



Spatial Aggregation Methods for Investigating the MAUP Effects in Migration Analysis

John Stillwell¹  · Konstantinos Daras² · Martin Bell³

Received: 12 December 2017 / Accepted: 5 August 2018 /

Published online: 13 August 2018

© The Author(s) 2018

Abstract

In this paper, we investigate the effects of scale and zone configuration on migration indicators and spatial interaction model parameters using a software system known as the IMAGE Studio. Internal migration flows in the United Kingdom and the local authority districts between which they move are aggregated into sets of increasingly fewer and larger polygons using alternative zone design algorithms. Indicators of migration intensity, impact and distance are revealed to vary significantly by scale but less so by zonation, whereas migration effectiveness and distance show greater scale independence but more sensitivity to zone configuration. Equal area and population optimised regions improve the quality of measures to a certain degree depending upon the imposition of shape constraints.

Keywords Zone aggregation · MAUP effects · Migration indicators · Spatial interaction modelling · IMAGE studio

Introduction

Spatial analysts are now familiar with the axiom that statistical indicators and model parameters that quantify different features of a particular human geographic phenomenon

✉ John Stillwell
j.c.h.stillwell@leeds.ac.uk

Konstantinos Daras
Konstantinos.daras@liverpool.ac.uk

Martin Bell
Martin.bell@uq.edu.au

¹ School of Geography, University of Leeds, Leeds LS2 9JT, UK

² School of Environmental Sciences, University of Liverpool, Roxby Building, Liverpool, UK

³ School of Earth and Environmental Sciences, University of Queensland, Chamberlain Building, Brisbane, Qld 4072, Australia

may vary with the spatial scale for which data are available and with the configuration (or shape) of the zones for which data are reported at each scale. This variation is attributable to the so-called ‘scale’ and ‘zonation’ effects of the Modifiable Areal Unit Problem (MAUP) that Openshaw (1984) documented carefully in his famous CATMOG publication and which has been addressed by a number of geographers since then, most recently by Lloyd (2014) and Manley (2014). Many studies of the MAUP effects have considered the impact of scale and zonation problems using attribute data in the form of stock variables measured for a limited set of scales and zonation systems. Our context is that of internal migration flows, where two geographies (of origin and destination) are involved and where individuals change usual address from one location to another during some period of time. Internal migration data are often released by the national statistical agencies as flows between the zones that constitute certain administrative or census geographies and in most cases, the geographies of origin and destination are equivalent. Migration flows in the 12 month period before the 2011 Census in the United Kingdom (UK), for example, are available in the form of symmetric origin-destination matrices at certain spatial scales (Duke-Williams et al. 2018) and consequently, the volume and intensity of migration between zones will be scale dependent. Thus, for example, the volume of migrants over 1 year of age between 404 local authority districts in the UK in the 12 months before the 2011 Census was around 2.8 million people and the crude migration intensity was 44.3 per thousand population, whereas only around 1.2 million individuals or 18.7 per thousand of the population moved between the 12 UK regions (2011 Census Special Migration Statistics¹ extracted from UK Data Service using WICID).

The aim of this paper is to investigate what are the MAUP implications for migration indicators and spatial interaction model parameters when we apply different zone design methods to a set of Basic Spatial Units (BSUs) for which we have data on inter-zonal migration flows such as the UK local authority districts mentioned above. We have chosen four alternative zone aggregation methods and our objective is to identify the variation in results of using each method, exposing some of the advantages and problems of each approach along the way. The algorithms are explained in detail in the third section of the paper, the data used in the analyses are introduced in the fourth section, and the results are reported in the fifth section. To begin with, however, we introduce the IMAGE project which has been the context in which this research has been undertaken and outline the structure and framework of the IMAGE Studio and its subsystems. The paper finishes with some conclusions and suggestions for further work.

The MAUP, the IMAGE Project and the IMAGE Studio

The MAUP

Whilst the MAUP was first identified by Gehlke and Biehl (1934), it remained relatively unexplored by geographers until Openshaw and Taylor (1979) demonstrated

¹ Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland Source: 2011 SMS Merged LA/LA [Origin and destination of migrants by age (broad grouped) by sex] - MM01CUK_all - Open.

how spatial data analysis using bivariate correlation methods might result in rather different coefficients depending on the number of spatial units (the scale) used to define the same area. These authors also identified an ‘aggregation problem’ as the second component of the MAUP, arising when the same number of zones were involved but their size and shape were allowed to vary. Subsequently, Openshaw and Rao (1995) used the example of Liverpool to demonstrate how the patterns of concentration of ethnic minority populations across 119 census wards in 1991 could be almost completely reversed by re-engineering the boundaries based on the underlying 2926 census enumeration districts into 119 zones of equal population.

Further explorations of the MAUP were reported in studies during the 1990s (e.g. Fotheringham and Wong 1991; Holt et al. 1996) and Marble (2000) challenged the research community to provide examples of situations in which the MAUP was an important problem. Flowerdew (2011), using bivariate correlation between pairs of variables from the 2001 Census for England, demonstrated that in many cases, the MAUP makes little or no difference but that there are some relationships where the effect is significant. Other studies (e.g. Holt et al. 1996; Tranmer and Steel 2001; Manley 2005) have provided measures that can be used to show the effect of the MAUP based on the variances of the variables concerned or within-area homogeneity. In the following sub-section, we explain the context in which an investigation of the MAUP has been imperative and outline the structure of the software system that has been developed to automate the procedures for identifying both the scale and zonation components.

IMAGE Project

The IMAGE (Internal Migration Around the GlobE) project² is an international research project funded by the Australian Research Council and based at the University of Queensland to facilitate cross-national comparisons of internal migration using a set of migration indicators that measure aspects of migration including intensity, distance, connectivity and impact (Bell et al. 2002) that can be used to advance understanding of the way that migration within countries varies around the world. Considerable effort has been spent on constructing a global inventory of internal migration data sources (Bell et al. 2015a) and creating a repository of migration and related (boundary and population) data sets (Bell et al. 2014). The IMAGE project had a number of objectives that derive from analysis of the data sets held in the repository, including the comparison of overall migration intensities in countries for which data are available or can be estimated (Bell et al. 2015b), the distances over which people migrate and the frictional effect of distance on migration (Stillwell et al. 2016) and the impact of migration on population distributions in different countries (Rees et al. 2016).

One of the key obstacles confronting cross-national comparison of migration indicators is the inequality or inconsistency in the geographical zones for which migration data are captured and collected in different countries. Every country has its own hierarchy of geographies; in some cases, such as small islands or principalities, there is only one spatial unit and no hierarchy; in other cases, data may be available for three or four tiers of geography with different numbers of spatial units in each level.

² <https://imageproject.com.au/>

However, the boundaries of each of these sets of zones define polygons that are unique in shape and size and the migration indicators associated with each geography in one country are not directly comparable with those relating to administrative or census geographies in other countries. In attempting to make comparisons of migration rates between, say, the NUTS 1 regions of the European Union (EU) countries, we encounter both components of the MAUP: there are different numbers of NUTS 1 regions in each country and the spatial configuration, i.e. the size and shape of each region, is different. Exactly the same problem applies when we attempt to make cross-national comparisons on a global level.

