Environment and Planning A 2007, volume 39, pages 2482 – 2499

DOI:10.1068/a38441

Building a spatial microsimulation-based planning support system for local policy making

Dimitris Ballas
Department of Geography, University of Sheffield, Sheffield, S10 2TN, England;
e-mail: d.ballas@sheffield.ac.uk

Richard Kingston
Planning and Landscape, School of Environment and Development, University of Manchester, Manchester, M13 9PL, England; e-mail: richard.kingston@manchester.ac.uk

John Stillwell, Jianhui Jin
School of Geography, University of Leeds, Leeds, LS2 9JT, England;
e-mail: j.c.h.stillwell@leeds.ac.uk, j.jin01@leeds.ac.uk
Received 28 October 2005; in revised form 5 January 2006

Abstract. This paper presents a spatial microsimulation modelling and predictive policy analysis system called Micro-MaPPAS, a Planning Support System (PSS) constructed for a local strategic partnership in a large metropolitan area of the UK. The innovative feature of this system is the use of spatial microsimulation techniques for the enhancement of local policy decision making in connection with the neighbourhood renewal strategy. The paper addresses the relevant data issues and technical aspects of the linkage of spatial microsimulation modelling frameworks to PSS and deals with the wider implications that such a linkage may have to local policy and planning procedures. Finally, the paper presents some illustrative examples of the policy relevance and policy analysis potential of the software.

1 Introduction

It has long been argued that spatial microsimulation has great potential for socioeconomic impact assessment (see, for instance, Ballas and Clarke, 2001a) and geographical analysis of the impacts of social policies (Ballas and Clarke, 2001b; Ballas et al, 2003). In this paper we attempt to demonstrate this potential further by presenting a system that incorporates a spatial microsimulation model within a planning support system (PSS) to assist decision making at a range of spatial scales for a large metropolitan district in the UK. We also present new spatial microsimulation model results, generated using data from the British Household Panel Survey (BHPS) and the 2001 UK Census of Population.

It can be argued that most of the spatial microsimulation work undertaken so far has been very important in identifying which particular microsimulation techniques are the most suitable for different purposes (see, for instance, Voas and Williamson, 2000; Williamson et al, 1998) and in demonstrating how spatial microsimulation can be used for policy analysis (see Ballas and Clarke, 2001b). However, there have been few if any attempts to develop a spatial microsimulation tool for spatial policy analysis or planning support. Thus, the aim of this paper is to describe an applied spatial microsimulation system and to demonstrate how it can be used to support planning practice. The paper is organised as follows: section 2 discusses the conceptual issues pertaining to microsimulation and PSSs and discusses how these components have been combined to create a Microsimulation Modelling and Predictive Policy Analysis System (Micro-MaPPAS). Further, section 3 outlines the policy background that led to the development of this model-based tool for planning support at Leeds City Council. Section 4 reviews the data and methods that underpin the Micro-MaPPAS system, whilst section 5 describes the system and its features in some detail. Section 6 shows how it is possible to add a projection capability to the system, whilst section 7 demonstrates the policy relevance and policy analysis potential of the software. Finally, section 8 offers some concluding comments.

2 Microsimulation and PSSs

Population microsimulation in the present context can be defined as a methodology for creating large population microdata sets which can subsequently be used for analysis of the impact of policy at the microlevel. It should be noted that, conceptually, microsimulation has much in common with cellular automata modelling as well as agent-based or multiagent-based modelling approaches (such as discussed by Batty, 2005) and it has long been argued that spatial microsimulation models could be converted to agent-based models by replacing the simulated units (for example, individuals or households) with adaptive rule-based agents (Williamson, 1999).

One of the main differences between agent-based modelling and microsimulation is that the former typically involves interaction between the simulated microunits or 'agents' on the basis of a set of prespecified rules, whereas the latter does not normally involve any interaction between the simulated entities. Moreover, agent-based modelling usually relies on rules that are specified on the basis of theories or assumptions pertaining to the organisation of cities and regions, whereas microsimulation typically recreates a simulated population of microunits on the basis of observed statistical data and derived conditional probabilities. In particular, spatial microsimulation techniques involve the merging of census and survey data to simulate a population of individuals within households (for different geographical units), whose characteristics are as close to the real population as it is possible to estimate (Ballas et al, 2005; Clarke, 1996). These models simulate virtual populations in given geographical areas, so that the characteristics of these populations are as close as possible to their 'real world' counterparts. The models are concerned with the creation of large datasets through estimating the attributes of individuals within households and are used to analyse the impacts of policy on these microunits (Clarke, 1996; Orcutt et al, 1986). Traditionally, confidentiality concerns have been the main reason why demographic and socioeconomic data on individuals, despite being collected from censuses and surveys, have not been available to researchers and policy makers. Ballas and Clarke (2003) have recently provided a detailed review of the development of the methodology from Orcutt's original work in the 1950s (Orcutt, 1957). Whilst the work presented here refers to people and households, microsimulation can also be used to simulate other entities [for example, the behaviour of firms (see Danson and Lavercombe, 1996; Moeckel, 2005; Van Wissen, 2000)].

The basic output from a population microsimulation model is a list of individuals or households with a series of attributes. Such a model can be evaluated in a historical context by comparing counts of subsets of individuals for small areas against observed data derived typically from the Census of Population. However, a major advantage of the microsimulation approach is the estimation of subsets of characteristics that are not directly available from existing sources and, consequently, the approach has enormous value as a substitute for conducting detailed surveys.

PSSs, on the other hand, are systems that incorporate a subset of computer-based instruments each of which incorporates a further unique set of components that planners utilise for a particular activity (Brail and Klosterman, 2001; Geertman and Stillwell, 2003; 2004). PSSs are beginning to play a key role in a variety of planning contexts (Geertman and Stillwell, 2003). In this paper we present a spatial microsimulation-based PSS for the Leeds Initiative, the city's Local Strategic Partnership (LSP), to assist with local neighbourhood regeneration and planning (Leeds Initiative, 2001). LSPs were set up by the national government in the UK to improve the coordination of the myriad of local services and agencies together with private businesses, community groups, and the voluntary sector. One of the main aims of LSPs is to deliver an improvement to people's lives at the community scale (DETR, 2001). The system presented in this paper is based on a geographical microsimulation model which is capable of constructing a list of approximately 715 000 individuals living within households along with their associated attributes for any point in time, past or future. Previously the software has been run from a DOS prompt and required the 'hard coding' of parameters and data tables together with some knowledge of Java programming—not a desirable task for the average policy or decision maker. With the development of Micro-MaPPAS, we now have an interface which is much more user friendly and suitable for decision makers to exploit the power of the spatial microsimulation methodology.

