.8e-13$	$9.5e-11$	$1.3e-09$	$4.0e-11$	0.00060	0.0056	$1.8e-04$
3.8	$1.2e-13$	$2.2e-12$	$6.8e-14$	$4.7e-11$	$8.9e-10$	$2.8e-11$	0.00030	0.0033	$1.1e-04$
3.9	$2.1e-13$	$5.8e-12$	$1.8e-13$	$1.8e-10$	$3.1e-09$	$9.6e-11$	0.00026	0.0027	$8.5e-05$
4.0	$7.9e-14$	$1.6e-12$	$5.1e-14$	$3.8e-10$	$8.6e-09$	$2.7e-10$	0.00080	0.0091	$2.9e-04$
	Most Dissimilar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$9.4e-01$	$1.1e-01$	$3.3e-03$	$4.3e-01$	$3.2e-01$	$1.0e-02$	0.33729	0.3126	0.00989
1.1	$5.8e-01$	$2.4e-01$	$7.7e-03$	$3.7e-01$	$3.2e-01$	$1.0e-02$	0.34071	0.3207	0.01014
1.2	$1.6e-01$	$1.4e-01$	$4.3e-03$	$3.0e-01$	$3.1e-01$	$9.9e-03$	0.29171	0.3035	0.00960
1.3	$4.0e-02$	$5.7e-02$	$1.8e-03$	$2.1e-01$	$2.9e-01$	$9.1e-03$	0.25999	0.3059	0.00967
1.4	$1.0e-02$	$2.0e-02$	$6.4e-04$	$1.4e-01$	$2.4e-01$	$7.6e-03$	0.22888	0.2924	0.00925
1.5	$2.3e-03$	$7.1e-03$	$2.2e-04$	$8.6e-02$	$1.9e-01$	$5.9e-03$	0.18922	0.2777	0.00878
1.6	$4.5e-04$	$1.4e-03$	$4.3e-05$	$4.2e-02$	$1.2e-01$	$3.9e-03$	0.15220	0.2530	0.00800
1.7	$1.2e-04$	$6.2e-04$	$2.0e-05$	$2.3e-02$	$8.5e-02$	$2.7e-03$	0.12174	0.2255	0.00713
1.8	$4.2e-05$	$2.6e-04$	$8.3e-06$	$1.5e-02$	$7.4e-02$	$2.3e-03$	0.10023	0.2057	0.00651
1.9	$1.2e-05$	$1.4e-04$	$4.4e-06$	$7.5e-03$	$5.3e-02$	$1.7e-03$	0.06721	0.1645	0.00520
2.0	$3.5e-06$	$3.4e-05$	$1.1e-06$	$1.8e-03$	$1.2e-02$	$3.9e-04$	0.06149	0.1656	0.00524
2.1	$2.2e-06$	$3.2e-05$	$1.0e-06$	$2.3e-03$	$2.6e-02$	$8.2e-04$	0.04457	0.1355	0.00429
2.2	$2.6e-07$	$2.8e-06$	$8.7e-08$	$4.5e-04$	$4.8e-03$	$1.5e-04$	0.04352	0.1338	0.00423
2.3	$4.1e-08$	$4.1e-07$	$1.3e-08$	$1.8e-04$	$2.7e-03$	$8.6e-05$	0.03445	0.1252	0.00396
2.4	$2.3e-08$	$3.5e-07$	$1.1e-08$	$2.1e-04$	$4.2e-03$	$1.3e-04$	0.02798	0.1151	0.00364
2.5	$7.6e-09$	$1.3e-07$	$4.0e-09$	$4.5e-05$	$5.8e-04$	$1.8e-05$	0.01616	0.0839	0.00265
2.6	$1.2e-08$	$3.4e-07$	$1.1e-08$	$3.2e-05$	$4.1e-04$	$1.3e-05$	0.01336	0.0659	0.00208
2.7	$1.1e-09$	$1.5e-08$	$4.7e-10$	$2.9e-05$	$5.3e-04$	$1.7e-05$	0.01041	0.0514	0.00163
2.8	$3.0e-10$	$4.3e-09$	$1.4e-10$	$2.1e-06$	$3.4e-05$	$1.1e-06$	0.00840	0.0568	0.00180
2.9	$9.7e-11$	$1.6e-09$	$5.2e-11$	$8.8e-07$	$1.5e-05$	$4.6e-07$	0.00629	0.0485	0.00153
3.0	$4.7e-11$	$6.3e-10$	$2.0e-11$	$1.2e-07$	$1.2e-06$	$3.8e-08$	0.00662	0.0467	0.00148
3.1	$4.2e-11$	$1.1e-09$	$3.6e-11$	$2.2e-06$	$6.4e-05$	$2.0e-06$	0.00485	0.0333	0.00105
3.2	$2.0e-11$	$3.4e-10$	$1.1e-11$	$6.3e-07$	$1.4e-05$	$4.3e-07$	0.00236	0.0233	0.00074
3.3	$1.5e-11$	$3.3e-10$	$1.0e-11$	$3.2e-08$	$9.0e-07$	$2.9e-08$	0.00342	0.0316	0.00100
3.4	$1.2e-12$	$2.1e-11$	$6.7e-13$	$1.5e-08$	$2.1e-07$	$6.7e-09$	0.00221	0.0204	0.00064
3.5	$1.8e-13$	$2.0e-12$	$6.4e-14$	$3.0e-08$	$9.0e-07$	$2.8e-08$	0.00462	0.0548	0.00173
3.6	$1.6e-13$	$2.0e-12$	$6.3e-14$	$1.1e-08$	$3.3e-07$	$1.1e-08$	0.00059	0.0075	0.00024
3.7	$5.9e-15$	$6.7e-14$	$2.1e-15$	$1.3e-09$	$4.0e-08$	$1.3e-09$	0.00095	0.0116	0.00037
3.8	$2.2e-14$	$2.9e-13$	$9.2e-15$	$1.6e-10$	$2.1e-09$	$6.5e-11$	0.00041	0.0041	0.00013
3.9	$4.2e-14$	$9.9e-13$	$3.1e-14$	$1.8e-10$	$4.0e-09$	$1.3e-10$	0.00064	0.0094	0.00030
4.0	$7.0e-14$	$1.4e-12$	$4.4e-14$	$1.1e-10$	$1.7e-09$	$5.4e-11$	0.00141	0.0162	0.00051

Table B.5: MSND (above) and MDND (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in mean.

B.2 Changing Unstructured Variation (Section 6.4.2)

	Mean Square Difference								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$2.2e-06$	$4.6e-07$	$1.5e-08$	$2.2e-08$	$4.6e-09$	$1.4e-10$	$2.2e-10$	$5.2e-11$	$1.6e-12$
1.1	$3.3e-06$	$7.1e-07$	$2.3e-08$	$2.2e-08$	$4.5e-09$	$1.4e-10$	$2.2e-10$	$5.5e-11$	$1.7e-12$
1.2	$6.6e-06$	$1.5e-06$	$4.6e-08$	$2.2e-08$	$4.8e-09$	$1.5e-10$	$2.2e-10$	$5.4e-11$	$1.7e-12$
1.3	$1.2e-05$	$2.9e-06$	$9.1e-08$	$2.3e-08$	$4.9e-09$	$1.6e-10$	$2.2e-10$	$5.2e-11$	$1.6e-12$
1.4	$2.0e-05$	$4.9e-06$	$1.6e-07$	$2.4e-08$	$5.1e-09$	$1.6e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.5	$3.0e-05$	$7.7e-06$	$2.4e-07$	$2.5e-08$	$5.3e-09$	$1.7e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.6	$4.3e-05$	$1.1e-05$	$3.4e-07$	$2.6e-08$	$5.5e-09$	$1.7e-10$	$2.2e-10$	$5.2e-11$	$1.6e-12$
1.7	$5.8e-05$	$1.5e-05$	$4.8e-07$	$2.8e-08$	$6.0e-09$	$1.9e-10$	$2.2e-10$	$5.7e-11$	$1.8e-12$
1.8	$7.5e-05$	$2.0e-05$	$6.3e-07$	$2.9e-08$	$6.4e-09$	$2.0e-10$	$2.3e-10$	$5.3e-11$	$1.7e-12$
1.9	$9.5e-05$	$2.6e-05$	$8.2e-07$	$3.2e-08$	$7.1e-09$	$2.3e-10$	$2.3e-10$	$5.5e-11$	$1.8e-12$
2.0	$1.2e-04$	$3.2e-05$	$1.0e-06$	$3.4e-08$	$7.7e-09$	$2.4e-10$	$2.2e-10$	$5.6e-11$	$1.8e-12$
2.1	$1.4e-04$	$3.7e-05$	$1.2e-06$	$3.6e-08$	$8.4e-09$	$2.7e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
2.2	$1.7e-04$	$5.1e-05$	$1.6e-06$	$4.0e-08$	$8.9e-09$	$2.8e-10$	$2.3e-10$	$5.6e-11$	$1.8e-12$
2.3	$2.1e-04$	$6.0e-05$	$1.9e-06$	$4.3e-08$	$1.1e-08$	$3.3e-10$	$2.2e-10$	$5.6e-11$	$1.8e-12$
2.4	$2.4e-04$	$7.2e-05$	$2.3e-06$	$4.6e-08$	$1.2e-08$	$3.7e-10$	$2.2e-10$	$5.4e-11$	$1.7e-12$
2.5	$2.8e-04$	$8.3e-05$	$2.6e-06$	$5.0e-08$	$1.3e-08$	$4.0e-10$	$2.3e-10$	$5.4e-11$	$1.7e-12$
2.6	$3.2e-04$	$9.5e-05$	$3.0e-06$	$5.4e-08$	$1.4e-08$	$4.3e-10$	$2.3e-10$	$5.5e-11$	$1.7e-12$
2.7	$3.6e-04$	$1.1e-04$	$3.6e-06$	$5.9e-08$	$1.5e-08$	$4.9e-10$	$2.3e-10$	$5.6e-11$	$1.8e-12$
2.8	$4.1e-04$	$1.3e-04$	$4.1e-06$	$6.4e-08$	$1.7e-08$	$5.3e-10$	$2.3e-10$	$5.2e-11$	$1.6e-12$
2.9	$4.6e-04$	$1.5e-04$	$4.6e-06$	$6.8e-08$	$1.8e-08$	$5.8e-10$	$2.3e-10$	$5.5e-11$	$1.7e-12$
3.0	$5.2e-04$	$1.7e-04$	$5.3e-06$	$7.4e-08$	$2.2e-08$	$6.9e-10$	$2.3e-10$	$5.7e-11$	$1.8e-12$
3.1	$5.8e-04$	$2.0e-04$	$6.2e-06$	$8.0e-08$	$2.2e-08$	$7.1e-10$	$2.3e-10$	$5.5e-11$	$1.8e-12$
3.2	$6.4e-04$	$2.2e-04$	$6.8e-06$	$8.6e-08$	$2.6e-08$	$8.4e-10$	$2.3e-10$	$5.7e-11$	$1.8e-12$
3.3	$7.0e-04$	$2.4e-04$	$7.6e-06$	$9.4e-08$	$2.9e-08$	$9.0e-10$	$2.3e-10$	$6.1e-11$	$1.9e-12$
3.4	$8.0e-04$	$2.8e-04$	$8.9e-06$	$1.0e-07$	$3.1e-08$	$9.8e-10$	$2.3e-10$	$5.9e-11$	$1.9e-12$
3.5	$8.5e-04$	$3.1e-04$	$9.8e-06$	$1.1e-07$	$3.4e-08$	$1.1e-09$	$2.4e-10$	$6.0e-11$	$1.9e-12$
3.6	$9.6e-04$	$3.6e-04$	$1.1e-05$	$1.2e-07$	$3.7e-08$	$1.2e-09$	$2.4e-10$	$6.0e-11$	$1.9e-12$
3.7	$1.0e-03$	$3.9e-04$	$1.2e-05$	$1.3e-07$	$4.5e-08$	$1.4e-09$	$2.4e-10$	$6.1e-11$	$1.9e-12$
3.8	$1.1e-03$	$4.3e-04$	$1.4e-05$	$1.3e-07$	$4.7e-08$	$1.5e-09$	$2.4e-10$	$6.0e-11$	$1.9e-12$
3.9	$1.2e-03$	$5.3e-04$	$1.7e-05$	$1.4e-07$	$4.7e-08$	$1.5e-09$	$2.4e-10$	$5.8e-11$	$1.8e-12$
4.0	$1.3e-03$	$4.9e-04$	$1.5e-05$	$1.6e-07$	$5.3e-08$	$1.7e-09$	$2.4e-10$	$6.2e-11$	$2.0e-12$

Table B.6: MSD summary measures (mean, standard deviation (sd), standard error (se)), for change in unstructured variation.

	Peak-to Signal Noise Ratio								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	1806.5	759.46	24.016	34.2	11.60	0.367	10.7	3.62	0.114
1.1	1328.4	552.19	17.462	35.7	12.35	0.391	10.7	3.92	0.124
1.2	737.6	288.12	9.111	36.8	13.54	0.428	10.6	3.65	0.116
1.3	452.7	157.99	4.996	37.3	13.29	0.420	10.6	3.77	0.119
1.4	304.3	95.39	3.016	38.4	13.40	0.424	10.7	3.90	0.123
1.5	224.2	67.78	2.143	39.3	13.00	0.411	11.1	4.04	0.128
1.6	174.8	51.95	1.643	39.4	13.72	0.434	10.5	3.52	0.111
1.7	142.1	39.66	1.254	39.7	13.74	0.434	11.0	3.85	0.122
1.8	119.5	31.55	0.998	40.0	13.15	0.416	10.9	4.06	0.128
1.9	105.0	28.64	0.906	41.1	13.26	0.419	11.0	3.82	0.121
2.0	93.0	24.45	0.773	39.6	12.04	0.381	10.9	3.77	0.119
2.1	84.2	21.85	0.691	40.5	13.75	0.435	10.9	4.00	0.127
2.2	76.2	21.02	0.665	40.1	12.65	0.400	11.0	4.03	0.127
2.3	70.4	18.85	0.596	39.8	12.54	0.397	11.0	4.01	0.127
2.4	65.5	16.87	0.534	40.0	12.05	0.381	11.0	3.99	0.126
2.5	61.5	15.77	0.499	39.3	12.08	0.382	11.1	3.89	0.123
2.6	58.1	14.84	0.469	39.4	12.13	0.384	11.1	4.23	0.134
2.7	55.4	15.13	0.478	38.0	10.87	0.344	11.0	3.87	0.122
2.8	52.2	13.56	0.429	38.1	10.94	0.346	11.2	3.95	0.125
2.9	50.0	12.77	0.404	37.0	10.57	0.334	11.2	4.15	0.131
3.0	49.2	12.85	0.406	36.4	10.34	0.327	11.0	3.95	0.125
3.1	46.6	12.05	0.381	36.4	10.51	0.332	11.0	3.94	0.125
3.2	45.1	11.50	0.364	36.1	10.29	0.325	11.3	4.44	0.141
3.3	44.1	11.28	0.357	35.8	10.26	0.324	11.3	4.14	0.131
3.4	42.6	10.82	0.342	35.5	10.27	0.325	11.2	4.11	0.130
3.5	41.1	10.79	0.341	34.5	9.50	0.300	11.4	4.10	0.130
3.6	40.6	10.65	0.337	34.7	9.85	0.311	11.3	4.07	0.129
3.7	39.5	10.80	0.342	33.9	9.83	0.311	11.4	4.08	0.129
3.8	38.9	10.25	0.324	34.1	9.50	0.301	11.4	4.32	0.137
3.9	38.1	9.83	0.311	33.6	9.44	0.299	11.4	4.28	0.135
4.0	36.9	9.85	0.311	33.8	9.63	0.305	11.9	4.38	0.139
	Structural Similarity Index Measure								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
1.0	0.99	0.0023	$7.3e-05$	0.59	0.098	0.0031	0.012	0.14	0.0046
1.1	0.99	0.0030	$9.3e-05$	0.61	0.095	0.0030	0.015	0.14	0.0045
1.2	0.98	0.0044	$1.4e-04$	0.62	0.100	0.0031	0.012	0.14	0.0043
1.3	0.97	0.0069	$2.2e-04$	0.62	0.096	0.0030	0.020	0.14	0.0044
1.4	0.95	0.0098	$3.1e-04$	0.63	0.091	0.0029	0.021	0.14	0.0046
1.5	0.94	0.0129	$4.1e-04$	0.63	0.089	0.0028	0.025	0.14	0.0045
1.6	0.92	0.0158	$5.0e-04$	0.63	0.091	0.0029	0.013	0.14	0.0044
1.7	0.89	0.0192	$6.1e-04$	0.62	0.087	0.0028	0.026	0.14	0.0046
1.8	0.87	0.0209	$6.6e-04$	0.62	0.088	0.0028	0.016	0.14	0.0043
1.9	0.85	0.0248	$7.8e-04$	0.63	0.081	0.0026	0.029	0.14	0.0044
2.0	0.83	0.0275	$8.7e-04$	0.62	0.083	0.0026	0.017	0.15	0.0046
2.1	0.80	0.0297	$9.4e-04$	0.61	0.084	0.0027	0.024	0.14	0.0044
2.2	0.78	0.0317	$1.0e-03$	0.60	0.078	0.0025	0.024	0.14	0.0045
2.3	0.76	0.0333	$1.1e-03$	0.59	0.081	0.0025	0.028	0.14	0.0045
2.4	0.73	0.0363	$1.1e-03$	0.58	0.079	0.0025	0.027	0.14	0.0045
2.5	0.71	0.0390	$1.2e-03$	0.57	0.080	0.0025	0.028	0.14	0.0044
2.6	0.69	0.0367	$1.2e-03$	0.56	0.082	0.0026	0.027	0.14	0.0045
2.7	0.67	0.0396	$1.3e-03$	0.55	0.077	0.0024	0.025	0.14	0.0044
2.8	0.65	0.0414	$1.3e-03$	0.54	0.074	0.0023	0.036	0.14	0.0043
2.9	0.63	0.0412	$1.3e-03$	0.53	0.078	0.0025	0.032	0.14	0.0043
3.0	0.61	0.0429	$1.4e-03$	0.52	0.072	0.0023	0.028	0.14	0.0044
3.1	0.60	0.0443	$1.4e-03$	0.51	0.075	0.0024	0.027	0.13	0.0042
3.2	0.58	0.0445	$1.4e-03$	0.50	0.074	0.0023	0.045	0.14	0.0044
3.3	0.56	0.0437	$1.4e-03$	0.48	0.070	0.0022	0.037	0.14	0.0044
3.4	0.54	0.0462	$1.5e-03$	0.47	0.073	0.0023	0.034	0.14	0.0045
3.5	0.53	0.0465	$1.5e-03$	0.46	0.070	0.0022	0.040	0.14	0.0046
3.6	0.51	0.0462	$1.5e-03$	0.45	0.070	0.0022	0.038	0.14	0.0045
3.7	0.50	0.0466	$1.5e-03$	0.44	0.072	0.0023	0.040	0.15	0.0046
3.8	0.48	0.0465	$1.5e-03$	0.43	0.070	0.0022	0.038	0.14	0.0044
3.9	0.47	0.0468	$1.5e-03$	0.42	0.072	0.0023	0.037	0.14	0.0044
4.0	0.45	0.0478	$1.5e-03$	0.42	0.070	0.0022	0.048	0.14	0.0045

Table B.7: PSNRR (above) and SSIM (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in unstructured variation.

