# Fuzzy Bayesian inference for mapping vague and place-based regions: a case study of sectarian territory

Huck, J.J.[1*], Whyatt, J.D.[2], Davies, G.[3], Dixon, J.[4], Sturgeon, B.[5], Hocking, B.[6], Tredoux, C.[7], Jarman, N.[6], Bryan, D.[4]

[1] *MCGIS, Department of Geography, The University of Manchester, UK*

[2] *Lancaster Environment Centre, Lancaster University, UK*

[3] *Map Action, UK*

[4] *Department of Psychology, The Open University, UK*

[5] *Institute of Irish Studies, Queens University, UK*

[6]*Independent Scholar*

[7]*Department of Psychology, University of Cape Town, South Africa*

* Corresponding author: jonathan.huck@manchester.ac.uk

# Fuzzy Bayesian inference for mapping vague and place-based regions: a case study of sectarian territory

The problem of mapping regions with socially-derived boundaries has been a topic of discussion in the GIS literature for many years. Fuzzy approaches have frequently been suggested as solutions, but none have been adopted. This is likely due to difficulties associated with determining suitable membership functions, which are often as arbitrary as the crisp boundaries that they seek to replace. This paper presents a novel approach to fuzzy geographical modelling that replaces the membership function with a possibility distribution that is estimated using Bayesian inference. In this method, data from multiple sources are combined to estimate the degree to which a given location is a member of a given set and the level of uncertainty associated with that estimate. The Fuzzy Bayesian Inference approach is demonstrated through a case study in which census data are combined with perceptual and behavioural evidence to model the territory of two segregated groups (Catholics and Protestants) in Belfast, Northern Ireland, UK. This novel method provides a robust empirical basis for the use of fuzzy models in GIS, and therefore has applications for mapping a range of socially-derived and otherwise vague boundaries.

**Keywords:** Fuzzy, Place, Vague, Segregation, Territory

## Introduction

### *The problem of social boundaries in GIS*

When describing a geographical area, there are two potential sources of uncertainty: the location of the area and its extent (Clementini and Di Felice, 1996). Here we are concerned with the latter case, in which it is clear *where* the area is located, but its boundaries cannot be unequivocally demarcated. This boundary uncertainty is referred to as *geographical vagueness*, the sources of which are enumerated by Montello (2003). There are several approaches to the categorisation of boundaries in GIS (e.g., Smith and Varzi, 2000; Montello, 2003), but here we adopt the high-level typology presented by Fisher (1996), who categorised boundaries as *real* (normally implying a physical delimiting feature such as a river), *perceived* (understood by individuals, but not delimited) and *imposed* (normally administrative zones, *i.e.*, all political and legal regions). *Imposed* boundaries are typically crisp, in that it is easy to unequivocally determine whether one is inside or outside the boundary. *Perceived* boundaries, however, exhibit geographical *vagueness* to some degree (Huck et al., 2014) and typically reflect the ways in which individuals understand the places that they occupy in their daily lives, such as the extent of their neighbourhood. *Real* boundaries can be either crisp or vague, often depending on the nature of the delimiting feature. Vague geographical regions are characterised by the fact that their bounds are not merely *undetermined* (i.e., merely unset or unconsidered), but rather are *indeterminate* (i.e., there are no unequivocal precise bounds that could be defined, even if it were so desired).

Where they are socially-derived, the boundaries of vague regions are determined by the daily experience of individuals and, whilst it is easy to select a location that is definitely either inside or outside them, there is no unequivocal manner by which a precise boundary may be drawn (Clementini and Di Felice, 1996; Fisher et al., 2004). Nevertheless, their meaning is fully understood and people can reason about them (Clementini and Di Felice, 1996). Even seemingly well understood social concepts can be revealed to be vague when carefully considered (e.g., *downtown*; Montello et al., 2003). Geographers typically account for vagueness in socially-derived regions by distinguishing between geometric *spaces*, which are universally defined and precisely allocated (e.g., a census zone); and geographical *places*, which inherently exhibit variation between individuals and through time, and which do not exhibit defined locations or boundaries (Goodchild, 2011).

Vague regions elude satisfactory representation using spatial primitives such as points, lines and polygons (Montello et al., 2003; Goodchild, 2011). This is because definitions of the perceived boundary must rely upon a multitude of perceived attributes of an area, which are often individualistic and may not relate to measurable phenomena (Carver et al., 2009). The representation must therefore consider not only the physical objects found therein, but also the meanings that those objects have for individuals and communities (Purves and Derungs, 2015). Vague geographical entities might therefore be considered as both synergistic and incommensurable, a condition that is captured by Tolkien (2001: 10): *"It is one of its qualities to be indescribable, though not imperceptible. It has many ingredients, but analysis will not necessarily discover the secret of the whole".* The representation of *place* remains of great importance in the field of GIS, and was described by Goodchild (2011) as one of the fundamental elements of our ability to deal with phenomena that are distributed in space.

It has frequently been suggested that vague or place-based regions lend themselves better to fuzzy or probabilistic representations, as opposed to precise geometric models (e.g., Leung, 1987; Montello et al., 2003; Evans and Waters, 2007). Fuzzy approaches, in which the degree of membership of a given set is determined by a membership function, have been applied in several areas of GIS research since early applications were first proposed by Burrough (1986), but there are few, if any, attempts to use such approaches for social phenomena. It is likely that this is a result of the difficulty associated with the robust definition of membership functions for social phenomena, in comparison with fields such as soil science and geomorphology where fuzzy approaches have been applied more widely. This paper therefore presents a Bayesian inference-based approach to fuzzy modelling that removes the need for a pre-defined membership function and so better lends itself to socially-derived boundaries. In support of this, we will briefly present an overview of the relationship between fuzzy and probabilistic methods, before describing our proposed method in more detail. We will then demonstrate our method using a case study relating to understanding patterns of intergroup segregation in Belfast, Northern Ireland.

## Literature review

### *The relationship between fuzziness and probability*

There has historically been a great deal of disagreement in the literature with respect to whether *fuzziness* is the same as *randomness* and hence whether fuzzy models are the same as probabilistic models. Key arguments 'for' and 'against' this position are given in Cheeseman (1985) and Kosko (1990) respectively. Though both authors present acceptable solutions to the debate, this paper will adopt the position that fuzziness and randomness are distinct, which is both the most straightforward position and the one most frequently adopted in the GIS literature. In this view, both fuzzy and probabilistic systems represent uncertainty numerically in the interval [0-1], but they differ substantially in interpretation and the problems to which they should be applied. In simple terms, fuzziness represents the degree to which an event occurs (*as a result of ambiguity in the event itself*), whereas probability represents uncertainty about whether the event occurs (*as a result of chance relating to the occurrence of the event*) (Kosko, 1990).

A formal distinction can be made by considering the extent to which 'the thing' ($A$) can be distinguished from its opposite ($A^c$), formally $A \cap A^c = \emptyset$. If $A$ and $A^c$ can be distinguished, then the event is probabilistic; otherwise it is fuzzy (Kosko, 1990). To provide examples in the context of a social phenomenon such as intergroup segregation: the question of the degree to which a location belongs to the territory of a given group is therefore fuzzy; whereas the question of whether one or more members of a given group are present in that location at a given time is probabilistic. The distinction here is clear, as the absence of an individual (i.e., the opposite of their presence) is readily discerned, whereas this is not the case for the degree to which a location is part of a territory. Though this distinction may appear unimportant at first glance: this difference in interpretation between fuzzy and probabilistic approaches is considerable and it is important that the two are not confused (Fisher, 1996). Indeed, Fisher (1996) identifies several instances in the literature, where prominent authors have erroneously taken the terms to be synonymous or interchangeable.

### *A Bayesian approach to fuzzy membership*

One of the key challenges in the application of fuzzy methods to geographical problems is the difficulty associated with defining suitable membership functions (Ahlqvist et al., 2000). For example, it is one thing to recognise a mountain as a vague entity (e.g., Fisher and Wood, 1998; Varzi, 2001; Smith and Mark, 2003), but quite another to determine a justifiable function with which a location can be evaluated for the degree to which it is part of the mountain (e.g., Fisher et al., 2004). This challenge is perhaps even greater in the case of socially defined *place*-based data because they are often defined by *perceived* bounds and so are inherently subjective, making it impossible to create a justifiable membership function.