One innovative feature of the Micro-MaPPAS system is the integration of spatial microsimulation techniques with geographical information systems (GIS) for the enhancement of local policy decision making. The practical application of the PSS in a local policy context is a key feature. Hitherto, economists have been involved in the development of microsimulation models that are capable of modelling the impact of national government policies (for recent examples of microsimulation applications by economists see Heady et al, 2001; Matsaganis et al, 2001; Mitton et al, 2000; Redmond et al, 1998; Sutherland and Piachaud, 2001). Adding spatial detail to traditional microsimulation involves creating geographically referenced microdata that refers to a particular locality, to a geographically defined and restricted area. Since there are very few sources of geographically detailed microdata, there is a need to create these data using spatial microsimulation techniques by merging census and survey data to simulate a population of individuals within households (for different geographical units) whose characteristics are as close to the real population as it is possible to estimate. They can then be used to answer questions such as:

- How does the quality of life of individuals and households vary across different regions, cities, or neighborhoods?
- What are the interdependencies of household characteristics with geographical factors such as the presence of hospitals, community centres, or schools in an area?
- To perform static 'what if' scenario analysis; that is, answer questions such as 'what would happen to personal accessibilities if the patterns of service provision change?'
- What would be the geographical impact of national social policies on personal incomes and how effective would they be compared with alternative area-based policies?

The remainder of this paper discusses in more detail how spatial microsimulation and PSSs have been combined to provide a powerful geographical information resource that can be used to address a number of important *local* policy questions.

3 Towards a PSS for tackling neighbourhood renewal through better information

Since 1997 the UK government has made a concerted effort to alleviate social exclusion through a series of specific Policy Action Teams (SEU, 2001). One of the overall aims of these policies has been to respond to the data demands of public sector organisations to help them tackle social exclusion and deprivation (SEU, 2000). A collaborative effort on the part of central and local government has resulted in many urban areas in England investing in geographical information technology and spatial data. This has led to several local authorities such as Leeds, Bradford, and Manchester Metropolitan District Council developing small area profiling systems to help describe conditions in neighbourhoods and assist in directing investment and service improvement. Such systems have been developed within a neighbourhood renewal strategy in line with central government policy; the Leeds Neighbourhood Renewal Strategy was launched in summer 2001 (Leeds Initiative, 2001). A key component of the strategy involved the development of a targeting framework which would enable the identification of neighbourhoods that may require comprehensive regeneration or remedial action to arrest decline or target action on specific issues.

The key issue for the research discussed in this paper, however, is the lack of the methodology that enables the city council to undertake predictive analysis of likely future trends based on past data and analysis of the likely impact of policy change, neighbourhood remodelling, or change to service provision on the composition and health of neighbourhoods. The Leeds Initiative therefore sought to develop a modelling tool in collaboration with the authors to enable current conditions and issues in neighbourhoods to be described, future trends in the composition and health of neighbourhoods to be predicted, the sensitivity analysis to be conducted so as to measure the likely impact of policy interventions such as the closure of a primary school or the demolition of public housing. The system presented in this paper utilises spatial microsimulation techniques to provide a planning support tool for use by officers in different departments of the local authority and partner agencies. Thus, the overall aims of the project underpinning the system presented here are as follows: (i) to develop a static microsimulation model to describe current conditions in Leeds; (ii) to develop a model to predict future conditions in Leeds under different policy scenarios; (iii) to enable the performance of 'what if?' analysis on a range of policy scenarios.

We seek to report progress in meeting the above aims and outline the associated difficulties and data issues. In particular, the paper focuses on the explanation of the method used for static microsimulation and the functions of the graphical user interface (GUI) but it also briefly describes the prediction and impact methods of the detailed analysis of model results.

The building of a PSS to meet the requirements of the Leeds Initiative has been split into four tasks. Task 1 aims to describe current neighbourhood composition and conditions by building a small area population microdata set for the Leeds Metropolitan District. In particular, the first challenge has been the construction of a static spatial microsimulation model utilising information from the 2001 Census Area Statistics (CAS) and other sources of small area data from Leeds Initiative, as well as survey data from sources such as the BHPS which collects information on a wide range of variables covering most aspects of life in Britain. The spatial microsimulation model adopted is capable of reweighting the records of a national survey database so that they would fit with small area statistics tables from the national census. The steps that are followed by the model are as follows: read in CAS tables from the census and other data sources; read in the records of the survey dataset; and apply a combinatorial optimisation algorithm to find the best fitting set of households (that is, reweighting the survey household records so that they fit the small area statistics tables).

Task 2 involved building a GUI that enables policy analysts to query the datasets and create thematic maps of different variables. One of the significant advantages of the spatial microsimulation approach is that it enables the user to query any combination of variables that is deemed desirable for policy analysis. Further, as noted above, survey data are used and therefore the analysis is not limited to traditional census variables, but also includes noncensus variables such as household income. It is therefore possible to identify disadvantaged areas in Leeds by submitting, for instance, a query on the spatial distribution of households living below half-median incomes.

By analysing a series of household and individual variables at the small area level, the model is capable of estimating which areas are disadvantaged and which may be 'on the edge' of declining into disadvantage. It is also possible to identify the areas that are most likely to be affected by a particular council policy. Further discussion of the GUI is provided in a later section of the paper.

Task 3 has involved predicting likely trends in neighbourhood composition and conditions by updating the small area microdataset in order to predict future socio-economic trends in different localities in Leeds. This has been achieved by building a model which projects into the future the small area descriptions used in the historical base model. It is then possible to rerun the static model based on the updated small area information for any year up to 2021 in order to produce the small area microdata. By repeating the static analysis using the updated microdata it will be possible to provide answers to questions such as: is polarisation of neighbourhoods reducing or increasing or likely to increase further? In order to enable the model to address such policy questions, the GUI can be updated and extended from that developed under task 1, so that it is possible for policy makers to project socioeconomic trends, based on the assumption that things continue as they are, get better, or get worse.

Task 4 involves assessing the impact of policy interventions and demographic change on neighbourhood composition and conditions following policy interventions or neighbourhood remodelling to assess the success of such changes. The main aim of this final task has been to further enhance the spatial modelling framework with policy modelling capabilities. In particular, the policy analysis options added to the GUI allow the user to rerun the model under different policy-intervention scenarios.