In response to this challenge, we have proposed a methodology which involves progressively aggregating a set of zones for any single country – called Basic Spatial Units (BSUs) – into larger and fewer zones – called Aggregated Spatial Regions (ASRs) – and generating multiple different configurations of zones at each level of aggregation or scale. Sets of migration indicators and model parameters are then computed at various levels for different configurations and summarised using measures of central tendency and deviation; variation in the value of a summary indicator from one level of ASRs to another can be identified as measuring the scale effect whilst variation in the indicator values between the zone configurations at any one level can be interpreted as the zonation effect. The IMAGE Studio has been constructed for automating the computation processes involved.

The IMAGE Studio

The IMAGE Studio is the software system that has been developed to facilitate the computation of migration indicators and model parameters for different zone systems. The framework of the Studio, illustrated in Fig. 1, involves four subsystems that are required for the: (i) initial preparation of data; (ii) aggregation of BSU polygons, migration flows and population counts; (iii) calculation of internal migration indicators; and (iv) calibration of a doubly constrained spatial interaction model (SIM). Each subsystem is autonomous, supporting standardised input/output data and executing any iterative function which is required for the analysis.

The Data preparation subsystem is where the various data sets are assembled and prepared for use in the other subsystems. The three data sets required are: (i) a matrix of migration flow counts with rows representing origin BSUs and columns representing destination BSUs and with BSU codes 1, 2, ... n in the first column and first row respectively; (ii) a vector of populations at risk with the equivalent numeric code for each BSU in the first column; and (iii) boundary data for the BSUs in the form of a shapefile containing the numeric BSU code for each polygon. A matrix of distances between BSUs (with BSU codes in the first row and column in same order as migration flows) can also be input if this is available from a particular source or has been estimated independently.

One of the key functions of the Data preparation subsystem is to generate a file of BSU contiguities from the raw boundary data since this is required for the aggregation routines in the Aggregation subsystem. The contiguity file which is generated provides critical information about which BSUs are adjacent to or tangential with other BSUs. The contiguities produced automatically by the subsystem can be visualised as lines on a map joining the polygon centroids. Figure 2 is a screenshot of the IMAGE Studio user

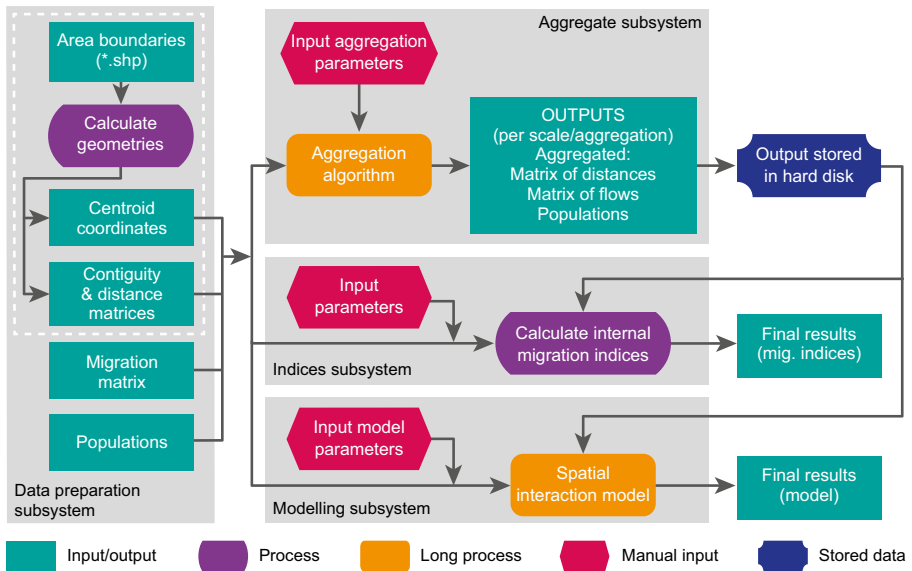


Fig. 1 The framework of the IMAGE studio. Source: Adapted from Stillwell et al. (2014)

interface showing the polygons that constitute parts of the UK in the map window and the red lines connecting centroids of adjacent polygons that have been automatically identified by the Data preparation subsystem.

It is necessary that every BSU polygon is deemed to be contiguous with at least one other polygon and that all ‘island’ polygons are joined to the rest of the system. This latter specification is important in countries where polygons are separated by stretches of water and no contiguous boundaries are present. In the UK, for example, it is clear from Fig. 2 that Northern Ireland and the Western Isles of Scotland are not ‘connected’ to the rest of mainland UK. This process is undertaken manually by adding to the contiguity file the codes of polygons that are most suitable for connection based on ferry routes or just proximity. It is necessary that contiguities are included for pairs of BSUs in both directions. A file of BSU centroids is also produced since these are the points representing the gravitational centres of all BSUs that are used to calculate distances between zones.

The Aggregation subsystem is required for the creation of spatial aggregations of BSUs into what we call Aggregated Spatial Regions (ASRs). The subsystem provides functionality for both single or multiple aggregation. In the case of the former, the user chooses the number of ASRs that are to be created from the initial BSUs and the number of required configurations of these ASRs at that one selected scale. If the raw data contained 400 BSUs, the user might want to aggregate the BSUs into 200 ASRs, for example, and produce 100 different configurations of these ASRs. Alternatively, with multiple aggregation, the user might specify a scale increment or step with which to aggregate BSUs on an iterative basis as well as the number of configurations at each scale. For example, if there are 100 BSUs and the user aggregates them using a scale step of 10 zones with 100 configurations, then the aggregations

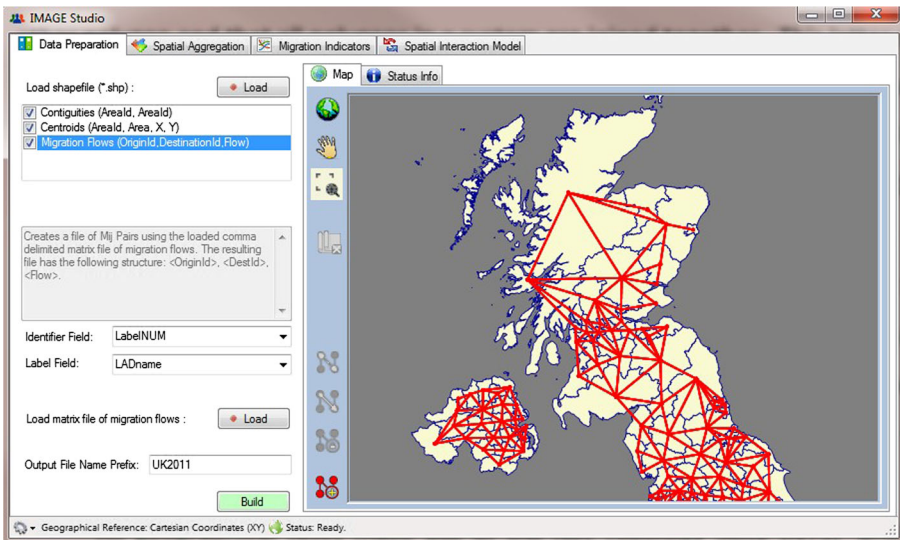


Fig. 2 IMAGE studio interface showing polygons which are defined as contiguous automatically by the data preparation subsystem

