
Article:
Lee, D., Minton, J. and Pryce, G. orcid.org/0000-0002-4380-0388 (2015) Bayesian inference for the dissimilarity index in the presence of spatial autocorrelation. SPATIAL STATISTICS, 11. pp. 81-95. ISSN 2211-6753

https://doi.org/10.1016/j.spasta.2014.12.001

Spatial Statistics 11 (2015) 81–95

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Bayesian inference for the dissimilarity index in the presence of spatial autocorrelation

Duncan Lee [a,*], Jon Minton [b], Gwilym Pryce [c]

[a] School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8QW, UK
[b] Urban Studies, School of Social and Political Sciences, University of Glasgow, Glasgow, G12 8RS, UK
[c] Sheffield Methods Institute, University of Sheffield, Sheffield, S10 2TN, UK

ARTICLE INFO

Article history:
Received 15 July 2014
Accepted 12 December 2014
Available online 13 January 2015

Keywords:
Bayesian hierarchical modelling
Conditional autoregressive models
Dissimilarity index
Spatial autocorrelation

ABSTRACT

The degree of segregation between two or more sub-populations has been studied since the 1950s, and examples include segregation along racial and religious lines. The Dissimilarity index is a commonly used measure to numerically quantify segregation, using population level data for a set of areal units that comprise a city or country. However, the construction of this index usually ignores the spatial autocorrelation present in the data, and it is also typically presented without a measure of uncertainty. Therefore we propose a Bayesian hierarchical modelling approach for estimating the Dissimilarity index and quantifying its uncertainty, which utilises a conditional autoregressive model to account for the spatial autocorrelation in the data. This modelling approach is motivated by a study of religious segregation in Northern Ireland, and allows us to quantify whether the dissimilarity index has exhibited a substantial change between 2001 and 2011.

1. Introduction

In the absence of legally enforced segregation, there are social processes at work that cause an uneven distribution of households by income, race and religion. Some argue (e.g. Cheshire, 2009) that

* Corresponding author. Tel.: +44 141 330 4047.
E-mail address: Duncan.Lee@glasgow.ac.uk (D. Lee).

http://dx.doi.org/10.1016/j.spasta.2014.12.001

self-segregation is no bad thing, because it gives rise to specialised communities generating greater neighbourhood variety in our cities. In addition, public health researchers (such as Whitley et al., 2006) have identified evidence of ethnic density effects, suggesting that minority ethnic groups living close together may have better health than if they integrate more with the majority population. The counter argument is that segregation may reduce affinity and understanding between social groups, and thereby undermine social cohesion. Whether or not segregation is rising or falling is therefore an important empirical question. For example, the publicity surrounding concerns expressed by Trevor Phillips (the former leader of the Commission for Racial Equality) in 2005 that Britain was sleepwalking into segregation – becoming more divided by race and religion – reflected wider anxieties about social fragmentation. However, such claims have been challenged by Simpson (2007) and Simpson and Finney (2010) and others Jivraj (2012), Catney (2013) and Johnston et al. (2013), who provide evidence that segregation may actually be falling. Parallel debates and concerns have occurred in USA, Europe and elsewhere, giving rise to a truly voluminous literature on the meaning and measurement of segregation (Clark, 1986; Glaster, 1988; Ihlanfeldt and Scafidi, 2002; Musterd, 2005; Semyonov and Glikman, 2009).

Measuring segregation numerically is an inherently difficult task, which is typically undertaken using population level data from a set of n non-overlapping areal units comprising a city or country. Typically, segregation measures quantify the extent to which two or more sub-populations are integrated and live together or are isolated and do not interact. Numerous different indices of segregation have been proposed in the literature, and the widely cited review by Massey and Denton (1988) in 1988 categorised segregation indices into five different dimensions: (i) evenness—the level of variation in the relative size of the minority sub-population across the n areal units; (ii) exposure—the extent of the interaction between the minority and majority sub-populations; (iii) concentration—the relative physical amount of space occupied by each sub-population; (iv) centralisation—the relative degrees to which each sub-population are based in the centre of the city; and (v) clustering—the degree to which each sub-population clusters together in geographically close areal units. Numerous extensions have been proposed to these indices in the literature since this seminal critique in 1988, including having more than two sub-populations (Reardon and Firebaugh, 2002; Reardon and O'Sullivan, 2004), and addressing the modifiable areal unit problem (MAUP, Wong, 2003 and Simpson, 2007).

In this paper we consider the Dissimilarity index (Duncan and Duncan, 1955), which is one of the most widely computed indices of residential segregation. We use this index purely to motivate the issues and modelling approaches discussed in this paper, but are in no way attempting to justify its use over alternative measures such as the Gini index. Rather, our view is given a desire to compute a particular index, what are the statistical issues that should be addressed when doing so. Specifically we focus on two such issues, which have largely been ignored by the existing segregation literature. The first is that the index is a purely descriptive summary statistic, and is typically presented without a corresponding measure of uncertainty. However, as argued by Leckie et al. (2012) the quantification of its uncertainty would enable researchers to determine whether observed differences in the Dissimilarity index over space or time correspond to real changes in segregation, or simply the result of random sampling variation. A small number of papers have attempted to address this issue, using either a bootstrapping algorithm (Brülhart and Traeger, 2005) or asymptotic theory (Cortese et al., 1976; Winship, 1977; Inman and Bradley, 1991). However, Mulekar et al. (2008) compared a number of these asymptotic theory approaches, and concluded that the proposals cannot be replied upon to yield correct confidence intervals.

The second issue we consider in this paper is the impact of spatial autocorrelation on the construction of the Dissimilarity index, a problem that to our knowledge is yet to be addressed in this context. We note that existing research has altered the algebraic form of the dissimilarity index to account for spatial features such as boundary effects (see Morrill, 1991 and Wong, 1993), but that is not the attempt of this paper. Instead, we consider the standard formula for the Dissimilarity index, and argue that its estimation and uncertainty quantification should be adjusted to allow for the spatial autocorrelation in the data. This is because the sample proportions used to compute the index are subject to sampling variation and other errors, and the true unknown proportions can be better estimated by using the spatial autocorrelation in the data to facilitate a borrowing of strength in the estimation, which should yield more reliable inference.

The application that motivated this paper is religious segregation in Northern Ireland and the extent to which it has changed between 2001 and 2011, an issue that has received a great deal of attention recently (see for example Shuttleworth and Lloyd, 2009 and Shuttleworth and Green, 2011). Specifically, we investigate whether the apparent fall in segregation in Northern Ireland reported in the Economist (2013) is indicative of a genuine decline in social fragmentation, or whether is it simply the result of sampling variation caused by random churn in the household moves. The Northern Ireland data we model in this paper are presented in Section 2, while Section 3 provides background to urban segregation indices and critiques their limitations from a spatial modelling perspective. In Section 4 of this paper we propose a solution to the spatial autocorrelation problem based on a Bayesian hierarchical model, and illustrate how it can be used to produce posterior predictive distributions and hence estimates and 95% credible intervals for segregation indices such as the Dissimilarity index. Section 5 quantifies the performance of the Bayesian spatial model proposed here, and compares it to the simple non-spatial estimate currently used. Section 6 presents the results of our Northern Ireland study, while Section 7 concludes the paper by summarising our key findings and suggesting future research avenues.