	Inter Region Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$8.7e-01$	$1.7e-01$	$5.4e-03$	$1.6e-01$	$2.6e-01$	$8.2e-03$	0.00588	0.0274	0.00087
1.1	$2.6e-01$	$2.3e-01$	$7.2e-03$	$1.7e-01$	$2.6e-01$	$8.2e-03$	0.00497	0.0253	0.00080
1.2	$1.9e-02$	$6.0e-02$	$1.9e-03$	$1.4e-01$	$2.5e-01$	$7.8e-03$	0.00506	0.0236	0.00075
1.3	$1.5e-03$	$1.6e-02$	$5.0e-04$	$1.2e-01$	$2.2e-01$	$7.1e-03$	0.00586	0.0258	0.00081
1.4	$1.6e-04$	$2.3e-03$	$7.2e-05$	$7.1e-02$	$1.9e-01$	$6.0e-03$	0.00638	0.0247	0.00078
1.5	$1.7e-05$	$3.5e-04$	$1.1e-05$	$4.9e-02$	$1.6e-01$	$5.0e-03$	0.00779	0.0347	0.00110
1.6	$1.4e-05$	$3.1e-04$	$9.7e-06$	$2.5e-02$	$1.1e-01$	$3.4e-03$	0.00587	0.0258	0.00082
1.7	$9.2e-08$	$1.7e-06$	$5.3e-08$	$1.2e-02$	$7.2e-02$	$2.3e-03$	0.00493	0.0242	0.00077
1.8	$4.2e-09$	$1.2e-07$	$3.7e-09$	$5.3e-03$	$5.4e-02$	$1.7e-03$	0.00593	0.0298	0.00094
1.9	$1.8e-15$	$3.9e-14$	$1.2e-15$	$2.6e-03$	$3.1e-02$	$9.7e-04$	0.00461	0.0214	0.00068
2.0	$9.3e-15$	$2.7e-13$	$8.4e-15$	$4.3e-04$	$5.1e-03$	$1.6e-04$	0.00685	0.0311	0.00098
2.1	$3.7e-16$	$1.2e-14$	$3.7e-16$	$7.1e-04$	$1.3e-02$	$4.1e-04$	0.00558	0.0258	0.00082
2.2	$2.6e-17$	$6.3e-16$	$2.0e-17$	$3.8e-06$	$5.5e-05$	$1.7e-06$	0.00612	0.0269	0.00085
2.3	$1.6e-13$	$5.0e-12$	$1.6e-13$	$6.0e-07$	$1.8e-05$	$5.8e-07$	0.00618	0.0266	0.00084
2.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$8.2e-06$	$2.6e-04$	$8.2e-06$	0.00636	0.0271	0.00086
2.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.4e-10$	$2.4e-09$	$7.7e-11$	0.00509	0.0237	0.00075
2.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$5.3e-07$	$1.6e-05$	$4.9e-07$	0.00482	0.0242	0.00077
2.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.4e-11$	$7.6e-10$	$2.4e-11$	0.00628	0.0284	0.00090
2.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$4.5e-13$	$1.4e-11$	$4.5e-13$	0.00660	0.0326	0.00103
2.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$6.7e-19$	$1.8e-17$	$5.7e-19$	0.00548	0.0251	0.00079
3.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.1e-16$	$6.5e-15$	$2.1e-16$	0.00744	0.0332	0.00105
3.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00735	0.0349	0.00110
3.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.1e-16$	$6.5e-15$	$2.1e-16$	0.00486	0.0236	0.00075
3.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00595	0.0280	0.00088
3.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.3e-14$	$1.0e-12$	$3.3e-14$	0.00552	0.0249	0.00079
3.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00655	0.0324	0.00102
3.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00591	0.0295	0.00093
3.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00597	0.0310	0.00098
3.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00448	0.0215	0.00068
3.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00481	0.0223	0.00070
4.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	$0.0e+00$	0.00400	0.0220	0.00070
	Most Similar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
1.0	0.4147	0.275	0.00870	0.3643	0.272	0.00859	0.53	0.30	0.0096
1.1	0.4068	0.276	0.00873	0.3771	0.279	0.00882	0.55	0.30	0.0096
1.2	0.3789	0.281	0.00889	0.3685	0.273	0.00865	0.55	0.30	0.0095
1.3	0.3514	0.276	0.00874	0.3674	0.280	0.00885	0.57	0.30	0.0094
1.4	0.3169	0.276	0.00873	0.3435	0.280	0.00886	0.52	0.30	0.0094
1.5	0.2896	0.270	0.00855	0.3286	0.270	0.00855	0.54	0.30	0.0094
1.6	0.2525	0.249	0.00786	0.2992	0.267	0.00844	0.55	0.31	0.0097
1.7	0.2084	0.243	0.00769	0.2900	0.268	0.00846	0.55	0.30	0.0094
1.8	0.1776	0.217	0.00685	0.2704	0.265	0.00838	0.55	0.30	0.0095
1.9	0.1521	0.202	0.00640	0.2502	0.259	0.00819	0.54	0.31	0.0097
2.0	0.1285	0.181	0.00572	0.2253	0.253	0.00799	0.54	0.30	0.0095
2.1	0.1075	0.167	0.00529	0.2081	0.236	0.00746	0.55	0.30	0.0094
2.2	0.0823	0.149	0.00471	0.1791	0.226	0.00714	0.55	0.30	0.0096
2.3	0.0678	0.126	0.00399	0.1613	0.214	0.00676	0.53	0.30	0.0096
2.4	0.0605	0.117	0.00368	0.1316	0.184	0.00582	0.55	0.30	0.0096
2.5	0.0547	0.117	0.00371	0.1209	0.184	0.00582	0.54	0.30	0.0096
2.6	0.0386	0.088	0.00278	0.1085	0.175	0.00552	0.54	0.30	0.0093
2.7	0.0370	0.088	0.00278	0.0936	0.163	0.00515	0.55	0.30	0.0096
2.8	0.0298	0.074	0.00235	0.0890	0.154	0.00485	0.54	0.30	0.0095
2.9	0.0259	0.076	0.00240	0.0696	0.129	0.00408	0.55	0.29	0.0093
3.0	0.0192	0.056	0.00178	0.0651	0.127	0.00402	0.54	0.30	0.0095
3.1	0.0182	0.056	0.00178	0.0569	0.116	0.00366	0.56	0.30	0.0096
3.2	0.0127	0.039	0.00124	0.0479	0.109	0.00344	0.53	0.29	0.0093
3.3	0.0097	0.030	0.00094	0.0438	0.106	0.00335	0.54	0.30	0.0094
3.4	0.0087	0.030	0.00095	0.0358	0.093	0.00295	0.54	0.30	0.0094
3.5	0.0097	0.045	0.00141	0.0303	0.078	0.00248	0.53	0.30	0.0094
3.6	0.0081	0.031	0.00097	0.0314	0.086	0.00273	0.51	0.30	0.0095
3.7	0.0065	0.036	0.00115	0.0252	0.068	0.00214	0.53	0.30	0.0096
3.8	0.0072	0.032	0.00101	0.0215	0.058	0.00185	0.54	0.30	0.0096
3.9	0.0059	0.034	0.00107	0.0212	0.061	0.00193	0.54	0.30	0.0094
4.0	0.0034	0.014	0.00045	0.0168	0.050	0.00158	0.54	0.31	0.0097

Table B.8: IRD (above) and MSDI (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in unstructured variation.

	Most Dissimilar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$7.5e-01$	$2.7e-01$	$8.6e-03$	$1.6e-01$	$2.6e-01$	$8.4e-03$	0.00735	0.069	0.00220
1.1	$3.4e-01$	$3.1e-01$	$9.8e-03$	$1.8e-01$	$2.8e-01$	$8.9e-03$	0.01213	0.096	0.00303
1.2	$8.2e-02$	$1.8e-01$	$5.7e-03$	$1.4e-01$	$2.4e-01$	$7.7e-03$	0.00846	0.076	0.00241
1.3	$2.2e-02$	$8.9e-02$	$2.8e-03$	$1.4e-01$	$2.5e-01$	$8.1e-03$	0.01463	0.106	0.00334
1.4	$6.8e-03$	$5.3e-02$	$1.7e-03$	$1.0e-01$	$2.2e-01$	$6.9e-03$	0.00912	0.080	0.00253
1.5	$2.7e-03$	$3.7e-02$	$1.2e-03$	$7.7e-02$	$1.9e-01$	$6.1e-03$	0.01241	0.094	0.00299
1.6	$4.1e-03$	$4.4e-02$	$1.4e-03$	$5.7e-02$	$1.7e-01$	$5.4e-03$	0.01171	0.093	0.00293
1.7	$1.1e-03$	$2.4e-02$	$7.7e-04$	$4.4e-02$	$1.5e-01$	$4.7e-03$	0.00704	0.066	0.00208
1.8	$3.8e-04$	$6.5e-03$	$2.0e-04$	$2.5e-02$	$1.0e-01$	$3.3e-03$	0.00937	0.087	0.00274
1.9	$1.7e-03$	$3.9e-02$	$1.2e-03$	$2.4e-02$	$1.1e-01$	$3.5e-03$	0.01283	0.101	0.00318
2.0	$3.0e-04$	$9.6e-03$	$3.0e-04$	$9.8e-03$	$6.6e-02$	$2.1e-03$	0.01174	0.093	0.00294
2.1	$2.7e-06$	$7.0e-05$	$2.2e-06$	$7.9e-03$	$6.3e-02$	$2.0e-03$	0.00677	0.066	0.00209
2.2	$1.4e-04$	$4.2e-03$	$1.3e-04$	$3.4e-03$	$4.1e-02$	$1.3e-03$	0.01093	0.085	0.00269
2.3	$1.2e-07$	$2.1e-06$	$6.6e-08$	$2.3e-03$	$3.1e-02$	$9.9e-04$	0.00883	0.077	0.00244
2.4	$3.0e-04$	$9.6e-03$	$3.0e-04$	$2.0e-03$	$3.5e-02$	$1.1e-03$	0.00847	0.074	0.00235
2.5	$4.4e-06$	$1.4e-04$	$4.4e-06$	$9.2e-04$	$1.6e-02$	$5.0e-04$	0.01110	0.086	0.00272
2.6	$1.6e-05$	$5.1e-04$	$1.6e-05$	$1.4e-04$	$2.7e-03$	$8.5e-05$	0.00690	0.063	0.00198
2.7	$5.8e-09$	$1.8e-07$	$5.8e-09$	$3.2e-04$	$9.6e-03$	$3.0e-04$	0.01333	0.100	0.00315
2.8	$2.8e-10$	$8.2e-09$	$2.6e-10$	$2.2e-04$	$6.4e-03$	$2.0e-04$	0.00932	0.083	0.00262
2.9	$8.8e-12$	$2.6e-10$	$8.2e-12$	$9.8e-04$	$2.4e-02$	$7.7e-04$	0.00920	0.076	0.00240
3.0	$8.4e-16$	$2.2e-14$	$7.1e-16$	$5.8e-04$	$1.8e-02$	$5.8e-04$	0.01074	0.089	0.00281
3.1	$3.1e-16$	$6.9e-15$	$2.2e-16$	$7.4e-09$	$1.8e-07$	$5.8e-09$	0.00417	0.047	0.00149
3.2	$4.8e-14$	$1.5e-12$	$4.8e-14$	$2.9e-05$	$9.1e-04$	$2.9e-05$	0.00847	0.072	0.00229
3.3	$2.9e-15$	$9.3e-14$	$2.9e-15$	$1.1e-10$	$2.8e-09$	$8.8e-11$	0.00773	0.077	0.00243
3.4	$4.1e-17$	$1.1e-15$	$3.4e-17$	$8.2e-09$	$1.9e-07$	$6.2e-09$	0.01176	0.091	0.00286
3.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.4e-10$	$8.5e-09$	$2.7e-10$	0.00928	0.075	0.00237
3.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.7e-11$	$8.4e-10$	$2.7e-11$	0.00847	0.077	0.00242
3.7	$7.0e-18$	$2.2e-16$	$7.0e-18$	$1.6e-11$	$3.6e-10$	$1.1e-11$	0.01227	0.095	0.00301
3.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.2e-06$	$7.0e-05$	$2.2e-06$	0.00943	0.084	0.00265
3.9	$6.7e-13$	$2.1e-11$	$6.7e-13$	$9.3e-08$	$2.9e-06$	$9.3e-08$	0.00419	0.054	0.00171
4.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.1e-11$	$2.7e-10$	$8.5e-12$	0.01080	0.091	0.00288
	Average Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
1.0	0.61111	0.2754	$8.7e-03$	0.3916	0.284	0.00897	0.45	0.32	0.0100
1.1	0.58282	0.2711	$8.6e-03$	0.3813	0.282	0.00893	0.45	0.31	0.0098
1.2	0.50018	0.2697	$8.5e-03$	0.3842	0.284	0.00898	0.46	0.31	0.0098
1.3	0.40554	0.2578	$8.2e-03$	0.3623	0.289	0.00913	0.46	0.32	0.0102
1.4	0.31220	0.2464	$7.8e-03$	0.3372	0.280	0.00885	0.46	0.32	0.0101
1.5	0.23853	0.2142	$6.8e-03$	0.3265	0.277	0.00876	0.45	0.31	0.0097
1.6	0.17341	0.1807	$5.7e-03$	0.2983	0.269	0.00851	0.47	0.32	0.0102
1.7	0.14235	0.1678	$5.3e-03$	0.2542	0.260	0.00823	0.44	0.31	0.0098
1.8	0.09994	0.1305	$4.1e-03$	0.2530	0.268	0.00846	0.45	0.32	0.0100
1.9	0.07309	0.1051	$3.3e-03$	0.2129	0.244	0.00771	0.45	0.31	0.0099
2.0	0.06240	0.0969	$3.1e-03$	0.1849	0.225	0.00711	0.45	0.32	0.0102
2.1	0.04271	0.0715	$2.3e-03$	0.1697	0.216	0.00683	0.46	0.32	0.0101
2.2	0.03520	0.0653	$2.1e-03$	0.1542	0.215	0.00681	0.43	0.31	0.0099
2.3	0.02686	0.0620	$2.0e-03$	0.1284	0.187	0.00591	0.45	0.32	0.0101
2.4	0.02005	0.0449	$1.4e-03$	0.1112	0.181	0.00573	0.46	0.31	0.0097
2.5	0.01530	0.0357	$1.1e-03$	0.0863	0.149	0.00473	0.44	0.31	0.0098
2.6	0.01101	0.0280	$8.9e-04$	0.0758	0.135	0.00427	0.44	0.31	0.0097
2.7	0.00845	0.0191	$6.0e-04$	0.0627	0.119	0.00377	0.44	0.31	0.0098
2.8	0.00815	0.0222	$7.0e-04$	0.0575	0.114	0.00360	0.45	0.32	0.0100
2.9	0.00604	0.0243	$7.7e-04$	0.0530	0.106	0.00336	0.46	0.31	0.0098
3.0	0.00380	0.0127	$4.0e-04$	0.0425	0.092	0.00292	0.45	0.32	0.0100
3.1	0.00360	0.0143	$4.5e-04$	0.0325	0.083	0.00263	0.44	0.31	0.0098
3.2	0.00234	0.0076	$2.4e-04$	0.0308	0.080	0.00254	0.43	0.31	0.0098
3.3	0.00222	0.0095	$3.0e-04$	0.0232	0.064	0.00203	0.46	0.31	0.0098
3.4	0.00202	0.0086	$2.7e-04$	0.0272	0.068	0.00214	0.44	0.30	0.0096
3.5	0.00123	0.0051	$1.6e-04$	0.0181	0.054	0.00172	0.44	0.32	0.0100
3.6	0.00147	0.0073	$2.3e-04$	0.0181	0.051	0.00160	0.44	0.31	0.0097
3.7	0.00099	0.0042	$1.3e-04$	0.0120	0.037	0.00117	0.44	0.32	0.0100
3.8	0.00089	0.0040	$1.3e-04$	0.0128	0.044	0.00139	0.42	0.31	0.0097
3.9	0.00049	0.0019	$5.9e-05$	0.0099	0.039	0.00123	0.45	0.31	0.0099
4.0	0.00040	0.0025	$7.8e-05$	0.0113	0.040	0.00128	0.43	0.31	0.0097

Table B.9: MDD (above) and AVND (below)Summary measures (mean, standard deviation (sd), standard error (se)), for change in unstructured variation.