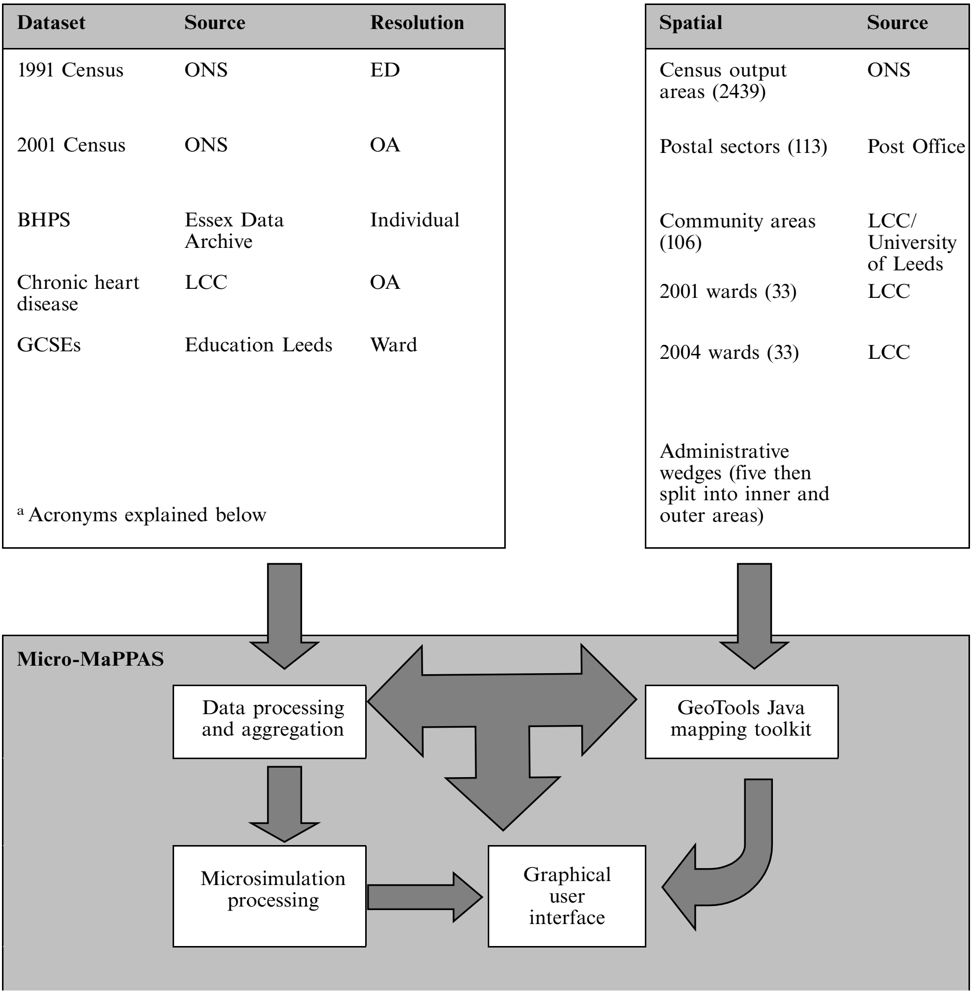
There are various different ways of calibrating the model but the results are particularly valuable because they combine data from different sources to provide estimates of the probabilities that individuals or households will have particular characteristics and thus create new population cross-classifications unavailable from published sources. So, for example, it becomes possible to estimate the numbers of individuals with the characteristics of being aged 18, a lone parent, unemployed, and living in private accommodation in an area prone to high levels of crime. Alternatively, households can be identified in the outer suburbs that contain five persons and have a head of household who is a professional working in another city and earning over £50 000 per year. Once the long list of individuals and their attributes has been simulated, the individuals and households (and the attributes which they possess) can be aggregated to any geographical scale which is deemed appropriate including census-based areas such as output areas and wards, postal sectors, user-defined community areas, or more specific areas designed for policy implementation such as neighbourhood regeneration areas.

The next section briefly describes the data used and introduces the combinatorial optimisation spatial microsimulation technique that has been adopted.

4 The Micro-MaPPAS framework

The PSS presented in this paper utilises a spatial microsimulation model which links a wide range of datasets, including 2001 Census data for output areas (OAs) and sample data from the BHPS. The framework of the Micro-MaPPAS system is illustrated in figure 1.

The datasets described in figure 1 are linked on the basis of a spatial microsimulation model which implements a combinatorial optimisation approach to generating spatially disaggregated population and household microdata sets at OA level for the metropolitan district of Leeds as defined in 2001. In particular, the modelling exercise involves the construction of microlevel population and household data using existing



[a] BHPS = British Household Panel Survey; ED = enumeration district; GCSE = General Certificate of Secondary Education; LCC = Leeds City Council; OA = output area; ONS = Office for National Statistics.

Figure 1. Datasets and system architecture of the Micro-MaPPAs project.

2001 Census Key Statistics (KS) tables for population and household characteristics of all 2439 OAs in Leeds together with sample data from the BHPS, a national panel microdata set of household or individual characteristics. The BHPS is a national annual survey of the adult population of the UK, drawn from a representative sample of over 5000 households (Taylor et al, 2001). The survey collects information on a wide range of variables covering most aspects of life in Britain. The BHPS contains certain data variables (such as income, educational attainment, weekly grocery bill, utility bills, mortgage and rental costs, pension, and childcare costs) for households and their occupants that are unavailable from the census and can be used to derive estimates of 'new' variables for OAs. The simulation technique we have chosen to adopt is 'simulated annealing' and is distinguished from other methods such as iterative proportional fitting (Ballas, 2001; Ballas and Clarke, 2000; Norman, 1999). Simulated annealing involves reweighting the microdata sample from the BHPS so that it fits OA data from the census for Leeds. In the first instance, the BHPS microdata set has been reweighted to estimate its parent population at the microspatial scale. The BHPS provides a detailed record for a sample of households and all of their occupants. The reweighting method can enable the sampling of this universe of records to find the set of household records that best matches the population described in the census tables for each OA.

The actual procedure works as follows. First, a series of census KS tables that describe the small area of interest must be selected. The next step is to identify the records of the BHPS microdata that best match these tables. However, there are a vast number of possible sets of households that can be drawn from the BHPS sample. Clearly, it would be impractical to exhaustively consider all possible sets so this is where simulated annealing (Ballas et al, 1999; 2003) is used to find a set that fits the target tables well. Annealing is a physical process in which a solid material is first melted in a heat bath by increasing the temperature to a maximum at which point all particles of the solid have high energies and the freedom to randomly arrange themselves in the liquid phase. This is then followed by a cooling phase, in which the temperature of the heat bath is slowly lowered. The particles of the material attempt to arrange themselves in a low energy state during the cooling phase. When the maximum temperature is sufficiently high and the cooling is carried out sufficiently slowly then all the particles of the material eventually arrange themselves in a state of high density and minimum energy (Dowsland, 1993; Kirkpatrick et al, 1983; Pham and Karaboga, 2000; Van Laarhoven and Aarts, 1987). In geography, simulated annealing has been applied in various contexts for different problems (see, for instance, Alvanides, 2000; Ballas, 2001; Openshaw and Rao, 1995; Openshaw and Schmidt, 1996; Williamson et al, 1998). The Micro-MaPPAS simulation model builds on previous computer software known as SimLeeds (Ballas, 2001) and uses the tenth wave of the BHPS to provide a detailed record for a sample of households and all of their occupants. It should be noted that this paper presents the first application of the simulated-annealing approach on BHPS data, building on extensive research that followed the first application of the technique in a population geography context as discussed by Williamson et al (1998).

A simple example can be described for clarification. Let us assume that, according to the census, in a particular OA there are 100 households, of which 60 are owner-occupiers, 10 are renting from an LA or HA, and the remaining 30 are renting privately. The simulated annealing procedure would select a combination of BHPS households that would have tenure characteristics as close as possible to the actual data. An exact match would be possible if the tenure KS table were the only 'constraint' in the procedure. However, the purpose of using a combinatorial optimisation technique is to select households that match several KS table constraints. Let us also assume that we introduce the number of cars by household KS table as a further constraint and that our OA contains 50 households with 1 car, 20 households with 2 or more cars, and 30 households with no car. The aim of the annealing would now be to find a set of 100 BHPS households that best fit both tenure and car ownership constraints. To do this, an initial random sample of records is selected from the BHPS until sufficient households are represented (that is, if there are 100 households in the OA, then 100 households will be selected at random). These records are used to create tables that match the selected target KS tables. An initial random selection of 100 BHPS households could result in the distribution described in the first row of table 1. A total absolute error of 88 is calculated as the sum of the differences between the simulated and the actual census values across both variables.

