UCD Urban Institute Ireland

Working Paper Series

2008

Calculating Population Density for MOLAND in the Dublin Urban Area

Ronan Foley, Martin Critchley, Sheila Convery

UCD UII 08/08

UCD Urban Institute Ireland

University College Dublin

www.ucd.ie/uii

UCD Urban Institute Ireland

WORKING PAPER SERIES

2008

Calculating Population Density for MOLAND in the Dublin Urban Area

Ronan Foley, Martin Critchley, Sheila Convery

University College Dublin

Working Paper Series: UCD Urban Institute Ireland

ISSN: 1649-3613

UCD Urban Institute Ireland
University College Dublin,
Clonskeagh Drive,
Clonskeagh,
Dublin 14
Ireland
www.ucd.ie/uii/

This working paper can be downloaded as a PDF at
www.uep.ie

Urban Institute Ireland Working Paper Series

Calculating Population Density for MOLAND in the Dublin Urban Area

Ronan Foley[a,*], Martin Critchley[b] and Sheila Convery[c]
[a]Department of Geography, National University of Ireland, Maynooth
[b]ERA Maptech Ltd
[c]UCD Urban Institute Ireland

Abstract

The EPA-funded Urban Environment Project seeks to apply MOLAND, a cellular automata based model, to predict a range of urban impacts in the Greater Dublin Region from 2006 onwards. One quantitative requirement, at least in the input stages of model building, is to establish average population densities for the different residential land use classes because the population must be allocated to cells in these classes. It has a separate output function when actually used within the MOLAND model, relating to the estimation of numbers of people affected by a particular scenario generated by the model. This working paper explores a series of approaches taken to identify population densities for a number of urban land-use categories. The results are documented against a number of differing spatial approaches to test for precision and applicability. Some suggestions are proposed as to the most effective method to use in developing the model.

Keywords: *population density estimates; MOLAND; land use; urban; enumerator area; enumerator districts; cellular-automata.*

1 Introduction

The Irish Environmental Protection Agency (EPA) has funded a group of partner organisations, led by the Urban Institute of Ireland at UCD Dublin and including NUI Maynooth, Trinity College Dublin and ERA Maptech Ltd., to model urban environments in the Greater Dublin Region (GDR). The project is funded from 2006-2010, with the aim of predicting changes to the urban environment up to 2025. Within the project, five broad domains around urban sprawl, air quality, transport, biodiversity and climate change are being developed as sub-projects. A sixth group is looking at specific modelling issues and this paper is drawn from research by members of that modelling group.

Technically, the project has been set up to make use of MOLAND (Monitoring Land Use / Cover Dynamics), a cellular automata model designed to produce spatially enabled predictions of future urban development patterns. Cellular automata were originally developed as an efficient tool for understanding dynamical systems, but have been adopted into spatial decision support systems for application in a range of planning-related environments (White and Engelen, 1993). The antecedents of the current project stem from the EU sponsored MURBANDY project, initiated in 1998 to dynamically model urban development across a number of European cities (Engelen et al., 2004). This project has been extended into the MOLAND project since 2004, which has been developed by the Institute for Environment and Sustainability at the EU's Joint Research Centre in Ispra, Italy (Lavalle and Barredo, 2004). The technical development of the cellular automata part of the project has been led by commercial partners from RIKS (Research Institute for Knowledge Systems Ltd.) based in Maastricht in the Netherlands.

The function of MOLAND is to attempt to predict, using a set of internal rules and a specified range of spatial, demographic, and economic data sets, future change scenarios in the GDR. The specific process by which it does this is not the core concern of this paper; however it is worth noting that the model is inherently spatial and generates output at a grid resolution of 200 m by 200 m. From

*Corresponding Author: Ronan Foley.
E: ronan.foley@nuim.ie. T: 00 353 1 7086024.
UCD Urban Institute Ireland Working Paper Series (08/08)
ISSN 1649-3613. 20081222

a GIS perspective, loose coupling is possible with the output and input to the model being developed in raster formats, transformed to match cell resolutions and locally projected co-ordinate systems (Longley et al., 2005). Essentially the model imports a set of spatial (raster) data sets such as land use; planning zoning; transport networks; and social and economic data at the county level to try and predict how current land-use in any raster cell will change in future scenarios as well as an assessment of what future land-uses will be. As part of the process the model is calibrated to ensure it performs in a robust way. One of the stages of that process is to identify a set of population densities for each residential land-use code as one of the required inputs to the model. This information also helps to predict population fluctuations against any modelled change and feed in to any impact assessments associated with that modelling.