will take place into sets of 10, 20, 30, 40, 50, 60, 70, 80 and 90 ASRs with 100 configurations at each scale. Since the initial BSUs are used for creating each configuration at each scale, this process can consume considerable amounts of computer time, so fewer configurations (e.g. 50) are often adopted in practice. Implementing the aggregation process involves choosing a spatial algorithm that is fed with the normalised data from the Data preparation subsystem to produce centroid coordinates, inter-centroid distances, contiguities, flow matrices and populations for each set of ASRs which can then be used in the migration indicators and modelling subsystems. This paper reports some results generated when using different zone design algorithms that are outlined in more detail in the next section.

The Migration indicators subsystem is where internal migration indicators are calculated for the set of initial BSUs or for each set of ASRs. The subsystem calculates the indicators at two levels: indicators at the global or system-wide level refer to measures for all BSUs or ASRs; indicators at the local level refer to measures for the individual BSUs. Local migration indicators for ASRs are not computed because each set of ASRs will be different from one scale to the next and therefore comparison of local indicators between scales will be compromised. The global indicators include basic descriptive counts: total population, population density, total migration flows and the mean, median, maximum and minimum values in the cells of the migration matrix together with various measures of migration intensity, effectiveness, connectivity and inequality. The local migration indicators computed for each BSU include those used for system-wide analysis and those capturing variation in out-migration and in-migration flows and in distance, turnover and churn. Full details of how each indicator is defined and calculated are available in the Image Studio manual (Daras 2014).

The fourth subsystem of the IMAGE Studio is the Spatial Interaction Modelling (SIM) subsystem, where an optimum distance decay parameter measuring the frictional effect of distance on migration is generated by calibrating a doubly constrained SIM of the type derived by Wilson (1970) from entropy-maximizing principles and expressed as:

$$M_{ij} = A_i O_i B_j D_j f(d_{ij}) \quad (1)$$

where M_{ij} is the migration flow between zones (BSUs or ASRs) i and j , O_i is the total out-migration from zone i and D_j is the total in-migration into each destination zone j , A_i and B_j are the respective balancing factors that ensure the out-migration and in-migration constraints are satisfied, d_{ij} is the Euclidian distance between zones i and j , and $f(d_{ij})$ is a distance term expressed as a negative power or exponential function to the power β where β is referred to as the distance decay parameter. The SIM code is an updated version of an original program written in Fortran IV (Stillwell 1983) and a user can choose to calibrate a single SIM for migration for one spatial system or multiple SIMs for the flows associated with the different configurations at various scales produced by the Aggregation subsystem.

Aggregation Methods and Indicators

Automated Aggregation Methods

Two Initial Random Aggregation (IRA) algorithms have been implemented in the IMAGE Studio: IRA and IRA-wave. The former provides a high degree of randomisation to ensure that the resulting aggregations are different during the iterations. Aggregation only takes place between contiguous zones and the algorithm is implemented following Openshaw's Fortran subroutine (Openshaw 1976). The latter aggregation algorithm is a hybrid version of the former with strong influences from the mechanics of the Breadth First Search (BFS) algorithm. If we require N aggregated zones, the first step of the IRA-wave algorithm is to select N BSUs randomly from the initial set and assign each one to an empty region (ASR). Using an iterative process until all the BSUs have been allocated to the N ASRs, the algorithm identifies the BSUs contiguous with each ASR, targeting only the BSUs without an assigned ASR and adds them to each ASR respectively. The advantages of using the IRA-wave algorithm include its speed in producing a large number of initial aggregations and the fact that it produces relatively well-shaped regions in comparison to the more irregular shapes derived using the IRA algorithm.

Since the initial aim of the Aggregation system was to provide the functionality of generating sets of alternative aggregations in order to identify the zonation effect, neither of these IRA algorithms involves an objective function. However, later versions of the Studio have included the options of running single or multiple aggregations with one of two objective functions: maximize equality or similarity. The equality function aims to generate a set of N ASRs with the aggregated values of the BSUs in each ASR being equivalent to or as close as possible to a targeted value T which is given prior to

each aggregation and is measured as the sum of the BSU attribute values, a_i , divided by the number of the ASRs, N :

$$T = \sum_i a_i / N \quad (2)$$

where a_i in this case refers to either the population or the area of BSU i and where there are n BSUs. Thus, the equality function in the IMAGE Studio is used for creating ASRs that either have equal populations or are of equivalent areal size. Although exact equality rarely occurs because of the constraints imposed by aggregating a limited set of BSU populations or areas, these options provide the opportunity to investigate the scale and zonation effects on internal migration while attempting to control for population or area size.

The similarity function is based on the calculation of attribute distance between two attribute values. In geometric space, the Euclidean distance (d_E) is the physical distance between two points A and B resulting from the sum of squared differences of their x , y coordinates:

$$d_E = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2} \quad (3)$$

whereas in non-geometric space, the notion of distance highlights the differences of attribute values and can be expressed as:

$$d_{AB} = \sqrt{(a_A - a_B)^2} \quad (4)$$

where, a_A and a_B are the values of attribute A and B respectively.

In the IMAGE Studio, the similarity function is structured on the basis of Eq. 4 and is the squared difference between the attribute value of each BSU (a_i) in ASR z and the mean value for ASR \bar{z} . Therefore, in the IMAGE Studio, the distance between the attribute of BSU i (a_i) and the mean value of the attribute for ASR (\bar{z}) is defined as:

$$d_i = \sqrt{(\bar{z} - a_i)^2} \quad (5)$$

where:

$$\bar{z} = \frac{\sum a_i}{n_z}, a_i \in z \quad (6)$$

and n_z is the number of BSUs in ASR z . The objective function (OF) for similarity is then calculated as the minimum value of the sum of the attribute distances divided by the number of ASRs N , expressed as:

$$OF_{Similarity} = \min \left(\sum_i d_i / N \right) \quad (7)$$

The minimisation of the attribute distances between the mean of the ASRs and their constituent BSUs produces homogeneous ASRs consisting of BSUs with similar values for the selected variable. The similarity function in the IMAGE Studio can be used for delivering two aggregation outputs, one based on minimising the differences in population density between ASRs which captures ASR urban/rural characteristics, and the other based on minimising the intra-ASR migration flows between the BSUs in each ASR and results in ASRs with higher/lower intra-ASR flows respectively.