Thus, the task in the simulated annealing procedure is to minimise the total absolute error. In order to do so, a record in the originally selected household set is selected at random and replaced with one chosen at random from the universe of records.

Table 1. Calculating the absolute error.

	Household car ownership characteristics			Household tenure characteristics		
	1 car	2+ cars	no car	owner-occupier	LA/HA rented [a]	other
Simulation	27	24	49	39	17	44
Census	50	20	30	60	10	30
Absolute error	23	4	19	21	7	14

[a] LA = local authority; HA = housing association.

The error is recalculated and the change in error (Δe) is calculated. If Δe is less than zero then there has been an improvement and the change is accepted; if not, then $\exp(-\Delta e/t)$ is compared with a random number between 0 and 1. If it is greater than the random number, then the change is accepted; otherwise the change is rejected and reversed. The method works as an analogy with the physical process of annealing that was briefly described above. The temperature in this case can be seen as a *control parameter* and is initially set high and then slowly lowered after a set number of iterations have taken place. As the temperature is lowered fewer uphill moves are accepted. This allows the method to recover from being stuck in local minimum by allowing a limited number of small moves uphill. In this implementation, if Δe is zero, the change is accepted to allow the exploration of a greater part of the solution space. If the new error is the best seen so far, then the set of households used is stored. The procedure is summarised in figure 2, which represents a refinement of the algorithmic process presented by Ballas (2001) and Ballas et al (1999).

The analogy of the above simulation algorithm to the thermodynamic annealing simulation is depicted in table 2.

So, in the example in table 1, a randomly selected household from the 100 originally sampled would be swapped with another household. Assuming, for example, that an LA (local authority) or HA (housing association) household is swapped with an owner-occupier household from the BHPS, this would result in the reduction of error by 1, as the new owner-occupier simulated total would be 40 (instead of 39) and therefore closer to the actual census total (60 households) and the LA/HA rented total would be 16 (instead of 17), which is closer to the census actual total of 10. This household swap would therefore be accepted. Conversely, if the change increased the error (other than temporarily) it would be rejected and another household selected. It is evident from this example that the simulated annealing model is computationally very intensive, particularly when several table constraints are introduced. The implication of this is that the microsimulation process for a system with many zones and table constraints can be extensive.

In the implementation of Micro-MaPPAS the following small area tables were selected to be used as constraints in the simulated annealing procedure:

- age groups (five-year) by sex;
- household type;
- economic activity by age;
- socioeconomic group;
- household tenure;
- ethnic group;
- limiting long-term illness; and
- educational qualifications.

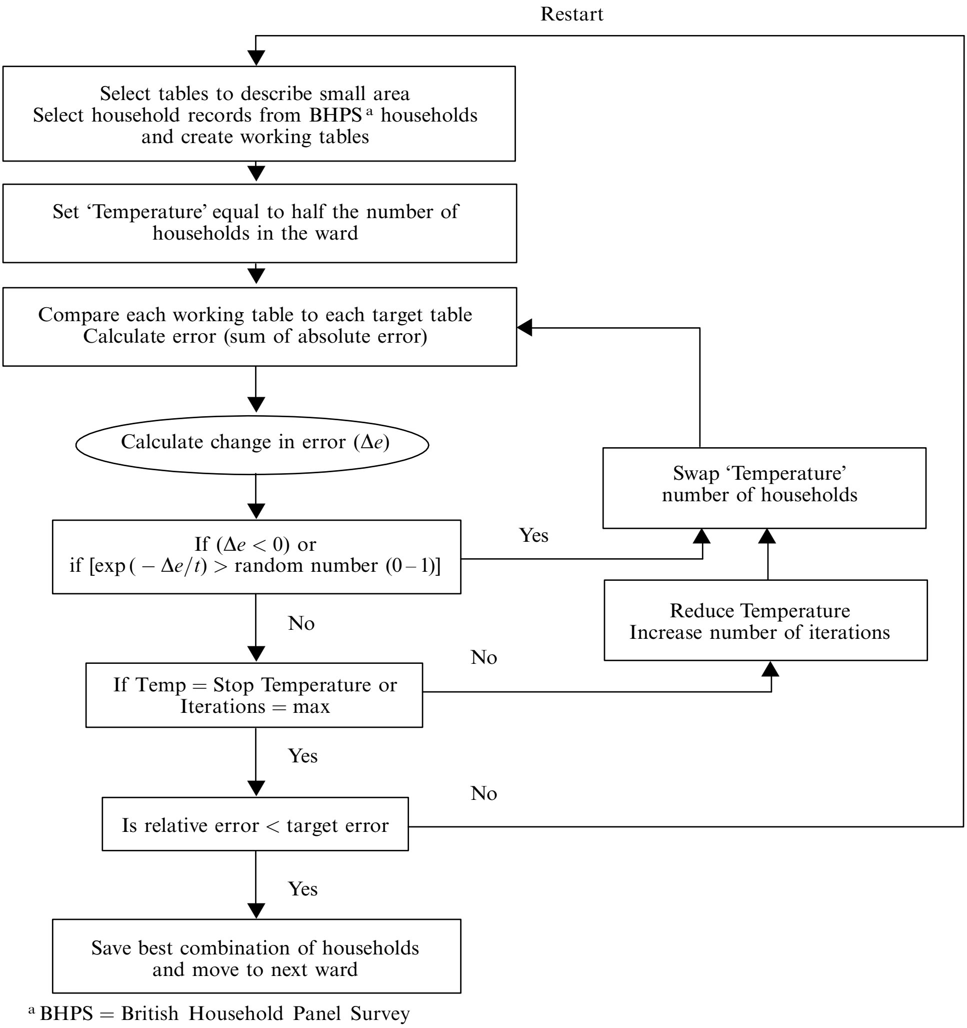


Figure 2. A simulated annealing approach to generating small area microdata.

Table 2. The Micro-MaPPAS analogy to the annealing simulation.

Thermodynamic simulated annealing	SimLeeds
System states	Feasible solutions
Energy	Total absolute error
Temperature	Control parameter
Frozen state	'Best' solution

It should be noted that the simulated annealing approach adopted here is very similar to that developed by Williamson et al (1998) and also used by Ballas et al (1999). However, the method was further refined to allow for a more extensive search for better solutions for areas where fitting the BHPS was particularly difficult. In particular, a new variable called 'restart' was introduced to allow for further searching, in the cases where the solution was not satisfactory. Further, the method developed and used by Micro-MaPPAS allows the user to indicate whether particular census variables are more important than others and weight them accordingly. For instance, if the user is interested in estimating household income as closely as possible, a census variable such as 'socioeconomic group' can be selected for use in the simulation. The respective 'weight' of this variable can be changed to ensure that the simulated annealing procedure will get the lowest possible error with regard to the selected variable or variables.