	Most Similar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$8.9e-01$	$1.5e-01$	$4.6e-03$	$3.3e-01$	$2.9e-01$	$9.2e-03$	0.229	0.25	0.0079
1.1	$5.5e-01$	$2.5e-01$	$7.9e-03$	$3.3e-01$	$2.9e-01$	$9.3e-03$	0.220	0.25	0.0079
1.2	$1.9e-01$	$1.6e-01$	$5.1e-03$	$3.0e-01$	$2.9e-01$	$9.2e-03$	0.224	0.25	0.0080
1.3	$5.5e-02$	$7.2e-02$	$2.3e-03$	$2.6e-01$	$2.9e-01$	$9.1e-03$	0.213	0.24	0.0075
1.4	$1.6e-02$	$2.8e-02$	$8.8e-04$	$1.8e-01$	$2.5e-01$	$7.9e-03$	0.220	0.24	0.0077
1.5	$5.1e-03$	$1.3e-02$	$4.2e-04$	$1.4e-01$	$2.3e-01$	$7.2e-03$	0.243	0.26	0.0083
1.6	$1.4e-03$	$4.7e-03$	$1.5e-04$	$9.3e-02$	$1.7e-01$	$5.5e-03$	0.237	0.26	0.0082
1.7	$4.9e-04$	$1.6e-03$	$5.0e-05$	$6.5e-02$	$1.5e-01$	$4.9e-03$	0.212	0.24	0.0077
1.8	$1.8e-04$	$1.1e-03$	$3.6e-05$	$4.3e-02$	$1.2e-01$	$3.7e-03$	0.216	0.25	0.0080
1.9	$6.2e-05$	$3.8e-04$	$1.2e-05$	$2.0e-02$	$6.6e-02$	$2.1e-03$	0.220	0.25	0.0078
2.0	$2.6e-05$	$1.5e-04$	$4.8e-06$	$1.4e-02$	$6.4e-02$	$2.0e-03$	0.234	0.26	0.0082
2.1	$6.8e-06$	$4.9e-05$	$1.5e-06$	$9.1e-03$	$5.7e-02$	$1.8e-03$	0.223	0.25	0.0079
2.2	$1.9e-06$	$1.3e-05$	$4.2e-07$	$5.7e-03$	$4.3e-02$	$1.4e-03$	0.226	0.26	0.0084
2.3	$6.6e-07$	$5.7e-06$	$1.8e-07$	$2.4e-03$	$2.5e-02$	$7.8e-04$	0.224	0.25	0.0079
2.4	$1.7e-06$	$3.4e-05$	$1.1e-06$	$1.4e-03$	$2.5e-02$	$7.8e-04$	0.224	0.25	0.0079
2.5	$4.1e-07$	$7.1e-06$	$2.2e-07$	$6.8e-04$	$1.0e-02$	$3.2e-04$	0.233	0.26	0.0081
2.6	$6.6e-08$	$7.3e-07$	$2.3e-08$	$5.4e-04$	$5.9e-03$	$1.9e-04$	0.215	0.25	0.0078
2.7	$6.4e-08$	$8.6e-07$	$2.7e-08$	$8.2e-05$	$1.1e-03$	$3.4e-05$	0.228	0.26	0.0082
2.8	$3.2e-08$	$5.5e-07$	$1.7e-08$	$4.2e-05$	$3.8e-04$	$1.2e-05$	0.226	0.25	0.0080
2.9	$4.5e-09$	$7.3e-08$	$2.3e-09$	$8.1e-05$	$1.3e-03$	$4.1e-05$	0.216	0.26	0.0081
3.0	$2.0e-08$	$2.8e-07$	$9.0e-09$	$2.0e-05$	$5.1e-04$	$1.6e-05$	0.229	0.26	0.0082
3.1	$1.3e-09$	$2.4e-08$	$7.7e-10$	$1.6e-04$	$4.2e-03$	$1.3e-04$	0.216	0.25	0.0079
3.2	$1.1e-09$	$2.4e-08$	$7.6e-10$	$4.2e-06$	$7.2e-05$	$2.3e-06$	0.216	0.25	0.0078
3.3	$3.3e-10$	$4.1e-09$	$1.3e-10$	$4.5e-06$	$8.0e-05$	$2.5e-06$	0.214	0.25	0.0078
3.4	$9.1e-10$	$2.4e-08$	$7.6e-10$	$1.8e-06$	$3.5e-05$	$1.1e-06$	0.225	0.25	0.0080
3.5	$7.8e-10$	$2.4e-08$	$7.5e-10$	$2.4e-07$	$3.5e-06$	$1.1e-07$	0.232	0.26	0.0081
3.6	$4.8e-11$	$8.9e-10$	$2.8e-11$	$2.3e-06$	$7.0e-05$	$2.2e-06$	0.199	0.25	0.0078
3.7	$2.9e-10$	$8.2e-09$	$2.6e-10$	$1.7e-07$	$3.1e-06$	$9.7e-08$	0.208	0.25	0.0079
3.8	$1.5e-11$	$2.8e-10$	$8.9e-12$	$6.0e-08$	$1.2e-06$	$3.9e-08$	0.215	0.25	0.0080
3.9	$3.2e-11$	$8.5e-10$	$2.7e-11$	$5.4e-08$	$7.3e-07$	$2.3e-08$	0.210	0.25	0.0080
4.0	$5.7e-12$	$8.5e-11$	$2.7e-12$	$1.0e-07$	$2.9e-06$	$9.3e-08$	0.194	0.25	0.0078
	Most Dissimilar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
1.0	$9.5e-01$	$9.5e-02$	$3.0e-03$	$4.2e-01$	$3.3e-01$	$1.0e-02$	0.34	0.31	0.0099
1.1	$6.7e-01$	$2.4e-01$	$7.5e-03$	$4.1e-01$	$3.3e-01$	$1.0e-02$	0.35	0.32	0.0100
1.2	$2.8e-01$	$2.0e-01$	$6.2e-03$	$3.6e-01$	$3.2e-01$	$1.0e-02$	0.35	0.33	0.0103
1.3	$9.5e-02$	$1.0e-01$	$3.2e-03$	$3.3e-01$	$3.2e-01$	$1.0e-02$	0.35	0.32	0.0100
1.4	$2.9e-02$	$4.5e-02$	$1.4e-03$	$2.5e-01$	$2.9e-01$	$9.3e-03$	0.35	0.32	0.0101
1.5	$8.6e-03$	$1.8e-02$	$5.8e-04$	$2.0e-01$	$2.7e-01$	$8.6e-03$	0.34	0.31	0.0099
1.6	$2.5e-03$	$6.7e-03$	$2.1e-04$	$1.4e-01$	$2.4e-01$	$7.5e-03$	0.36	0.32	0.0101
1.7	$9.3e-04$	$3.4e-03$	$1.1e-04$	$9.6e-02$	$1.9e-01$	$6.1e-03$	0.33	0.31	0.0099
1.8	$3.1e-04$	$1.3e-03$	$4.1e-05$	$7.7e-02$	$1.8e-01$	$5.6e-03$	0.33	0.32	0.0100
1.9	$1.3e-04$	$8.2e-04$	$2.6e-05$	$3.6e-02$	$1.1e-01$	$3.4e-03$	0.34	0.31	0.0098
2.0	$4.0e-05$	$3.2e-04$	$1.0e-05$	$2.7e-02$	$9.1e-02$	$2.9e-03$	0.34	0.32	0.0100
2.1	$1.2e-05$	$9.3e-05$	$2.9e-06$	$1.5e-02$	$6.2e-02$	$2.0e-03$	0.33	0.31	0.0099
2.2	$1.1e-05$	$2.6e-04$	$8.2e-06$	$9.7e-03$	$4.9e-02$	$1.5e-03$	0.34	0.32	0.0101
2.3	$8.1e-07$	$4.8e-06$	$1.5e-07$	$3.9e-03$	$3.1e-02$	$9.7e-04$	0.35	0.32	0.0101
2.4	$4.0e-07$	$3.3e-06$	$1.0e-07$	$2.0e-03$	$2.0e-02$	$6.2e-04$	0.35	0.31	0.0099
2.5	$4.8e-07$	$8.0e-06$	$2.5e-07$	$1.7e-03$	$1.8e-02$	$5.8e-04$	0.35	0.31	0.0098
2.6	$1.0e-07$	$1.3e-06$	$4.2e-08$	$4.6e-04$	$3.6e-03$	$1.1e-04$	0.32	0.30	0.0096
2.7	$2.3e-08$	$2.2e-07$	$7.1e-09$	$3.1e-04$	$6.5e-03$	$2.1e-04$	0.33	0.32	0.0101
2.8	$2.6e-08$	$3.8e-07$	$1.2e-08$	$1.6e-04$	$2.0e-03$	$6.2e-05$	0.34	0.32	0.0101
2.9	$3.6e-09$	$4.7e-08$	$1.5e-09$	$1.4e-04$	$1.9e-03$	$6.0e-05$	0.33	0.31	0.0098
3.0	$5.2e-09$	$7.1e-08$	$2.2e-09$	$2.3e-05$	$3.0e-04$	$9.6e-06$	0.35	0.32	0.0103
3.1	$1.0e-09$	$9.9e-09$	$3.1e-10$	$3.2e-05$	$4.6e-04$	$1.4e-05$	0.33	0.31	0.0098
3.2	$1.1e-09$	$1.9e-08$	$5.9e-10$	$1.3e-05$	$2.6e-04$	$8.3e-06$	0.32	0.31	0.0097
3.3	$2.4e-10$	$4.1e-09$	$1.3e-10$	$6.9e-06$	$1.4e-04$	$4.4e-06$	0.34	0.32	0.0101
3.4	$3.2e-10$	$4.6e-09$	$1.4e-10$	$1.3e-05$	$1.8e-04$	$5.5e-06$	0.34	0.32	0.0100
3.5	$1.1e-10$	$1.7e-09$	$5.3e-11$	$1.7e-06$	$3.3e-05$	$1.0e-06$	0.34	0.31	0.0099
3.6	$1.7e-10$	$4.1e-09$	$1.3e-10$	$9.6e-07$	$1.4e-05$	$4.5e-07$	0.32	0.32	0.0100
3.7	$3.8e-11$	$1.1e-09$	$3.6e-11$	$2.5e-07$	$5.8e-06$	$1.8e-07$	0.34	0.32	0.0101
3.8	$3.5e-12$	$8.1e-11$	$2.6e-12$	$3.0e-07$	$5.9e-06$	$1.9e-07$	0.33	0.32	0.0100
3.9	$7.7e-12$	$1.2e-10$	$3.8e-12$	$2.3e-07$	$3.4e-06$	$1.1e-07$	0.32	0.32	0.0101
4.0	$6.2e-12$	$1.2e-10$	$3.6e-12$	$9.9e-08$	$1.4e-06$	$4.3e-08$	0.29	0.30	0.0096

Table B.10: MSND (above) and MDND (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in unstructured variation.

B.3 Changing Structured Variation (Section 6.4.2)

	Mean Square Difference								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$2.2e-06$	$4.7e-07$	$1.5e-08$	$2.2e-08$	$4.7e-09$	$1.5e-10$	$2.2e-10$	$5.1e-11$	$1.6e-12$
1.1	$2.9e-06$	$7.3e-07$	$2.3e-08$	$2.2e-08$	$4.7e-09$	$1.5e-10$	$2.2e-10$	$5.2e-11$	$1.7e-12$
1.2	$5.0e-06$	$1.9e-06$	$6.0e-08$	$2.2e-08$	$4.9e-09$	$1.6e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.3	$8.5e-06$	$3.8e-06$	$1.2e-07$	$2.3e-08$	$4.9e-09$	$1.5e-10$	$2.2e-10$	$5.5e-11$	$1.7e-12$
1.4	$1.4e-05$	$8.1e-06$	$2.6e-07$	$2.4e-08$	$5.2e-09$	$1.6e-10$	$2.2e-10$	$5.4e-11$	$1.7e-12$
1.5	$2.0e-05$	$1.1e-05$	$3.3e-07$	$2.4e-08$	$5.4e-09$	$1.7e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.6	$2.7e-05$	$1.4e-05$	$4.4e-07$	$2.5e-08$	$5.4e-09$	$1.7e-10$	$2.2e-10$	$5.2e-11$	$1.7e-12$
1.7	$3.7e-05$	$2.2e-05$	$6.9e-07$	$2.6e-08$	$5.9e-09$	$1.9e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.8	$4.9e-05$	$3.0e-05$	$9.6e-07$	$2.7e-08$	$6.4e-09$	$2.0e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
1.9	$6.2e-05$	$3.6e-05$	$1.1e-06$	$2.9e-08$	$7.9e-09$	$2.5e-10$	$2.2e-10$	$5.3e-11$	$1.7e-12$
2.0	$7.6e-05$	$5.1e-05$	$1.6e-06$	$3.0e-08$	$7.7e-09$	$2.4e-10$	$2.2e-10$	$5.5e-11$	$1.7e-12$
2.1	$9.3e-05$	$6.0e-05$	$1.9e-06$	$3.1e-08$	$9.0e-09$	$2.9e-10$	$2.3e-10$	$5.5e-11$	$1.7e-12$
2.2	$1.1e-04$	$8.4e-05$	$2.7e-06$	$3.3e-08$	$9.7e-09$	$3.1e-10$	$2.2e-10$	$5.6e-11$	$1.8e-12$
2.3	$1.3e-04$	$8.6e-05$	$2.7e-06$	$3.5e-08$	$1.2e-08$	$3.7e-10$	$2.2e-10$	$5.5e-11$	$1.7e-12$
2.4	$1.5e-04$	$1.1e-04$	$3.6e-06$	$3.7e-08$	$1.5e-08$	$4.9e-10$	$2.3e-10$	$5.5e-11$	$1.7e-12$
2.5	$1.8e-04$	$1.3e-04$	$4.2e-06$	$4.1e-08$	$1.6e-08$	$5.2e-10$	$2.2e-10$	$5.6e-11$	$1.8e-12$
2.6	$2.1e-04$	$1.6e-04$	$5.1e-06$	$4.2e-08$	$1.5e-08$	$4.9e-10$	$2.3e-10$	$5.7e-11$	$1.8e-12$
2.7	$2.3e-04$	$1.7e-04$	$5.2e-06$	$4.6e-08$	$2.4e-08$	$7.7e-10$	$2.2e-10$	$5.4e-11$	$1.7e-12$
2.8	$2.6e-04$	$2.2e-04$	$6.8e-06$	$4.8e-08$	$2.4e-08$	$7.5e-10$	$2.3e-10$	$5.6e-11$	$1.8e-12$
2.9	$3.1e-04$	$2.7e-04$	$8.7e-06$	$5.2e-08$	$2.5e-08$	$7.8e-10$	$2.3e-10$	$5.7e-11$	$1.8e-12$
3.0	$3.3e-04$	$2.7e-04$	$8.5e-06$	$5.7e-08$	$3.7e-08$	$1.2e-09$	$2.3e-10$	$5.6e-11$	$1.8e-12$
3.1	$3.8e-04$	$3.8e-04$	$1.2e-05$	$6.0e-08$	$3.4e-08$	$1.1e-09$	$2.3e-10$	$5.5e-11$	$1.7e-12$
3.2	$4.1e-04$	$3.2e-04$	$1.0e-05$	$6.4e-08$	$4.1e-08$	$1.3e-09$	$2.3e-10$	$5.7e-11$	$1.8e-12$
3.3	$4.4e-04$	$3.6e-04$	$1.1e-05$	$7.0e-08$	$5.1e-08$	$1.6e-09$	$2.3e-10$	$5.5e-11$	$1.8e-12$
3.4	$5.0e-04$	$4.4e-04$	$1.4e-05$	$7.4e-08$	$4.8e-08$	$1.5e-09$	$2.3e-10$	$5.9e-11$	$1.9e-12$
3.5	$5.6e-04$	$6.7e-04$	$2.1e-05$	$7.5e-08$	$4.2e-08$	$1.3e-09$	$2.3e-10$	$5.6e-11$	$1.8e-12$
3.6	$6.1e-04$	$6.5e-04$	$2.1e-05$	$8.0e-08$	$5.6e-08$	$1.8e-09$	$2.3e-10$	$5.4e-11$	$1.7e-12$
3.7	$6.4e-04$	$5.7e-04$	$1.8e-05$	$9.7e-08$	$1.3e-07$	$4.0e-09$	$2.3e-10$	$5.8e-11$	$1.8e-12$
3.8	$7.4e-04$	$7.1e-04$	$2.3e-05$	$9.2e-08$	$6.3e-08$	$2.0e-09$	$2.3e-10$	$5.9e-11$	$1.9e-12$
3.9	$8.2e-04$	$1.0e-03$	$3.2e-05$	$1.1e-07$	$1.2e-07$	$3.8e-09$	$2.3e-10$	$5.9e-11$	$1.9e-12$
4.0	$8.7e-04$	$8.1e-04$	$2.6e-05$	$1.1e-07$	$1.1e-07$	$3.4e-09$	$2.3e-10$	$6.0e-11$	$1.9e-12$

Table B.11: MSD summary measures (mean, standard deviation (sd), standard error (se)), for change in structured variation.