2. Northern Ireland study

Data from the 2001 and 2011 censuses were used to estimate and compare the levels of religious segregation in Northern Ireland, and were downloaded from the Northern Ireland Neighbourhood Information Service (NINIS) website. The religion of the head of household was recorded in both years. The responses were divided into different numbers of mutually exclusive categories in each census, but the category 'Catholic' was common to both years of data. The responses were therefore converted into the binary categories Catholic/Not Catholic to allow comparability between the censuses. The number of heads of household who identified as Catholic and the total number of heads of households were reported for the 890 super output area (SOA) that comprise Northern Ireland.

Fig. 1 displays the raw proportion of people in each SOA who are Catholic in 2001 and 2011, and in both cases the left plot is for Northern Ireland while the right plot zooms in on Belfast. The sample proportions visually exhibit substantial spatial autocorrelation, with areal units close together tending to exhibit similar proportions. This is confirmed by a Moran's I permutation test (Moran, 1950) for spatial autocorrelation based on 100,000 permutations, which yields highly significant I statistics equal to 0.691 (p-value 0.00001) and 0.719 (p-value 0.00001) for 2001 and 2011 respectively. The existence of this high spatial autocorrelation suggests that the estimation and uncertainty quantification of the set of true unknown proportions in each SOA should account for the spatial autocorrelation in the observed data.

However, Fig. 1 also suggests that the spatial autocorrelation is localised, because there are pairs of areal units that are geographically adjacent but have very different catholic proportions. These locations where predominately catholic and protestant populations neighbour each other violate the assumption of spatial autocorrelation, as the spatial surface of the proportion of catholics exhibits a step change. These step changes are known as boundaries in the related field of disease mapping (see for example Lee and Mitchell, 2012), and suggest that commonly used global spatial smoothing models would be inappropriate to capture the complex spatial structure in these data. Therefore, in Section 4 we discuss a commonly used global spatial smoothing model, and then propose an extension that can capture both sub-regions of localised smoothness and step changes between different communities. However, before that we provide a brief critique of the Dissimilarity index in Section 3.

3. Dissimilarity index

Segregation indices in their most basic form are used to summarise the level of segregation or mixing between two sub-populations, one of which is typically considered to be the *minority sub-population* while the other is the *majority sub-population*. In this paper we focus on the *Dissimilarity index* (Duncan and Duncan, 1955), which is the most commonly used index of segregation and is a measure of the evenness of the distribution of the minority population across a city. We note that the

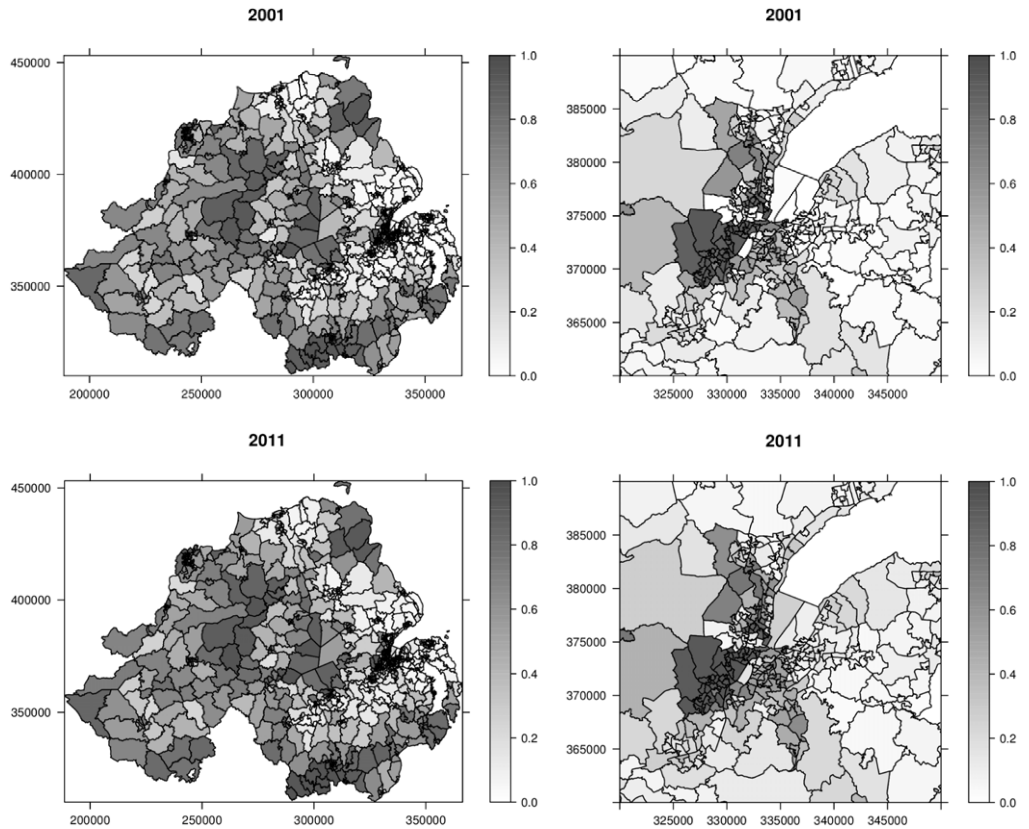

Fig. 1. Maps of the sample Catholic proportions for each Super Output Area in Northern Ireland (left column) and Belfast specifically (right column) for 2001 (top row) and 2011 (bottom row).

methodology we propose could easily be applied to other indices such as the *Gini coefficient*, but in the interests of brevity we only focus on one index here. The data relate to a study region that has been partitioned into n non-overlapping areal units, which for the Northern Ireland application are SOAs. The data are denoted by $\mathbf{Y} = (Y_1, \ldots, Y_n)$ and $\mathbf{N} = (N_1, \ldots, N_n)$, which respectively denote the sizes of the minority population and the total population for each of the n areal units. Letting $\mathbf{p} = (p_1, \ldots, p_n)$ denote the true minority proportion in each areal unit, the *Dissimilarity index* is given by

$$D = \sum_{k=1}^{n} \frac{N_k |p_k - p|}{2Np(1-p)}, \tag{1}$$

where $N = \sum_{k=1}^{n} N_k$ and p are the total population and minority proportion for the entire study region. The value of D lies in the interval [0, 1], where 0 represents complete evenness (i.e. $p_k = p \ \forall k$) and 1 represents complete segregation where each p_k equals zero or one. The unknown true proportions $\mathbf{p}$ are typically estimated by their sample equivalents, that is $\hat{p}_k = Y_k/N_k$ and $\hat{p} = (\sum_{k=1}^{n} Y_k)/(\sum_{k=1}^{n} N_k)$. This approach was used to compute D for the Northern Ireland data described in Section 2, and values of $\hat{D} = 0.5807$ (2001) and $\hat{D} = 0.5672$ (2011) were obtained. Thus the level of religious segregation has reduced by 0.014 on the D scale, but it is unclear if this is a statistically significant decline in segregation over the 10 years or just due to sampling variation in the observed data $(\mathbf{Y}, \mathbf{N})$.

The existence of sampling variation is clear if the data (Y_k, N_k) are minority and total population counts from a survey, because they only relate to a sample of the total population in each areal unit. For census data, as in the case of the motivating application, (Y_k, N_k) relate to the entire population, and thus in theory $\hat{p}_k = p_k$ exactly. However, we argue that it is still appropriate to consider Y_k as a random variable rather than a fixed and known constant, which means that $\hat{p}_k$ is only an estimate of the true unknown p_k. This in turn means that D in (1) is also unknown, and that the use of $\hat{p}_k = p_k$ is simply method of moments estimation. We distinguish between two different sources of variation and uncertainty that affect $\mathbf{Y}$. The first is due to error in the computation of $\mathbf{Y}$, which can come from participants filling out the census incorrectly, processing errors in digitising the paper census questionnaires, and different people interpreting the census questions in different ways. In the motivating religious segregation example some people might regard themselves as catholic if they were brought up as Catholic even though they no longer attend church, whilst others may feel that regular attendance is a necessary condition.