5 The Micro-MaPPAS system

The Micro-MaPPAS software is written in the Java programming language version 1.4, which means that it can be installed and operated on any computer system and platform. A default set of simulations generated for OAs is loaded when the system is booted up. The architecture of the Micro-MaPPAS system is illustrated in figure 1. Through a GUI, the user has access to various modules: the base simulator, the model diagnostics, the data analyser, the mapping controls, the projection scenarios, and the impact scenarios.

The base simulator module allows the user to set the parameters for a small area simulation. The user can change the default values of the simulated annealing parameters that were described above. In particular, the user can change the temperature, the number of model iterations, and the number of restarts, and also apply weights to input tables using slider bars. A simulation can take several hours to run if results are required for all OAs in Leeds and if a relatively high temperature and a large number of iterations or restarts are selected. Therefore the system allows the user to run the simulation model for OAs contained within individual wards or a selection of wards.

The model diagnostics module provides details of the accuracy of the spatial microsimulation. It compares the simulated data for OAs with the actual census variables and produces a set of basic statistics (minimum, maximum, absolute mean, mean, standard deviation) together with the percentages of values that have been overpredicted and underpredicted. Each simulation generated has a corresponding model diagnostics table.

The mapping controls module allows the user to select a variable from a query and map the results at any of the geographical scales of OA, community area, ward, or postal sector. Community areas in Leeds have been designed as an intermediate level of spatial units (106 in total), between OAs (2349) and wards (33), that endeavour to represent functional communities recognisable by Leeds residents and planners (Unsworth and Stillwell, 2004). Figure 3 illustrates the mapping of a query relating to persons aged 20 – 30 with incomes of £20 000 – 30 000 at the ward level. The software provides some basic mapping functions which include panning and zooming and symbology editing. The mapping capability in the software is provided by the GeoTools (http://www.geotools.org) open-source Java mapping library which has been written by a group of researchers independent of this project. GeoTools is a versatile Java library which conforms to the Open GIS Consortium standard specifications in relation to GIS open operability. The library can be adapted to work in any Java-based GUI or web-based Applet.

One of the principal priorities in the development of Micro-MaPPAS was to render the system capable of designing and running simulations for scenarios based upon different assumptions about population change in the future, and also to undertake some evaluation of 'what if?' scenarios. Thus, two scenario modules have been developed known as the projection and impact modules.

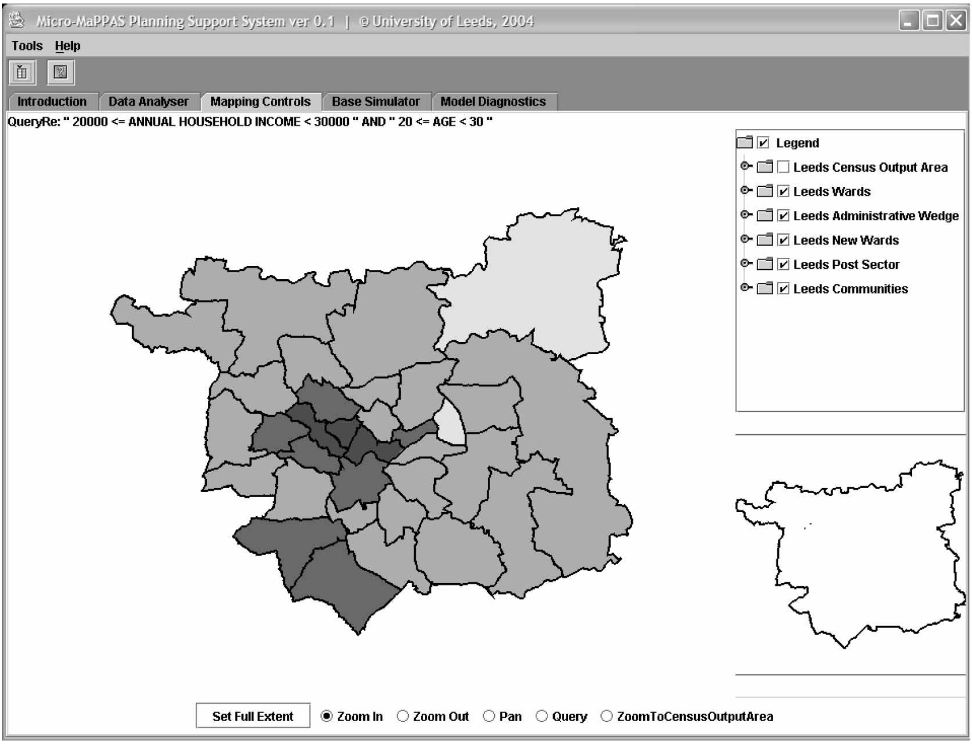


Figure 3. The results of the query as shown in the mapping controller.

6 Projecting the population into the future

In addition to modelling the population in 2001, microsimulation can be used to predict the characteristics of the Leeds population in future years. The method that is currently adopted is based on a crude projection model for aggregate ward populations and on using these totals to constrain more detailed projections generated by the Micro-MaPPAS model. In other words, the Micro-MaPPAS reweighting methodology will be applied to readjust the weights of BHPS households so that they fit small area constraint data in any selected year.

It is important to recognise that, currently, there is no provision of small area projections by a single agency in Britain. The Office for National Statistics is responsible for creating subnational population projections for England but these are only produced for local authority areas and on a relatively infrequent basis. However, it should be recognised that even short-term projections for relatively large areas may be error prone and it might not be wise to use the 1996-based district projections as constraints. Consequently, the first step in our methodology is to project the numbers of households and individuals for every year into 2021 by ward within the district independently. In order to do this, we make the simple assumption that the annual rate of change between 1991 and 2001 will continue until 2021. Therefore, we calculate for each ward the annual rates of change between 1991 and 2001 for households and individuals and apply these to successive years after 2001 to give ward-based population totals up to 2021.

The projection of the disaggregated counts of the population in future years according to marital status, socioeconomic group, number of cars owned, and so on, is undertaken by applying annual rates of change between 1991 and 2001 in exactly the same way as with the aggregate populations. The projection of these counts is necessary given that these variables are used as constraints in the simulated annealing household reweighting procedure. So, for example, having projected the car ownership characteristics by household, the next step involves calculating the proportions of all car ownership categories in each ward. These proportions are then applied to the projected numbers of households by ward in each future year. In this way, we ensure that the sum of all cars by household categories adds up to the aggregate household projection. The same method can be applied for all other household (eg, tenure) and individual (eg, ethnic group) variables.

The above discussion implicitly assumes that the 1991 Census recorded accurately the populations living in Leeds wards. However, over 2% of the population was missed overall in 1991 and this underenumeration did not occur uniformly across all areas or age – sex groups (ONS, 2001). Further, the 1991 Census did not record the number of students, which is quite large in some electoral wards such as Headingley. In order to deal with these problems, the following strategies were adopted. To tackle the problem of the undercount in 1991, the ward populations in 1991 can be readjusted on the basis of alternative assumptions about the extent of the undercount. For instance, if it is assumed that the Leeds population in 1991 was underestimated by 4%, the 1991 population numbers can be increased by this rate and the projection procedure described above is applied using the annual rates of change recomputed on this basis. A reasonable solution to the problem of not counting the students in the 1991 Census is to estimate their numbers on the basis of 2001 proportions. For instance, according to the 2001 Census, students in Headingley comprise 54.5% of the total population and this proportion can be added to the published 1991 population total.