	Peak-to Signal Noise Ratio								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	1848.4	894.9	28.299	34.5	12.46	0.394	10.6	3.75	0.119
1.1	1446.1	552.6	17.476	35.8	12.39	0.392	10.6	3.74	0.118
1.2	916.9	320.7	10.140	36.0	12.47	0.394	10.6	3.48	0.110
1.3	593.0	202.7	6.410	35.8	12.12	0.383	10.8	3.72	0.118
1.4	417.2	139.6	4.413	37.2	13.78	0.436	11.0	4.21	0.133
1.5	310.1	108.1	3.420	37.7	14.05	0.444	10.8	3.96	0.125
1.6	240.4	78.9	2.494	38.3	14.14	0.447	10.8	3.89	0.123
1.7	192.6	64.2	2.032	37.0	12.77	0.404	10.8	3.83	0.121
1.8	161.7	49.5	1.566	38.7	14.11	0.446	10.8	3.95	0.125
1.9	138.7	44.0	1.391	38.1	13.58	0.429	10.9	3.99	0.126
2.0	122.3	38.2	1.207	38.5	12.97	0.410	10.8	3.88	0.123
2.1	111.2	36.6	1.158	39.5	14.34	0.454	10.6	3.77	0.119
2.2	99.8	30.7	0.970	38.5	13.11	0.415	11.0	3.99	0.126
2.3	89.3	26.0	0.823	38.8	13.60	0.430	10.7	3.76	0.119
2.4	84.6	25.6	0.811	39.3	13.35	0.422	10.8	3.85	0.122
2.5	77.3	23.0	0.728	37.9	13.16	0.416	10.9	3.95	0.125
2.6	73.3	22.4	0.709	37.8	12.87	0.407	10.9	3.79	0.120
2.7	68.5	20.9	0.662	38.6	13.78	0.436	10.9	3.76	0.119
2.8	64.2	18.8	0.595	38.1	12.78	0.404	10.8	3.95	0.125
2.9	61.8	18.1	0.572	37.7	12.88	0.407	10.9	3.94	0.125
3.0	58.2	16.5	0.523	37.0	11.96	0.378	10.9	3.71	0.117
3.1	56.4	15.6	0.493	38.0	12.35	0.391	10.8	3.80	0.120
3.2	54.6	15.9	0.504	36.6	11.90	0.376	11.2	4.11	0.130
3.3	52.7	15.7	0.496	37.3	12.32	0.390	11.3	4.18	0.132
3.4	51.0	14.6	0.462	36.8	11.99	0.379	11.2	3.94	0.125
3.5	48.8	13.6	0.430	35.9	11.23	0.355	11.1	4.12	0.130
3.6	46.8	13.3	0.422	36.1	11.60	0.367	11.2	4.23	0.134
3.7	46.0	13.3	0.422	36.0	11.29	0.357	11.2	4.25	0.134
3.8	44.9	12.9	0.406	35.1	10.86	0.343	11.1	4.08	0.129
3.9	43.6	12.5	0.396	35.4	11.95	0.378	11.1	4.10	0.130
4.0	43.6	12.4	0.391	34.6	10.42	0.330	11.1	4.24	0.134
	Structural Similarity Index Measure								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	0.99	0.0024	$7.7e-05$	0.59	0.103	0.0033	0.022	0.14	0.0045
1.1	0.99	0.0024	$7.7e-05$	0.61	0.101	0.0032	0.015	0.14	0.0043
1.2	0.99	0.0033	$1.1e-04$	0.61	0.102	0.0032	0.017	0.14	0.0044
1.3	0.98	0.0059	$1.9e-04$	0.61	0.096	0.0030	0.019	0.14	0.0045
1.4	0.97	0.0089	$2.8e-04$	0.62	0.096	0.0031	0.020	0.14	0.0045
1.5	0.95	0.0129	$4.1e-04$	0.62	0.100	0.0032	0.017	0.15	0.0046
1.6	0.94	0.0156	$4.9e-04$	0.62	0.096	0.0030	0.019	0.14	0.0045
1.7	0.92	0.0206	$6.5e-04$	0.61	0.098	0.0031	0.017	0.14	0.0043
1.8	0.91	0.0243	$7.7e-04$	0.61	0.092	0.0029	0.019	0.14	0.0044
1.9	0.89	0.0289	$9.1e-04$	0.61	0.095	0.0030	0.024	0.14	0.0044
2.0	0.87	0.0336	$1.1e-03$	0.60	0.095	0.0030	0.018	0.14	0.0044
2.1	0.85	0.0372	$1.2e-03$	0.60	0.089	0.0028	0.014	0.14	0.0044
2.2	0.83	0.0424	$1.3e-03$	0.59	0.086	0.0027	0.031	0.14	0.0045
2.3	0.81	0.0425	$1.3e-03$	0.58	0.089	0.0028	0.026	0.13	0.0042
2.4	0.79	0.0471	$1.5e-03$	0.58	0.088	0.0028	0.022	0.14	0.0044
2.5	0.77	0.0504	$1.6e-03$	0.57	0.086	0.0027	0.026	0.14	0.0046
2.6	0.75	0.0533	$1.7e-03$	0.57	0.083	0.0026	0.021	0.14	0.0044
2.7	0.73	0.0561	$1.8e-03$	0.56	0.083	0.0026	0.022	0.14	0.0045
2.8	0.71	0.0601	$1.9e-03$	0.55	0.085	0.0027	0.022	0.14	0.0046
2.9	0.69	0.0626	$2.0e-03$	0.54	0.083	0.0026	0.023	0.14	0.0045
3.0	0.68	0.0627	$2.0e-03$	0.53	0.085	0.0027	0.027	0.14	0.0046
3.1	0.66	0.0655	$2.1e-03$	0.53	0.086	0.0027	0.019	0.14	0.0044
3.2	0.64	0.0656	$2.1e-03$	0.51	0.083	0.0026	0.032	0.14	0.0045
3.3	0.63	0.0651	$2.1e-03$	0.50	0.085	0.0027	0.031	0.14	0.0045
3.4	0.61	0.0671	$2.1e-03$	0.49	0.083	0.0026	0.028	0.14	0.0044
3.5	0.60	0.0679	$2.1e-03$	0.48	0.082	0.0026	0.029	0.15	0.0047
3.6	0.58	0.0717	$2.3e-03$	0.48	0.087	0.0028	0.027	0.14	0.0045
3.7	0.56	0.0703	$2.2e-03$	0.47	0.084	0.0026	0.026	0.15	0.0046
3.8	0.55	0.0711	$2.2e-03$	0.46	0.083	0.0026	0.030	0.14	0.0044
3.9	0.53	0.0749	$2.4e-03$	0.45	0.089	0.0028	0.027	0.14	0.0045
4.0	0.52	0.0732	$2.3e-03$	0.44	0.084	0.0027	0.028	0.14	0.0044

Table B.12: PSNRR (above) and SSIM (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in structured variation.

	Inter Region Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$8.7e-01$	$1.7e-01$	$5.3e-03$	$1.8e-01$	$2.7e-01$	$8.6e-03$	0.0067	0.031	0.00099
1.1	$5.6e-01$	$3.0e-01$	$9.4e-03$	$1.7e-01$	$2.7e-01$	$8.6e-03$	0.0065	0.030	0.00096
1.2	$1.8e-01$	$2.2e-01$	$6.9e-03$	$1.7e-01$	$2.7e-01$	$8.5e-03$	0.0075	0.041	0.00129
1.3	$4.2e-02$	$9.9e-02$	$3.1e-03$	$1.4e-01$	$2.5e-01$	$7.8e-03$	0.0074	0.035	0.00111
1.4	$1.1e-02$	$4.4e-02$	$1.4e-03$	$1.3e-01$	$2.4e-01$	$7.5e-03$	0.0055	0.026	0.00082
1.5	$2.4e-03$	$1.6e-02$	$5.0e-04$	$9.5e-02$	$2.0e-01$	$6.4e-03$	0.0051	0.025	0.00079
1.6	$5.5e-04$	$5.1e-03$	$1.6e-04$	$6.9e-02$	$1.8e-01$	$5.5e-03$	0.0056	0.023	0.00072
1.7	$4.3e-04$	$1.0e-02$	$3.3e-04$	$6.2e-02$	$1.7e-01$	$5.4e-03$	0.0053	0.024	0.00077
1.8	$1.1e-05$	$1.4e-04$	$4.4e-06$	$4.8e-02$	$1.6e-01$	$5.0e-03$	0.0058	0.028	0.00090
1.9	$3.2e-05$	$9.6e-04$	$3.0e-05$	$3.4e-02$	$1.3e-01$	$4.2e-03$	0.0065	0.030	0.00095
2.0	$1.2e-05$	$3.3e-04$	$1.1e-05$	$1.9e-02$	$9.2e-02$	$2.9e-03$	0.0059	0.028	0.00090
2.1	$1.2e-07$	$2.3e-06$	$7.1e-08$	$1.7e-02$	$9.2e-02$	$2.9e-03$	0.0052	0.025	0.00078
2.2	$3.6e-07$	$1.1e-05$	$3.6e-07$	$1.0e-02$	$6.8e-02$	$2.1e-03$	0.0060	0.031	0.00097
2.3	$6.1e-13$	$1.3e-11$	$4.0e-13$	$5.8e-03$	$5.0e-02$	$1.6e-03$	0.0056	0.026	0.00082
2.4	$3.2e-12$	$6.6e-11$	$2.1e-12$	$3.3e-03$	$3.8e-02$	$1.2e-03$	0.0060	0.032	0.00100
2.5	$9.7e-08$	$3.1e-06$	$9.7e-08$	$2.3e-03$	$3.4e-02$	$1.1e-03$	0.0059	0.028	0.00088
2.6	$4.9e-14$	$1.1e-12$	$3.6e-14$	$2.9e-03$	$4.3e-02$	$1.4e-03$	0.0047	0.021	0.00068
2.7	$2.2e-11$	$6.8e-10$	$2.2e-11$	$1.6e-03$	$2.3e-02$	$7.3e-04$	0.0071	0.029	0.00091
2.8	$9.5e-13$	$3.0e-11$	$9.5e-13$	$1.1e-03$	$2.1e-02$	$6.6e-04$	0.0058	0.026	0.00083
2.9	$5.7e-17$	$1.5e-15$	$4.7e-17$	$1.7e-05$	$3.2e-04$	$1.0e-05$	0.0046	0.020	0.00064
3.0	$3.7e-17$	$1.1e-15$	$3.4e-17$	$4.3e-05$	$1.0e-03$	$3.3e-05$	0.0057	0.025	0.00078
3.1	$0.0e+00$	$0.0e+00$	$0.0e+00$	$7.8e-05$	$2.2e-03$	$7.0e-05$	0.0065	0.032	0.00100
3.2	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.3e-05$	$3.4e-04$	$1.1e-05$	0.0058	0.025	0.00079
3.3	$0.0e+00$	$0.0e+00$	$0.0e+00$	$4.9e-08$	$1.5e-06$	$4.9e-08$	0.0066	0.034	0.00109
3.4	$0.0e+00$	$0.0e+00$	$0.0e+00$	$6.4e-06$	$2.0e-04$	$6.4e-06$	0.0054	0.024	0.00077
3.5	$0.0e+00$	$0.0e+00$	$0.0e+00$	$2.2e-06$	$6.8e-05$	$2.2e-06$	0.0047	0.019	0.00059
3.6	$0.0e+00$	$0.0e+00$	$0.0e+00$	$1.4e-07$	$4.3e-06$	$1.4e-07$	0.0057	0.026	0.00081
3.7	$0.0e+00$	$0.0e+00$	$0.0e+00$	$3.0e-10$	$6.6e-09$	$2.1e-10$	0.0048	0.022	0.00070
3.8	$0.0e+00$	$0.0e+00$	$0.0e+00$	$8.5e-12$	$2.7e-10$	$8.5e-12$	0.0043	0.020	0.00062
3.9	$0.0e+00$	$0.0e+00$	$0.0e+00$	$4.3e-12$	$1.3e-10$	$4.1e-12$	0.0057	0.027	0.00085
4.0	$0.0e+00$	$0.0e+00$	$0.0e+00$	$7.1e-16$	$2.1e-14$	$6.6e-16$	0.0051	0.023	0.00072
	Most Similar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	0.415	0.283	0.0090	0.383	0.283	0.0090	0.54	0.30	0.0095
1.1	0.395	0.277	0.0088	0.381	0.279	0.0088	0.54	0.30	0.0095
1.2	0.408	0.285	0.0090	0.379	0.283	0.0090	0.54	0.30	0.0094
1.3	0.380	0.285	0.0090	0.356	0.272	0.0086	0.54	0.30	0.0096
1.4	0.387	0.274	0.0087	0.375	0.283	0.0090	0.54	0.30	0.0094
1.5	0.351	0.280	0.0088	0.355	0.278	0.0088	0.53	0.30	0.0096
1.6	0.326	0.273	0.0086	0.343	0.269	0.0085	0.53	0.30	0.0096
1.7	0.289	0.263	0.0083	0.326	0.276	0.0087	0.54	0.30	0.0095
1.8	0.274	0.260	0.0082	0.327	0.273	0.0086	0.54	0.29	0.0093
1.9	0.269	0.264	0.0083	0.302	0.271	0.0086	0.54	0.30	0.0095
2.0	0.234	0.257	0.0081	0.290	0.272	0.0086	0.55	0.30	0.0095
2.1	0.203	0.234	0.0074	0.288	0.269	0.0085	0.56	0.30	0.0094
2.2	0.191	0.234	0.0074	0.255	0.254	0.0080	0.55	0.30	0.0095
2.3	0.173	0.218	0.0069	0.262	0.265	0.0084	0.54	0.29	0.0093
2.4	0.159	0.211	0.0067	0.236	0.251	0.0079	0.53	0.31	0.0096
2.5	0.140	0.196	0.0062	0.227	0.251	0.0079	0.53	0.30	0.0094
2.6	0.117	0.179	0.0057	0.196	0.235	0.0074	0.56	0.30	0.0095
2.7	0.104	0.160	0.0051	0.195	0.233	0.0074	0.57	0.30	0.0094
2.8	0.086	0.141	0.0045	0.178	0.219	0.0069	0.54	0.30	0.0094
2.9	0.086	0.147	0.0046	0.165	0.218	0.0069	0.53	0.30	0.0093
3.0	0.072	0.141	0.0045	0.147	0.208	0.0066	0.53	0.30	0.0096
3.1	0.061	0.123	0.0039	0.142	0.205	0.0065	0.54	0.30	0.0095
3.2	0.061	0.128	0.0040	0.127	0.190	0.0060	0.55	0.30	0.0096
3.3	0.049	0.115	0.0036	0.120	0.188	0.0059	0.53	0.30	0.0095
3.4	0.045	0.098	0.0031	0.102	0.172	0.0054	0.55	0.30	0.0094
3.5	0.039	0.096	0.0031	0.096	0.169	0.0053	0.57	0.30	0.0095
3.6	0.033	0.090	0.0029	0.094	0.166	0.0052	0.54	0.30	0.0094
3.7	0.030	0.074	0.0024	0.089	0.162	0.0051	0.53	0.30	0.0094
3.8	0.024	0.066	0.0021	0.075	0.143	0.0045	0.54	0.30	0.0094
3.9	0.021	0.063	0.0020	0.069	0.138	0.0044	0.52	0.31	0.0097
4.0	0.022	0.069	0.0022	0.055	0.112	0.0035	0.56	0.30	0.0094

Table B.13: IRD (above) and MSDI (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in structured variation.