One of the most widely used methods for evaluating optimising functions is the steepest descent or greedy algorithm (Luenberger 1973). Given a function $F(x)$, the steepest descent optimisation targets the direction in which $F(x)$ is optimised locally. This method proceeds along one of two directions: minimising $F(x)$ or maximising $F(x)$. Although maximisation of $F(x)$ is feasible, minimisation of $F(x)$ is the most common implementation of a steepest descent algorithm. For example, if we want to construct a method of equality in a number of ASRs (m), then a steepest descent function could be formulated as the minimisation of the sum of differences between each ASR (x_i) and the target value (T). The generic formulation of such a function is:

$$\min F(x) = \sum_{i=1}^m |T - x_i| \quad (8)$$

In a zone design context, the way to proceed from an existing aggregation to a better one is by swapping areal units at the borders of the ASRs, while optimising an objective function. During these swaps, it is possible for one ASR to lose its contiguity and therefore a method of holding contiguity intact is essential. For example, Openshaw's Automated Zoning Procedure (AZP) tackled this problem by tracing an adjacency matrix using the Depth First Search (DFS) algorithm. The method of maintaining ASR contiguities should be as simple as possible, avoiding complicated structures that may lead to an exponential increase of processing time, during the iterative zone design procedure.

Additional zone design properties could be identified as equally important, such as the initial aggregation algorithm, the starting point for a zone design system. An initial aggregation targeting the criteria directly is avoided as the main zone design procedure is likely to be trapped into local optima and end the process, thus providing an inadequate solution. Hence, Openshaw (1977, 1978) suggested the use of an IRA algorithm focusing on the principle of contiguous zones as an appropriate first aggregation, which provides a high degree of randomisation to ensure that the resulting aggregations differ during each iteration. It has been implemented in the IMAGE Studio with object-oriented principles, thus avoiding the sustained sequential processes and resulting in much quicker random aggregation (Daras 2014). However, the alternative IRA-wave algorithm, a hybrid version of the original IRA algorithm and the BFS algorithm, provides a swifter solution and is often preferred when further optimisation is not required.

Although the three characteristics of a zone design system: the objective function, the contiguity checking algorithm and the initial aggregation are structurally important, it is possible to introduce further criteria in order to influence the shape (compactness) of ASRs. Evidently, each criterion applied to zone design acts as a constraint on the optimum solution with an additional increase of processing time. Therefore, extensive use of criteria should be avoided if the study does not require such constraints.

In the IMAGE Studio, we make use of the Local Spatial Dispersion (LSD) method for controlling the shape of the ASRs which is a type of location-allocation problem (Alvanides and Openshaw 1999). This method controls the shape compactness by calculating the distance between the centroids of BSUs in each ASR and their output ASR centroid. Generally, the LSD algorithm is developed using the geometrical features of BSUs and ASRs. For example, for a given aggregation, the LSD measure is calculated by computing the Euclidian distances between the centroid of each BSU i and the centroid of its ASR z . Mathematically, it is expressed as follows:

$$LSD = \sum_{i \in z} \sqrt{(\bar{x}_z - x_i)^2 + (\bar{y}_z - y_i)^2} / n_z \quad (9)$$

where \bar{x}_z and \bar{y}_z are the coordinates of the centroid of ASR z , x_i and y_i are the coordinates of the centroid of BSU i and n_z is the number of BSUs in ASR z .

During the aggregation process, the BSUs constantly change ASR membership while attempting to achieve an optimum solution. Therefore, every time such a change occurs, it is necessary to recalculate the ASR centroid. Consequently, in the IMAGE Studio, the LSD approach is implemented using only the centroid coordinates of each BSU. The developed LSD approach derives the coordinates of each ASR by calculating the mean of the coordinates of the BSU centroids in ASR z :

$$\bar{x}_z = \frac{\sum_{i \in z} x_i}{n_z}, \quad \bar{y}_z = \frac{\sum_{i \in z} y_i}{n_z} \quad (10)$$

The ASR centroid coordinates are then used in Eq. 9 to provide the final LSD measure for the selected ASR. The minimisation of all LSD measures during the aggregation process results in the output of spatially compact ASRs.

Internal Migration Indicators

Bell et al. (2002) suggest a number of system-wide indicators across four domains of internal migration – intensity, impact, distance and connectivity – that can be used for comparative analysis of migration in different countries, where data are available. In this paper, we have selected five variables that are representative of the first three of these domains in order to identify the MAUP components and explore the consequences of using different types of aggregation based on data for the UK. The first of these indicators is a measure of the Crude Migration Intensity (CMI) and is expressed as a rate of migration per 100 population by dividing the total number of inter-zonal migrants in a time period by the total population as follows:

$$CMI = 100 \left(\frac{\sum_{ij} M_{ij}}{\sum_i P_i} \right) \quad (11)$$

where M_{ij} is the migration flow from zone i to zone j and P_i is the population of zone i . Zone i may be either an initial BSU or an ASR at a particular spatial scale. The second indicator is a measure of migration impact called the Migration Effectiveness Index

(*MEI*), defined by expressing the sum of the absolute value of the net migration balances for all zones in the system as a percentage of the sum of the migration turnovers in all zones as follows:

$$MEI = 100 \left(\sum_i |D_i - O_i| / \sum_i (D_i + O_i) \right) \quad (12)$$

where D_i is the total in-migration into zone i and O_i is the total out-migration from zone i . The third indicator is the Aggregate Net Migration Rate (*ANMR*) which is defined as half the sum of the absolute net changes across all zones and standardised by the population at risk:

$$ANMR = 100 (0.5) \left(\sum_i |D_i - O_i| / \sum_i P_i \right) \quad (13)$$

The *ANMR* therefore measures the overall impact of internal migration on the population distribution but can also be defined as the product of the *CMI* and the *MEI* as follows:

$$ANMR = 100 (CMI * MEI) \quad (14)$$

Thus, a high migration impact might result from high levels of both *CMI* and *MEI* or a high value of one component offsetting a low value of the other. The variation in the relationship between these two components has been explained by Rees et al. (2016).

The fourth and fifth indicators are both related to the distance over which individuals migrate. The fourth is the Mean Migration Distance (*MMD*) which is computed as:

$$MMD = \left(\frac{\sum_{ij} M_{ij} d_{ij}}{\sum_{ij} M_{ij}} \right) \quad (15)$$

where the d_{ij} term is a measure of the Euclidian distance between the centroids of origin zone i and destination zone j for the initial set of BSUs and is a composite measure of the distances between BSUs within ASRs at different levels of aggregation. The fifth and final indicator is the beta (β) parameter calibrated using a spatial interaction model (Eq. 1) that provides a measure of distance deterrence. The calibration method, which uses a Newton Raphson search routine to identify the optimum decay parameter, is explained more fully in Stillwell (1990).