The second source of variation is temporal sampling variation, which arises because census data represent a snapshot of the level of segregation on a single day in the year. Thus, if the census was conducted on a different day then $\mathbf{Y}$ would be different due to people continually moving house, meaning that the true average minority proportions $\mathbf{p}$ across the whole year are unknown. Finally, if two areal units (j, k) exhibit the same true propensity for containing the minority population, then the observed proportions $(\hat{p}_j, \hat{p}_k)$ from the census on a single day would almost certainly be different due to sampling variation. An extended discussion of this point can be found in Leckie et al. (2012). The simple estimate $\hat{p}_k$ based on sample proportions is both the method of moments estimator and the maximum likelihood estimator under the model $Y_k \sim \text{Binomial}(N_k, p_k)$, which assumes the data in each areal unit are independent. This was shown not to be true for the Northern Ireland data presented in Section 2 which are spatially autocorrelated, so in the next section we propose a Bayesian spatial model for segregation data and use it to compute the posterior predictive distribution of D in (1).

4. Modelling

Spatially autocorrelated areal unit data are typically modelled in a hierarchical Bayesian setting, with inference based on Markov chain Monte Carlo (McMC) simulation. Conditional autoregressive (CAR) models are commonly used to model the spatial autocorrelation in these data (see for example Banerjee et al., 2004 and Wakefield, 2007), and are a special case of a Gaussian Markov Random Field (GMRF). We consider two distinct models here, the first assumes the true underlying probability surface $\mathbf{p}$ is globally spatially smooth, while the second assumes local spatial smoothness by allowing geographically adjacent areal units to have very similar or very different minority proportions. In both cases an estimate and a 95% credible interval for the Dissimilarity index can be obtained, by computing the posterior predictive distribution of D. Both models can be implemented using the CARBayes package (Lee, 2013) in the R software environment, and an illustration on simulated data is given in the supplementary material (see Appendix A).

4.1. *A globally smooth model for* $\mathbf{p} = (p_1, \ldots, p_n)$

The global smoothing model we propose for these data is a binomial generalised linear mixed model (GLMM), where the set of random effects are spatially autocorrelated. A binomial sampling model is used here for consistency with existing research (Goldstein and Noden, 2003; Mulekar et al., 2008; Leckie et al., 2012), and the full model is given by:

$$\begin{aligned} &Y_k \sim \text{Binomial}(N_k, p_k), \\ &\ln\left(\frac{p_k}{1-p_k}\right) = \beta_0 + \phi_k, \\ &\boldsymbol{\phi} \sim \text{N}(\mathbf{0}, \tau^2 Q(\rho, W)^{-1}), \\ &\beta_0 \sim \text{N}(0, 100), \end{aligned} \tag{2}$$

$\tau^2 \sim \text{Inverse-Gamma}(a, b)$,

$\rho \sim \text{Uniform}(0, 1)$.

The random effects $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)$ are included in (2) to model the spatial autocorrelation in the data, and are represented by a CAR prior distribution. CAR priors induce spatial autocorrelation by means of a binary $n \times n$ adjacency matrix $W = (w_{ki})$, which is based on the contiguity structure of the n areal units. Element $w_{ki} = 1$ if and only if areal unit k shares a border with areal unit i, otherwise $w_{ki} = 0$, and also $w_{kk} = 0\ \forall k$. Based on this proximity matrix CAR priors take the form of a zero-mean multivariate Gaussian distribution, where spatial autocorrelation is induced via the precision matrix that depends on W. The first CAR prior proposed in the literature was the intrinsic model (Besag et al., 1991) for strong spatial autocorrelation, although recently it has been shown to induce too much spatial smoothing in certain circumstances. Therefore, the model we use here was proposed by Leroux et al. (1999) and allows the strength of the autocorrelation to be estimated from the data. The precision matrix for this model is given by $Q(\rho, W) = \rho(\text{diag}(W\mathbf{1}) - W) + (1 - \rho)I$, where I is an $n \times n$ identity matrix, $\mathbf{1}$ is an $n \times 1$ vector of ones, and $\text{diag}(W\mathbf{1})$ is a diagonal matrix with elements equal to the row sums of W. This matrix is proper if $\rho \in [0, 1)$, and the spatial structure amongst $\boldsymbol{\phi}$ can be observed more clearly from the univariate full conditional distributions:

$$\phi_k|\boldsymbol{\phi}_{-k} \sim \text{N}\left(\frac{\rho\sum_{i=1}^{n} w_{ki}\phi_i}{\rho\sum_{i=1}^{n} w_{ki} + 1 - \rho}, \frac{\tau^2}{\rho\sum_{i=1}^{n} w_{ki} + 1 - \rho}\right). \tag{3}$$

In the above equation, $\boldsymbol{\phi}_{-k}$ denotes the vector of random effects except for ϕ_k. From (3) it is clear that ρ controls the spatial autocorrelation structure, with $\rho = 1$ corresponding to the intrinsic CAR prior proposed by Besag et al. (1991) for strong spatial autocorrelation, while $\rho = 0$ corresponds to independent random effects with constant mean and variance. The remaining parameters are assigned weakly informative uniform, inverse-gamma or Gaussian priors, so that the data play the dominant role in determining their values. The prior variance is assigned an inverse-gamma ($a = 0.001,\ b = 0.001$) prior, and a sensitivity analysis in Section 6 shows that changing these values has no impact on the results.

The posterior predictive distribution for the Dissimilarity index D given by (1) can be computed from this model using M McMC samples from the posterior distribution $\{\boldsymbol{\Theta}^{(j)}\}_{j=1}^{M}$, where $\boldsymbol{\Theta}^{(j)} = (\boldsymbol{\phi}^{(j)}, \beta_0^{(j)}, \tau^{2^{(j)}}, \rho^{(j)})$. These posterior samples are used to construct samples $\mathbf{p}^{(j)} = (p_1^{(j)}, \ldots, p_n^{(j)})$, using the inverse logit transform $p_k^{(j)} = \exp(\beta_0^{(j)} + \phi_k^{(j)})/[1 + \exp(\beta_0^{(j)} + \phi_k^{(j)})]$. Using these samples the jth sample from the posterior predictive distribution of D is constructed as

$$D^{(j)} = \sum_{k=1}^{n} \frac{N_k|p_k^{(j)} - p^{(j)}|}{2Np^{(j)}(1 - p^{(j)})} \quad \text{for } j = 1, \ldots, M, \tag{4}$$

where $p^{(j)} = (\sum_{k=1}^{n} N_k p_k^{(j)})/(\sum_{k=1}^{n} N_k)$. Thus D can be estimated by the median of $\{D^{(1)}, \ldots, D^{(M)}\}$, while a 95% credible interval is obtained from the [2.5, 97.5] percentiles of this distribution. We note that (1) and (4) utilise the same algebraic form for computing D, which is an a-spatial measure applied to spatial data. Here, (4) allows for the spatial autocorrelation in the data when estimating $\mathbf{p}$, while the standard approach of using the sample proportions to estimate $\mathbf{p}$ naively assumes independence.