	Most Dissimilar Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$7.6e-01$	$2.7e-01$	$8.4e-03$	$1.6e-01$	$2.7e-01$	$8.5e-03$	0.01004	0.080	0.00253
1.1	$5.0e-01$	$3.4e-01$	$1.1e-02$	$1.8e-01$	$2.8e-01$	$8.8e-03$	0.00958	0.076	0.00239
1.2	$2.4e-01$	$3.0e-01$	$9.4e-03$	$1.6e-01$	$2.7e-01$	$8.5e-03$	0.00924	0.085	0.00269
1.3	$9.8e-02$	$2.0e-01$	$6.2e-03$	$1.4e-01$	$2.5e-01$	$7.8e-03$	0.00941	0.084	0.00266
1.4	$4.8e-02$	$1.4e-01$	$4.5e-03$	$1.4e-01$	$2.5e-01$	$7.8e-03$	0.00552	0.063	0.00199
1.5	$1.8e-02$	$7.9e-02$	$2.5e-03$	$1.2e-01$	$2.4e-01$	$7.5e-03$	0.01039	0.081	0.00257
1.6	$1.2e-02$	$6.7e-02$	$2.1e-03$	$8.8e-02$	$2.1e-01$	$6.5e-03$	0.00900	0.079	0.00248
1.7	$6.2e-03$	$4.8e-02$	$1.5e-03$	$9.1e-02$	$2.2e-01$	$6.8e-03$	0.01125	0.093	0.00294
1.8	$2.4e-03$	$2.2e-02$	$7.1e-04$	$7.2e-02$	$2.0e-01$	$6.2e-03$	0.00863	0.075	0.00237
1.9	$4.0e-03$	$4.6e-02$	$1.4e-03$	$6.4e-02$	$1.8e-01$	$5.7e-03$	0.01200	0.097	0.00305
2.0	$5.3e-04$	$5.7e-03$	$1.8e-04$	$4.7e-02$	$1.6e-01$	$5.0e-03$	0.00929	0.078	0.00248
2.1	$9.6e-04$	$2.1e-02$	$6.6e-04$	$3.4e-02$	$1.3e-01$	$4.1e-03$	0.01086	0.082	0.00259
2.2	$9.7e-04$	$2.0e-02$	$6.3e-04$	$3.0e-02$	$1.3e-01$	$4.0e-03$	0.01021	0.084	0.00266
2.3	$4.7e-05$	$1.0e-03$	$3.3e-05$	$2.0e-02$	$1.0e-01$	$3.2e-03$	0.01141	0.086	0.00272
2.4	$1.1e-04$	$2.7e-03$	$8.4e-05$	$1.5e-02$	$8.5e-02$	$2.7e-03$	0.00922	0.079	0.00249
2.5	$7.5e-04$	$2.4e-02$	$7.4e-04$	$1.5e-02$	$8.7e-02$	$2.7e-03$	0.00609	0.061	0.00191
2.6	$9.6e-06$	$2.7e-04$	$8.6e-06$	$1.6e-02$	$9.2e-02$	$2.9e-03$	0.00926	0.080	0.00254
2.7	$2.7e-07$	$4.5e-06$	$1.4e-07$	$7.0e-03$	$5.7e-02$	$1.8e-03$	0.00991	0.083	0.00263
2.8	$4.9e-07$	$1.5e-05$	$4.9e-07$	$6.6e-03$	$5.8e-02$	$1.8e-03$	0.01219	0.092	0.00291
2.9	$4.9e-07$	$1.5e-05$	$4.9e-07$	$4.3e-03$	$4.7e-02$	$1.5e-03$	0.00927	0.083	0.00264
3.0	$3.5e-10$	$8.5e-09$	$2.7e-10$	$2.8e-03$	$3.8e-02$	$1.2e-03$	0.01023	0.086	0.00271
3.1	$1.5e-08$	$4.8e-07$	$1.5e-08$	$2.8e-03$	$4.0e-02$	$1.3e-03$	0.01195	0.094	0.00297
3.2	$5.1e-10$	$9.3e-09$	$3.0e-10$	$1.2e-03$	$2.0e-02$	$6.3e-04$	0.00773	0.072	0.00226
3.3	$8.6e-11$	$2.7e-09$	$8.4e-11$	$6.5e-04$	$1.2e-02$	$3.7e-04$	0.00499	0.054	0.00172
3.4	$2.7e-11$	$8.4e-10$	$2.7e-11$	$9.9e-04$	$2.4e-02$	$7.5e-04$	0.00859	0.077	0.00245
3.5	$8.4e-10$	$2.4e-08$	$7.6e-10$	$1.9e-03$	$3.1e-02$	$9.8e-04$	0.01211	0.098	0.00309
3.6	$2.4e-12$	$7.5e-11$	$2.4e-12$	$2.1e-04$	$3.8e-03$	$1.2e-04$	0.01084	0.086	0.00272
3.7	$8.1e-12$	$2.6e-10$	$8.1e-12$	$2.0e-04$	$4.3e-03$	$1.4e-04$	0.01035	0.086	0.00273
3.8	$1.2e-14$	$3.8e-13$	$1.2e-14$	$6.1e-04$	$1.4e-02$	$4.3e-04$	0.01023	0.087	0.00276
3.9	$1.8e-13$	$5.8e-12$	$1.8e-13$	$5.1e-05$	$1.6e-03$	$5.0e-05$	0.00912	0.082	0.00259
4.0	$2.7e-11$	$8.4e-10$	$2.7e-11$	$1.1e-04$	$2.8e-03$	$8.7e-05$	0.01024	0.091	0.00286

	Average Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	0.592	0.27	0.0086	0.390	0.29	0.0092	0.44	0.31	0.0099
1.1	0.608	0.27	0.0085	0.369	0.29	0.0090	0.45	0.32	0.0100
1.2	0.577	0.27	0.0087	0.389	0.28	0.0088	0.45	0.32	0.0101
1.3	0.533	0.28	0.0089	0.389	0.28	0.0088	0.45	0.32	0.0100
1.4	0.525	0.28	0.0088	0.399	0.29	0.0093	0.46	0.31	0.0099
1.5	0.500	0.28	0.0089	0.404	0.30	0.0094	0.43	0.31	0.0099
1.6	0.452	0.28	0.0089	0.367	0.29	0.0092	0.44	0.31	0.0098
1.7	0.428	0.28	0.0089	0.383	0.28	0.0089	0.43	0.31	0.0098
1.8	0.392	0.28	0.0088	0.372	0.28	0.0089	0.46	0.31	0.0097
1.9	0.375	0.27	0.0085	0.355	0.28	0.0087	0.46	0.31	0.0098
2.0	0.349	0.28	0.0088	0.357	0.28	0.0090	0.44	0.32	0.0101
2.1	0.337	0.27	0.0086	0.344	0.28	0.0088	0.45	0.32	0.0100
2.2	0.294	0.25	0.0079	0.335	0.28	0.0087	0.46	0.31	0.0099
2.3	0.276	0.26	0.0082	0.340	0.28	0.0088	0.47	0.32	0.0101
2.4	0.271	0.26	0.0082	0.340	0.28	0.0088	0.44	0.31	0.0098
2.5	0.243	0.25	0.0078	0.321	0.27	0.0085	0.46	0.32	0.0101
2.6	0.232	0.24	0.0075	0.332	0.28	0.0088	0.46	0.32	0.0101
2.7	0.226	0.24	0.0076	0.319	0.28	0.0089	0.46	0.31	0.0097
2.8	0.198	0.23	0.0072	0.295	0.28	0.0088	0.46	0.31	0.0098
2.9	0.187	0.22	0.0070	0.280	0.27	0.0085	0.44	0.31	0.0097
3.0	0.166	0.21	0.0065	0.279	0.27	0.0086	0.45	0.31	0.0098
3.1	0.161	0.20	0.0064	0.263	0.26	0.0084	0.44	0.31	0.0098
3.2	0.142	0.19	0.0060	0.245	0.25	0.0080	0.45	0.32	0.0101
3.3	0.130	0.18	0.0058	0.231	0.25	0.0078	0.46	0.32	0.0101
3.4	0.116	0.17	0.0054	0.232	0.25	0.0080	0.44	0.32	0.0100
3.5	0.107	0.16	0.0052	0.218	0.25	0.0078	0.44	0.32	0.0101
3.6	0.092	0.15	0.0048	0.213	0.24	0.0077	0.46	0.31	0.0097
3.7	0.087	0.14	0.0045	0.205	0.24	0.0075	0.44	0.31	0.0098
3.8	0.084	0.14	0.0044	0.186	0.22	0.0071	0.45	0.30	0.0096
3.9	0.068	0.12	0.0038	0.180	0.22	0.0070	0.46	0.32	0.0101
4.0	0.069	0.13	0.0041	0.165	0.21	0.0067	0.46	0.31	0.0098

Table B.14: MDD (above) and AVND (below) summary measures (mean, standard deviation (sd), standard error (se)), for change in structured variation.

	Most Similar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$9.0e-01$	0.14062	$4.4e-03$	$3.3e-01$	0.28943	$9.2e-03$	0.23	0.26	0.0082
1.1	$8.5e-01$	0.17149	$5.4e-03$	$3.3e-01$	0.29154	$9.2e-03$	0.22	0.25	0.0080
1.2	$7.3e-01$	0.22896	$7.2e-03$	$3.4e-01$	0.28573	$9.0e-03$	0.23	0.26	0.0082
1.3	$5.8e-01$	0.26753	$8.5e-03$	$3.3e-01$	0.29169	$9.2e-03$	0.24	0.27	0.0084
1.4	$4.5e-01$	0.25851	$8.2e-03$	$3.3e-01$	0.29454	$9.3e-03$	0.23	0.25	0.0080
1.5	$3.2e-01$	0.23429	$7.4e-03$	$3.3e-01$	0.29130	$9.2e-03$	0.22	0.26	0.0081
1.6	$2.3e-01$	0.20298	$6.4e-03$	$2.9e-01$	0.28820	$9.1e-03$	0.22	0.26	0.0081
1.7	$1.7e-01$	0.17513	$5.5e-03$	$2.9e-01$	0.28609	$9.0e-03$	0.22	0.25	0.0081
1.8	$1.2e-01$	0.15556	$4.9e-03$	$2.6e-01$	0.27500	$8.7e-03$	0.23	0.25	0.0080
1.9	$8.3e-02$	0.11898	$3.8e-03$	$2.6e-01$	0.28176	$8.9e-03$	0.22	0.24	0.0077
2.0	$6.0e-02$	0.10496	$3.3e-03$	$2.1e-01$	0.26440	$8.4e-03$	0.22	0.25	0.0080
2.1	$3.7e-02$	0.07616	$2.4e-03$	$2.1e-01$	0.27137	$8.6e-03$	0.24	0.26	0.0082
2.2	$2.7e-02$	0.06048	$1.9e-03$	$1.7e-01$	0.24451	$7.7e-03$	0.22	0.25	0.0080
2.3	$1.7e-02$	0.04976	$1.6e-03$	$1.5e-01$	0.22809	$7.2e-03$	0.24	0.26	0.0083
2.4	$1.2e-02$	0.03294	$1.0e-03$	$1.4e-01$	0.21336	$6.7e-03$	0.22	0.25	0.0078
2.5	$7.5e-03$	0.02365	$7.5e-04$	$1.2e-01$	0.19935	$6.3e-03$	0.22	0.25	0.0077
2.6	$5.5e-03$	0.01937	$6.1e-04$	$1.1e-01$	0.19439	$6.1e-03$	0.23	0.26	0.0083
2.7	$4.6e-03$	0.01757	$5.6e-04$	$8.3e-02$	0.17419	$5.5e-03$	0.23	0.25	0.0079
2.8	$2.6e-03$	0.00969	$3.1e-04$	$8.0e-02$	0.18036	$5.7e-03$	0.23	0.25	0.0079
2.9	$2.7e-03$	0.01645	$5.2e-04$	$5.9e-02$	0.14486	$4.6e-03$	0.23	0.26	0.0081
3.0	$9.2e-04$	0.00610	$1.9e-04$	$4.6e-02$	0.12435	$3.9e-03$	0.23	0.26	0.0083
3.1	$1.0e-03$	0.00660	$2.1e-04$	$4.3e-02$	0.12157	$3.8e-03$	0.23	0.25	0.0080
3.2	$5.0e-04$	0.00506	$1.6e-04$	$3.2e-02$	0.10360	$3.3e-03$	0.24	0.26	0.0082
3.3	$3.9e-04$	0.00297	$9.4e-05$	$2.0e-02$	0.06481	$2.0e-03$	0.23	0.25	0.0078
3.4	$2.4e-04$	0.00174	$5.5e-05$	$2.3e-02$	0.08717	$2.8e-03$	0.24	0.26	0.0082
3.5	$2.2e-04$	0.00285	$9.0e-05$	$1.5e-02$	0.06057	$1.9e-03$	0.23	0.26	0.0081
3.6	$7.2e-05$	0.00043	$1.4e-05$	$1.5e-02$	0.07415	$2.3e-03$	0.23	0.25	0.0080
3.7	$4.9e-05$	0.00059	$1.9e-05$	$1.0e-02$	0.05516	$1.7e-03$	0.21	0.24	0.0077
3.8	$2.2e-05$	0.00015	$4.8e-06$	$8.2e-03$	0.04875	$1.5e-03$	0.23	0.25	0.0079
3.9	$3.6e-05$	0.00038	$1.2e-05$	$6.0e-03$	0.03415	$1.1e-03$	0.23	0.26	0.0082
4.0	$1.4e-05$	0.00012	$3.7e-06$	$3.9e-03$	0.02410	$7.6e-04$	0.22	0.25	0.0080
	Most Dissimilar Neighbour Differences								
	$\mu = \frac{1}{10}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{1000}, sdu = 0.1, sdv = 0.1$			$\mu = \frac{1}{100000}, sdu = 0.1, sdv = 0.1$		
Multiple	mean	sd	se	mean	sd	se	mean	sd	se
1.0	$9.4e-01$	0.10121	$3.2e-03$	$4.1e-01$	0.3178	$1.0e-02$	0.33	0.32	0.0100
1.1	$9.2e-01$	0.12742	$4.0e-03$	$4.1e-01$	0.3284	$1.0e-02$	0.33	0.32	0.0101
1.2	$8.4e-01$	0.18772	$5.9e-03$	$4.0e-01$	0.3143	$9.9e-03$	0.36	0.32	0.0101
1.3	$7.0e-01$	0.24317	$7.7e-03$	$4.1e-01$	0.3226	$1.0e-02$	0.35	0.32	0.0101
1.4	$5.8e-01$	0.27054	$8.6e-03$	$4.2e-01$	0.3316	$1.0e-02$	0.35	0.32	0.0101
1.5	$4.5e-01$	0.25819	$8.2e-03$	$4.2e-01$	0.3334	$1.1e-02$	0.35	0.32	0.0101
1.6	$3.5e-01$	0.25131	$7.9e-03$	$3.8e-01$	0.3281	$1.0e-02$	0.34	0.32	0.0101
1.7	$2.5e-01$	0.22142	$7.0e-03$	$3.7e-01$	0.3228	$1.0e-02$	0.35	0.32	0.0101
1.8	$1.9e-01$	0.19866	$6.3e-03$	$3.4e-01$	0.3214	$1.0e-02$	0.35	0.32	0.0102
1.9	$1.4e-01$	0.17850	$5.6e-03$	$3.2e-01$	0.3196	$1.0e-02$	0.35	0.32	0.0101
2.0	$1.0e-01$	0.14922	$4.7e-03$	$2.9e-01$	0.3149	$1.0e-02$	0.35	0.32	0.0100
2.1	$6.6e-02$	0.11232	$3.6e-03$	$2.8e-01$	0.3113	$9.8e-03$	0.36	0.32	0.0101
2.2	$4.7e-02$	0.09187	$2.9e-03$	$2.5e-01$	0.2925	$9.3e-03$	0.35	0.31	0.0099
2.3	$3.1e-02$	0.07215	$2.3e-03$	$2.2e-01$	0.2844	$9.0e-03$	0.35	0.32	0.0100
2.4	$2.3e-02$	0.05080	$1.6e-03$	$2.1e-01$	0.2736	$8.7e-03$	0.32	0.31	0.0099
2.5	$1.5e-02$	0.04576	$1.4e-03$	$1.8e-01$	0.2627	$8.3e-03$	0.36	0.32	0.0103
2.6	$1.1e-02$	0.03076	$9.7e-04$	$1.6e-01$	0.2447	$7.7e-03$	0.33	0.31	0.0099
2.7	$8.9e-03$	0.02891	$9.1e-04$	$1.3e-01$	0.2181	$6.9e-03$	0.35	0.32	0.0101
2.8	$6.1e-03$	0.02335	$7.4e-04$	$1.3e-01$	0.2229	$7.0e-03$	0.34	0.30	0.0096
2.9	$4.9e-03$	0.02072	$6.6e-04$	$1.0e-01$	0.2026	$6.4e-03$	0.35	0.32	0.0101
3.0	$2.3e-03$	0.01332	$4.2e-04$	$8.0e-02$	0.1724	$5.5e-03$	0.34	0.32	0.0100
3.1	$2.6e-03$	0.01544	$4.9e-04$	$7.5e-02$	0.1708	$5.4e-03$	0.34	0.32	0.0101
3.2	$1.6e-03$	0.01933	$6.1e-04$	$6.1e-02$	0.1567	$5.0e-03$	0.35	0.32	0.0101
3.3	$7.4e-04$	0.00397	$1.3e-04$	$4.3e-02$	0.1203	$3.8e-03$	0.35	0.31	0.0098
3.4	$7.8e-04$	0.00513	$1.6e-04$	$4.3e-02$	0.1235	$3.9e-03$	0.33	0.31	0.0100
3.5	$5.7e-04$	0.00395	$1.2e-04$	$3.4e-02$	0.1070	$3.4e-03$	0.33	0.31	0.0098
3.6	$5.2e-04$	0.00723	$2.3e-04$	$2.6e-02$	0.0901	$2.8e-03$	0.34	0.31	0.0099
3.7	$9.1e-05$	0.00052	$1.6e-05$	$2.1e-02$	0.0757	$2.4e-03$	0.35	0.33	0.0103
3.8	$2.5e-04$	0.00433	$1.4e-04$	$1.6e-02$	0.0689	$2.2e-03$	0.34	0.31	0.0098
3.9	$7.4e-05$	0.00056	$1.8e-05$	$1.3e-02$	0.0563	$1.8e-03$	0.35	0.32	0.0100
4.0	$5.2e-05$	0.00042	$1.3e-05$	$1.1e-02$	0.0512	$1.6e-03$	0.35	0.31	0.0099

Table B.15: MSDN (above) and MDND (below) Summary measures (mean, standard deviation (sd), standard error (se)), for change in structured variation.

Appendix C

R Codes

C.1 Chapter 1, 4, 5, 6 and 7 Codes

C.1.1 Dot Map (Section 1.3.1)

`leukaemia` is a text file which contains locations of disease addresses in x, y coordinates in two columns.

```
leukaemia<-read.table("leukaemia.txt",header=TRUE)
```

`coordinates` is a shape file with Humberside boundaries

```
coordinates<-readShapePoly("coordinates.shp")
plot(coordinates)
par("new"=TRUE)
plot(leukaemia,pch=20,xlab="",ylab="",axes=FALSE,frame=TRUE)
```

C.1.2 Choropleth Map

To use function *leglabs* in R to make the legend breaks:

```
library(maptools)
```

To create quantiles for susceptibility rates to use to allocate colour to each quantile:

```
brks <- round(quantile(susceptible,probs=seq(0, 1, 0.2)),digits=2)
```

where `susceptible` is a vector of rates.