The aim of this brief paper is to identify a set of working population densities for specified urban land-uses. This will be achieved by looking at five different methods to identify population densities using spatial data sets at differing scales within a GIS environment. A full technical description of the processes will be outlined below but will essentially move from a small-scale set of indicators to a large-scale set. The final outcome of the investigation will be to suggest an optimal approach from those tested which will strike an effective balance between speed, efficiency and precision. As part of the investigation, a critical appraisal will also be made around the quality of the datasets themselves as well as the relationship between the modelled population densities and the structure of the MOLAND model.

2 Data & Method

2.1 Data

The foundation for the land-use data is CORINE, an EU-wide harmonised digital land-use classification system. CORINE data has been derived from remotely sensed satellite imagery for three time periods, 1990, 2000 and 2006. Within MOLAND, the same land-use classifications have been adopted using five broad categories such as artificial surfaces, agricultural areas and wetlands. Within the artificial surfaces category (1) a sub-division relates to the urban fabric (1.1) which in turn is sub-divided into two further categories: continuous urban fabric (1.1.1) and discontinuous urban fabric (1.1.2). It should be noted there are additional urban land-uses as well such as industrial, commercial etc., but this paper will focus solely on the residential sub-classes. Up until 2000, these were the classes used for residential land use in urban areas. In the most recent iteration of MOLAND, a number of further sub-categories have been created (see Table 1.). These further sub-divisions are used in this paper. Specifically, the categories 1.1.1.1 (RCDU), 1.1.1.2 (RCMDU), 1.1.2.1 (RDU) and 1.1.2.2 (RDSU) will be modelled and tabulated.

Other data sets required for the modelling related to the background population data from which the estimated population densities were extracted. There were four in total. The first two of these were population data gathered from the Irish Censuses of 2002 and 2006. These were made available by the Central Statistics Office (CSO) at two spatial levels, Electoral Division (ED) and Enumerator Area (EA). EDs have average populations nationally of around 1,100 though this was found to be closer to 3,000 in the GDR. EAs are sub-divisions of EDs and typically consist of approximately three to four per ED with population averages in the region of 7-800. In terms of nationally validated areal data, these were the most precise data available within Ireland. A third dataset that was available for estimation of population was GeoDirectory, the national address database. The dataset is jointly produced on a quarterly basis by An Post, the national postal operator and Ordnance Survey Ireland (OSi). It identifies 1.8 million delivery points with additional information on the estimated number of households at each point. These can range for example from a large individual detached house to a block of flats under multiple occupancy. A fourth dataset was provided on request from the CSO to ERA Maptech and contains real data at individual street level for some selected sample areas in the city against which a full count of residents could be used for density estimation purposes. While this fourth dataset is clearly the most accurate, it is an ad-hoc dataset, which has limited scientific potential for the model in terms of its availability. It is used here simply as an indicator of the effectiveness of the other three methods and as a potential validator of those methods.

2.2 Method

Once the relevant land-use codes were identified, ERA Maptech provided a set of polygons in GIS vector format to act as a set of sample areas for the modelling. These were provided for the four MOLAND classifications, namely RCDU, CRMDU, RDU and RDSU. The fifth category, residential urban blocks, was too small to model and was not originally applied within Dublin and was excluded for those reasons. There were between three and four sample polygons available for each of the four

Table 1 – MOLAND Sub-Codes for Residential Land-Uses

1.1.1 Continuous Urban Fabric	1.1.1.1 Residential continuous dense urban fabric (RCDU)
	1.1.1.2 Residential continuous medium dense urban fabric (RCMDU)
	1.1.1.3 Informal settlements*
1.1.2 Discontinuous Urban Fabric	1.1.2.1 Residential discontinuous urban fabric (RDU)
	1.1.2.2 Residential discontinuous sparse urban fabric (RDSU)
	1.1.2.3 Residential urban blocks (RUB)*
	1.1.2.4 Informal discontinuous residential structures*

* *These land-uses are not considered in this work*

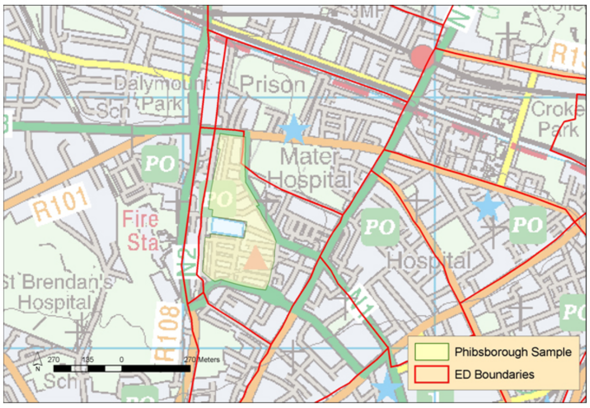

Figure 1 – Visualisation of Population Density Estimation Method 1 (ED)