Sources of Internal Migration Data and Spatial Units

Internal migration data are collected in countries around the world using various different collection instruments; in England and Wales, for example, the national statistical agency – the Office for National Statistics (ONS) – retains a migration question in its decadal census but estimates annual migration between censuses by comparing the addresses of National Health Service (NHS) patient registers from 1 year to the next, and also draws on the Labour Force Survey (LFS) for samples of data on migrants whose behaviour is linked to the labour market.

In this paper, we use internal migration flows for the UK obtained from the 2011 Census Special Migration Statistics (SMS) to illustrate results from the IMAGE Studio.

The data format is a matrix of the flows between 404 local authority districts (LADs) in the UK for the 12 month period prior to the 2011 Census. There are three national statistical agencies in the UK – for England and Wales, Scotland and Northern Ireland respectively – each of which undertakes an independent but partially harmonized census. One consequence of this division of labour is that the ONS has to compile a full set of sub-national migration flows between LADs in the UK. This synthesis is only undertaken with single-year census data once a decade.

Populations at risk are required if the user wishes to compute migration intensities or use the population equality algorithm; in this instance, usually resident populations of LADs across the UK in 2011 are extracted from the 2011 Census using the InFuse interface to Aggregate Data on the UK Data Service web site. These end-of-period populations are not the ideal populations at risk for migration rates in the previous 12 months but since no start-of-period populations are available, and therefore no mid-period populations can easily be derived, the end-of-period populations are deemed to be the most suitable. Finally, the boundaries of these LAD administrative units have been sourced from the UK Data Service repository of Boundary Data using the EasyDownload facility. While the Studio's Data Preparation Subsystem automatically ensures that all mainland LADs are contiguous with at least one other mainland LAD, contiguities between each of the Isle of Wight, Belfast (Northern Ireland), Western Isles, Orkneys and Shetlands and their respective nearest neighbours on the mainland are added to the contiguities file that is created by the Data Preparation Subsystem for use in the subsequent aggregation. Having explained where the data come from, we turn our attention to reporting the results of running the various aggregation approaches available in the Studio with the UK 2011 Census internal migration flow data.

Results

Choice of Aggregation Algorithm

In order to investigate the speed at which the alternative processes in the IMAGE Studio produce solutions, we have experimented by selecting the LAD administrative units and the UK 2011 Census internal migration flow data to aggregate the 404 BSUs in steps of 1, 10, 20 and 50 with 10, 100, 500, 1000 aggregation iterations generated from random seeds at each step. Figure 3 shows the respective greediness running time for the IRA algorithm, the IRA-wave algorithm, the aggregation of attribute data for the new ASRs and the calculation of migration indicators. Remembering that a logarithmic scale is used to display time on the vertical axis, the graph indicates that both the data aggregation and the calculation of indicators are the costliest processes and we should consider these in particular when choosing which algorithm to use, when setting the scale step size and when specifying the number of iterations at each scale. Also, Fig. 3 shows that the use of single step aggregations under any number of iterations requires extreme run times e.g. about 2.8 h (9,746 s) overall run time using only 10 iterations. Given that time available imposes a constraint on our choice of optimal approach, the use of single-step aggregations would mean that only a limited number of boundary configurations for each scale could be generated and this would diminish the extent to which we could explore the sensitivity of migration indicators to the different zonations.

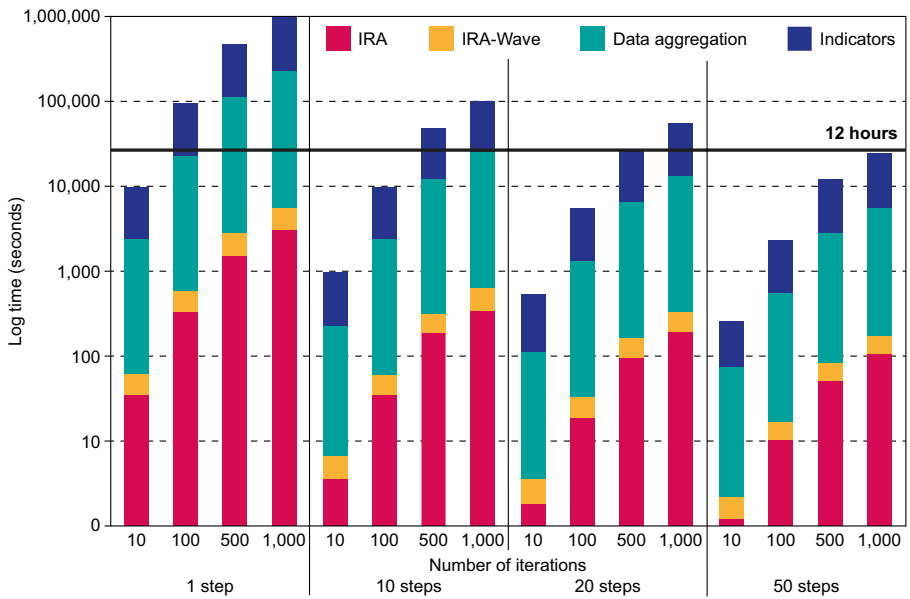


Fig. 3 Time processing costs for 404 LADs using different step size and number of iterations

Scale and Zonation Effects for Selected Indicators Using the IRA Wave Algorithm

In this section, the results produced by the IMAGE Studio and presented in Fig. 4 are based on the computation of *CMI*, *MEI* and *ANMR* values for aggregations into ASRs of the initial matrix of flows between 404 LADs in

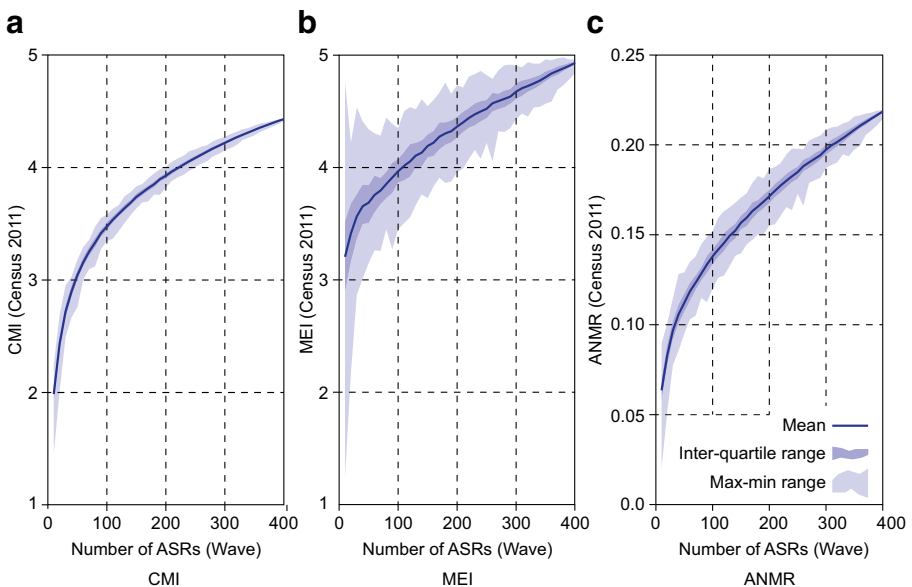


Fig. 4 Mean values of CMI (a), MEI (b) and ANMR (c) by number of ASRs (scale)