Fuzzy approaches have previously been combined with Bayesian methods because they lend themselves well to the formulation and analysis of subjective concepts (Taheri and Behboodian, 2001), though this has not yet been the case in the GIS literature and no spatial applications have previously been presented. Nevertheless, this relationship clearly has great potential in a geographical context, as Bayesian inference can be used to combine multiple types of evidence to determine the degree membership of a vague region (which is represented as a fuzzy set). The generalisation of Bayesian statistics to fuzzy data approaches has previously been referred to as ‘Fuzzy Bayesian Inference’ in the mathematical literature (Frühwirth-Schnatter, 1993); and we will adopt this terminology for the spatial implementation presented here.

In Fuzzy Bayesian Inference (FBI), the membership functions upon which fuzzy methods usually rely are replaced with *possibility distributions*, which are more directly relatable to probability theory and so allow for the use of Bayes' theorem (Gentili, 2021; Bacani and de Barros, 2017). The distinction between *possibility* and *probability* is the same as that between fuzziness and randomness (as described in the preceding section). The possibility distribution function therefore represents the state of knowledge of an agent, returning a value based on the current evidence ranging between 0 when a state is impossible and 1 when a state is totally possible (Gentili, 2021). To illustrate this, consider Bayes' theorem in **Equation 1**:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

In the conventional probabilistic format, Bayes' theorem comprises the following terms:

- 'posterior probability', *P(A|B)*: the updated probability of state *A*, given data *B.*
- 'likelihood function', *P(B|A)*: the probability of the data *B*, assuming that state *A* is true
- 'prior probability', *P(A)*: our current belief about the probability of state *A*, before accounting for data *B*
- 'plausibility' *P(B)*: the probability of observing the data *B* irrespective of state A, which simply serves to normalise the resulting value to the scale 0-1.

In our fuzzy (*possibilistic*) case, where *A* is the state and *B* is the supporting evidence, the *possibility distribution* is the likelihood function (Gentili, 2021), the prior *probability* is replaced by the prior *possibility,* the *plausibility* remains to serve to normalise the resulting membership value and the posterior *probability* is replaced by the posterior *possibility distribution*, from which we can derive a degree of membership on the scale [0-1]. To provide an example once again in the context of intergroup segregation, state *A* would be that a location was part of the territory of a given group and B would be data that will be used to update the posterior possibility of that state (e.g., the presence of individuals from either that or another group).

***Case study: modelling segregation and inter-group boundaries***

In order to provide an example of the benefits that FBI can have for analysis of socially-derived geographical phenomena, we will provide a case study concerning the spatial analysis of intergroup segregation. There is a rich literature of the quantitative measurement of segregation, with a variety of indices and measures, the overwhelming majority of which are based on the use of administrative tessellations (e.g. census zones; Catney, 2018). Such approaches typically take these zones (comprising *imposed* boundaries) as a proxy for ‘neighbourhoods’ or similar socially-derived areas (which typically exhibit *perceived* boundaries) and carry the implicit assumption that aggregated data can be used to infer individual experiences (Wong and Shaw, 2011; Farber et al., 2012). This approach carries with it three fundamental methodological challenges that are widely understood in the literature and equally apply to many other areas of research into geographical patterns of social phenomena. The first of these is the *Ecological Fallacy* (Robinson, 1950), whereby all locations and members of the population within each zone are implicitly (and incorrectly) assumed to share common characteristics. Such datasets and the analyses arising from them implicitly assume relationships between all individuals within a given zone and no relationship at all between members of different zones, effectively removing each zone (and by extension each individual) from its spatial context (Farber et al., 2012). The second issue is the *Modifiable Areal Unit Problem* (MAUP; Openshaw, 1981). The MAUP comprises two interrelated issues: (i) a *zoning effect*, representing a variance in results caused by the use of alternative areal unit delineations; and (ii) a *scale effect*, reflecting the sensitivity of the results and inferences to areal unit size. The former effect is clearly demonstrated in the context of the present research by Huck et al. (2019), who identify a ‘Small Area’ (a standard census unit for Northern Ireland) that appears to be mixed, but in fact represents two highly segregated communities aggregated into the same zone (also Davies et al., 2019). The latter effect is examined in detail by Wong (1997), who notes that simply using smaller zones can increase measured levels of segregation, as a result of the positive spatial autocorrelation of people in the same groups. The third issue is the *Uncertain Geographic Context Problem* (UGCoP; Kwan, 2012), which is a phenomenon arising due to the spatial and temporal uncertainty in the zones and the way in which they deviate from the actual areas they are intended to represent. A simple example of this problem is the use of census zones as a proxy for ‘neighbourhoods’, which is a common and convenient solution for many types of spatial analysis (Labib et al., 2020), but which fails to acknowledge that the *imposed* boundaries defined by administrative agencies are unlikely to map on to the individualistic notion of ‘neighbourhood’ that is understood by the people living within it (Grannis, 2005; Goodchild, 2011; Wong and Shaw, 2011).

Approaches such as those described above also carry a fourth challenge that is specific to the study of segregation: an overwhelming focus upon only *residential* patterns of segregation (i.e., measures of segregation that consider only where people live, not where they work or spend their leisure time). However, there is an increasing recognition that individuals may experience different levels of segregation across their various socio-geographical spaces, not only residential spaces. As a result, authors such as Schnell and Yoav (2001); Wong and Shaw (2011); Farber et al. (2012); and Huck et al. (2019) have begun to develop a range of methods that allow for more individualistic ‘activity space’ approaches to understanding patterns of segregation. Understanding behaviour at the individual level is of vital importance to understanding dynamic patterns of segregation (Dixon et al., 2020a; Dixon and McKeown, 2021).

Approaches based on administrative tessellations have been extremely valuable in understanding the fundamental characteristics of urban residential segregation and the associated negative socioeconomic consequences (Dixon et al., 2020a). Nevertheless, such research has provided an incomplete understanding of the nature of segregation, ignoring the time that people spend outside the home (Wong and Shaw, 2011). A more complete understanding of segregation must account for the time that people spend at work, in places of education and in everyday activity spaces such as street corners, parks, markets and leisure facilities; as well as the time that they spend travelling in cars, on public transport and on foot (Dixon et al., 2020a; 2022). We will therefore present a spatial application of FBI to a geographical model of territory between segregated groups in Belfast, UK. In doing so, we will demonstrate how this approach addresses all four of the above challenges in order to provide a deeper understanding of patterns of sectarian territory in the region.

## Materials and methods

### *Study area*

This study is based in a region of North Belfast, Northern Ireland, UK. Segregation in Northern Ireland principally occurs between the two main communities: *Catholics* and *Protestants* (Merrilees et al., 2018; Roulston and Young, 2013). The nature of the conflict is, however, far more complex than this religious nomenclature suggests: the chief driver of the conflict is ethno-political, with *Unionist* Protestants tending to identify as British and wishing to remain part of the United Kingdom and *Nationalist* Catholics tending to identify as Irish and wishing to unify with the Republic of Ireland (Mac Ginty et al., 2007; Merrilees et al., 2018; Roulston and Young, 2013). Segregation and sectarianism are everyday realities for many residents of Northern Ireland (Roulston and Young, 2013) and despite the conflict officially ending with the ‘Good Friday Agreement’ in 1998, daily routines, practices and mobilities of individuals in North Belfast remain significantly impacted by the ongoing effects of sectarianism (Hamilton et al., 2008; Dixon et al., 2020b). Notably, residential patterns in this part of the city persist in a distinctive ‘checkerboard’ pattern in which nationalist and unionist communities exist in close proximity yet remain divided in their everyday activities and use of space, with divisions often enforced by physical barriers known as ‘peace walls’. Specifically, we will focus on five pairs of adjacent Catholic/Protestant communities, which are indicated on **Figure 1**.