7 The potential of Micro-MaPPAS for planning support and policy analysis

The overall aim of developing the system presented above is to provide data estimates and intelligence at the small area level that can be used to inform policy making. As pointed out by Martin (2004), there is a growing interest in neighbourhood-based policy in England. The need for neighbourhood-level data has been addressed to a certain extent by the 2001 UK Census of population which released information for very small geographical areas, output areas, for the first time. Martin (2004) discusses the characteristics of this new census geography and the extent to which the reporting of census data at that level addresses the growing need for neighbourhood-based information and suggests that there are major research challenges to 'delivering the information base necessary for truly effective neighbourhood policy" (page 144). It can be argued that one of these research challenges is to estimate policy-relevant noncensus variables, such as household income, on the basis of census area data and survey data. One of the advantages of spatial microsimulation is the ability to combine small area information from the census with survey data such as the BHPS, in order to provide policy relevant information for very small areas. In particular, it is possible to add geography to survey data such as the BHPS on the basis of small area census statistics. For instance, spatial microsimulation can be used to combine small area data from different censuses with survey data in order to simulate trends in poverty and income inequality at different geographical scales (see, for instance, Ballas, 2004).

As noted above, the Micro-MaPPAS system has been designed to create sets of simulations for 2001 as the base period. In addition, as noted in the introduction, spatial microsimulation can provide estimates of the distribution of socioeconomic and demographic variables that can be extremely policy relevant (Ballas, 2004; Ballas et al, 2005).

For instance, the Micro-MaPPAS software tool can be used in order to estimate and map geographical distributions of variables such as:

- average earned household income;
- unemployed individuals with no qualifications;
- households with below half-median income;
- number of children living in households below half-median income;
- children of school age and estimated school performance; and
- elderly individuals with limiting long-term illness (useful for health service planning).

The above variables are currently not available at small area level from any published data sources. The software is capable of estimating these datasets and of producing thematic maps for use in policy analysis. In particular, it is possible to use the geographical database in order to perform 'what if?' policy analysis, addressing policy relevant questions such as those described above.

Two examples of subgroup distributions generated by the Micro-MaPPAS queries are presented by way of illustration. These examples are both produced from a model simulation run that gives equal weight to all the tables used in the simulation. The first query provides more detailed information about the distribution of female lone parents with dependent children, where the household income is less than £10 000 per year (figure 4). The query has been carried out at the scale of community areas and the distribution demonstrates the concentration of this particular subgroup in areas to the east and south of the city centre.

The second example involves individuals at the other end of the social spectrum. Figure 5 shows the distribution by output area (with community area boundaries shown)

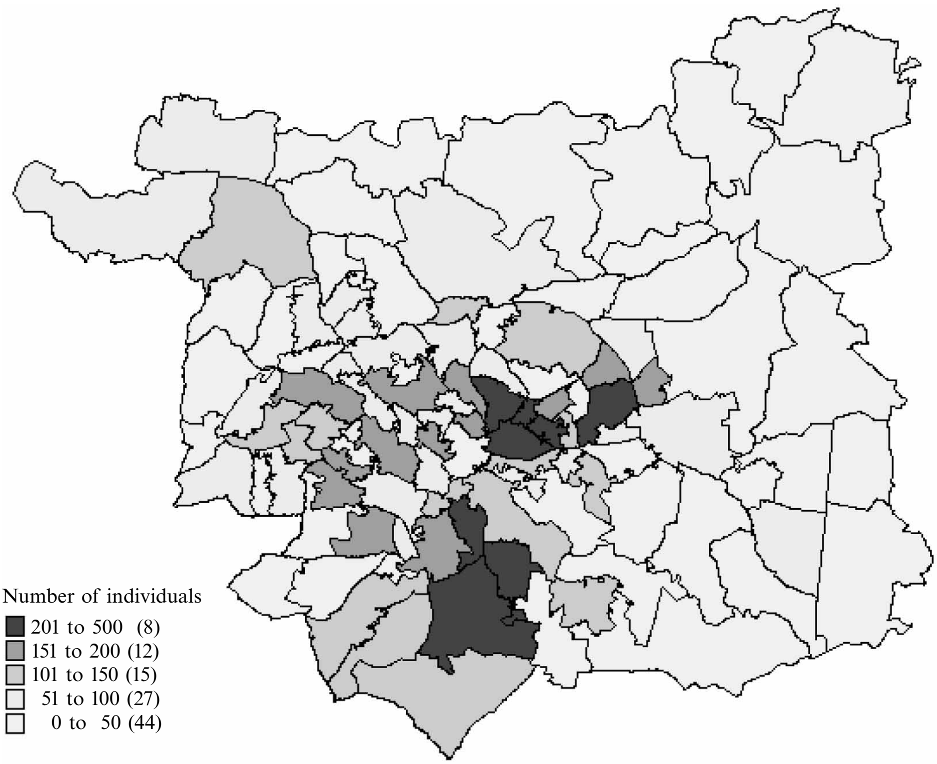


Figure 4. Simulated distribution of female lone parents with dependent children in households with low income by community area, 2001.

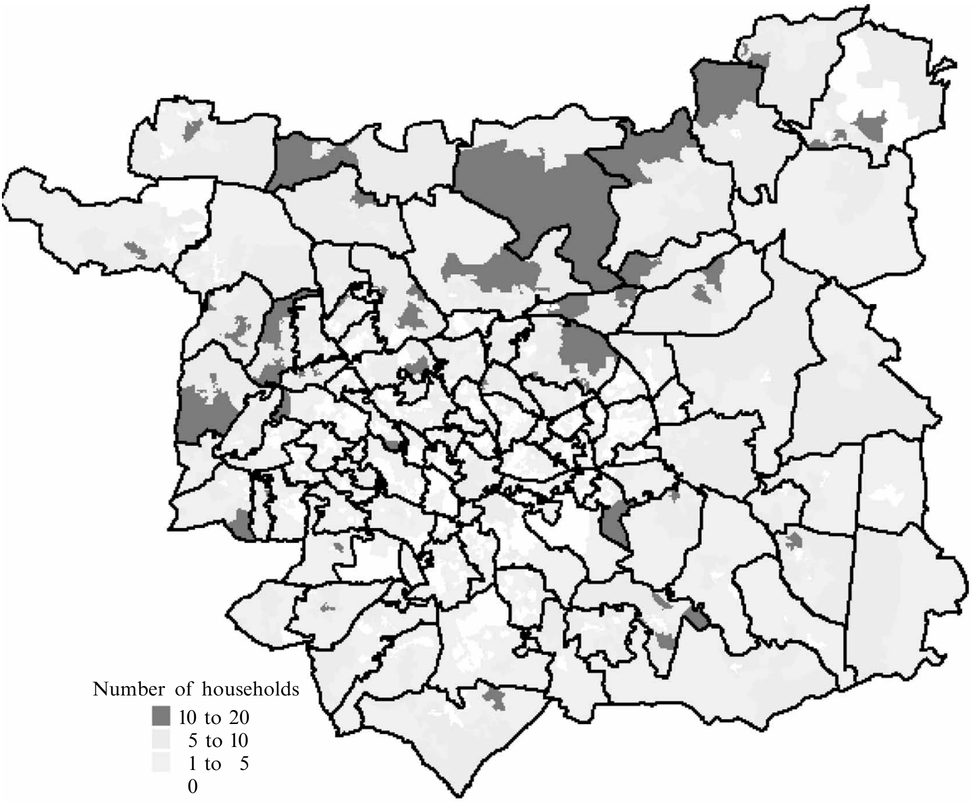


Figure 5. Simulated distribution of two-person households without dependent children with high household income by output area, 2001.

of those individuals living in two-person households with no dependent children and with a household income in excess of £50 000. These are the so-called DINKies (double income no kids) and their spatial incidence is observed across parts of the north of the district but also in pockets elsewhere.