the UK in scale steps of 10 and with 200 alternative configurations of ASRs at each scale. The central line in each graph therefore connects the mean value of the respective indicator at different spatial scales as the number of ASRs increases from left to right on the horizontal axis. The minimum number of ASRs is 10 and the maximum is 400. This enables us to visualise the scale effect associated with each indicator and compare the trajectories of the mean values for each indicator, although the *ANMR* will have a much smaller value than either of its component variables (as indicated on the vertical axis in Fig. 4c). A scale effect is most apparent for the *CMI*, which decreases progressively as the number of ASRs is reduced and the individual ASRs get larger, and least evident for the *MEI*, which appears more scale independent, a finding that is in line with results reported for several other countries by Bell et al. (2015b). The trajectory of the mean *ANMR* indicates a significant scale effect suggesting that in the UK, in 2011, it is the *CMI* that is more influential on population redistribution than the *MEI*. Aggregations were performed initially using both the IRA and the IRA-wave algorithms but the differences were barely noticeable so the results presented here are those based on the much speedier IRA-wave algorithm.

The shaded areas around the lines of central tendency reflect the variation due to alternative configurations or shapes of ASRs as measured by the inter-quartile range (darker shading) and the full range (lighter shading). The shaded areas give a useful visualisation of the zonation effect of the MAUP, an effect which is most apparent for the *MEI* indicator and least evident for the *CMI*. Thus, we observe that whereas the number of zones is important in measuring the intensity and the overall impact of migration on the population distribution, the shape and configuration of zones is more important when measuring how effective migration is as a process of population redistribution.

Figure 5 provides evidence of how the mean migration distance and distance decay parameter changes with scale and zonation. Whereas the analysis of migration intensity and impact requires a matrix of flows between districts, intra-district flows can be included in the SIM runs that generate the distance indicators, where intra-district (BSU) distance is measured as the square root of the radius of a circle whose area is equivalent to that of the district concerned. The effect of scale on migration distance is pronounced with *MMD* increasing at an increasing rate as ASRs get larger (Fig. 5a) but the zonation effect is relatively insignificant. The *MMD* is reduced by around 50 km when the intra-district flows are included and this difference is preserved at all scales. In contrast, whilst the beta parameter increases as the spatial units get larger when all migration flows are modelled, the frictional effect of distance on migration appears scale independent when only inter-zonal migration is included (Fig. 5b). Moreover, the configuration of ASRs appears to have a relatively low effect on the decay parameter until around 50 ASRs, when the range of values from alternative zonations gets wider. The stability of the beta parameter across scales has been reported for other countries in Stillwell et al. (2016).

Scale Effects Using the IRA Algorithm with Alternative Functions

The remaining sets of results report on the implications of using the algorithms that generate optimal sets of ASRs based on satisfying certain objective

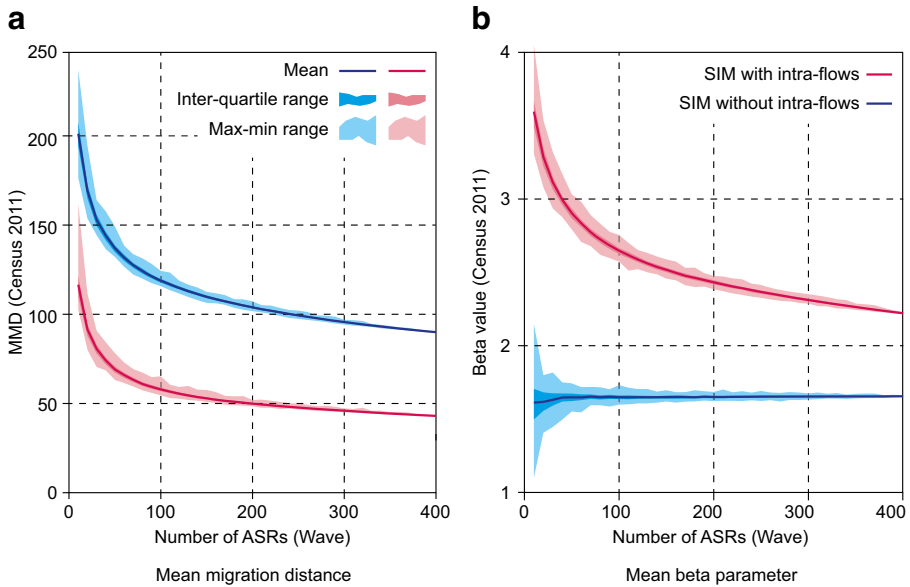


Fig. 5 Mean migration distance (a) and mean decay parameter (b) by number of ASRs

functions relating to the approximate equality according to ASR area, population, population density and intra-ASR migration, and to ASR similarity based on population density or intra-ASR flows. In each case, it is possible to show only scale effects because just one optimised set of ASRs is derived at each scale. In Fig. 6, we have chosen to show the trajectories over scale for the overall migration impact indicator, the ANMR, overlaid on the trajectory of the

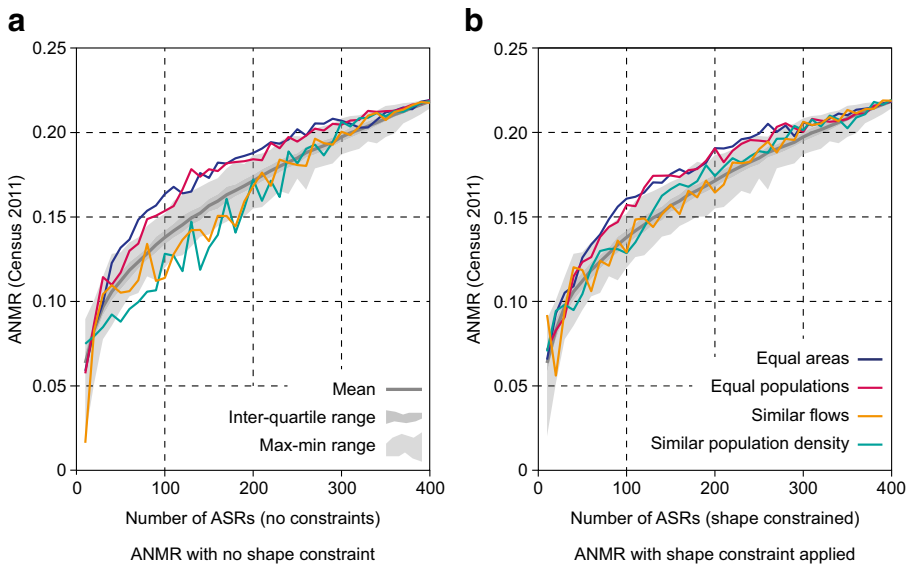


Fig. 6 Mean values of ANMR by scale without (a) and with (b) a shape constraint applied

mean ANMR values and their ranges (the shaded area) derived from the IRA wave algorithm and shown in Fig. 4. The graph in Fig. 6a shows the ANMR values derived at each spatial scale using the different objective functions without a shape constraint whereas the graph in Fig. 6b shows results when a shape constraint is imposed. It is clear that the trajectories of the optimised ANMR under all scenarios reduce in a less uniform and more erratic manner as ASR size increases in comparison with the mean ANMR values derived by the IRA wave algorithm. When no shape constraint is applied, the equal area and equal population alternatives generate similar optimised ANMR values which tend to be higher than the other options for much of the scale gradient and outside the range of the values derived using the basic IRA wave algorithm. The two alternatives based on similarity, however, show the most erratic behaviour and across most of the scale gradient, fall below the mean ANMR value.