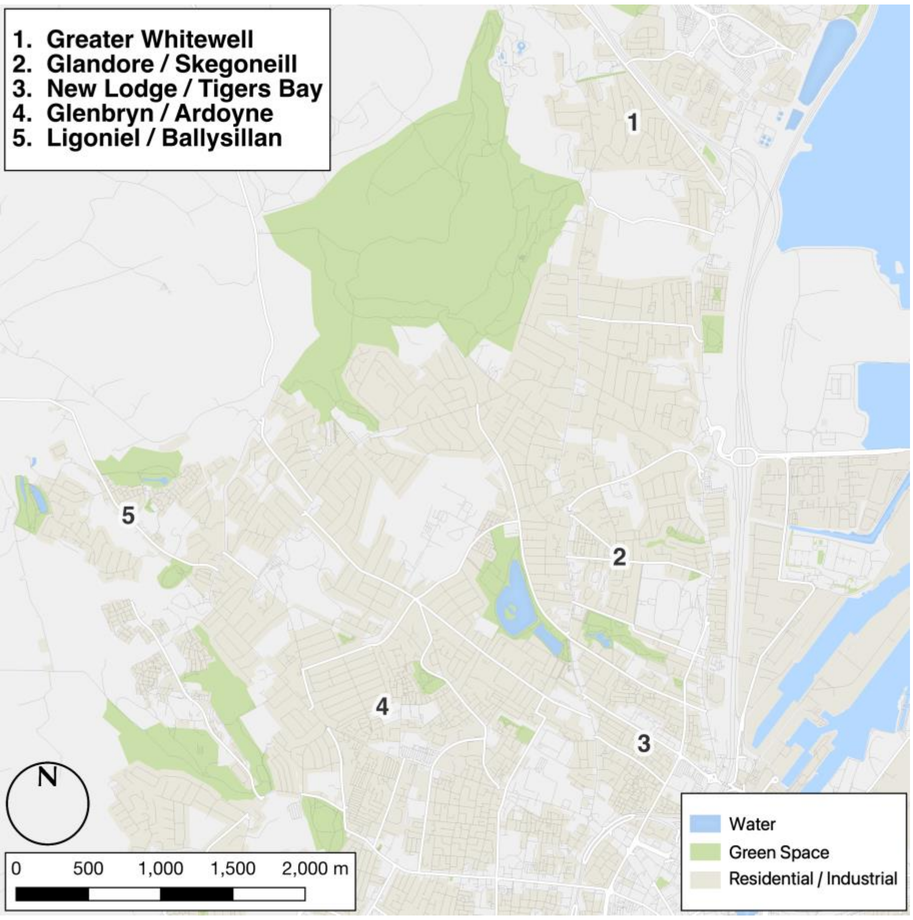

**Figure 1:** The study area and North Belfast communities upon which this research focuses. Base map data © OpenStreetMap Contributors.

### *Data collection*

Our approach comprises a spatial application of FBI to combine evidence from three primary sources and one secondary source in order to produce fuzzy surfaces in which each cell represents the degree of membership of the territory of a given group (Catholic or Protestant). The secondary dataset is the Northern Ireland Small Area (SA) data from the 2011 census, which was obtained from NISRA (2011) and includes the percentage of Catholic and Protestant residents in each area, from which the overall ratio can be calculated and the areas classified (illustrated in **Figure 2A**). This approach to representing territory only accounts for residential patterns and suffers from the problems described earlier, but provides a useful basis for our *prior possibility*, to which further evidence can then be added.

All three primary datasets were collected during a continuous campaign between February and December 2016. The first primary dataset used in this research comprise survey data collected from 488 residents of the study area, of which 242 were Catholic and 246 Protestant; 196 were male, 291 female and 1 did not disclose a gender. Participants were asked a range of questions relating to their experience of segregation, but here we only use whether they classify themselves as Catholic or Protestant and the location of their home. These data on Catholic and Protestant residential locations are illustrated in **Figure 2B** and the dataset and methodology are described in more detail in Dixon et al. (2020a).

The second primary dataset comprises Participatory GIS (PGIS) data collected using the Map-Me platform[1], which uses a 'spraycan' (or 'airbrush') interface for users to add data to the maps (Huck et al., 2014). This is intended to better capture the vagueness inherent in the data avoiding the imposition of '*artificial precision*' (after Montello et al., 2003) by forcing place-based data into fixed boundaries. Participants use the zoom level of the map to control the level of precision and density of the spray; and this is therefore often used as a proxy for

[1] http://map-me.org

‘strength of feeling’ (Huck et al., 2019). Data were collected from 33 residents of the study area, of which 14 were Catholic, 17 Protestant and 2 ‘Other’; 21 were male and 12 female. Participants were asked to use the ‘spraycan’ interface to ‘spray paint’ onto a Google Map in response to the following prompts: “*Please spray the areas you would consider to be Catholic”*, “*Please spray the areas you would consider to be Protestant”* and “*Please spray the areas you would consider to be Mixed”* (i.e., not segregated). This PGIS survey was conducted in the form of a one-to-one mapping exercise with each participant to ensure that the data reflected participants’ intentions (i.e., the data were not affected by mistakes or difficulties using the platform etc.). The resulting dataset is illustrated in **Figure 2C** and described in detail in Huck et al. (2019), whilst the software is described in Huck et al. (2014).

The third primary dataset comprises GNSS[2] (Global Navigation Satellite System) traces collected for a period of up to 14 days from 196 participants, of which 93 were Catholic, 91 Protestant and 12 ‘Other’; 79 male and 117 female. Data were collected using a custom Android mobile phone application, which recorded participants’ location at 4-second intervals and uploaded them to a server along with a timestamp and estimate of accuracy. Participants could pause the app for defined periods of time, but otherwise the application continued to track even if it was closed or the device restarted, ensuring high levels of data capture. The raw GNSS traces (comprising approximately 21.7 million data points) were processed as described in Davies et al. (2017). This dataset is illustrated in **Figure 2D** and is described in more detail in Hocking et al. (2018) and Dixon et al. (2020a).

[2] GNSS is the generic term for satellite-based navigational systems, prominent examples of which include GPS, GLONASS and Galileo.

These four datasets comprise the evidence that will be combined using FBI in order to estimate a possibility distribution for each location in the study area (each cell in the surface), from which we can derive values for both membership and uncertainty. We will do this for both Catholic and Protestant territories, yielding two output surfaces: one in which each cell contains a value representing the degree to which a given location is part of a Catholic territory; and one in which each cell contains a value representing the degree to which a given location is part of a Protestant territory (we will refer to these values as 'territoriality'). Note that both are required as the two datasets are not necessarily perfectly inverse of each other due to the presence of the 'mixed' and 'other' classifications in the input datasets.

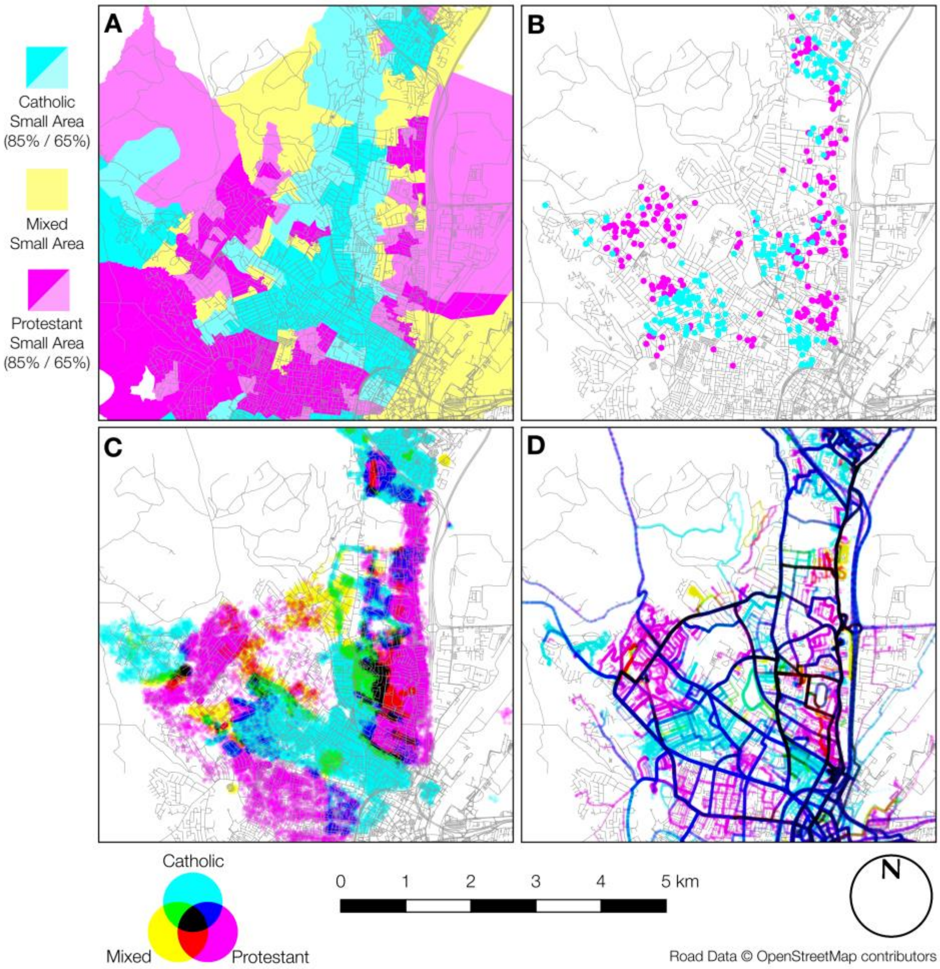
A
B
C
D
Catholic Small Area (85% / 65%)
Mixed Small Area
Protestant Small Area (85% / 65%)
Catholic
Mixed
Protestant
0
1
2
3
4
5 km
N
Road Data © OpenStreetMap contributors

**Figure 2:** Key Datasets: (A) Small Area Census Zones, including percentage of Catholic and Protestant residents; (B) Location of 488 Catholic and Protestant households that were surveyed; (C) PGIS data from 33 participants relating to perceptions of Catholic and

Protestant territory, as well as mixed areas; (D) GNSS Traces for 196 Catholic and Protestant participants.