4.2. A locally smooth model for $\mathbf{p} = (p_1, \ldots, p_n)$

Model (2) assumes the true proportions $\mathbf{p}$ exhibit a single global level of spatial smoothness, which is controlled by ρ. However, as illustrated by Fig. 1 the Northern Ireland data exhibit localised rather than global smoothness, with some pairs of geographically adjacent areal units exhibiting very similar sample proportions while others exhibit very different values. The global smoothing model is unlikely to be flexible enough to capture this localised autocorrelation, so here we propose an alternative

model which allows for either a smooth transition or a large jump in the true proportions between adjacent areal units. A number of different approaches have been proposed for extending model (2) to allow for localised spatial smoothness, including spatially varying variances (Brewer and Nolan, 2007), treating elements in W as additional parameters to be updated (Lee and Mitchell, 2013), and augmenting the CAR model with a non-smooth component (Lawson and Clark, 2002).

Here we follow the latter general approach, and model the probability surface $\mathbf{p}$ as locally smooth by combining the random effects in (2) with a piecewise constant intercept surface that allows sudden jumps in the spatial surface between geographically adjacent areal units. This piecewise constant intercept term has q distinct values $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)$, and if adjacent units (k, i) are in different groups their probabilities (p_k, p_i) may exhibit very different values. The model is given by:

$$\begin{aligned}
&Y_k \sim \text{Binomial}(N_k, p_k), \\
&\ln\left(\frac{p_k}{1-p_k}\right) = \beta_{Z_k} + \phi_k, \\
&Z_k \sim \text{Multinomial}(1; 1/q, \ldots, 1/q), \\
&\boldsymbol{\phi} \sim \text{N}(\mathbf{0}, \tau^2 Q(\rho, W)^{-1}), \\
&\beta_i \sim \text{Uniform}(\beta_{i-1}, \beta_{i+1}) \quad \text{for } i = 1, \ldots, q, \\
&\tau^2 \sim \text{Inverse-Gamma}(a, b), \\
&\rho \sim \text{Uniform}(0, 1),
\end{aligned} \tag{5}$$

where the CAR model is as above and $\beta_0 = -\infty$ and $\beta_{q+1} = \infty$. The vector of q area specific intercepts $\boldsymbol{\beta}$ has been constrained so that $\beta_1 < \beta_2 < \cdots < \beta_q$, which prevents the label switching problem common in mixture models. The vector $\mathbf{Z} = (Z_1, \ldots, Z_n)$ comprises a set of group or cluster indicators, where each indicator $Z_k \in \{1, \ldots, q\}$. A multinomial prior is specified for each indicator with equal probabilities, which specifies our prior ignorance as $\mathbb{P}(Z_k = j) = 1/q$ for all areal units k for $j = 1, \ldots, q$. In terms of notation the number 1 before the semi-colon in the multinomial model above indicates there is only one 'trial' in the multinomial distribution, as each areal unit can only have one intercept term. We note that we do not put any spatial smoothing constraints on the indicator vector $\mathbf{Z}$, because from the data in Fig. 1 it is clear that similar minority proportions occur at opposite ends of the study region. The piecewise constant intercept surface is thus inherently non-spatial, while spatial autocorrelation is accounted for by the CAR prior for $\boldsymbol{\phi}$.

The model is dependent on q, the number of different groups or clusters in the piecewise constant intercept term, which is assumed known. This is clearly unrealistic, and one approach would be to use a reversible jump McMC algorithm (Green, 1995) to choose q within the model fitting procedure. However, such algorithms can be slow to converge, so here we take the model comparison approach adopted by Choi et al. (2012). Specifically, we fit the model with values of $q = 1, \ldots, 5$, where $q = 1$ corresponds to the global spatial smoothing model with a common intercept term given by (2). Following Choi et al. (2012), we base model selection on the conditional predictive ordinate (CPO), which for observation k is given by

$$\text{CPO}_k = f(Y_k|\mathbf{Y}_{-k}) = \int_{\boldsymbol{\Theta}} f(Y_k|\boldsymbol{\Theta}) f(\boldsymbol{\Theta}|\mathbf{Y}_{-k}) \mathrm{d}\boldsymbol{\Theta},$$

where $\mathbf{Y}_{-k}$ denotes all data points except Y_k and $\boldsymbol{\Theta} = (\boldsymbol{\beta}, \mathbf{Z}, \boldsymbol{\phi}, \tau^2, \rho)$ denotes all model parameters. Following Congdon (2005) this can be approximated based on M McMC samples as

$$\widehat{\text{CPO}}_k = \left[\frac{1}{M}\sum_{j=1}^{M} \frac{1}{f(Y_k|\boldsymbol{\Theta}^{(j)})}\right]^{-1},$$

where $\boldsymbol{\Theta}^{(j)}$ is the jth posterior sample. Then a summary of the CPO for all n data points is the Log Marginal Predictive Likelihood (LMPL), which is given by

$$\text{LMPL} = \sum_{k=1}^{n} \ln(\widehat{\text{CPO}}_k),$$

and is maximised in this paper to select q. Once the best value of q has been selected the posterior predictive distribution of the Dissimilarity index given by (4) can be computed as described in Section 4.1, where under this model $p_k^{(j)} = \exp(\beta_{Z_k^{(j)}}^{(j)} + \phi_k^{(j)})/[1 + \exp(\beta_{Z_k^{(j)}}^{(j)} + \phi_k^{(j)})]$.

5. Simulation study

This section presents a simulation study, which compares the three approaches described in Sections 3 and 4 for estimating the Dissimilarity index and quantifying its uncertainty via a 95% uncertainty interval. Model **M1** is the standard approach for constructing D, which estimates p_k in (1) by $\hat{p}_k = Y_k/N_k$, the simple method of moments estimator. A 95% confidence interval is computed for this estimate of D using a bootstrapping approach, similar to that proposed by Brülhart and Traeger (2005). The interval is constructed by first creating 10,000 pseudo data sets of size n, by resampling the n data points (Y_k, N_k) with replacement. An estimate of the dissimilarity index D is then computed for each resampled data set, and a 95% confidence interval is constructed by calculating the 2.5 and 97.5 percentiles of the set of D values. Model **M2** is the global spatial smoothing model given by (2), and constructs a point estimate and a 95% credible interval for D from its posterior predictive distribution as outlined in Section 4.1. Finally, model **M3** uses the same approach to estimation as model **M2**, except that the localised spatial smoothing model given by (5) is used in place of the global smoothing model (2).

5.1. Data generation and study design

Simulated segregation data are generated for the set of $n = 890$ SOAs that comprise Northern Ireland, which is the study region for the motivating application described in Section 2. The true proportions of the minority populations in each areal unit $(p_1, \ldots, p_n)$ are generated as described below, and the total population sizes $(N_1, \ldots, N_n)$ are assumed known but are varied in this study to determine their impact on model performance. From these true proportions and population totals binomial sampling is used to generate the simulated counts $(Y_1, \ldots, Y_n)$, which mimics the realistic situation where the sample proportions are subject to sampling variation. Five hundred simulated data sets are generated under four different scenarios, each of which are summarised below.

Scenario 1—low segregation with little spatial variation: The true minority proportions across the n areal units are equal to 0.15 on average and exhibit low variation across the region. The logit transform of the proportions are generated from a multivariate Gaussian distribution with a mean of -1.73, a variance of 0.2 and an exponential correlation matrix with decay parameter 0.02.