To define colours:

```
colours <-c("lightblue","skyblue2","skyblue3","blue3","blue4")
```

To plot boundaries of the map:

`plot(coordinates,type="l",asp=0.85,axes=FALSE,frame=TRUE,xlab="",ylab="")`
where `coordinates` is a file containing boundary coordinates in two columns (eastings and northings).

To add colours to the map:
`polygon(coordinates,col=colours[findInterval(susceptible,brks,all.inside=TRUE)])`

To add a legend:
`legend(x=1*10^5, y=7.5*10^5, legend=leglabs(brks),fill=colours,cex=0.6,bty="n")`

To add a title:
`title(main=paste("Susceptibility Map"))`

C.1.3 Kriging (Section 4.4)

Load package for kriging:
`library(geoR)`

To carry out kriging:
`geo <- as.geodata(smrcoords,coords.col=1:2,data.col=3)`
`bin<-variog(geo)` where `smrcoords` is a file containing susceptibility ratios (observed/expected counts) and centroids for the regions, where column 1 is x-coordinates, and column 2 is y-coordinates for the centroids and column 3 is the susceptibility ratios.

Estimate spherical variogram using weighted least square method:

```
wls<-variofit(bin,fix.nugget=FALSE,cov.model="spherical",max.dist=bin$max.dist)
pred<-expand.grid(seq(15000,400000,l=15),seq(600000,900000,l=15))
krig<-krige.conv(geo,loc=pred, krige=krige.control(type.krige="ok",obj.m=
wls))
plot(scot$X,scot$Y,axes = FALSE, frame = TRUE,xlab="",ylab="")
contour(krig,nlevel=18,labcex=0.7,lty = "solid",axes = FALSE, frame = TRUE,
xlab="",ylab="",add=TRUE)
```

`title("Kriging Map")` Plot semi-variogram:
`plot(bin ,main="Plot of Semi-Variogram")`

```
lines(wls)
```

C.1.4 Kernel Smoothing (Section 4.4)

Load package with Nadaraya-Watson kernel smoother:

```
library(JLLprod)
```

`smrcoords.txt` is a file which contains the regions centroids and measles susceptibility ratios in three columns

```
smrcoords<-read.table("smrcoords.txt",sep=",",header=TRUE)
```

Kernel smoothing for 2000:

```
kern<-Blocc(smrcoords$X,smrcoords$Y,smrcoords$ratios)
plot(scot$X,scot$Y,type="l",asp=0.9,axes = FALSE,xlab="",ylab="")
contour(kern$xxe,kern$zze,kern$r,nlevel=10,labcex=0.5,lty = "solid",axes
= TRUE, frame = TRUE,xlab="",ylab="",add=TRUE)
title("Kernel Smoothing Map")
```

C.1.5 Pseudo-Colour Map

`x1` and `x2` are numeric vectors of rates for map 1 and map 2.

Find 5th and 95th percentile of each one (xq5, yq5 say and xq95 and yq95), using type 4 quantiles:

```
xq5<- quantile(x1, 0.05, type=4)
xq95<-quantile(x1, 0.95, type=4)
yq5<- quantile(y1, 0.05, type=4)
yq95<-quantile(y1, 0.95, type=4)
```

Take the minimum of the two 5th percentiles and the maximum of the two 95th percentiles:

```
newmin<-min(xq5, yq5)
newmax<-max(xq95, yq95)
```

Scale each vector to give new rate vectors `scaledx` and `scaledy`, so each new one ranges from 0 to 255 (the usual intensity range for images):

```
range<-newmax-newmin
```

```
scaledx <-(x1- newmin)*255/range
scaledx[scaledx<0]=0
scaledx[scaledx>255]=255
scaledx<-round(scaledx,0)

scaledy <-(y1- newmin)*255/range
scaledy[scaledy<0]=0
scaledy[scaledy>255]=255
scaledy=round(scaledy,0)
```

Allocate colours to new rate vectors, using `scaledx` (map 1) to define red and blue, and `scaledy` (map 2) to define green:

```
ycol2 <- rgb(scaledx,scaledy,scaledx,maxColorValue=255)
```

Produce a pseudo-colour map and add an interpretative legend:

```
plot(coord,type="l",asp=0.85,axes=FALSE,frame=TRUE,xlab="",ylab="")
polygon(coord,col=ycol2)
legend(3.05*10^5, 9.8*10^5,c("Gray-Maps Similar","Red-Map 1 > Map 2", "Green-Map
2 > Map 1"),cex=0.568,bty="n")
title(main=paste("Pseudo-Colour Map"))
```

C.1.6 Image Analysis based Methods

Function `IMAGE` computes image analysis based measures to compare two maps.

Function input:
`x1` and `x2` are numeric vectors of relative risks for map 1 and map 2.

Function output:
`MSD` is the mean square difference of relative risks between two maps. (A value of zero indicates that similar maps and large values indicate that maps differ greatly).

`PSNRR` is a form of peak-to-signal-noise ratio given by P^2/MSD, where P is the difference between the maximum of the two sets of relative risks and the minimum of the two set of relative risks. (A value of zero indicates that maps are very different and large values indicate similar maps).

`PSNRM` is a form of peak-to-signal-noise ratio given by P^2/MSD where P is the max-

imum absolute difference between the two sets of relative risks. (A value of zero indicates that maps are very different and large values indicate similar maps).

SSIM/SSIMM/SSIMR are a form of structural similarity index given by $(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)/(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)$, where μ_x and μ_y are the means for relative risks in maps 1 and 2 respectively, σ_{xy} is correlation between the two sets of relative risks, and σ_x^2 and σ_y^2 are the variances of relative risks in maps 1 and 2 respectively. $C_i = (K_i * L)^2, i = 1, 2$, and K_i is a very small arbitrary constant used to avoid unstable results when $(\mu_x^2+\mu_y^2)$ or $(\sigma_x^2+\sigma_y^2)$ are near zero.

For SSIM, constants are set to zero. For SSIMM, L is taken as the maximum absolute difference between the two sets of relative risks for the ith region. For SSIMR, L is the difference between the maximum of the two set of relative risks and the minimum of the two sets of relative risks.

```
IMAGE <- function(x1,y1) {
w<-(is.numeric(x1))&(is.numeric(y1))
if(!w)
stop("Error!  data must be numeric vectors")
k<-(length(x1)==length(y1))
if(k)
n<-length(x1)
else stop("Error!  vectors must be of the same length")

#For MSD:
alldiff<-x1-y1
alldiffsq<-(alldiff)^2
MSD<-sum(alldiffsq)/n
if(is.na(MSD)) stop("Obtained NaN value for MSD")
if(is.infinite(MSD)) stop("Obtained infinite value for MSD")

#For PSNRR:
Rangevalues<-max(x1,y1)-min(x1,y1)
Rangevaluessq<-(Rangevalues)^2
PSNRR<-Rangevaluessq/MSD
if(is.na(PSNRR)) stop("Obtained NaN value for PSNRR")
```

```
if(is.infinite(PSNRR)) stop("Obtained infinite value for PSNRR")

#For PSNRM:
Mvalues<-abs(x1-y1)
R<-max(Mvalues)
Rsq<-R^2
PSNRM<-Rsq/MSD
if(is.na(PSNRM)) stop("Obtained NaN value for PSNRM")
if(is.infinite(PSNRM)) stop("Obtained infinite value for PSNRM")

#For SSIM:
mpart<-(2*mean(x1)*mean(y1))/((mean(x1))^2+(mean(y1))^2)
vpart<-(2*sd(x1)*sd(y1))/(var(x1)+var(y1))
cpart<-(cov(x1,y1))/(sd(x1)*sd(y1))
SSIM<-mpart*vpart*cpart
if(is.na(SSIMR)) stop("Obtained NaN value for SSIMR")
if(is.infinite(SSIMR)) stop("Obtained infinite value for SSIMR")

#For SSIMR:
mpartR<-((2*mean(x1)*mean(y1))+(0.001*(max(x1,y1)-min(x1,y1)))^2)
/((mean(x1))^2+ (mean(y1))^2+(0.001*(max(x1,y1)-min(x1,y1)))^2)
vpartR<-((2*sd(x1)*sd(y1))+(0.002*(max(x1,y1)-min(x1,y1)))^2)/
((var(x1)+var(y1))+ (0.002*(max(x1,y1)-min(x1,y1)))^2)
cpartR<-(cov(x1,y1)+(0.001*(max(x1,y1)-min(x1,y1)))^2)/
((sd(x1)*sd(y1))+(0.001*(max(x1,y1)-min(x1,y1)))^2)
SSIMR<-mpartR*vpartR*cpartR
if(is.na(SSIMR)) stop("Obtained NaN value for SSIMR")
if(is.infinite(SSIMR)) stop("Obtained infinite value for SSIMR")

#For SSIMM:
mpartM<-((2*mean(x1)*mean(y1))+(0.001*max(abs(x1-y1)))^2)/((mean(x1))^2+
(mean(y1))^2+(0.001*
max(abs(x1-y1)))^2)
vpartM<-((2*sd(x1)*sd(y1))+(0.002*max(abs(x1-y1)))^2)/(var(x1)+var(y1)+
(0.002*max(abs(x1-y1)))^2)
```

```
cpartM<-(cov(x1,y1)+(0.001*max(abs(x1-y1)))^2)/((sd(x1)*sd(y1))+
(0.001*max(abs(x1-y1)))^2) SSIMM<-mpartM*vpartM*cpartM
if(is.na(SSIMM)) stop("Obtained NaN value for SSIMM")
if(is.infinite(SSIMM)) stop("Obtained infinite value for SSIMM")

c(MSD,PSNRR,PSNRM,SSIMR,SSIMM,SSIM)
}
```

C.1.7 Point Process based Methods

Function `IED` computes vectors of absolute differences from the relative risks of regions in a map.

Function input:
`RR`, a vector of length `n` containing relative risks of `n` regions of a map.
`wm`, an $n \times n$adjacency weight matrix with entries 1 (when regions are neighbours) and 0 (when regions are not neighbours).

Function output a list with the following difference components:
`diff`, a vector of length $\frac{1}{2}n(n-1)$ of absolute differences between all possible distinct pairs of relative risks in a map.
`mindiff`, a vector of length `n` with the ith entry as the absolute difference between the ith relative risk and the most similar relative risk in the map.
`maxdiff`, a vector of length `n` with the ith entry as the absolute difference between the ith relative risk and the most dissimilar relative risk in the map.
`aveneigh`, a vector of length `n` with the ith entry as the average difference between the relative risk of region i and the relative risks of its neighbouring regions.
`maxneigh`, a vector of length `n` with the ith entry as the absolute difference between the relative risk of ith region and the most dissimilar relative risk of its neighbouring regions.
`minneigh`, a vector of length `n` with the ith entry as the absolute difference between the relative risk of ith region and the most similar relative risk of its neighbouring regions.

```
IED<-function(RR,wm)
{
```

```
n <-length(RR)
Check2<-(length(RR)==dim(wm)[1])
if (!Check2) stop("Error!  weight matrix has wrong dimension")
```

#Replicate each risk value n times

```
m.RR<-matrix(rep(RR,n),ncol=n,byrow=T)
```

#Obtain absolute difference:

```
diff.RR<- abs(RR - m.RR)
```

#Extract the Lower triangular matrix from diff.R (only need this):

```
m.diff.RR<-diff.RR[lower.tri(diff.RR)]
```

#Obtain matrix excluding the difference between a value and itself (diagonal elements):

```
keep.RR<-matrix(T,nrow=n,ncol=n)
diag(keep.RR)<- F
m.keep.RR<-matrix(c(diff.RR)[c(keep.RR)],nrow=n-1,ncol=n)
# Obtain (nearest value) of absolute differences:
min.diff.RR<-apply(m.keep.RR,2,min)
```

Obtain (farthest value) of absolute differences:

```
max.diff.RR<-apply(m.keep.RR,2,max)
```

Obtain absolute difference for neighbours:

```
neigh.RR←sweep(wm,1,diff.RR,FUN="*")
```

Obtain average of absolute differences from neighbours:

```
sumnear<-apply(neigh.RR,2,sum)
sumweight<-apply(wm,2,sum)
aveneigh.RR<-sumnear/sumweight
# Obtain maximum absolute differences of neighbours
max.neigh.RR<-apply(neigh.RR,2,max)
```

Exclude islands: k a vector of indices of any regions (islands) to be omitted

```
max.neigh.RR<-max.neigh.RR[-k]
```

Obtain minimum absolute differences of neighbours:

```
wm[wm==0]<-NA
neigh.RR<-sweep(wm,1,diff.RR,FUN="*")
min.neigh.RR<-apply(neigh.RR,2,function(x) min(x[!is.na(x)]))

# Exclude islands:
min.neigh.RR<-min.neigh.RR[-i]
list(diff=m.diff.RR,mindiff=min.diff.RR,maxdiff=max.diff.RR,aveneigh=aveneigh.RR,
minneigh=min.neigh.RR,maxneigh=max.neigh.RR)
}
```

Function `compare` uses `ks.test` from library `Hmisc` to apply the two-sided Kolmogorov Smirnoff test to compare two empirical cdfs. It plots two empirical cdfs and adds the K-S p-value to the plot.

Function input:
`xx`, any of the vectors of the absolute differences computed from map 1 by function IED.
`yy`, any of the vectors of absolute differences computed from map 2 by function IED.

```
compare<-function(xx,yy)
{
library(Hmisc)
KS<- ks.test(xx,yy)
Ecdf(xx,xlab="2000 (blue), 2001 (red)",xlim=range(xx,yy),col="blue")
Ecdf(yy,col="red",add=TRUE,text((max(xx)+max(yy))/2.4,0.4,
paste("KS p-value=",signif(KS$p,2) )))
}
```

Function `comp.all6` uses function `compare` to compare each of the difference components, obtained from function IED, for two maps. It then extracts the p-values from the results.

Function input:
`xx`, any of the vectors of the absolute differences computed from map 1 by function IED.
`yy`, any of the vectors of absolute differences computed from map 2 by function IED.

Function output:

the p-values for the Kolmogorov-Smirnoff test for each difference component.

```
comp.all6<-function(xx,yy)
{
Check1<-(length(xx)==length(yy))
if (!Check1) stop("Error!  numeric vectors have different lengths")
IRD<-compare(xx,yy,"IRD")
MSDI<-compare(xx,yy,"MSDI")
MDD<-compare(xx,yy,"MDD")
MSND<-compare(xx,yy,"MSND")
MDND<-compare(xx,yy,"MDND")
AVND<-compare(xx,yy,"AVND")

c(IRD$p,MSDI$p,MDD$p,MSND$p,MDND$p,AVND$p)
}
```

Function `bigf` takes two vectors of relative risks (for two maps) and calls function `IED` twice to obtain difference components for each map. Function `comp.all6` is then used to compare the maps by comparing the empirical cdfs of each of the difference components using the Kolmogorov-Smirnoff test.

Function input:

`X1`, a vector of relative risks for map 1.

`X2`, a vector of relative risks for map 2.

`wm`, an adjacency $n \times n$ weight matrix with entries 1 (when when the regions are neighbours) and 0 (when when the regions are not neighbours).

Function output:

Kolmogorov-Smirnoff test p-values for each of the difference components.

```
bigf<-function(X1,Y1,wm)
{
Check<-(length(X1)==dim(wm)[1]) & (length(Y1)==dim(wm)[2])
if (!Check) stop("Error!  weight matrix has wrong dimension")
y1<-IED(X1,wm)
y2<-IED(Y1,wm)
```

```
y3<-comp.all6(y1,y2)
y3
}
```

C.1.8 Simulation Code

Function `simul` simulates a number of times two observed samples, one from each of two maps, each generated using a Poisson distribution, and compares the maps using methods based on image analysis and point process methodology. It calls function `IMAGE` and function `bigf`.

For the first sample, the Poisson mean is $\underline{N} * \underline{\theta}$ where $\underline{N}$ is the vector of the number of people in each region of the map, simulated from the Uniform distribution, and $\underline{\theta}$ is a vector of risks obtained from $log(\theta_i) = \alpha + U_i + V_i$, i.e. $\theta_i = exp(\alpha + U_i + V_i)$. Three values of α are used to represent three cases: common disease, rare disease and very rare disease. $\underline{U}$ is simulated from the CAR-normal model with precision lambda.st and $\underline{V}$ is simulated from a Normal distribution with mean 0 and standard deviation sd.un which measures unstructured variability. For the second sample, different multiples of the precision of $\underline{U}$ are used to vary the level of structured heterogeneity. The observed sample counts are divided by N to obtain proportions (rates) which are used to compare the two maps.

Function input:
`n`, the no. of regions
`n.comp`, the number of multiples to use
`mu.x`, the disease rate
`lambda.st`, the precision inverse matrix for structured variability
`sd.un`, the standard deviation for unstructured variability
`mu.inflate`, a constant controlling multiples of the disease rate
`sd.un.inflate`, a constant controlling multiples of the standard deviation of unstructured variability
`sd.st.inflate`, a constant controlling multiples of the precision of structured variability
`n.sims`, number of simulations performed
`min.pop`, the minimum population size

max.pop, the maximum population size

neighb, a vector of length n containing number of neighbours for each region

wm, an $n \times n$ adjacency weight matrix with entries 1 (when the regions are neighbours) and 0 (when the regions are not neighbours).