***Fuzzy Bayesian inference***

Our approach seeks to construct one FBI model for each cell in a surface that covers the full extent of the study area. Each location will therefore have a separate possibility distribution, which is initialised with a prior possibility using the residential ratios calculated from the Small Area census data. This approach is described as the use of an *informed prior* (i.e., one based on pre-existing knowledge, information or belief), which is preferable to starting with a *flat prior* (i.e., starting with no information). Despite the issues that we have described arising from sole reliance upon this dataset as a measure of territory, it nevertheless provides a sensible starting point for an inference model. It also has the advantage of spatial contiguity, meaning that there will be at least some data for all locations in the study area. We then iteratively add evidence from each of the three primary datasets to determine a new posterior possibility distribution. Each iteration of our spatial implementation of FBI comprises two distinct steps. The first step requires the use of an *evidence function* to gather spatial data (evidence) from the dataset in question (either survey, PGIS and GNSS data) and process this into a distance-weighted *evidence value*. The second step then actually incorporates this *evidence value* into the model for a given location to estimate the posterior possibility distribution.

In the first step, the evidence for a given cell being a member of catholic or protestant territory is gathered using the *evidence function*, the specifics of which are likely to vary between applications of FBI. The simplest version of this would be a count of data points within a given distance of the cell location, though in practice it is likely that some form of distance weighting function will be applied to ensure that data closest to the cell location are privileged over more peripheral data. In this example, our *evidence function* comprises an inverse distance weighted squared (IDW$^2$) score based on an attribute-weighted count of all data points relating to each group within a certain distance of the cell centre (the bandwidth), as per **Equations 2** and **3**. Data are weighted twice in this instance: once based on their location (using IDW$^2$ as described above), and once based on their attributes (depending on the dataset, as described below). This provides a good 'general' example of an *evidence function* that could be applied to a range of datasets and applications. As with many GIS algorithms that require a bandwidth parameter, there is rarely clear empirical evidence to support the selection of a specific bandwidth value. Effort should be made to ensure that this value reflects the nature of the phenomena in question insofar as is possible, though the impact of minor changes in bandwidth will be limited due to the IDW$^2$ weighting.

$$v_g = \{ m_g \cap c_r \} \tag{2}$$

Where: $v_g$ is the evidence relating to group $g$ (Catholic or Protestant) for a given cell location; calculated as the intersection between: $m_g$, which is the complete set of data points (evidence) relating to group $g$; and $c_r$, which is a circle of radius $r$ (the selected bandwidth) centred at the given cell location.

$$e_g = \sum_{i=0}^{|v_g|} w_i \left( 1 - \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{r} \right)^2 \tag{3}$$

Where: $e_g$ is the *evidence value* for group $g$ at a given cell location; $|v_g|$is the length of the evidence set; $x$ and $y$ represent the projected coordinates of the given cell location; $x_i$ and $y_i$ are the projected coordinates of data point $i$; and $w_i$ is the weighting value for data point $i$, which may be dataset specific and can vary either by data point or by dataset. In this case, weighting values ($w_i$) for the PGIS data were calculated on a scale of 0-1, determined by the zoom level at which the PGIS data were created (as a proxy for strength of feeling; Huck et al., 2019), with data produced at the smallest scale given the lowest weight and vice versa. For the GNSS data, $w_i$ was the number of individuals that produced the data, with a larger number of individuals resulting in a greater weighting. In absence of evidence to support weighting the Survey data, $w_i$ was fixed at 1 for this dataset.

In the second step, the evidence values for each group and location ($e_g$ for Catholic and Protestant territory) are then used to construct a Multinomial *possibility distribution,* with the parameter vector drawn from a Dirichlet distribution. The relationship between these distributions is well established in Bayesian methods, with the latter representing the conjugate prior of the former. This simply means that, for a *possibility distribution* that conforms to a Multinomial distribution, the *prior possibility distribution* will conform to a Dirichlet distribution, which is useful as knowing the distributions beforehand significantly reduces the required amount of computation in the FBI process.

When applied to probability distributions and continuous parameters, the denominator in Bayes formula (**Equation 1**) often becomes either computationally or analytically intractable for all but the most trivial models (Blei et al., 2017). It is therefore common practice in Bayesian inference to estimate the posterior distribution using a Markov Chain Monte Carlo (MCMC) approach, which allows us to sample from (and therefore estimate) the posterior distribution without the need to solve **Equation 1** directly. Once the evidence distribution has been calculated from all three evidence datasets, we therefore use the No U-Turn Sampler (an efficient MCMC algorithm; Hoffman and Gelman, 2014) to estimate the parameters of the *posterior possibility distribution* based on 1000 sample draws. Once we have estimated our *posterior probability distribution*, we simply report the mean (our value for the degree to which a location is a member of the given group territory) and the width of the 95% credible interval, which provides our value for the uncertainty associated with this membership value. This process is repeated with a separate model for every cell in the output raster surface for each community. All analysis was undertaken in Python 3.8 and the source code repository, complete with example data, is given at the end of this manuscript.

## Results

To illustrate the proposed method, fuzzy surfaces describing the membership of a given location to the Catholic and Protestant territories were calculated using a 20m resolution (cell size) and 40m bandwidth (radius). Both surfaces are illustrated in **Figure 3**.

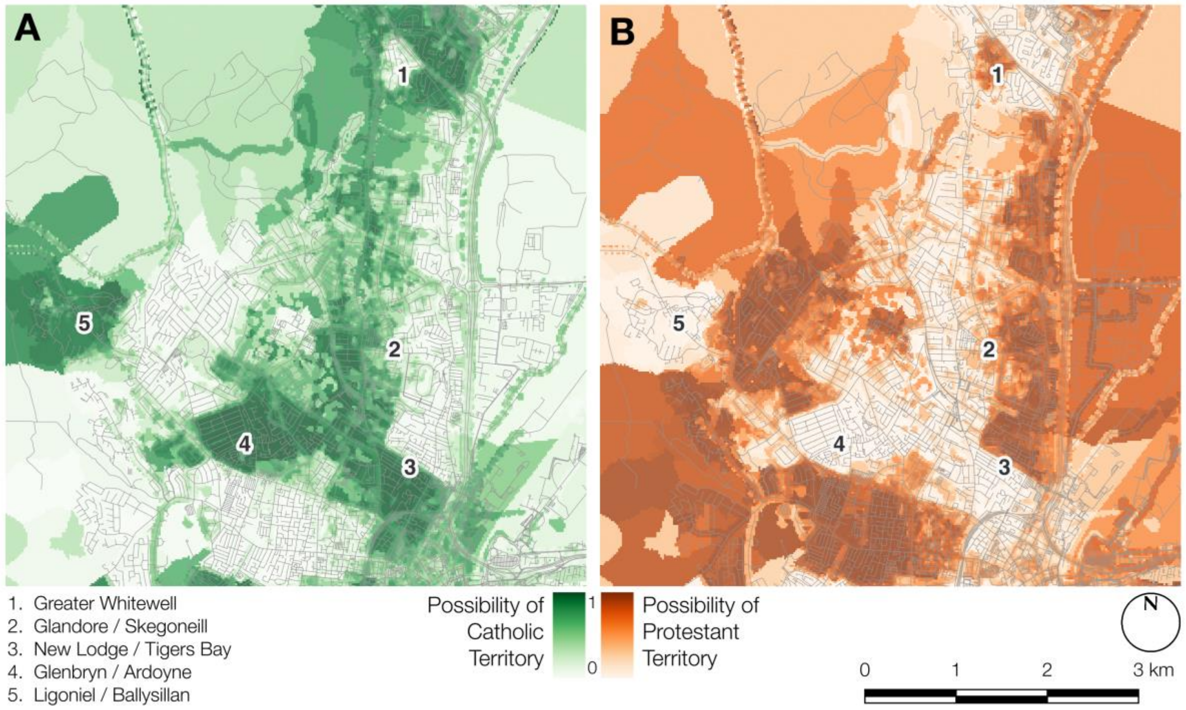
A
B
1
2
3
4
5
1. Greater Whitewell
2. Glandore / Skegoneill
3. New Lodge / Tigers Bay
4. Glenbryn / Ardoyne
5. Ligoniel / Ballysillan
Possibility of Catholic Territory
1
0
Possibility of Protestant Territory
N
0
1
2
3 km