The imposition of a shape constraint, as shown in Fig. 6b, has the effect of reducing the variation in any one option between scales and also of bringing all the alternatives closer together and closer to the mean derived from the basic IRA wave aggregation with 200 iterations at each scale. Similar sets of results are derived when plotting the optimised *MMD* and *beta* values for the different options in Figs. 7 and 8 respectively. It is the two similarity options, involving population density and intra-ASR flows, which generate higher *MMD* and *beta* values for the middle section of the scale gradient when no shape constraint is applied, whereas the equal area and equal population options tend to generate lower *MMD* and *beta* values. Once again the use of the shape constraint has the effect of reducing the variation between the alternatives and giving similar scale effects for both these indicators.

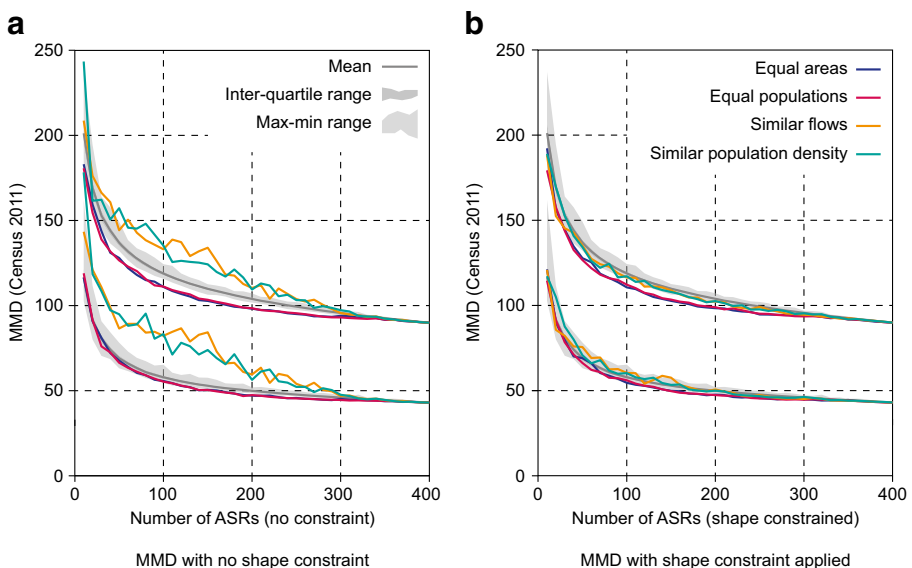


Fig. 7 Mean migration distance by scale without (a) and with (b) a shape constraint applied

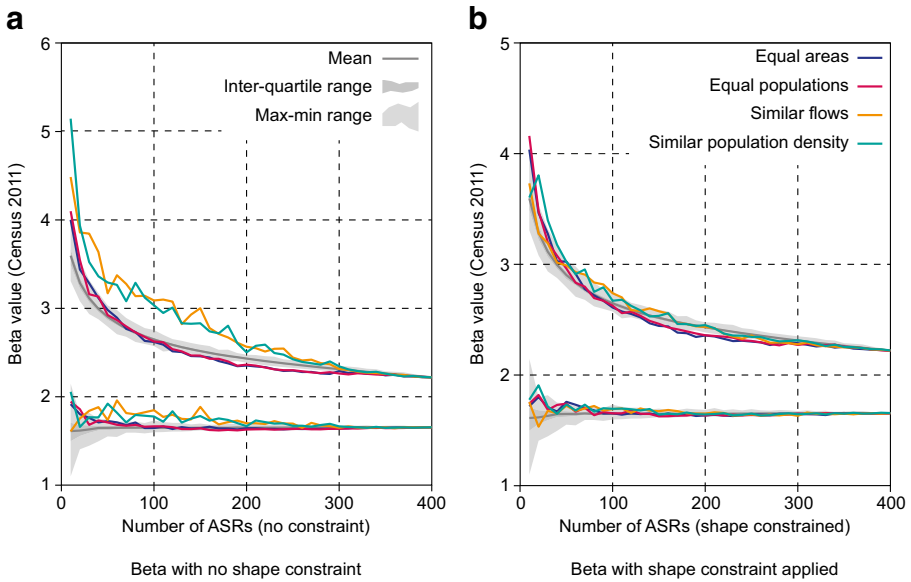


Fig. 8 Beta value by scale without (a) and with (b) a shape constraint applied

Conclusions

The redistribution of the population through internal migration has become increasingly important as a component of population change in many countries around the world, including the UK, yet most research studies are based on data on migration flows between one set of administrative or statistical zones at one particular spatial scale. This poses intractable problems for policy-makers who want to compare internal migration in different countries using one or more indicators and to understand the relationship between migration and development. As we have shown in the case of the UK, the intensity at which people move between regions depends upon the size and shape of the regions concerned and it is only when all internal migrations are included in an aggregate *CMI* that direct comparisons between countries can be made and national league tables constructed as reported in Bell et al. (2015b). However, we contend that the IMAGE Studio provides researchers and practitioners with a means to develop a much better understanding of how different migration indicators are affected by the scale and zonation components of the MAUP. In this paper we have looked at selected indicators of migration intensity, effectiveness, impact, distance and distance deterrence and shown that whilst intensity, impact and distance are revealed to vary significantly by scale but less so by zonation, migration effectiveness and distance show greater scale independence but more sensitivity to zone shape.

Whilst these results are based on analysis of multiple zone configurations across a range of scales, the paper has also reported the scale effects when zones are optimised at different scales using the alternative algorithms available in the Studio that maximise certain objective functions subject to the constraints of contiguity. There are subtle differences in the scale gradients for particular indicators with the zone shape constraint serving to reduce the variations between the results from using different algorithms in all cases. We also observe that an optimized indicator at a particular scale may fall

outside the range of values computed when the IRA wave algorithm is adopted. This finding is not unexpected because we explore only a fraction of possible configurations using the IRA-wave aggregations (200 iterations per scale) under the shape constraints of adjacent regions. Fundamentally, the full exploration of possible configurations is a large computational problem and even today an exhaustive algorithm is only applicable to small aggregation problems (Keane 1975). One interesting conclusion that emerges from using the IMAGE Studio is that a migration indicator such as the *CMI* calculated on the basis of published migration data at one spatial scale does not necessarily reflect the ‘true’ migration rate because it is reflecting particular size and shape characteristics of the zones in the country that have been used to collect the migration data in the first place. The mean value of the *CMI* computed from many configurations at any one scale, i.e. with the same number of zones, will offer a better measure. Further research is required using countries where data are available on migration at different spatial scales to compare published rates with estimated means derived using the IMAGE Studio from configurations based on lower level spatial units.