Scenario 2—low segregation with moderate spatial variation: The true minority proportions across the n areal units are equal to 0.15 on average and exhibit moderate variation across the region. The logit transform of the proportions are generated from a multivariate Gaussian distribution with a mean of -1.73, a variance of 0.8 and an exponential correlation matrix with decay parameter 0.04.

Scenario 3—high segregation with little spatial variation: The true minority proportions are equal to 0.15 in 56% of the region and 0.85 in 44% of the region on average, and exhibit low variation across the region. The spatial structure for this segregation follows that observed in the real data analysed in Section 6, and is the two-group cluster structure observed for 2011 in Fig. 3. The logit transform of the proportions are generated from a multivariate Gaussian distribution with a mean of -1.73 (where $p_k = 0.15$) or 1.75 (where $p_k = 0.85$), a standard deviation of 0.2 and an exponential correlation matrix with decay parameter 0.01.

Scenario 4—high segregation with moderate spatial variation: The true minority proportions are equal to 0.15 in 56% of the region and 0.85 in 44% of the region on average, and exhibit moderate variation across the region. In common with scenario 3 the spatial structure is based on the template given in Fig. 3. The logit transform of the proportions are generated from a multivariate Gaussian distribution with a mean of -1.73 (where $p_k = 0.15$) or 1.75

(where $p_k = 0.85$), a standard deviation of 0.8 and an exponential correlation matrix with decay parameter 0.04.

Scenarios 1 and 2 have true proportion surfaces that are equal to 0.15 on average in every areal unit, which corresponds to an even distribution of the minority population. The difference between these scenarios is that the level of spatial variation around that 0.15 average proportion within each simulated data set is larger under scenario 2 than under scenario 1, but in both cases the proportion surface evolves smoothly over space. In contrast, scenarios 3 and 4 represent segregated communities, where the true average minority proportions are 0.15 in 56% of areal units and 0.85 in 44% of the units. Again, the difference between these scenarios is in the level of spatial variation around these average proportions, which is larger for scenario 4 compared with scenario 3. Data are generated under these scenarios for values of N_k of 100, 500 and 2500, to see how well the estimates of the Dissimilarity index perform for data with small, medium and large population sizes. The bootstrapped 95% confidence intervals computed to accompany model **M1** are based on 10,000 bootstrap samples, while for the Bayesian models inference is based on 10,000 McMC samples, which were collected after a burn-in period of 10,000 samples. The local smoothing model is applied to the data with q between 2 and 5 ($q = 1$ corresponds to the global smoothing model), and the final model is chosen by minimising the LMPL.

5.2. Results

The accuracy with which D is estimated is quantified by the root mean square errors (RMSE) from the three models, while the appropriateness of the 95% uncertainty intervals are quantified by their coverage probabilities and average widths. These results are displayed in Table 1, where the top panel displays the root mean square error for D (as a percentage of its true value), the middle panel displays the coverage probabilities for the 95% uncertainty intervals, while the bottom panel summarises the mean widths of the uncertainty intervals. The table shows that model **M1** performs poorly throughout, as its RMSE values are either much worse or about the same as those from model **M3**, while its 95% uncertainty intervals are never close to attaining their nominal coverage levels of 95%. Its RMSE value is up to 7 times larger than that of model **M3** when the population size is small ($N_k = 100$), which is because the sample proportions used to estimate (1) are based on the least data. In contrast, the spatial models suffer less from this problem, as they utilise the autocorrelation in the data to borrow strength over neighbouring data points in the estimation. As N_k increases the performance of model **M1** improves in terms of RMSE, and has the same values as model **M3**. However, the coverage probabilities from the data re-sampling bootstrap in Model **M1** are less than 35% for 3 combinations of scenario and population size and greater than 99% for the remaining 9 cases, which suggests that this is not an adequate approach for constructing uncertainty intervals in this context. The bottom panel of the table shows that the bootstrapped uncertainty intervals from model **M1** are nearly always much wider than those from models **M2** and **M3**, with intervals that are up to 4 times wider. This is because unlike **M1**, models **M2** and **M3** spatially smooth the proportion surface, which reduces the amount of variability in its estimate and hence in (1).

In contrast, the localised spatial smoothing model **M3** performs consistently well across all scenarios, and is the model of choice for constructing point estimates and 95% uncertainty intervals for D. It has RMSE values that are lower or the same as those from the other models, while its coverage probabilities are close to their nominal levels ranging between 85.2% and 97.6%. For the latter there is a slight drop in performance when N_k is small and the data exhibit little spatial variation (scenario 1), as the coverage probabilities drop slightly below 90%. This is likely due to the small sample sizes and the fact that the spatial model is trying to model spatial variation in the data when very little is present. When the true proportion surface is spatially smooth (scenarios 1 and 2) models **M2** and **M3** provide similar results, which is not surprising as both can capture spatially smooth variation. However, when the true proportion surface exhibits segregation (scenarios 3 and 4) then model **M3** vastly outperforms **M2**, as the latter is trying to smooth over the spatial discontinuities in the true proportion surface which is not appropriate. This results in coverage probabilities between 0 and 90% and RMSE

Table 1
Results of the simulation study for all three models **M1**, **M2** and **M3**. The top panel displays the root mean square error for D (as a percentage of its true value), the middle panel displays the coverages for the 95% uncertainty intervals, while the bottom panel summarises the mean widths of the uncertainty intervals.

Scenario	$N = 100$			$N = 500$			$N = 2500$		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
RMSE (%)									
1	32.82	4.43	4.43	7.62	1.85	1.85	1.69	0.76	0.76
2	6.73	1.64	1.64	1.52	0.66	0.68	0.38	0.29	0.33
3	0.33	1.62	0.33	0.16	0.38	0.16	0.07	0.09	0.07
4	0.46	1.24	0.44	0.16	0.28	0.16	0.07	0.09	0.08
Coverage (%)									
1	0.2	85.2	85.2	31.4	86.8	86.8	99.4	92.2	92.2
2	33.2	89.4	89.4	100	93.4	93.0	100	95.6	91.4
3	100	0.0	96.5	100	42.4	94.1	100	85.9	95.3
4	100	20.0	92.4	100	73.5	97.6	100	90.0	94.1
Interval width									
1	0.017	0.017	0.017	0.014	0.007	0.007	0.014	0.003	0.003
2	0.029	0.015	0.015	0.028	0.007	0.007	0.027	0.003	0.003
3	0.021	0.010	0.009	0.018	0.004	0.004	0.017	0.002	0.002
4	0.030	0.010	0.010	0.030	0.004	0.005	0.029	0.002	0.002

values that are up to 5 times larger than those from **M3**. In contrast model **M3** does not suffer from these problems, as the piecewise constant intercept term allows for neighbouring areas to have very different estimated proportions. Additionally, model **M2** performs worse than **M1** in these scenarios, which is because the latter does not make inappropriate global smoothing assumptions on the true proportion surface. An extension to this study with the proportion surfaces having 3 (low, medium and high) rather than 1 (low as in scenarios 1 and 2) or 2 (low and high as in scenarios 3 and 4) distinct catholic proportions is presented in the supplementary material (see Appendix A), and the results are similar to the two group case presented here.