**Figure 3:** Fuzzy surfaces describing the degree to which a given location is a member of (A) Catholic and (B) Protestant territory. *Road data © OpenStreetMap Contributors.*

### *Comparison with the small area census data*

To explore the benefits of this approach, we will provide a brief comparison with traditional zonal (census-based) approaches to understanding territory. Because census data only report data relating to residences but are spatially contiguous and so also incorporate non-residential areas, their use in attempts to understand territory can be misleading. This occurs, for example, where Small Area zones contain non-residential facilities that are shared, but which are misclassified in conventional analyses because the census data only accounts for residences (which are segregated). This is an example of the Ecological Fallacy (Robinson, 1950), whereby all of the area inside the zone is erroneously considered to match the characteristics of the residential part. In the study area, shared facilities are often located in the vicinity of residential areas as part of attempts to promote integration, such as through the creation of 'integrated schools' (i.e., schools intended to be attended by both Catholic and Protestant children) and other 'shared' facilities that are created to promote interaction between members of both groups. One such example is Cliftonville Integrated Primary School, which is located in the Catholic Cliftonville community and is accordingly represented as such in the Census data (>92% Catholic, Figure **4Ai**). In the fuzzy surfaces, however, the school is recognised as a mixed 'island' (i.e., a relatively low possibility of membership to either community) in this otherwise strongly Catholic territory (black dot, figure **4Aii** and **4Aiii**). There are many examples of such facilities (integrated schools, parks, shopping centres etc.) that are 'lost' due to the spatial aggregation inherent in the production of administrative tessellations.

A second example occurs where census zones are plotted across multiple neighbourhoods, causing them to be aggregated and misrepresent the actual underlying patterns. This is an example of the MAUP (Openshaw, 1981) as well as the UGCoP (Kwan, 2012), in that the implicit assumption that suitably small zones can act as a proxy for neighbourhoods fails either spatially, temporally, or both. An illustration of this effect occurs in the Bellevue ward (**Figure 4B**), where there are multiple examples of Small Area zones that are plotted across perceived community boundaries, normally where community boundaries are quite small. In the case of the Bellevue ward, this has created an erroneous ‘mixed’ area in the census data (**Figure 4Bi,** 54% Catholic and 43% Protestant), whereas in fact the area is strongly Catholic (**Figure 4Bii** and **4Biii**). In this case, it is likely that the Small Area encloses two highly segregated communities, as opposed to one mixed one, which is an example of both the MAUP and the spatial component of the UGCoP, as the assumption that the Small Area is representative of a neighbourhood is flawed. However, as is clear from **Figure 4Bii** and **4Biii**, this is no longer the case and the surrounding area is now predominantly Catholic. This is likely due to temporality component of the UGCoP, as the study area contains many areas of housing that will change territory over time because of demographic shifts in the area. For example, there is a widely held perception that wealthier and more socially mobile members of the Protestant community tend to leave the area for more desirable areas, resulting in a so-called 'greening' of North Belfast as their places are taken by Catholics, leading to shifts in local territorial boundaries and increases in social tensions.

Finally, we should also consider situations in which a fuzzy model might not be expected to perform as well, such as where there are very sharp boundaries between territories (normally accompanied by a physical boundary such as a peace wall), which *does* coincide with a Small Area zone boundary, meaning that the census data provides a good representation. Such an example can be found in the boundary between the Duncairn (Protestant) and New Lodge (Catholic) communities, which occurs along Duncairn Gardens (a main road) and is defined by a series of peace walls that provide a precise physical boundary (**Figure 4Ci**). Here, the fuzzy surfaces have performed well, maintaining relatively sharp edges whilst allowing for the ‘mixed’ behaviour on the road itself and the comparatively less segregated industrial buildings that open onto Duncairn Gardens, some of which are separated from the more segregated residential areas by the peace walls (**Figure 4Cii** and **4Ciii**).

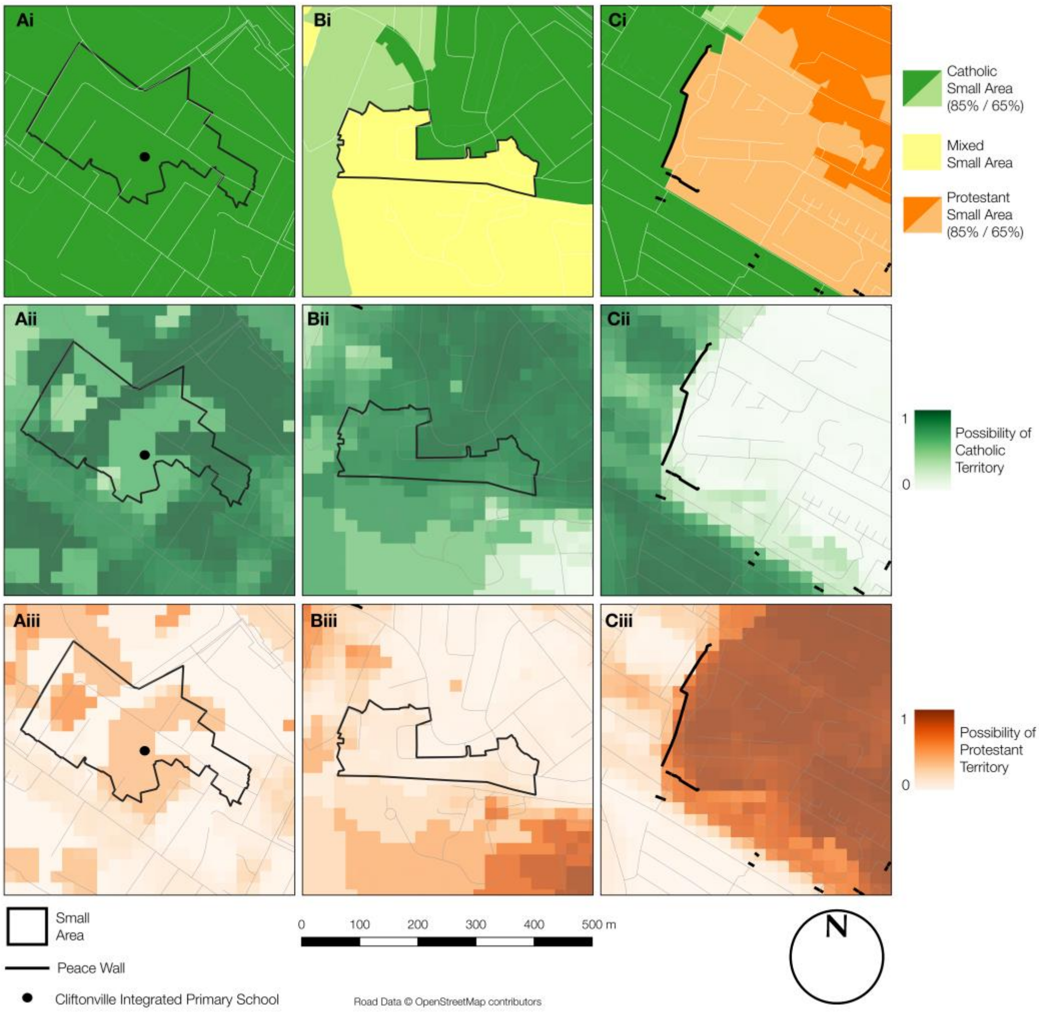
Ai
Bi
Ci
Catholic Small Area (85% / 65%)
Mixed Small Area
Protestant Small Area (85% / 65%)
Aii
Bii
Cii
1
0
Possibility of Catholic Territory
Aiii
Biii
Ciii
1
0
Possibility of Protestant Territory
Small Area
Peace Wall
Cliftonville Integrated Primary School
0
100
200
300
400
500 m
Road Data © OpenStreetMap contributors
N

**Figure 4:** Comparison of Small Area census data **(i)** with fuzzy possibility surfaces for Catholic (**ii**) and Protestant (**iii**) territory, illustrating all three datasets for three locations to demonstrate how FBI addresses **(A)** the Ecological Fallacy (Cliftonville Integrated Primary School), **(B)** The UGCoP (Bellevue) and **(C)** sharp boundary (between New Lodge and Duncairn).