Function output:

a list Z3.st containing:

Z1.st, values for the image analysis methods

Z2.st, a list of p-values for the difference based methods

XM.st, a matrix of disease rates for the first sample

YM.st, a matrix of disease rates for the second sample

V.st, is a matrix of unstructured random effects

NM.st, is a matrix of total population in each region

neighb, is a vector of size n with number of neighbours for each region

wm, is a weight matrix

ZZ.st, is a matrix containing values each of the 12 measures with Z1.st and Z2.st for each simulation

```
simul<-function(n,n.comp,mu.x,lambda.st,sd.un,mu.inflate,sd.un.inflate,
sd.st.inflate, n.sims,min.pop=50000,max.pop=150000,neighb,wm)
{
library(MASS)
# Define arrays
Z1.st<-matrix(NA,nrow=n.comp,ncol=6)
Z2.st<-matrix(NA,nrow=n.comp,ncol=12)
XM.st<-matrix(NA,nrow=n,ncol=n.comp)
YM.st<-matrix(NA,nrow=n,ncol=n.comp)
NM.st<-matrix(NA,nrow=n,ncol=n.comp)
theta.x<-matrix(NA,nrow=n,ncol=n.comp)
theta.y<-matrix(NA,nrow=n,ncol=n.comp)
I<-diag(n) #identity matrix
#Check input
s<-(length(neighb)==dim(I)[1])
if (!s) stop("Error!  vector neighb has wrong length")
```

```
D<-diag(neighb*lambda.st)
t<-(length(neighb)==dim(wm)[1])
if (t)
C<-wm/neighb
else stop("Error!  vector neighb and weight matrix wm have different dimensions")
#Calculating variance matrix for structured heterogeneity
Q<-D*(I-C)
Q.x<-Q[1:n-1,1:n-1], Upper left (n-1) by (n-1) matrix of Q.x
for (i in seq(1,n.comp,by=1))
{
N.st<-round(runif(n,min.pop,max.pop),0)
V.un<-rnorm(n,mean=0,sd=sd.un)
x.val<-mvrnorm(1,rep(0,n-1),solve(Q.x)),
#Variance matrix is the inverse of Q.x
Un<-sum(x.val)/n
U.x<-x.val-Un
U<-c(U.x,Un)
theta.st.x<-exp(log(mu.x)+U+V.un)

#FOR MEAN
#Simulation based on changing mean level.
theta.st.y<-exp(log((1+(i-1)*mu.inflate)*mu.x)+U+V.un)
#FOR UNSTRUCTURED VARIATION
#Simulation based on changing unstructured variation.
theta.st.y<-exp(log(mu.x)+(1+(i-1)*sd.st.inflate)*U+V.un)
#FOR STRUCTURED VARIATION
#Simulation based on changing unstructured variation.
theta.st.y<-exp(log(mu.x)+U+(1+(i-1)*sd.un.inflate)*V.un)

#Proportions:

#Obtain counts for two maps from the Poisson distribution, get proportions and use
#image analysis and point processes methods to compare the maps
X.st[i,]<-rpois(n,N.st*theta.st.x)/N.st
Y.st[i,]<-rpois(n,N.st*theta.st.y)/N.st
```

```
theta.x[i,]<-theta.st.x
theta.y[i,]<-theta.st.y
XM.st[i,]<-X.st
YM.st[i,]<-Y.st
NM.st[i,]<-N.st
Z1.st[i,]<-IMAGE(X.st,Y.st)
Z2.st[i,]<-bigf(X.st,Y.st,wm)
}
#Output list(Z1.st=Z1.st,Z2.st=Z2.st,XM.st=XM.st,YM.st=YM.st,NM.st=NM.st)
}
#end of function

#Set parameters for simulations
set.seed(123)
n<-53
n.sims<-1000
n.comp<-31
mu.x<-1/1000
sd.un<-0.1
lambda.st<-100
mu.inflate<-1
sd.st.inflate<-1
sd.un.inflate<-0.10

#CARRY OUT SIMULATIONS
ZZ.st<-array(NA,c(n.sims,n.comp,12)), #array for results
dimnames(ZZ.st)<-list(NULL,NULL,c("MSD","PSNRR","PSNRM","SSIMR",
"SSIMM","SSIM","IRD","MSDI","MDD","AVND","MSND","MDND"))
for (i in 1:n.sims) {
Z3.st<-simul(n,n.comp,mu.x,lambda.st,sd.un,mu.inflate,sd.un.inflate,
sd.st.inflate,neighb=neighb,wm=wm)
ZZ3<-cbind(Z3.st$Z1.st, Z3.st$Z2.st)
ZZ.st[i,,]<-ZZ3
}
```

Appendix D

WinBUGS Codes

D.1 Chapter 4

D.1.1 Poisson-Gamma model

```
model{
for (i in 1:N)
```

where N is the number of regions.

Poisson likelihood for observed counts

```
y[i]~ dpois(mu[i])
mu[i]<-e[i]*theta[i]
```

Calculate relative risks

```
theta[i] ~ dgamma(a,b)
```

Calculate residuals

```
Residuals[i]<-(y[i]-e[i]*theta[i])/sqrt(e[i]*theta[i])
```

Calculate residual sum of squares

```
RSS[i]<-inprod(Residuals[],Residuals[])
}
```

Prior distribution for population parameters

```
ã dexp(0.1)
b̃ dexp(0.1)
```

Population mean and variance

```
mean<-a/b
var <-a/pow(b,2)
}
```

D.1.2 Log-normal model

```
model {
for (i in 1:N)
{
```

Poisson likelihood for observed counts

```
y[i]∼dpois(mu[i])
log(mu[i])<-log(e[i])+alpha+u[i]+v[i]
```

Relative risks

```
theta[i]<-exp(alpha+u[i]+v[i])
```

Residuals

```
Residuals[i]<-(y[i]-e[i]*theta[i])/sqrt(e[i]*theta[i])
```

Residual sum of squares

```
RSS[i]<-inprod(Residuals[],Residuals[])
```

Prior on uncorrelated heterogeneity

```
v[i]∼dnorm(0,tau.v)
}
eps<-1.0E-6
```

Prior on correlated heterogeneity where `sumNumNeigh` is the sum of number of neighbours for all regions.

```
u[1:N]∼car.normal(adj[],weights[],num[],tau.u)
for (k in 1:sumNumNeigh)
{
weights[k]<-1
}
```

Improper prior for the mean relative risk in the study region

```
alpha∼dflat()
mean<-exp(alpha)
varu<-1/tau.u
varv<-1/tau.v
stdeu<-sqrt(varu)
stdev<-sqrt(varv)
```

Hyperprior on inverse variances

```
tau.u∼dgamma(0.1,0.001)
tau.v∼dgamma(0.1,0.001)
}
```

D.1.3 Logistic model

```
model {
for (i in 1:N) {
```

Binomial likelihood for observed counts `y[i] dbin(p[i],n[i])`

```
logit(p[i])<-alpha+u[i]+v[i]
Residuals[i]<-(y[i]-n[i]*p[i])/sqrt(n[i]*p[i])
RSS[i]<-inprod(Residuals[],Residuals[])
```

Prior on uncorrelated heterogeneity

```
v[i]∼dnorm(0,tau.v)
}
eps<-1.0E-6
```

Prior on correlated heterogeneity where `sumNumNeigh` is the sum of number of neighbours for all regions

```
u[1:N]∼car.normal(adj[],weights[],num[],tau.u)
for (k in 1:sumNumNeigh)
{
weights[k]<-1
}
```

Improper prior for the mean relative risk in the study region

```
alpha∼dflat()
mean<-exp(alpha)
```

Variability and standard deviations for correlated and uncorrelated heterogeneity.

```
varu<-1/tau.u
varv<-1/tau.v
stdeu<-sqrt(varu)
stdev<-sqrt(varv)
```

Prior for the inverse variances

```
tau.u∼dgamma(0.1,0.001)
tau.v∼dgamma(0.1,0.001)
}
```

D.1.4 Logistic model with variables

```
model {
for (i in 1:N)
{
```

Binomial likelihood for observed counts with census variables born other EU (bornoeu), working in agriculture (workagr), no car (nocar), overcrowding (overcr), unemployment (unemp) and low social class (lowsc)

```
y[i] dbin(p[i],n[i])
logit(p[i])<-alpha+u[i]+v[i]+beta1*bornoeu[i]+
beta2*workagr[i]+beta3*nocar[i]+beta4*overcr[i]+beta5*unemp[i]
+beta6*lowsc[i]
```

Prior on uncorrelated heterogeneity

```
v[i]∼dnorm(0,tau.v)
}
eps<-1.0E-6
```

Prior on correlated heterogeneity where `sumNumNeigh` is the sum of number of neighbours for all regions

```
u[1:N]~car.normal(adj[],weights[],num[],tau.u)
for (k in 1:sumNumNeigh)
{
weights[k]<-1
}
```

Prior on regression coefficients

```
beta1~dnorm(0.0,1.0E-6)
beta2~dnorm(0.0,1.0E-6)
beta3~dnorm(0.0,1.0E-6)
beta4~dnorm(0.0,1.0E-6)
beta5~dnorm(0.0,1.0E-6)
beta6~dnorm(0.0,1.0E-6)
```

Improper prior for the mean relative risk in the study region

```
alpha~dflat()
mean<-exp(alpha)
```

Variability and standard deviations for correlated and uncorrelated heterogeneity

```
varu<-1/tau.u
varv<-1/tau.v
stdeu<-sqrt(varu)
stdev<-sqrt(varv)
```

Hyperprior on inverse variances

```
tau.u~dgamma(0.1,0.001)
tau.v~dgamma(0.1,0.001)
}
```

D.1.5 Waller et al. (1997) space-time model

```
model {
for (t in 1:T)
```

where `T` is total time points.

```
for (i in 1:N)
{
```

Binomial likelihood for observed counts

```
y[i,t]∼dbin(p[i,t],n[i,t])
logit(p[i,t])<-alpha+u[i,t]+v[i,t]
}
}
```

Prior on correlated heterogeneity

```
u1[1:N]∼car.normal(adj[],weights[],num[],tau.u[1])
u2[1:N]∼car.normal(adj[],weights[],num[],tau.u[2])
u3[1:N]∼car.normal(adj[],weights[],num[],tau.u[3])
u4[1:N]∼car.normal(adj[],weights[],num[],tau.u[4])
u5[1:N]∼car.normal(adj[],weights[],num[],tau.u[5])
u6[1:N]∼car.normal(adj[],weights[],num[],tau.u[6])
for (i in 1:N)
u[i,1]<-u1[i]
u[i,2]<-u2[i]
u[i,3]<-u3[i]
u[i,4]<-u4[i]
u[i,5]<-u5[i]
u[i,6]<-u6[i]
```

Prior on uncorrelated heterogeneity

```
} for (i in 1:N)
{
v1[i]∼dnorm(0,tau.v[1])
v2[i]∼dnorm(0,tau.v[2])
v3[i]∼dnorm(0,tau.v[3])
v4[i]∼dnorm(0,tau.v[4])
v5[i]∼dnorm(0,tau.v[5])
v6[i]∼dnorm(0,tau.v[6])
}
for (i in 1:N)
{
v[i,1]<-v1[i]
```

```
v[i,2]<-v2[i]
v[i,3]<-v3[i]
v[i,4]<-v4[i]
v[i,5]vv5[i]
v[i,6]<-v6[i]
}
for (k in 1:sumNumNeigh)
{ weights[k]<-1
```

Improper prior for the mean relative risk in the study region

```
alpha∼dflat()
mean<-exp(alpha)
```

Hyperprior on inverse variances

```
for (i in 1:T)
{
tau.u[i]∼dgamma(0.1,0.001)
tau.v[i]∼dgamma(0.1,0.001)
```

Variability and standard deviations for correlated and uncorrelated heterogeneity.

```
varu[i]<-1/tau.u[i]
varv[i]<-1/tau.v[i]
stdeu[i]<-sqrt(varu[i])
stdev[i]<-sqrt(varv[i]) }
```

Bibliography

[1] Acland HW (1856), *Memoir of the Cholera at Oxford in the Year 1854*. London: J.Churchill.

[2] Baird G, Pickles A, Simonoff E, Charman T, Sullivan P, Chandler S, Loucas T, Meldrum D, Afzal M, Thomas B, Jin L and Brown D (2008), Measles vaccination and antibody response in autism spectrum disorders. *Archives of Disease in Childhood* **93**, 832-837.

[3] Baker R (1833), Report of Leeds Board of Health, Leeds, UK.

[4] Beaglehole R, Bonita R and Kjellstrom T (1993), *Basic Epidemiology*. World Health Organisation.

[5] Berke O (2004), Exploratory disease mapping: kriging the spatial risk function from regional count data. *International Journal of Health Geographics* **3**, 3:18.

[6] Bernardinelli LD, Clayton DG and Montomoli C (1992), Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine* **11**, 983-1007.

[7] Bernardinelli LD, Clayton DG, Pascutto C, Montomoli C, Ghislandi M and Songini M (1995a), Bayesian analysis of space-time variation in disease risk. *Statistics in Medicine* **14**, 2433-2443.

[8] Bernardinelli LD, Clayton DG and Montomoli C (1995b), Bayesian estimates of disease maps: how important are priors? *Statistics in Medicine* **14**, 2411-2431.

[9] Besag J, York J and Mollié A (1991), Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**, 1-59.

[10] Besag J and Kooperberg C (1995), On conditional and intrinsic autoregressions. *Biometrika* **82**, 733-746.

[11] Best N, Richardson S and Thomson A (2005), A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research* **14**, 35-59.

[12] Biggeri A, Dreassi E, Lagazio C and Bohning D (2003), A transitional non-parametric maximum pseudo-likelihood estimator for disease mapping. *Computational Statistics and Data Analysis* **41**, 617-629.

[13] Böhning D, Dietz E and Schlattmann P (2000), Space-time mixture modelling of public health data. *Statistics in Medicine* **19**, 2333-2344.

[14] Bouzerdoum A, Havstad A and Beghdadi A (2004), Image quality assessment using a neural network approach. *Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology*, 330-333. *http://ro.uow.edu.au/infopapers/43*.

[15] Bowman AW and Azzalini A (1997), *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-plus Illustrations*. New York: Oxford University Press.

[16] Breslow N and Day N (1987), *Statistical Methods in Cancer Research. Volume 2. The Design and Analysis of Cohort Studies*. Lyon, International Agency for Research on Cancer.

[17] Breslow NE and Clayton DG (1993), Approximate inference in generalized linear mixed models. *Journal of the American Statististical Association* **88**, 9-25.

[18] Brillinger DR (1990), Spatio-temporal modelling of spatially aggregated birth data. *Survey Methodology* **16**, 255-269.

[19] Brooks S and Gelman AE (1998), General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**, 434-455.

[20] Carlin BP and Louis TA (1996), *Bayes and Empirical Bayes Methods for Data Analysis*. London, Chapman and Hall.

[21] Carrat F and Valleron A-J (1992), Epidemiologic mapping using 'kriging' method: application to an influenza-like illness epidemic in France. *American Journal of Epidemiology* **135**, 1293-1300.

[22] Cartwright SA (1826), A series of essays on the causes, symptoms, morbid anatomy, and treatment of some of the principal diseases of the southern state. *The Medical Recorder* **9** (3-44, 225-267), 4-7. Cited by Stevenson (1965) *op. cit.*

[23] Casella G and George EI (1992), Explaining the Gibbs sampler. *The American Statistician* **46**, 167-174.

[24] Chen M, Shao Q and Ibrahim J (2000), *Monte Carlo Methods in Bayesian Computation*. New York: Springer Verlag.

[25] Clayton DG and Kaldor J (1987), Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. *Biometrics* **43**, 671-691.

[26] Copperthwaite NH (1972), *Mortality or Morbidity Mapping: Some Examples from Yugoslav Macedonia*. In McGlashan ND (eds), *Medical Geography: Techniques and Field Studies*. London: Methuen.

[27] Cook D and Pocock S (1983), Multiple regression in geographical mortality studies with spatially correlated errors. *Biometrics* **39**, 361-371.

[28] Cressie N (1993), *Statistics for Spatial Data*. New York: John Wiley and Sons.

[29] Croner CM and De Cola L (2001), Visualization of disease surveillance data with geostatistics. *www.unece.org/stats/documents/2001/09/gis/25.e.pdf* .

[30] Dabney AR and Wakefield JC (2005), Issues in the mapping of two diseases. *Statistical Methods in Medical Research* **15**, 83-112.

[31] Dean CB, Ugarte MD and Militino AF (2004), Penalized quasi-likelihood with spatially correlated data. *Computational Statistics and Data Analysis* **45**, No. 2, 235-248.

[32] Demichell J, Jefferson T, Rivetti A, and Price D (2005), Vaccines for measles, mumps and rubella in children. *Cochrane Database of Systematic Reviews* **4**, CD004407.

[33] Dempster AP, Laird NM and Rubin DB (1977), Maximum likelihood for incomplete data via the EM algorithm (with Discussion). *Journal of the Royal Statistical Society, B* **39**, 1-38.

[34] Deneke T (1895), Nachtrgliches zur Hamburger cholera-Epidemie von 1892. *Munchener Medicinische Wochenschrift* **41**, 957-961.

[35] Diggle PJ (1993), Point process modelling in environmental epidemiology. In: Barnett V and Turkman K (eds), *Statistics in the Environment (SPRUCE)*. New York: John Wiley and Sons, vol. 1.

[36] Diggle P, Tawn J and Moyeed R (1998), Model-based geostatistics. *Journal of the Royal Statistics Society C* **47**, 299-350.

[37] Diggle PJ (2003), *Statistical Analysis of Spatial Point Patterns*. London: Academic Press Inc.

[38] Donnelly CA (1995), The spatial analysis of covariates in a study of environmental epidemiology. *Statistics in Medicine* **14**, 2393-2409.

[39] Dorling D (2008), Looking from outside the goldfish bowl. *Significance* **5**, No.3 122-125 (4).

[40] Durbin J (1973), *Distribution theory for tests based on the sample distribution function*. Philadelphia: SIAM.