Finally, the LMPL statistic chose the correct value of q in model **M3** for between 92.2% and 100% of the simulated data sets in 9 out of the 12 combinations of scenario and population size. However, it struggled to choose the correct value of q when the population size is large ($N_k = 2500$) or there is moderate levels of spatial variation in the proportion surface (Scenarios 2 and 4). In these cases it chose q correctly for between 27% and 66.8% of simulated data sets, and always chose too many rather than too few groups (clusters). These deficiencies in the LMPL occurred when the variation in the spatial proportion surfaces was largest, and it is thus not surprising that it choose models with overly large values of q. However, in these cases the estimation of D and the quantification of its uncertainty were not affected. Similar performance of the LMPL statistic was observed in the three group simulation case in the supplementary material (see Appendix A).

6. Northern Ireland religious segregation study

We now present the results of applying models **M1**, **M2** and **M3** to the Northern Ireland religious segregation study described in Section 2 that motivated this paper.

6.1. Estimation and inference for D

Estimates and 95% uncertainty intervals for the Dissimilarity index from all three models are presented in Table 2, for both 2001 and 2011. The table shows that for model **M3** the LMPL statistic selects $q = 2$ groups in both 2001 and 2011, but that the estimates and 95% credible intervals for D are similar for $q = 2, \ldots, 5$ and also for model **M2** which corresponds to $q = 1$. The estimates of D from model **M1** and the two spatial models are almost identical in 2011 but differ slightly in 2001, with the latter having $\hat{D} = 0.5807$ from **M1** and $\hat{D} = 0.6022$ from **M3**. These similarities in

Table 2
Estimates and uncertainty intervals for the Dissimilarity index D from models **M1**, **M2** and **M3**. In addition, the log marginal predictive likelihood (LMPL) is presented as a model comparison tool for the Bayesian spatial models.

Model	2001 data		2011 data	
	LMPL	D	LMPL	D
M1	–	0.5807 (0.5594, 0.6002)	–	0.5672 (0.5470, 0.5856)
M2-$q = 1$	−4289.5	0.6020 (0.6009, 0.6031)	−4417.0	0.5669 (0.5658, 0.5680)
M3-$q = 2$	−4287.1	0.6022 (0.6010, 0.6034)	−4414.1	0.5673 (0.5658, 0.5688)
M3-$q = 3$	−4306.3	0.6020 (0.6009, 0.6031)	−4416.4	0.5670 (0.5658, 0.5682)
M3-$q = 4$	−4304.5	0.6020 (0.6009, 0.6032)	−4426.1	0.5672 (0.5660, 0.5685)
M3-$q = 5$	−4298.5	0.6020 (0.6009, 0.6031)	−4431.6	0.5669 (0.5657, 0.5681)

point estimation were also seen in the simulation study, where models **M1** and **M3** produced similar RMSE values for large and segregated populations. The 95% credible intervals from the Bayesian spatial models are much narrower than the bootstrapped 95% confidence intervals, a result that was also observed for the simulated data. The simulation study showed that the former are more likely to have the correct nominal coverage levels, while the latter are likely to be too conservative. Based on this observation and model **M3**, religious segregation in Northern Ireland has undergone a small but significant decline between 2001 and 2011, as the estimates of D are 0.6022 (2001) and 0.5672 (2011) and the corresponding 95% credible intervals do not overlap.

6.2. Cluster structure in the data

The piecewise constant intercept (cluster) structures $(\hat{\beta}_1, \ldots, \hat{\beta}_q)$ estimated by model **M3** for $q = 2, \ldots, 5$ are displayed in Fig. 2, where the results are based on posterior medians. The figure shows density estimates for the estimated proportion of people who are Catholic in each group, and the left column is for 2001 while the right column is for 2011. The number of areal units in each group is displayed in the top right of each figure, where the groups are numbered from left to right. The figure shows that the posterior densities for the groups rarely overlap, and when $q = 2$ the areal units are partitioned into a low Catholic proportion group and a medium to high group. As q increases the lowest Catholic proportion group always has the smallest range of values, while the groups in the middle of the interval are typically the widest. As the number of groups increases they become more unimodal, and with a few exceptions (such as the 5 group model in 2001) typically contain similar numbers of observations.

Fig. 3 displays the spatial pattern in the estimated group structure for $q = 2$ for both 2001 and 2011, which was the structure chosen to be most appropriate by the LMPL for both years. In the figure the grey areas relate to the medium to high proportion Catholic areas, while the white areas are low proportion Catholic areas. The figure shows that the groups are spatially structured, and follow the general pattern seen in the data in Fig. 1. We note that model **M3** does not enforce the groups to be spatially structured, but does obviously allow this to occur. The spatial patterns are similar for both years, as 93% of Super Output Areas are classified in the same group in both years. The number of areas classified as having medium to high Catholic proportions has dropped from 50.1% in 2001 to 43.9% in 2011, which is a result of the differing thresholds shown in Fig. 2 rather than a decrease in the proportion of Catholics between the two years. In fact the mean proportion of Catholics has increased from 39% in 2001 to 44% in 2011.

6.3. Sensitivity analyses

A sensitivity analysis was conducted to assess the robustness of the results, by changing the Inverse-Gamma ($a = 0.001,\ b = 0.001$) prior distribution specified for the random effects variance τ^2. Models were fitted with (i) $a = 0.01,\ b = 0.01$, (ii) $a = 0.1,\ b = 0.1$, (iii) $a = 0.5,\ b = 0.0005$ and (iv) $f(\tau) \propto 1$, and the estimates and 95% credible intervals for the Dissimilarity index changed by less than 0.0001.