### *Uncertainty in the fuzzy surfaces*

The uncertainty for each of the surfaces in **Figure 3** is derived from the 95% credible interval of the possibility distribution for each cell in the surface, the result of which is shown in **Figure 5**. In spatial applications of FBI, it is essential to report uncertainty maps alongside the results to ensure that results are expressed with the appropriate level of confidence, allowing the quality of evidence to be evaluated. Areas of high uncertainty can therefore either be discarded or further evidence can be collected to reduce uncertainties to acceptable levels.

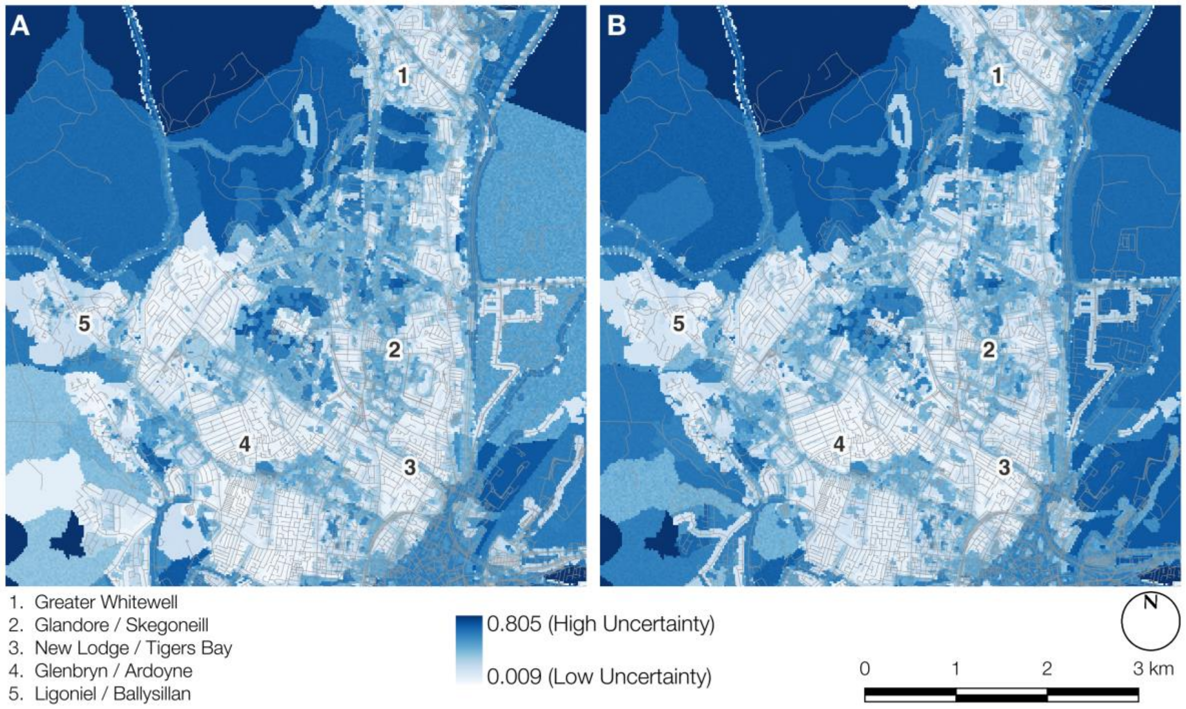

**Figure 5:** Surfaces describing the level of uncertainty (width of the 95% credible interval of the possibility distribution) at each location in (A) the Catholic territorial surface and (B) the Protestant territorial surface. *Road data © OpenStreetMap Contributors.*

As expected, uncertainties are lowest in the communities from which participants were drawn (identified in **Figure 1**) and highest in the peripheral zones of the map, where we did not collect sufficient data to make a reliable estimate of territorial membership. The visibility of Small Area boundaries to the northwest of the **Figure 5**, for example, demonstrates that little or no further evidence has been added to the prior possibility in these regions. The uncertainties within our target communities are generally very small, with the largest occurring either at locations in which we collected less data such as Cliftonville golf club (approx. 2km west of Glandore / Skegoneill, labelled 2 on the map) and the cluster of schools and churches between Fort William Park and Somerton road (approx. 1km north of Glandore / Skegoneill, labelled 2 on the map); or in locations that are shared by both groups, such as Cityside Retail Park (approx. 1km south west of New Lodge / Tigers Bay, labelled 3 on the map) and Hillview Retail Park (immediately south east of Glenbryn / Ardoyne, labelled 4 on the map).

### *Composite fuzzy territoriality map*

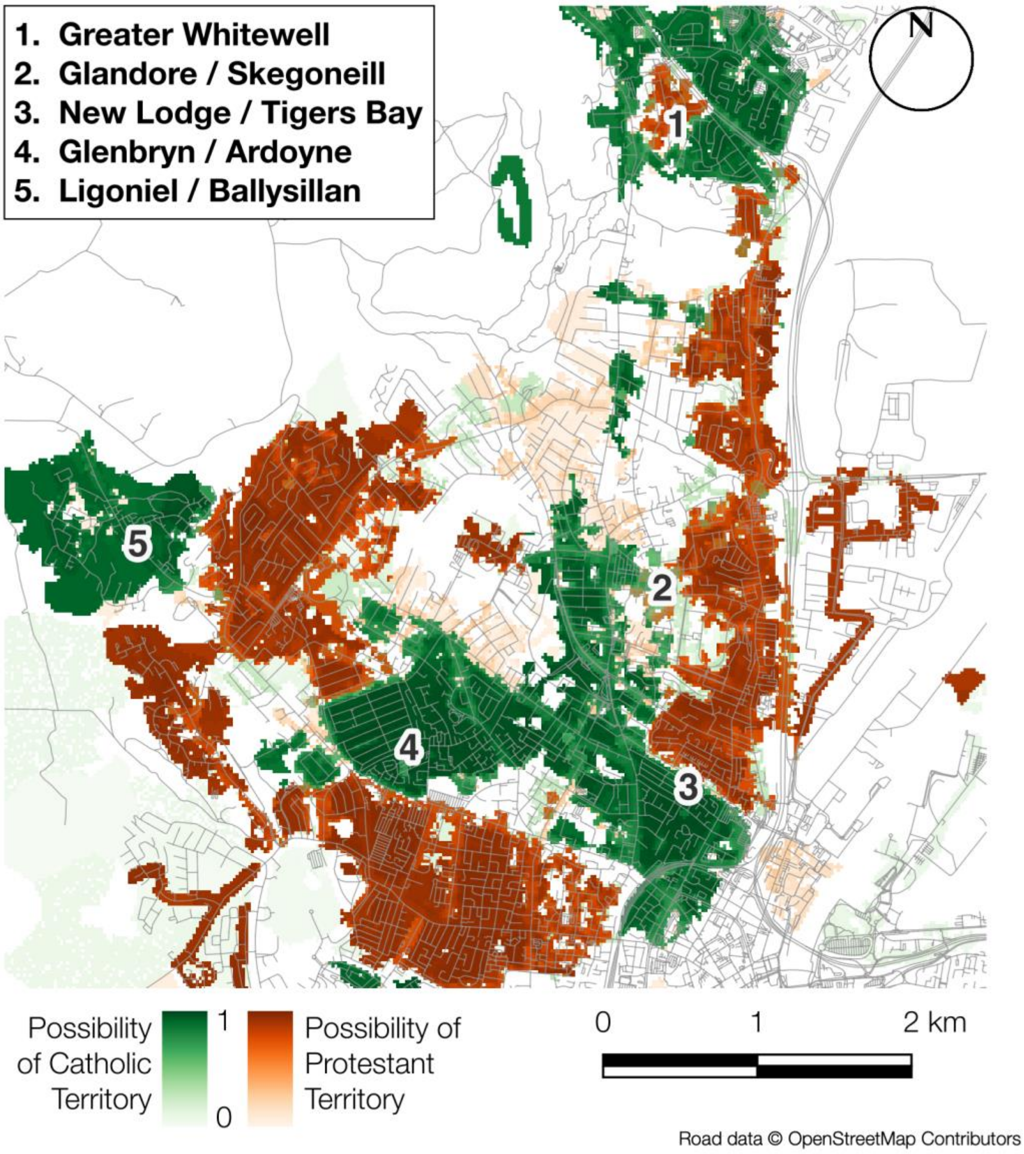
1. Greater Whitewell
2. Glandore / Skegoneill
3. New Lodge / Tigers Bay
4. Glenbryn / Ardoyne
5. Ligoniel / Ballysillan
N
1
2
3
4
5
Possibility of Catholic Territory
1
0
Possibility of Protestant Territory
0
1
2 km
Road data © OpenStreetMap Contributors

**Figure 6:** A fuzzy map of territoriality demonstrating the degree to which each part of the area is a member of either Catholic or Protestant territory. Areas with the darkest colours are

members of their respective territory to the greatest degree, whereas areas with the lighter colours are more likely to be shared. Areas for which the level of uncertainty was greater than 0.1 are excluded.