[41] Fan J and Gijbels I (1992), Variable bandwidth and local linear regression smoothers. *The Annals of Statistics* **4**, 2008-2036.

[42] Friederichs V, Cameron JC and Robertson C (2006), Impact of adverse publicity on MMR vaccine uptake: A population based analysis of vaccine uptake records for one million children, born 1987-2004. *Archives of Disease in Childhood* **91**, 465-468.

[43] Fortin M, Dale MRT and ver Hoef J (2002), Spatial analysis in ecology. *Encyclopedia of Environmetrics* **4**, 2051-2058. Chichester: John Wiley and Sons.

[44] Gardner MJ, Winter PD, Taylor CP and Acheson ED (1983), *Atlas of Cancer Mortality in England and Wales 1968-1978*. Chichester: John Wiley and Sons.

[45] Gardner MJ, Winter PD and Barker DJP (1984), *Atlas of Disease Mortality from Selected Diseases in England and Wales 1968-1978*. Chichester: John Wiley.

[46] Gebhart F (1998), Survey on cluster tests for spatial area data. *dhf.ddc.moph.go.th/abstract/s25.pdf*.

[47] Gelman AE and Rubin D (1992), Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science* **7**, 457-511.

[48] Gelman A (2002), Prior distributions. In: El-Shaarawi AH and Piegorsch WW (eds), *Encyclopedia of Environmetrics*. Chichester: John Wiley and Sons, Vol.3, 1634-1637.

[49] Gelman A, Carlin JB, Stern HS and Rubin DB (2004), *Bayesian Data Analysis*. London: Chapman and Hall.

[50] Gelman A (2006), Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* **1**, No. 3, 515-533.

[51] Geman S and Geman D (1984), Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-741.

[52] Gilbert EW (1958), Pioneer maps of health and disease in England. *Geographical Journal* **124**, 172-183.

[53] Gilks WR, Richardson S and Spiegelhalter DJ (eds) (1996), *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

[54] Glasbey CA and Horgan GW (1995), *Image Analysis for the Biological Sciences*. Chichester: John Wiley and Sons.

[55] Glick BJ (1979), Mortality mapping-origins.

http://zappa.nku.edu/∼longa/geomed/ modules/av/lab/study.html.

[56] Green PJ and Silverman BW (1994), *Nonparametric Regression and Generalised Linear Models*, 2nd edn. London: Chapman and Hall.

[57] Green PJ and Richardson S (2002), Hidden Markov models and disease mapping. *Journal of the American Statistical Association* **97**, 1055-1070.

[58] Hagen A (2002), Comparison of maps containing nominal data. *Statistics in Medicine* **12**, 1895-1913.

[59] Hagen A (2007), Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science* **17**, No.3, 235-249.

[60] Härdle W (1990), *Applied Nonparametric Regression.* New York: Cambridge University Press.

[61] Hargrove WW, Hoffman FM and Hessburg PF (2006), Mapcurves: a quantitative method for comparing categorical maps. *Journal of Geographical Systems* **8**, 187-208.

[62] Harty W (1820), *A Historic Sketch of the Causes, Progress, Extent and Mortality of the Contagious Fever Epidemic in Ireland During the Years 1817, 1818 and 1819.* Dublin: Hodges and McArthur.

[63] Harville DA (1977), Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Society* **72**, 320-340.

[64] Hastings WK (1970), Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97-109.

[65] Haviland A (1875), *The Geographical Distribution of Heart Disease and Dropsy, Cancer in Females and Phthisis in Females in England and Wales*, 2nd edn. London: Swan Sonnenschein.

[66] Heisterkamp SH, Doornbos G and Gankema M (1993), Disease mapping using empirical Bayes and Bayes methods on mortality statistics in the Netherlands. *Statistics in Medicine* **12**, 1895-1913.

[67] Held L, Natário I, Fenton SE, Rue H and Becker N (2005), Towards joint disease mapping. *Statistical Methods in Medical Research* **14**, No.1, 61-82.

[68] Howe GM (1963), *National Atlas of Disease Mortality in the United Kingdom.* London: Nelson.

[69] Howe GM (1986), Disease Mapping. In: Pacione M (ed), *MEDICAL GEOGRAPHY: Progress and Prospect*. Britain: Biddles Ltd, Guildford and King's Lynn.

[70] Howe GM (1971), The mapping of disease in history. In: Clarke E (ed.), *Modern Methods in the History of Medicine*. London: University of London, Athlone Press, Ch. 20.

[71] Howe GM (1972), *Man, Environment and Disease in Britain. A Medical Geography through the Ages*. New York: Barnes and Noble.

[72] Hsiao CK, Tzeng J and Wang C (2000), Comparing the performance of two indices for spatial model selection: Application to two mortality data. *Statistics in Medicine* **19**, 1915-1930.

[73] Jackson C, Best N and Richardson S (2006), Improving ecological inference using individual-level data. *Statistics in Medicine* **25**, 2136-2159.

[74] Kafadar K (1996), Smoothing geographical data, particularly rates of disease. *Statistics in Medicine* **15**, 2539-2560.

[75] Kavanagh K and Robertson C (2008), Spatial variation in NHS24 call uptake and NHS24 reports of the rate of influenza-like illness and diarrhoea and vomiting. *http://www.stams.strath.ac.uk/ kim/SpatialWriteUp2.pdf*

[76] Kelsall J and Diggle PJ (1998), Spatial variation in risk of disease: a nonparametric binary regression approach. *Applied Statistics* **47**, 559-573.

[77] Kelsall J and Wakefield J (2002), Modelling spatial variation in disease risk: a geostatistical approach. *Journal of the American Statistical Association* **97**, 692-701.

[78] Kemp I, Boyle P, Smans M and Miur C (1985), *Atlas of Cancer in Scotland 1975-1980: Incidence and Epidemiological Perspective*. Lyon: IARC Scientific Publications.

[79] Knorr-Held L and Rasser G (2000), Bayesian detection of clusters and discontinuities in disease maps. *Biometrics* **56**, 13-21.

[80] Knorr-Held L and Best N (2001), A shared component model for detecting joint and selective clustering of two disease. *Journal of the Royal Statistical Society, A* **164**, 73-85.

[81] Lajaunie C (1991), Local risk estimation for a rare non-contagious disease based on observed frequencies. Note N-36/91/G. Fontainebleau: Centre de Gostatistique, Ecole des Mines de Paris.

[82] Lancaster P and Salkauskas K (1986), *Curve and Surface Fitting: An Introduction*. London: Academic Press.

[83] Last JM (1988), *A Dictionary of Epidemiology*, 2nd edn. Oxford: Oxford University Press.

[84] Lawson AB, Biggeri AB, Boehning D, Lesaffre E, Viel J-F, Clark A, Schlattmann P and Divino F (2000), Disease mapping models: An empirical evaluation. *Statistics in Medicine* **19**, 2217-2241.

[85] Lawson AB (2001), *Statistical Methods in Spatial Epidemiology*. Chichester: John Wiley and Sons.

[86] Lawson AB and Williams FLR (2001), *An Introductory Guide to Disease Mapping*. Chichester: John Wiley and Sons.

[87] Lawson AB and Clark A (2002), Spatial mixture relative risk models applied to disease mapping. *Statistics in Medicine* **21**, 359-370.

[88] Lawson AB, Browne WJ and Vidal Rodeiro CL (2003), *Disease Mapping with WinBUGS and MLwiN*. Chichester: John Wiley and Sons.

[89] Learmonth ATA (1972), Atlas in Medical Geography 1950-1970. A Review. In: McGlashan ND (ed), *Medical Geography: Techniques and Field Studies*, pp. 133-152. London: Methuen.

[90] Leonard T (1975), Bayesian estimation methods for two-way contingency tables. *Journal of the Royal Statistical Society*, B **37**, 23-37.

[91] Leroux BG (2000), Modelling spatial disease rates using maximum likelihood. *Statistics in Medicine* **19**, 2321-2332.

[92] Leyland AH and Davies CA (2005), Empirical Bayes methods for disease mapping. *Statistical Methods in Medical Research*, **14**, 17-34.

[93] Lilienfeld DE and Stolley PD (1994), Foundations of Epidemiology, 3rd edn. New York: Oxford University Press.

[94] Lloyd OL, Williams FLR, Berry WG and Florey CD (1987), *An Atlas of Mortality in Scotland*. London: Croom Helm.

[95] Lyman O (1988), *An Introduction to Statistical Methods and Data Analysis*, 3rd ed. Boston: PWS Kent Publishing.

[96] MacEarcheran AM (1995), *How Maps Work: Representation, Visualisation and Design*. New York: Guildford Press.

[97] MacNab YC and Dean CB (2001), Autoregressive spatial smoothing and temporal smoothing B-splines for mapping rates. *Biometrics* **57**, 949-956.

[98] MacNab YC and Dean CB (2002), Spatio-temporal modelling of rates for the construction of disease maps. *Statistics in Medicine* **21**, 347-358.

[99] McNeill L (1991), Interpolation of smoothing of binomial data for the Southern African bird atlas project. *South African Statistical Journal* **25**, 129-136.

[100] Manton K, Woodbury M and Stallard E (1981), A variance component approach to categorical data models with heterogeneous mortality rates in North Carolina counties. *Biometrics* **37**, 259-269.

[101] Marshall RJ (1991), Mapping disease and mortality rates using empirical Bayes estimators. *Applied Statistics* **40**, 283-294.

[102] Marshall RJ (1991), A review of methods for the statistical analysis of spatial patterns of disease. *Journal of the Royal Statistical Society A* **154**, 421-441.

[103] May JM (1955), World Atlas of Diseases. *American Geographical Society*.

[104] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH and Teller E (1953), Equations of state calculations by fast computing machine. *Journal of Chemical Physics* **21**, 1087-1091.

[105] Militino AF, Ugarte MD and Dean CB (2001), The use of mixture models for identifying high risks in disease mapping. *Statistics in Medicine* **20**, 2035-2049.

[106] Monmonier M (1996), *How to Lie with Maps*, 2nd edn. London: University of Chicago Press.

[107] Mulopulos GP, Hernandez AA and Gasztonyi LS (2003), Peak signal to noise performance comparison of JPEG and JPEG 2000 for various medical image modalities. *Symposium on Computer Applications*.

[108] Odland J (1988), *Spatial Autocorrelation*, Newbury Park: SAGE Publications, Inc.

[109] Oliver MA, Muir KR, Webster R, Parkes SE, Cameron AH, Stevens MCG and Mann JR (1992), A geostatistical approach to the analysis of pattern in rare disease. *Journal of Public Health Medicine* **14**, 280-289.

[110] Oliver MA, Webster R, Lajaunie C and Muir KR (1998), Binomial cokriging for estimating and mapping the risk of childhood cancer. *Journal of Mathematics Applied in Medicine and Biology* **15**, 279-297.

[111] Ormerod WP (1848), *On the Sanatory [sic] Condition of Oxford*. Oxford: The Ashmolean Society.

[112] Oslon JM (1976), Noncontiguous area cartograms. *Professional Geographer* **28**, 371-380.

[113] Pearce A, Law C, Elliman D, Cole TJ and Bedford H (2008), Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study. *British Medical Journal* **336**, 754-757.

[114] Petermann AH (1852), *Cholera Map of the Britain Isles, Showing the Districts Attacked in 1831, 1832 and 1833*. London: John Betts.

[115] Pickle L and Herman D (1995), *Cognitive aspects of statistical mapping. CAM program volume*. Working Paper Series 18, NCHS Office of Research and Methodology, pp 323.

[116] Pontius RG (2002), Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing* **68**, No. 10, 1041-1049.

[117] Richardson S (2003), Spatial models in epidemiological applications. In: Green P, Hjort N and Richardson S (eds), *Highly Structured Stochastic Systems*. London: Oxford University Press.

[118] Richardson S, Thomson A, Best N and Elliot P (2004), Interpreting posterior relative risk estimates in disease-mapping studies. *Environmental Health Perspectives* **15**, 385-407.

[119] Richardson S, Abellan JJ and Best N (2006), Bayesian spatial-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). *Statistical Methods in Medical Research* **112**, 1016-1025.

[120] Ripley BD (1981), *Spatial Statistics*. New York: John Wiley and Sons.

[121] Roberts GO (1995), Markov chain concepts related to sampling algorithms. In: Gilks WR, Richardson S and Spiegelhalter DJ (eds) (1996), *Markov Chain Monte Carlo in Practice*, 45-57. London: Chapman and Hall.

[122] Robert CP and Casella G (1999), *Monte Carlo Statistical Methods*. New York: Springer Verlag.

[123] Rodenwaldt E and Jusatz HH (1961), *Weltseuchatlas* (Atlas of Epidemic disease). Hamburg: Falk.

[124] Rosenberg MS, Sokal RR, Oden NL and DiGiovanni D (1999), Spatial autocorrelation of cancer in Wetern Europe. *European Journal of Epidemiology* **15**, 15-21.

[125] Samet A, Ayed MAB, Masmoudi N and Khriji L (2005), New perceptual image quality assessment metric. *Asian Journal of Information Technology* **11**, 996-1000.

[126] Schlattmann P and Böhning D (1993), Mixture models and disease mapping. *Statistics in Medicine* **12**, 1943-1950.

[127] Schulman J, Selvin S and Merrill DW (1988), Density equalised map projections: a method for analysing clustering around a fixed point. *Statistics in Medicine* **7**, 491-505.

[128] Seaman V (1798), An enquiry into the cause of the prevalence of the yellow fever in New York. *The Medical Repository, New York*, **1**, 315-372.

[129] Shapter T (1849), *The History of the Cholera in Exeter in 1832.* London: J. Churchill.

[130] Shnayderman A, Gusev A and Eskicioglu AM (2003), An SVD-based grayscale image quality measure for local and global assessment. *IEEE Transactions on Image Processing* **15**, No.2, 422-429.

[131] Simonoff JS (1996), *Smoothing Methods in Statistics.* New York: Springer-Verlag.

[132] Snow J (1854), *On the Mode of Communication of Cholera, 2nd edn.* London: Churchill Livingstone.

[133] Sokal RR and Uytterschaut H (1987), Cranial variation in European populations: a spatial autocorrelation study at three time periods. *American Journal of Physical Anthropology* **74**, 21-38.

[134] Spiegelhalter DJ, Thomas A, Best NG, Gilks WR and Lunn D (1994, 2003), BUGS: Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, England. *www.mrc-bsu.cam.ac.uk/bugs/.*

[135] Spiegelhalter DJ, Best NG, Gilks WR and Inskip H (1996), Hepatitis B: a case study in MCMC methods. In: Gilks WR, Richardson S and Spiegelhalter DJ (eds) (1996) *Markov Chain Monte Carlo in Practice*, 21-43. London: Chapman and Hall.

[136] Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A (2002), Bayesian deviance, the effective number of parameters and the comparison of arbitrarily complex models. *Journal of the Royal Statistical Society, B*, **64**, 583-640.

[137] Steel DG and Holt D (1996), Analysing and adjusting aggregation effects: The ecological fallacy revisited. *International Statistical Review*, **64**, 39-60.

[138] Steel DG, Tranmer M and Holt D (2006), Unravelling ecological analysis. *Journal of Applied Mathematics and Decision Sciences*, 1-8.

[139] Stehman SV (1999), Comparing thematic maps based on map value. *International Journal of Remote Sensing* **20**, No. 12, 2347-2366.

[140] Stevenson L (1965), Putting disease on the map: the early use of spot maps in the study of yellow fever. *Journal of the History of Medicine* **20**, 227-261.

[141] Tierney L (1995), Introduction to general state-space Markov chain theory. In: Gilks WR, Richardson S and Spiegelhalter DJ (eds) (1996) *Markov Chain Monte Carlo in Practice*, 59-74. London: Chapman and Hall.

[142] Tsutakawa RK (1988), Mixed model for analysing geographic variability in mortality rates. *Journal of the American Statistical Association* **83**, 37-42.

[143] Ugarte MD, Ibaanez B and Militino AF (2006), Modelling risks in disease mapping. *Statistical Methods in Medical Research* **15**, 21-35.

[144] Visser H and de Nijs T (2006), The map comparison kit. *Environmental Modelling and Software* **21**, 346-358.

[145] Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillon AP, Thomson MA, Harvey P, Valentine A, Davies SE and Walker-Smith JA, (1998), Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* **351**, 637-641.

[146] Wakefield J (2006), Disease Mapping and Spatial Regression with Count Data *http://www.bepress.com/uwbiostat/paper286*.

[147] Waller LA and Gotway CA (2004), *Applied Spatial Statistics for Public Health Data*, New Jersey: John Wiley and Sons.

[148] Waller LA, Carlin BP, Xia H and Gelfand AE (1997), Hierarchical spatiotemporal mapping of disease rates. *Journal of the American Statistical Association* **92**, 607-617.

[149] Walter S (1993), Visual and statistical assessment of spatial clustering in mapped data. *Statistics in Medicine* **12**, 1275-1291.

[150] Wand MP and Jones MC (1995), *Kernel Smoothing*, London: Chapman and Hall.

[151] Wang Z and Bovik AC (2002), A universal image quality index. *IEEE Signal Processing Letters* **9**, 81-84.

[152] Wang Z, Bovik AC, Sheikh HR and Simoncelli EP (2004), Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **14**, 600-611.

[153] Webster R, Oliver MA, Muir KR, Mann JR (1994), Kriging the local risk of a rare disease from a register of diagnoses. *Geographical Analysis* **26**, 168-185.

[154] White R (2006), Pattern based map comparison. *Journal of Geographical Systems* **8**, No.2, 145-164.

[155] Wright JA and Polack C (2006), Understanding variation in measles mumps rubella immunisation coverage. A population based study. *European Journal of Public Health* **16**, No.2, 137-142.