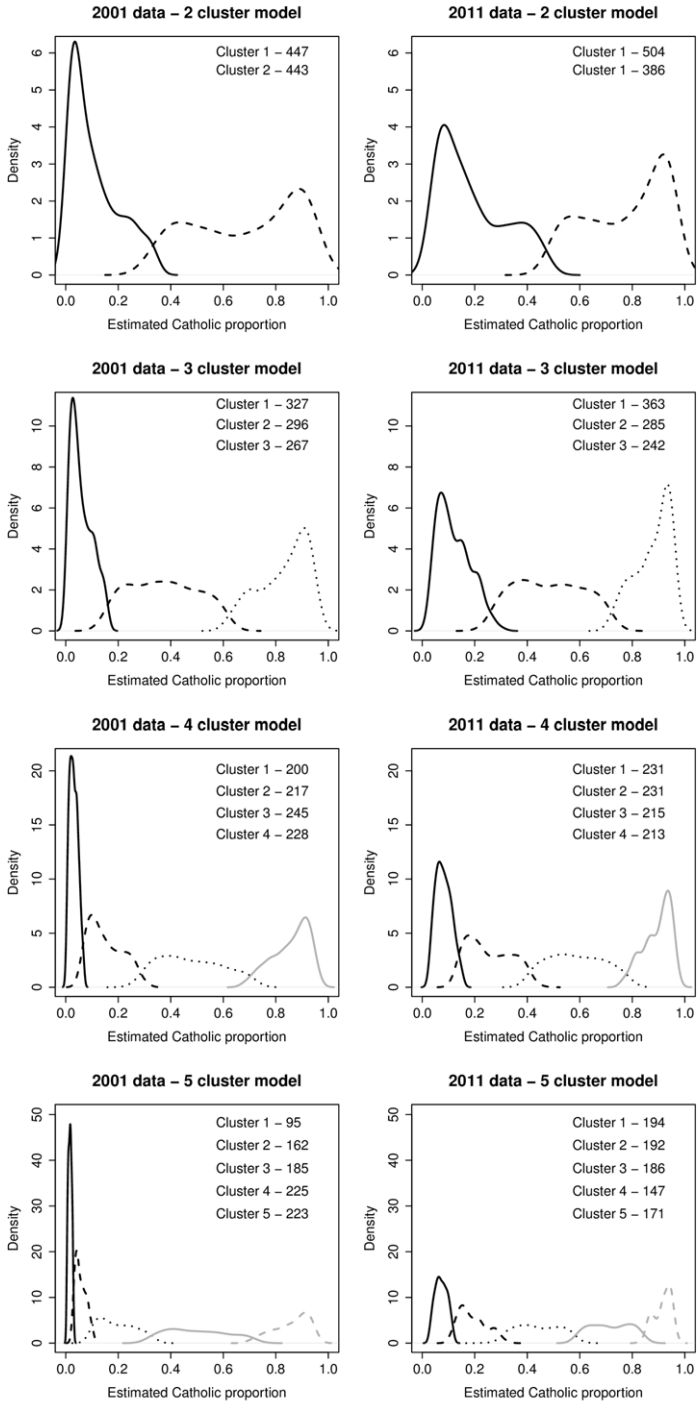

Fig. 2. Density estimates of the group structure estimated by model **M3** for 2001 (left) and 2011 (right) for $q = 2, \ldots, 5$. The number of areal units in each group is also displayed.

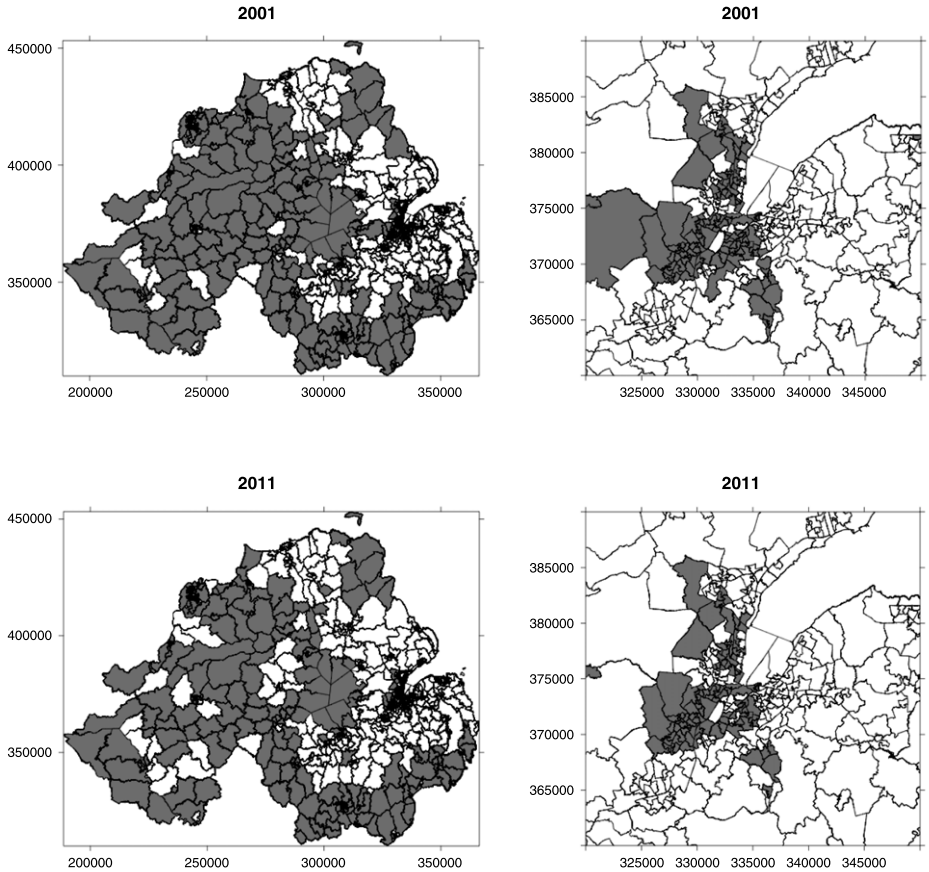

Fig. 3. Maps of the estimated cluster structure from model **M3** with $q = 2$ for 2001 and 2011 for Northern Ireland (left column) and Belfast specifically (right column). The grey areas are medium to high proportion of Catholic areas while the white ones have low Catholic proportions.

7. Discussion

The measurement of residential segregation has been an active research topic since the 1950s, and was born out of the desire to quantify the level of racial segregation in cities in USA. The data used for quantifying segregation are inherently spatial, even if the Dissimilarity index commonly used to quantify segregation is an a-spatial measure. To our knowledge this paper is the first to allow for spatial autocorrelation in the data when estimating the true proportion surface, upon which the Dissimilarity index is based. We note that other authors have adjusted the Dissimilarity index to allow for the spatial nature of the data (see Morrill, 1991 and Wong, 1993), but here we have used the standard formulation and argued that it is the estimation of the proportion surface rather than the index that should be adjusted to account for spatial autocorrelation. This is also one of the few papers to address the issue of uncertainty quantification for the Dissimilarity index, and builds on pioneering work by Mulekar et al. (2008) and Leckie et al. (2012) in this regard.

The simulation study in Section 4 demonstrates that a simple bootstrap approach to calculating a 95% uncertainty interval for the Dissimilarity index is completely inappropriate, and that a locally smooth Bayesian hierarchical modelling approach yields credible intervals which generally have close to their nominal coverage levels. This approach also leads to either improved or identical point estimation for D in terms of RMSE, with the biggest gains being made when the population in each

areal unit is small. This suggests that the localised smoothing model **M3** is the method of choice for estimating D and quantifying its uncertainty, and is likely to be applicable to other segregation indices such as the Gini index. Finally, the simulation study has shown that model **M3** is robust to the number of groups (q) specified, as model performance did not retard when the LMPL statistic choose the wrong value. However, this does suggest that if the aim of the study was cluster detection, then the LMPL statistic will not be a reliable tool for identifying the correct number of clusters.

The results from the Northern Ireland religious segregation study presented in Sections 2 and 6 show a small but clear decline in segregation between 2001 and 2011, as the Dissimilarity index has reduced from 0.6022 in 2001 to 0.5673 in 2011, and the corresponding 95% credible intervals from model **M3** do not overlap. This confirms the potentially important finding reported in the media that segregation has indeed fallen in Belfast and the surrounding areas, which is likely to be down to a combination of factors. Two possibilities in this regard are the continuation of the peace process and the in-migration of Polish Catholic families locating in Protestant areas of Belfast, although a full analysis of the causes are beyond the scope of the current paper. Nevertheless, our findings will provide some reassurance for those who view segregation as indicative of social fragmentation.

Thus there are three key conclusions resulting from our paper. The first is that the creation of uncertainty intervals are a cornerstone of statistical inference and should be routinely constructed for segregation measures, which at present they are not. This is an important omission because of the uncertainty implicit in estimates of dissimilarity arising from sampling variation caused by random churn in household moves and from potential errors in data recording and coding. Secondly, the proportion surface upon which the dissimilarity index is based should be estimated allowing for the spatial autocorrelation present in the data, rather than using simple method of moments estimation based on the naïve assumption of independence. This was evident in the Northern Ireland data used in the paper for which Moran's I tests confirmed the presence of spatial autocorrelation. The third is that the simulation study presented in the paper shows that the spatial model **M3** performs consistently well in terms of both point estimation (RMSE) and uncertainty quantification (coverage probabilities) across a range of scenarios, whereas the simpler approach commonly used does not. Thus the extra computational cost of spatial models is clearly worth the effort here. This is the first paper to our knowledge that uses such spatial modelling techniques in this context, and will hopefully lead to such methods being used in future.