Based on the FBI and associated uncertainty surfaces above, it is a simple matter to extract only those areas for which there is a high degree of confidence in the territoriality value to create a fuzzy representation of the 'core' community territories. **Figure 6** presents a composite surface of areas for which the degree of uncertainty was ≤0.1. A standard 'sieve' operation has been used to remove extremely small areas of territory. Darker colours indicate territorial membership to a greater degree, whereas the lighter colours indicate areas that are shared. As with the individual layers presented in **Figure 3**, variations in membership both between (inter-territorial) and within (intra-territorial) territories are clear, as are the patterns of segregation and sharing.

## Discussion and conclusion

This paper is not intended as a criticism of the use of administrative tessellations in research, as they are well suited to the purposes for which they are intended. However, the uncritical use of crisply defined polygons as a proxy for socially-derived regions is widely understood to be inadequate and there is a clear need for new approaches. This paper demonstrates one such approach, which allows researchers to gain deeper insight into geographical areas with perceived boundaries by combining information from multiple data sources. Fuzzy approaches have long been recognised as a potential solution for modelling vague geographical entities, but applications of fuzzy methods to social boundaries are extremely rare in the literature. One isolated example that has similarities to the approach presented here is provided by Gao et al. (2017), who present two approaches to mapping *cognitive regions* through the synthesis of multiple geotagged datasets from web and social media sources. These approaches, however, which are based on a grid-based and point-clustering approaches combined with an evaluation of agreement between the data sources, provide neither the formal analytical framework, nor the ability to evaluate uncertainties afforded by FBI.

A major contributing factor to the lack of uptake of fuzzy approaches to social boundaries is the challenge in producing justifiable membership functions for vague geographical entities (i.e., those with perceived boundaries), which can often prove to be as arbitrary as the administrative zones that they seek to replace. Here, we address this issue by demonstrating how a spatial implementation of FBI can be used to produce *possibility distributions* that describe both the expected degree of membership of a given location to a set (i.e., Catholic or Protestant territory) and the level of uncertainty associated with that value. The ability to combine primary and secondary evidence from multiple official (e.g., census), empirical (e.g., GNSS Traces, Survey) and qualitative (e.g., PGIS) datasets, means that we can capture multiple facets of what ‘territory’ means to both those who are included and excluded from it, which in turn facilitates a more sophisticated and nuanced analysis. In the context of the case study, this sophistication includes the ability to look beyond only residential patterns of segregation and instead provide a holistic model that accounts for both residential and activity space patterns, as well as both official and individual views.

There are two key limitations to this spatial application of FBI that might impact upon adoption. First is the requirement for extensive data collection. Most work on residential segregation exploits government census data, which is readily available and the product of a huge investment of time and money. The FBI case study presented here required the collection of large amounts of additional primary data (the survey, PGIS and GNSS data) to provide evidence for the inference of the *possibility distribution*, which may limit applications to only those projects that are adequately resourced to undertake such work. Clearly, shifts toward open data (including the web- and social media-based datasets such as those used by Gao et al., 2017) will help with this in some areas by preventing the duplication of efforts, but it is likely that this will remain a barrier to adoption for some applications. As with any inference-based approach, the quality and representivity of the input datasets are also of great importance, as systematic biases in the data will inevitably be reproduced in the resulting maps, so rigorous collection techniques are required to ensure meaningful outputs.

The second key limitation is that FBI is highly computationally intensive due to the use of the MCMC simulation to estimate the *posterior possibility distribution*, which is computationally intensive. In most cases, it would not be possible to determine the uncertainties associated with the resulting surfaces without this simulation step, though alternative approaches such as quadratic approximation would provide a more efficient estimate for some simple models (see McElreath, 2020). To address this issue, the model presented here was calculated in 40 subregions that were processed in parallel using a cluster computing facility and then stitched together at the end. Where such facilities are not available, the computational burden of this approach would be significant, which could limit the extent of the study area or the resolution of the output surfaces. However, given the increasing popularity of Bayesian methods in recent years, there are promising examples of ways in which to increase the computational efficiency of MCMC processes (e.g., Rajabi and Ataie-Ashtiani, 2016). It is likely that such advancements will continue to improve the computational efficiency of MCMC, thus reducing the computational burden of spatial applications of FBI.

The outputs presented here were primarily intended as inputs to other types of model, such as Agent Based Models that require agents to have a detailed understanding of the environment that cannot be satisfactorily obtained from census data alone. However, the output maps clearly hold substantial value in their own right, providing new perspectives on patterns of segregation, territoriality and the use of shared spaces, which would have great value in the formulation of future policy. FBI also creates rich possibilities for the exploration of the different lived and perceived landscapes of spatial division. For example, data from different groups of participants could be analysed separately to understand similarities and differences in the ways in which boundaries are understood. For example, in the context of our case study, FBI would permit the exploration of questions such as whether young people who grew up in the years following the Good Friday Agreement view sectarian boundaries in north Belfast in the same way as older residents who lived through 'the Troubles'. Factoring in varying and perhaps even contested boundary perceptions into Bayesian models might yield quite different maps of the divided city and thus reveal deeper and more understanding of how perceptions vary between groups existing within the study area.

Similarly, if evidence is continually fed into an FBI model with a temporally weighted evidence function, then it would permit these patterns to be tracked over time. This could provide authorities, NGOs and communities to gain deeper insight into the patterns that have such a significant impact upon their daily lives and allowing policies and interventions to be evaluated by observing temporal changes in attitude and behaviour. From this perspective, FBI could prove to be a valuable tool for providing up-to date understandings of territories, which would have a significant impact upon the effective targeting of activities intended to promote integration in the study's area. This might include, for example, supporting on-going efforts to remove approximately 100 peace walls that currently exist within the study area (DoJ, 2019) by supporting decision making around which walls to remove and evaluating of the impact of the removal of walls upon territorial boundaries. Further research should seek to apply datasets produced using FBI to models of segregation, including the development of new and improved segregation metrics and the adoption of the temporal approach described above. The FBI method should also be applied to a range of other vague social regions, such as the determination of 'communities of interest' for use in electoral mapping (e.g., Phillips and Montello, 2017).

In his discussion of *place* in GIS, Goodchild (2011) identifies that, whilst GIS has been accused of taking an excessively simplistic view of many complex geographical ideas, there are clear benefits to this approach with respect to the ease with which the resulting data can be analysed, visualised and modelled. However, as Pickles (1995) recognised, there has also been much discussion throughout the history of GIS around the extent to which technologies bias, filter or otherwise intrude on the interactions between people and their environment (also Montello et al., 2003). This discussion is of great relevance here, as it is common for researchers to simply adopt census data as a proxy for neighbourhoods, communities, territories and similar simply because they are readily available as a secondary data source. FBI provides an approach that permits the modelling of vague geographical areas that can illustrate both inter- and intra-territorial differences without falling foul of widely understood issues such as the *Ecological Fallacy*, the *MAUP* or the *UGCoP*, whilst also navigating the challenges associated with fuzzy methods that required defined membership functions. We contend that convenience should not dictate the approaches taken in scientific applications, and that methodological innovations such as those presented in this paper can provide novel alternatives that enable deeper insights to be gained into a range of vague geographical entities.

## Data and codes availability statement

All Python code and data for the FBI analysis is available under an Open Source / Open Data License at https://github.com/jonnyhuck/fuzzy-bayesian-inference. The spatial data includes from the PGIS (real) GNSS tracking (simulated) and Survey (simulated). The real GNSS and Survey data could not be published as there is a significant risk of identifying individual participants. Small Area census data are available from NISRA at https://www.nisra.gov.uk/support/geography/northern-ireland-small-areas. The Android Application used for GNSS data collection is available at https://github.com/jonnyhuck/bmp-pathways-app.

## Acknowledgements

This research, conducted as part of the Belfast Mobility Project, was funded by the ESRC (ES/L016583/1). The authors would like to acknowledge the use of the Computational Shared Facility at The University of Manchester, which is provided by the Research IT team. We also gratefully acknowledge the open source Software and Python libraries that we used in this research, including the authors of QGIS (https://qgis.org), PyMC (https://www.pymc.io), GeoPandas (https://geopandas.org), Shapely (https://shapely.readthedocs.io), Rasterio (https://rasterio.readthedocs.io) Numpy (https://numpy.org) and the many other open source projects upon which these libraries depend in turn. Finally, thank the residents of North Belfast who contributed so openly and generously to this study.