Whilst the results of the IMAGE project have reported the use of the Studio for comparative analysis of internal migration in different countries around the world (Bell et al. 2015b; Rees et al. 2016; Stillwell et al. 2016) where zone systems are very different, there is also the potential in using the Studio to explore how scale and zonation effects might vary by demographic (age, sex, ethnicity) or socio-economic (occupation, tenure, health status) group in any single country (see Stillwell et al. 2018, for an initial study of variations by age group in the UK). A further avenue of investigation might be to explore the relationship between migration indicators and explanatory variables at different spatial scales using correlation analysis of the type that was employed to investigate the MAUP effects in earlier studies of stock variables. Moreover, the aggregation algorithms in Studio might be usefully adapted to provide an automated system for aggregating explanatory variables and generating summary measures.

Acknowledgements The IMAGE Studio was developed as part of the Discovery Project, DP11010136, Comparing Internal Migration around the World (2011–2015), funded by the Australian Research Council.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alvanides, S., & Openshaw, S. (1999). Zone design for planning and policy analysis. In J. Stillwell, S. Geertman, & S. Openshaw (Eds.), *Geographical information and planning* (pp. 299–315). Berlin: Springer.
- Bell, M., Blake, M., Duke-Williams, O., Rees, P., Stillwell, J., & Hugo, G. (2002). Cross-national comparison of internal migration: issues and measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(3), 435–464.

- Bell, M., Bernard, A., Ueffing, P., & Charles-Edwards, E. (2014). The IMAGE repository: A user guide. *Working Paper No 2014/01*, Queensland Centre for Population Research, School of Geography, Planning and Environmental Management, The University of Queensland, Brisbane.
- Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., & Zhu, Y. (2015a). Internal migration data around the world: assessing contemporary practice. *Population, Space and Place*, 21(1), 1–17.
- Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M., & Kupiszewska, D. (2015b). Internal migration and development: comparing migration intensities around the world. *Population and Development Review*, 41(1), 33–58.
- Daras, K. (2014). IMAGE studio 1.4.2 user manual. School of Geography, University of Leeds, Leeds.
- Duke-Williams, O., Routis, V., & Stillwell, J. (2018). Census interaction data and the means of access. In J. Stillwell (Ed.), *The Routledge handbook of census resources, methods and applications: Unlocking the UK 2011 census* (pp. 110–125). London: Routledge.
- Flowerdew, R. (2011). How serious is the modifiable areal unit problem for analysis of English census data? *Population Trends*, 145, 1–13.
- Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23, 1025–1044.
- Gehlke, C. E., & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29(185A), 169–170.
- Holt, D., Steel, D. G., & Tranter, M. (1996). Area homogeneity and the modifiable areal unit problem. *Geographical Systems*, 3, 181–200.
- Keane, M. (1975). The size of the region-building problem. *Environment and Planning A*, 7, 575–577.
- Lloyd, C. D. (2014). *Exploring spatial scale in geography*. Chichester: Wiley.
- Luenberger, D. (1973). *Introduction to linear and non-linear programming*. Boston: Addison-Wesley.
- Manley, D. (2005). The Modifiable areal unit phenomenon: an investigation into the scale effect using UK census data. *Unpublished PhD Thesis*, School of Geography and Geosciences, University of St Andrews.
- Manley, D. (2014). Scale, aggregation, and the modifiable areal unit problem. In M. M. Fischer & P. Nijkamp (Eds.), *Handbook of regional science* (pp. 1157–1171). Berlin: Springer-Verlag.
- Marble, D. F. (2000). Some thoughts on the integration of spatial analysis and geographic information systems. *Journal of Geographical Systems*, 2, 31–35.
- Openshaw, S. (1976). A regionalisation procedure for a comparative regional taxonomy. *Area*, 8, 149–152.
- Openshaw, S. (1977). Algorithm 3: a procedure to generate pseudo-random aggregations of N spatial units into M spatial units, where M is less than N. *Environment and Planning A*, 9, 1423–1428.
- Openshaw, S. (1978). An optimal zoning approach to the study of spatially aggregated data. In I. Masser & P. J. B. Brown (Eds.), *Spatial representation and spatial interaction* (pp. 93–113). Leiden: Martinus Nijhoff.
- Openshaw, S. (1984). *The modifiable areal unit problem, CATMOG 38*. Norwich: Geo Books.
- Openshaw, S., & Rao, L. (1995). Algorithms for re-engineering 1991 census geography. *Environment and Planning A*, 27, 425–446.
- Openshaw, S., & Taylor, P. (1979). A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In N. Wrigley (Ed.), *Statistical applications in the spatial sciences* (pp. 127–144). London: Pion.
- Rees, P., Bell, M., Kupiszewski, M., Kupiszewska, D., Ueffing, P., Bernard, A., Charles-Edwards, E., & Stillwell, J. (2016). The impact of internal migration on population redistribution: an international comparison. *Population, Space and Place*, Published online in Wiley Online Library, 23. <https://doi.org/10.1002/psp.2036>.
- Stillwell, J. (1983) SPAINT: A computer program for spatial interaction model calibration and analysis. *Computer Manual 14*, School of Geography, University of Leeds, Leeds.
- Stillwell, J. (1990). Spatial interaction models and the propensity to migrate over distance. In J. Stillwell & P. Congdon (Eds.), *Migration models: Macro and micro approaches* (pp. 34–56). London: Belhaven Press.
- Stillwell, J., Daras, K., Bell, M., & Lomax, N. (2014). The IMAGE studio: a tool for internal migration analysis and modelling. *Applied Spatial Analysis and Policy*, 7(1), 5–23.
- Stillwell, J., Bell, M., Ueffing, P., Daras, K., Charles-Edwards, E., Kupiszewski, M., & Kupiszewska, D. (2016). Internal migration around the world: comparing distance travelled and its frictional effect. *Environment and Planning A*, 48(8), 1657–1675.
- Stillwell, J., Lomax, N., & Chatagnier, S. (2018). Changing intensities and spatial patterns of internal migration in the United Kingdom. In J. Stillwell (Ed.), *The Routledge handbook of census resources, methods and applications: Unlocking the UK 2011 census* (pp. 362–376). London: Routledge.
- Tranter, M., & Steel, D. (2001). Using local census data to investigate scale effects. In N. J. Tate & P. M. Atkinson (Eds.), *Modelling scale in geographical information science* (pp. 105–122). Chichester: Wiley.
- Wilson, A. G. (1970). *Entropy in urban and regional Modelling*. London: Pion.