For us we are working on a number of future developments, the first of which is to see whether the Bayesian spatial modelling paradigm proposed here can be applied to other segregation measures such as the Gini index. Secondly, the models developed in this paper were implemented in the R programming environment (R Core Team, 2013), which is not easy to use for non-statisticians. Therefore we plan to develop software with a user friendly web-based interface to allow social scientists to apply our methods to their own data. Finally, although not relevant for the data considered here, many segregation data sets are available for multiple consecutive years, which lends itself to extending the statistical models developed here to the spatio-temporal domain.

Acknowledgements

The authors are grateful for the comments of two reviewers, which have improved the motivation for and content of this paper. In addition they would like to thank Alastair Rushworth for helping with the implementation of the models via the Rcpp software. This work was funded by the Economic and Social Research Council through the Applied Quantitative Methods Network: Phase II, grant number ES/K006460/1. The data were provided by the Northern Ireland Neighbourhood Information Service.

Appendix A. Supplementary material

This paper includes supplementary material that includes extra simulation results and code to run the models considered in this paper.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.spasta.2014.12.001.

References

Banerjee, S., Carlin, B., Gelfand, A., 2004. Hierarchical Modelling and Analysis for Spatial Data, first ed. Chapman and Hall.

Besag, J., York, J., Mollie, A., 1991. Bayesian image restoration with two applications in spatial statistics. Ann. Inst. Statist. Math. 43, 1–59.

Brewer, M.J., Nolan, A.J., 2007. Variable smoothing in Bayesian intrinsic autoregressions. Environmetrics 18 (8), 841–857.

Brülhart, M., Traeger, R., 2005. An account of geographic concentration patterns in Europe. Reg. Sci. Urban Econ. 35, 597–624.

Catney, G., Has neighbourhood ethnic segregation decreased? Manchester ESRC Centre on Dynamics of Ethnicity Bulletin, 2013, pp. 1–4.

Cheshire, P., 2009. Policies for mixed communities: faith-based displacement activity? Int. Reg. Sci. Rev. 32, 343–375.

Choi, J., Lawson, A., Cai, B., Hossain, M., Kirby, R., Liu, J., 2012. A Bayesian latent model with spatio-temporally varying coefficients in low birth weight incidence data. Stat. Methods Med. Res. 21, 445–456.

Clark, W., 1986. Residential segregation in American cities: a review and interpretation. Popul. Res. Policy Rev. 5, 95–127.

Congdon, P., 2005. Bayesian Models for Categorical Data, first ed. John Wiley and Sons.

Cortese, C., Falk, R., Cohen, J., 1976. Further considerations on the methodological analysis of segregation indices. Amer. Sociol. Rev. 41, 630–637.

Duncan, O., Duncan, B., 1955. A methodological analysis of segregation index. Amer. Sociol. Rev. 20, 210–217.

Glaster, G., 1988. Residential segregation in American cities: a contrary review. Popul. Res. Policy Rev. 7, 93–112.

Goldstein, H., Noden, P., 2003. Modelling social segregation. Oxf. Rev. Educ. 29, 225–237.

Green, P., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 711–732.

Ihlanfeldt, K., Scafidi, B., 2002. Black self-segregation as a cause of housing segregation: evidence from the multi-city study of urban inequality. J. Urban Econ. 51, 366–390.

Inman, H., Bradley, E., 1991. Approximations to the mean and variance of the index of dissimilarity in $2 \times c$ tables under a random allocation model. Sociol. Methods Res. 20, 242–255.

Jivraj, S., How has ethnic diversity grown 1991–2001–2011? Manchester ESRC Centre on the Dynamics of Ethnicity Bulletin, 2012, pp. 1–4.

Johnston, R., Poulsen, M., Forrest, J., 2013. Multiethnic residential areas in a multiethnic country? A decade of major change. Environ. Plann. A 45, 753–759.

Lawson, A., Clark, A., 2002. Spatial mixture relative risk models applied to disease mapping. Stat. Med. 21, 359–370.

Leckie, G., Pillinger, R., Jones, K., Goldstein, H., 2012. Multilevel modelling of social segregation. J. Educ. Behav. Stat. 37, 3–30.

Lee, D., 2013. CARBayes: an R package for Bayesian spatial modelling with conditional autoregressive priors. J. Stat. Softw. 55, 13.

Lee, D., Mitchell, R., 2012. Boundary detection in disease mapping studies. Biostatistics 13, 415–426.

Lee, D., Mitchell, R., 2013. Locally adaptive spatial smoothing using conditional autoregressive models. J. Roy. Statist. Soc. Ser. C 62, 593–608.

Leroux, B., Lei, X., Breslow, N., 1999. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: Halloran, M., Berry, D. (Eds.), Statistical Models in Epidemiology, the Environment and Clinical Trials. Springer-Verlag, New York, pp. 135–178. (Chapter).

Massey, D., Denton, N., 1988. The dimensions of residential segregation. Soc. Forces 67, 281–315.

Moran, P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.

Morrill, R., 1991. On the measure of geographic segregation. Geogr. Res. Forum 11, 25–36.

Mulekar, M., Knutson, J., Champanerkar, J., 2008. How useful are approximations to mean and variance of the index of dissimilarity? Comput. Statist. Data Anal. 52, 2098–2109.

Musterd, S., 2005. Social and ethnic segregation in Europe: levels, causes, and effects. J. Urban Aff. 27, 331–348.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: http://www.R-project.org.

Reardon, S., Firebaugh, G., 2002. Measures of multigroup segregation. Sociol. Methodol. 32, 33–67.

Reardon, S., O'Sullivan, D., 2004. Measures of spatial segregation. Sociol. Methodol. 34, 121–162.

Semyonov, M., Glikman, A., 2009. Ethnic residential segregation, social contacts, and anti-minority attitudes in European societies. Eur. Sociol. Rev. 25, 639–708.

Shuttleworth, I., Green, A., 2011. Spatial mobility intentions, the labour market and incapacity benefit claimants. Urban Stud. 48, 911–927.

Shuttleworth, I., Lloyd, C., 2009. Are Northern Ireland's communities dividing? Evidence from geographically consistent census of population data, 1971–2001. Environ. Plann. A 41, 213–229.

Simpson, L., 2007. Ghettos of the mind: the empirical behaviour of indices of segregation and diversity. J. Roy. Statist. Soc. Ser. A 170, 405–424.

Simpson, L., Finney, N., 2010. Parallel lives and ghettos in Britain: facts or myths? Geography 95, 124–131.

Wakefield, J., 2007. Disease mapping and spatial regression with count data. Biostatistics 8, 158–183.

Whitley, R., Prince, M., McKenzie, K., Stewart, R., 2006. Exploring the ethnic density effect: a qualitative study of a london electoral ward. Int. J. Soc. Psychiatry 52, 376–391.

Winship, C., 1977. A revaluation of indexes of residential segregation. Soc. Forces 55, 1058–1066.

Wong, D., 1993. Spatial indices of segregation. Urban Stud. 30, 559–572.

Wong, D., 2003. Spatial decomposition of segregation indices: a framework toward measuring segregation at multiple levels. Geogr. Anal. 35, 179–194.