## References

Ahlqvist O, Keukelaar J and Oukbir K (2000) Rough classification and accuracy assessment. *International Journal of Geographical Information Science* 14(5): 475-496.

Bacani F and de Barros LC (2017) Application of prediction models using fuzzy sets: A Bayesian inspired approach. *Fuzzy sets and systems* 319: 104-116.

Blei DM, Kucukelbir A and McAuliffe JD (2017) Variational inference: A review for statisticians. *Journal of the American Statistical Association* 112(518): 859-877.

Burrough PA (1986) *Principles of geographical information systems for land resources assessment.* Oxford, UK: Oxford University press.

Carver S, Watson A, Waters T, et al. (2009) Developing computer-based participatory approaches to mapping landscape values for landscape and resource management. *Planning support systems best practice and new methods*. Springer, pp.431-448.

Catney G (2018) The complex geographies of ethnic residential segregation: Using spatial and local measures to explore scale-dependency and spatial relationships. *Transactions of the Institute of British Geographers* 43(1): 137-152.

Cheeseman PC (1985) In Defense of Probability. *IJCAI.* 1002-1009.

Clementini E and Di Felice P (1996) An algebraic model for spatial objects with indeterminate boundaries. *Geographic objects with indeterminate boundaries*. CRC Press, pp.155-169.

Davies G, Dixon J, Tredoux CG, et al. (2019) Networks of (Dis)connection: Mobility Practices, Tertiary Streets, and Sectarian Divisions in North Belfast. *Annals of the American Association of Geographers* 109(6): 1729-1747.

Davies G, Huck J, Whyatt D, et al. (2017) Belfast mobility: Extracting route information from GPS tracks.

Dixon J and McKeown S (2021) Negative contact, collective action, and social change: Critical reflections, technological advances, and new directions. *Journal of Social Issues*.

Dixon J, Sturgeon B, Huck J, et al. (2022) Navigating the divided city: Place identity and the time-geography of segregation. *Journal of Environmental Psychology* 84: 101908.

Dixon J, Tredoux C, Davies G, et al. (2020a) Parallel lives: Intergroup contact, threat, and the segregation of everyday activity spaces. *Journal of personality and social psychology* 118(3): 457-480.

Dixon J, Tredoux C, Sturgeon B, et al. (2020b) 'When the walls come tumbling down': The role of intergroup proximity, threat, and contact in shaping attitudes towards the removal of Northern Ireland's peace walls. *British Journal of Social Psychology* 59(4): 922-944.

DoJ (2019) Interfaces Programme: A Framework Document. Reportno. Report Number|, Date. Place Published|: Institution|.

Evans A and Waters T (2007) Mapping vernacular geography: web-based GIS tools for capturing'fuzzy'or'vague'entities. *International Journal of Technology, Policy and Management* 7(2): 134-150.

Farber S, Páez A and Morency C (2012) Activity spaces and the measurement of clustering and exposure: A case study of linguistic groups in Montreal. *Environment and Planning A* 44(2): 315-332.

Fisher P (1996) Boolean and fuzzy regions. *Geographic objects with indeterminate boundaries*. CRC Press, pp.87-94.

Fisher P and Wood J (1998) What is a mountain? Or the Englishman who went up a Boolean geographical concept but realised it was fuzzy. *Geography*. 247-256.

Fisher P, Wood J and Cheng T (2004) Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers* 29(1): 106-128.

Frühwirth-Schnatter S (1993) On fuzzy Bayesian inference. *Fuzzy sets and systems* 60(1): 41-58.

Gao S, Janowicz K, Montello DR, et al. (2017) A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science* 31(6): 1245-1271.

Gentili PL (2021) Establishing a New Link between Fuzzy Logic, Neuroscience, and Quantum Mechanics through Bayesian Probability: Perspectives in Artificial Intelligence and Unconventional Computing. *Molecules* 26(19).

Goodchild MF (2011) Formalizing place in geographic information systems. *Communities, neighborhoods, and health*. Springer, pp.21-33.

Grannis R (2005) T-Communities: pedestrian street networks and residential segregation in Chicago, Los Angeles, and New York. *City & Community* 4(3): 295-321.

Hamilton J, Hansson U, Bell J, et al. (2008) Segregated Lives. *Social division, sectarianism and everyday life in Northern Ireland. Belfast: Institute for Conflict Research*.

Hocking BT, Sturgeon B, Whyatt D, et al. (2018) Negotiating the ground: 'mobilizing' a divided field site in the 'post-conflict' city. *Mobilities* 13(6): 876-893.

Hoffman MD and Gelman A (2014) The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 15(1): 1593-1623.

Huck J, Whyatt J and Coulton P (2014) Spraycan: A PPGIS for capturing imprecise notions of place. *Applied geography* 55: 229-237.

Huck JJ, Whyatt JD, Dixon J, et al. (2019) Exploring segregation and sharing in Belfast: A PGIS approach. *Annals of the American Association of Geographers* 109(1): 223-241.

Kosko B (1990) Fuzziness vs. probability. *International Journal of General System* 17(2-3): 211-240.

Kwan M-P (2012) The uncertain geographic context problem. *Annals of the Association of American Geographers* 102(5): 958-968.

Labib SM, Lindley S and Huck JJ (2020) Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. *Computers, Environment and Urban Systems* 82: 101501.

Leung Y (1987) On the imprecision of boundaries. *Geographical analysis* 19(2): 125-151.

Mac Ginty R, Muldoon OT and Ferguson N (2007) No war, no peace: Northern Ireland after the agreement. *Political psychology* 28(1): 1-11.

McElreath R (2020) *Statistical rethinking: A Bayesian course with examples in R and Stan.* CRC press.

Merrilees CE, Taylor LK, Baird R, et al. (2018) Neighborhood effects of intergroup contact on change in youth intergroup bias. *Journal of youth and adolescence* 47(1): 77-87.

Montello DR (2003) Regions in geography: Process and content. In: Duckham M, Goodchild MF and Worboys M (eds) *Foundations of geographic information science*. CRC Press.

Montello DR, Goodchild MF, Gottsegen J, et al. (2003) Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition & Computation* 3(2-3): 185-204.

NISRA (2011) *Northern Ireland Small Areas*. Available at: https://www.nisra.gov.uk/support/geography/northern-ireland-small-areas.

Openshaw S (1981) The modifiable areal unit problem. *Quantitative geography: A British view*. 60-69.

Phillips DW and Montello DR (2017) Defining the community of interest as thematic and cognitive regions. *Political Geography* 61: 31-45.

Pickles J (1995) *Ground truth: The social implications of geographic information systems.* Guilford Press.

Purves RS and Derungs C (2015) From space to place: Place-based explorations of text. *International Journal of Humanities and Arts Computing* 9(1): 74-94.

Rajabi MM and Ataie-Ashtiani B (2016) Efficient fuzzy Bayesian inference algorithms for incorporating expert knowledge in parameter estimation. *Journal of Hydrology* 536: 255-272.

Robinson WS (1950) Ecological Correlations and the Behavior of Individuals. *American sociological review* 15(3): 351-357.

Roulston S and Young O (2013) GPS tracking of some Northern Ireland students–patterns of shared and separated space: divided we stand? *International Research in Geographical and Environmental Education* 22(3): 241-258.

Schnell I and Yoav B (2001) The sociospatial isolation of agents in everyday life spaces as an aspect of segregation. *Annals of the Association of American Geographers* 91(4): 622-636.

Smith B and Mark DM (2003) Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design* 30(3): 411-427.

Smith B and Varzi AC (2000) Fiat and bona fide boundaries. *Philosophical and phenomenological research*. 401-420.

Taheri SM and Behboodian J (2001) A Bayesian approach to fuzzy hypotheses testing. *Fuzzy sets and systems* 123(1): 39-48.

Tolkien JRR (2001) On Fairy Stories. *Tree and Leaf.* London: Harper Collins.

Varzi AC (2001) Vagueness in geography. *Philosophy & Geography* 4(1): 49-65.

Wong DW and Shaw S-L (2011) Measuring segregation: An activity space approach. *Journal of Geographical Systems* 13(2): 127-145.

Wong DWS (1997) Spatial Dependency of Segregation Indices. *The Canadian Geographer / Le Géographe canadien* 41(2): 128-136.