Another important policy-relevant feature of the system is to construct interactive facilities to assess the impact of particular scenarios that may be policy related. What is the employment impact across the city of the closure of a factory or the opening of a new supermarket, for example? This type of scenario requires some measure of the interaction between where people work and where they live. In order to model economic scenarios such as these, the system uses census interaction statistics from Special Workplace Statistics (Stillwell and Duke-Williams, 2003). In particular, all simulated individuals are assigned workplaces (at ward level) on the basis of Monte Carlo sampling from probabilities based on the census travel-to-work flows.

This module allows the user to specify the numbers of new jobs or job losses for a particular ward, identify the occupation of the individuals who would lose their jobs under a particular scenario, or experiment with assumptions about the occupational categorisation of individuals in a new company. Once these parameters have been specified, it is possible to estimate the direct employment effect of job losses or job creations by area. Further, it is also possible to estimate the income effect on the basis of the microsimulated income distributions and travel-to-work flows.

Amongst the innovative features of the Micro-MaPPAS system is its use to support local policy decision making and interactive policy analysis. By providing new information at a detailed small area scale, Micro-MaPPAS can be used to help identify areas for funding initiatives or regeneration. In addition, it has the potential to allow comparisons of the spatial impacts of national social policies with the respective impacts of area-based policies, as social policies can be seen as alternatives to area-based policies. Moreover, the Micro-MaPPAS system can be used to analyse social policy in a geographically oriented proactive fashion, ie, to identify deprived localities in which poor individuals and households are overrepresented. It has the potential to answer questions such as: what kind of social policy could be applied which, all else being equal, would be most likely to improve the quality of life of residents in the inner-city localities of a city? In other words, new social policies can be formulated on the basis of the Micro-MaPPAS outputs. Spatially oriented social policies can be seen as a substitute or an alternative to traditional area-based policies and direct comparisons of their efficiency and effectiveness can be made.

Another innovation-related aspect of Micro-MaPPAS is its potential for use in the measurement and analysis of income and wealth distribution at the local level. In particular, it could be used to create a new set of deprivation indicators based on household poverty, wealth, and social exclusion that can replace or be used in tandem with traditional census-based proxies for income and wealth such as car ownership and household tenure. One of the ways of evaluating the impact of social policy on the welfare of society is to compute indices such as income inequality measures and to then study the change of these indices after a social policy programme is applied. It can be argued that the mapping of inequality measures can highlight the degree of heterogeneity of income, household types, and lifestyles within small areas. This would also allow estimation of the degree of social cohesion at various geographical scales.

Thus the Micro-MaPPAS system can be used to address a large number of social and economic problems and stimulate debate about the use of GIS and microsimulation methods for policy analysis. The system can be employed to monitor trends in socioeconomic polarisation and inequalities and evaluate the sustainability of local socioeconomic systems if past trends were to continue. In particular, one of the problems that social policies aim to tackle is poverty, yet, as McCormick and Philo (1995) point out, much poverty is hidden in the sense that poor people and localities are largely invisible. They argue that poverty in these localities is not only the result of economic decline reflected as a shift in demand for specific labour-market skills, but it is also the cause of the decline. Changes in national social policies have different and major implications for household incomes in different areas. Using Micro-MaPPAS, it is possible to assess a number of policy changes addressing societal problems in terms of geographical and socioeconomic distributive impacts and fiscal costs.

8 Concluding comments

By providing new information at a detailed small area scale, the Micro-MaPPAS system can be used to support decision making and policy formulation in relation to area regeneration. Moreover, systems such as Micro-MaPPAS can play a very important role in the ongoing debates on the role of potential new technologies to promote local democracy and new forms of decision making (6 P, 2004). The system has been designed to be as flexible as possible so that it can be adapted to work for any other city or region by changing the appropriate spatial and attribute data to match the appropriate locations.

It should be noted that, although spatial microsimulation models have had a relatively successful history so far, they have been inaccessible to policy makers and the general public in a practical sense. Albert Einstein once said "It is not enough for a handful of experts to attempt the solution of a problem, solve it, and then apply it. The restriction of knowledge to an elite group destroys the spirit of society and leads to its intellectual impoverishment" (Einstein, 1931, quoted in Craig et al, 2002).

This paper attempts to address similar concerns that microsimulation and related spatial modelling methods have been too complicated, difficult to explain, and restricted to elite groups. In particular, it is hoped that the development of an open source software model such as Micro-MaPPAS demystifies spatial microsimulation and allows for the transfer of knowledge to support further research.

It should be noted that the Micro-MaPPAS GUI has been built in Java as a standalone application. Whilst at the moment the software is linked to internal datasets, future development will enable the software to be linked to central Leeds City Council databases across an intranet. This will provide a data integration framework that will enable individual policy makers and services to use a common database structure thus reducing data replication. The advantage of building the application in Java is that it will operate on any computer system and is easily adaptable to end-users' needs. The GUI will incorporate the modelling algorithms and a mapping system built in-house called GeoTools. It must be stressed that the software product for this project will be a unique system built from scratch but will be flexible enough to be capable of reading spatial data in ESRI shapefile format and attribute data in dbf format. The modular format of the software means that it can be applied to any geographical area, at any scale and with an interchangeable range of census and BHPS data variables.

Acknowledgements. The research presented in this paper was funded by Leeds Initiative as part of their Futures Group. All census and digitised boundary data used in this paper are Crown Copyright. The BHPS data were obtained from the UK Data Archive (University of Essex). The constructive comments of the editor and three anonymous referees are greatly appreciated.

References

6 P, 2004 *Rewiring Local Decision Making for Political Judgment* (New Local Government Network, London)

Alvanides S, 2000 *Zone Design Methods for Application in Human Geography* unpublished PhD Thesis, School of Geography, University of Leeds

Ballas D, 2001 *A Spatial Microsimulation Approach to Local Labour Market Policy Analysis* unpublished PhD Thesis, School of Geography, University of Leeds

Ballas D, 2004, "Simulating trends in poverty and income inequality on the basis of 1991 and 2001 census data: a tale of two cities" *Area* **36** 146 – 163

Ballas D, Clarke G P, 2000, "GIS and microsimulation for local labour market policy analysis" *Computers, Environment and Urban Systems* **24** 305 – 330

Ballas D, Clarke G P, 2001a, "Towards local implications of major job transformations in the city: a spatial microsimulation approach" *Geographical Analysis* **33** 291 – 311

Ballas D, Clarke G P, 2001b, "Modelling the local impacts of national social policies: a spatial microsimulation approach" *Environment and Planning C: Government and Policy* **19** 587 – 606

Ballas D, Clarke G P, 2003, "Microsimulation and regional science: 30 years of spatial microsimulation of populations", paper presented at the 50th Annual North American Meeting of the Regional Science Association, Philadelphia, USA, 19 – 22 November, copy available from Dr Ballas

Ballas D, Clarke G P, Turton I, 1999, "Exploring microsimulation methodologies for the estimation of household attributes", paper presented at the 4th International Conference on GeoComputation, Fredericksburg, Virginia, USA, 25 – 28 July, http://www.geocomputation.org/1999/index.html

Ballas D, Clarke G P, Turton I, 2003, "A spatial microsimulation model for social policy micro-spatial analysis", in *Modelling Geographical Systems: Statistical and Computational* Eds B Boots, A Okabe, R Thomas (Kluwer, Dordrecht) pp 143 – 168

Ballas D, Rossiter D, Thomas B, Clarke G P, Dorling D, 2005 *Geography Matters: Simulating the Local Impacts of National Social Policies* (Joseph Rowntree Foundation, York)

Batty M, 2005, "Agents, cells, and cities: new representational models for simulating multiscale urban dynamics" *Environment and Planning A* **37** 1373 – 1394
Brail R K, Klosterman R E, (Eds), 2001 *Planning Support Systems: Integrating Geographic Information Systems, Models and Visualization Tools* (ESRI Press, Redlands, CA)
Clarke G P, 1996, "Microsimulation: an introduction", in *Microsimulation for Urban and Regional Policy Analysis* Ed. G P Clarke (Pion, London) pp 1 – 9
Craig W, Harris T M, Weiner D, (Eds), 2002 *Community Empowerment, Public Participation and Geographic Information Science* (Taylor and Francis, London)
Danson M, Lavarcombe A, 1996, "Simulating changes in the production process of a firm", in *Microsimulation for Urban and Regional Policy Analysis* Ed. G Clarke (Pion, London) pp 187 – 200
DETR, 2001 *Local Strategic Partnerships: Government Guidance* Department of the Environment, Transport and the Regions, London
Dowsland K, 1993, "Simulated annealing", in *Heuristic Techniques for Combinatorial Problems* Ed. C Reeves (Blackwell, Oxford) pp 20 – 69
Geertman S, Stillwell J, (Eds), 2003 *Planning Support Systems in Practice* (Springer, London)
Geertman S, Stillwell J, 2004, "Planning support systems: an inventory of current practice" *Computers, Environment and Urban Systems* **28** 291 – 310
Heady C, Mitrakos T, Tsakloglou P, 2001, "The distributional impact of social transfers in the EU: evidence from the ECHP" *Fiscal Studies* **22** 547 – 565
Kirkpatrick S, Gelatt C D, Vecchi M P, 1983, "Optimization by simulated annealing" *Science* **220** 671 – 680
Leeds Initiative, 2001 *A Neighbourhood Renewal Strategy for Leeds* Neighbourhood and Communities Partnership, Leeds City Council, Leeds
McCormick J, Philo C, 1995, "Where is poverty? The hidden geography of poverty in the United Kingdom", in *Off the Map: The Social Geography of Poverty in the UK* Ed. C Philo (Child Poverty Action Group, London) pp 1 – 22
Martin D, 2004, "Neighbourhoods and area statistics in the post 2001 census era" *Area* **36** 136 – 145
Matsaganis M, Papadopoulos F, Tsakloglou P, 2001, "Eliminating extreme poverty in Greece" *Journal of Income Distribution* **10** 40 – 57
Mitton L, Sutherland H, Weeks M (Eds), 2000 *Microsimulation Modelling for Policy Analysis: Challenges and Innovations* (Cambridge University Press, Cambridge)
Moeckel R, 2005, "Microsimulation of firm location decisions", in *Proceedings of the 9th International Conference in Computers in Urban Planning and Urban Management (CUPUM)* London, 29th June – 1 July, available from http://www.cupum.org
Norman P, 1999, "Putting iterative proportion fitting on the researcher's desk", WP 99/03, School of Geography, University of Leeds
ONS, 2001 *Census 2001: A Guide to the One Number Census* Office for National Statistics, http://www.statistics.gov.uk/census2001/pdfs/oncguide.pdf
Openshaw S, Rao L, 1995, "Algorithms for reengineering 1991 Census geography" *Environment and Planning A* **27** 425 – 446
Openshaw S, Schmidt J, 1996, "Parallel simulated annealing and genetic algorithms for re-engineering zoning systems" *Geographical Systems* **3** 201 – 220
Orcutt G H, 1957, "A new type of socio-economic system" *The Review of Economics and Statistics* **39** 116 – 123
Pham D T, Karaboga D, 2000 *Intelligent Optimisation Techniques: Genetic Algorithms, Tabu Search, Simulated Annealing and Neural Networks* (Springer, London)
Redmond G, Sutherland H, Wilson M, 1998 *The Arithmetic of Tax and Social Security Reform: A User's Guide to Microsimulation Methods and Analysis* Cambridge University Press, Cambridge
SEU, 2000 *National Strategy for Neighbourhood Renewal* Report of Policy Action Team 18, Social Exclusion Unit (The Stationary Office, London)
SEU, 2001 *A New Commitment to Neighbourhood Renewal—National Strategy Action Plan* Social Exclusion Unit, London
Stillwell J, Duke-Williams O, 2003, "A new web-based interface to British census of population origin – destination statistics" *Environment and Planning A* **35** 113 – 132
Sutherland H, Piachaud D, 2001, "Reducing child poverty in Britain: an assessment of government policy 1997 – 2001" *The Economic Journal* **111** 85 – 101
Taylor M F, Brice J, Buck N, Prentice-Lane E, 2001 *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices* (University of Essex, Colchester)

Unsworth R, Stillwell J, (Eds), 2004 *Twenty-first Century Leeds: Geographies of a Regional City* (Leeds University Press, Leeds)
Van Laarhoven P J M, Aarts E H L, 1987 *Stimulated Annealing: Theory and Applications* (Kluwer, Dordrecht)
Van Wissen L, 2000, "A micro-simulation model of firms: applications of concepts of the demography of the firm" *Papers in Regional Science* **79** 111 – 134
Voas D W, Williamson P, 2000, "An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata" *International Journal of Population Geography* **6** 349 – 366
Williamson P, 1999, "Microsimulation: an idea whose time has come? ", paper presented at the 39th European Regional Science Association Congress, University College Dublin, 23 – 27 August; copy available from the author, Department of Geography, University of Liverpool
Williamson P, Birkin M, Rees P, 1998, "The estimation of population microdata by using data from small area statistics and samples of anonymised records" *Environment and Planning A* **30** 785 – 816