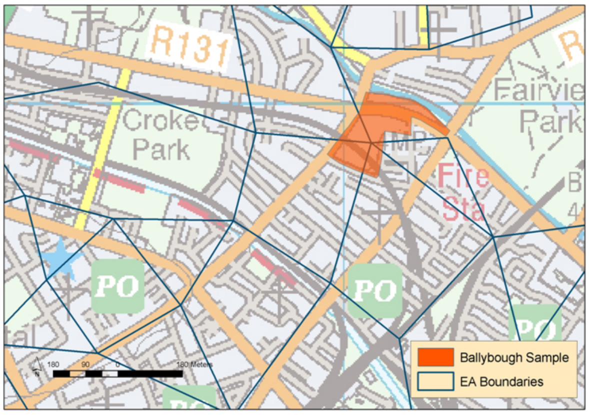

Figure 2 – Visualisation of Population Density Estimation Method 2 (EA)

categories. Around five of these fell within the Urban Environment Project (UEP) local area of the Malahide Corridor, though all of the others fell within the model's broader regional catchment. While only one example for each classification was chosen for the modelling, there were a number of issues related to the decision to use a sample approach that will be problematised more fully below.

Method 1: Population Estimation using ED demographic and extent data

For this method, the individual representative land use polygons were overlaid with the ED boundary layer. While the extent to which the sample polygons fell completely within an ED varied and some crossed EDs, in general they fell within a single ED. The population density for that ED, or the ED into which the largest proportion of the sample area fell, was then calculated from the area and 2006 population census counts and applied to that sample polygon. The method is visualised in Figure 1, which shows the sample polygon for Phibsborough overlaid with the surrounding ED.

Method 2: Population Estimation using EA demographic and extent data

A similar approach to that used for EDs was also taken with EAs. However, in the case of EAs, the sample polygons were much more likely to cross the EA boundaries, given the smaller size of those units when compared to EDs. In some cases they fell across five or six EAs. In those cases the individual population densities (based on area and 2006 population counts) for each EA were averaged out to provide a representative population density estimate which would arguably be more accurate than that of the bigger ED. Clearly there will be some modifiable areal unit (MAUP) problems with this process but these are in part addressed in the next model, though they remain problematic within all the methods tested (Charlton, 2007). The method is visualised below in Figure 2, which shows the sample polygon for Ballybough overlaid with the surrounding EAs.

Method 3: Population Estimation using GeoDirectory points

As a third approach, GeoDirectory was specifically used to firstly, get past MAUP problems and secondly, to provide a point (rather than areal) estimation approach. For each of the sample poly-

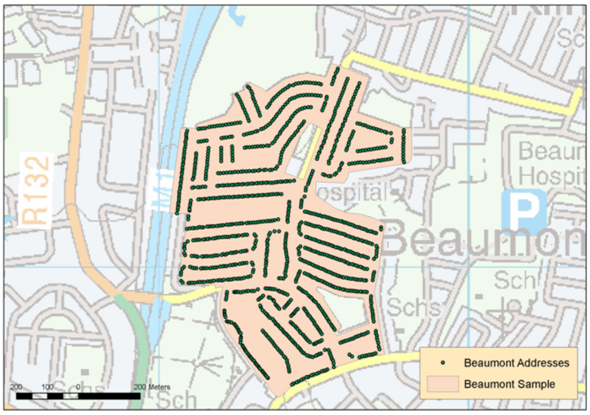

Figure 3 – Visualisation of Population Density Estimation Method 3 (GeoDirectory)

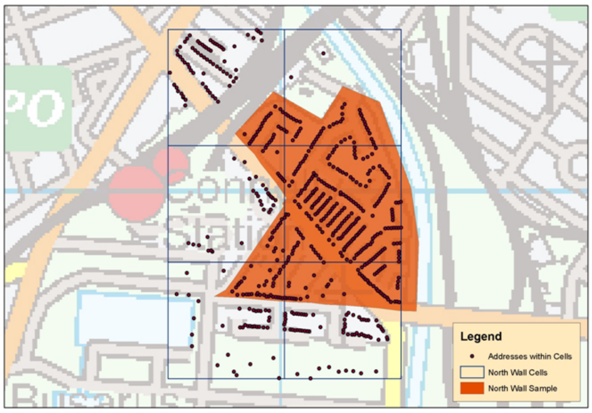

Figure 4 – Visualisation of Population Density Estimation Method 4 (GeoDirectory/GridCell)

gons, the GeoDirectory was spatially queried (usually using the postal district code for the specific part of Dublin) and clipped within a GIS to strip out only those points that fell within the polygons. Within GeoDirectory there was a field that identified a count of the number of residential delivery points, an effective proxy for numbers of households. This was in turn, modelled with the relevant local ED level data for the number of persons per household and then multiplied by the modelled number of households to provide an estimate of population. Finally, this was weighted by the polygon's total area to provide an estimate of population density derived from individual household points. The method is visualised in Figure 3, which shows the sample polygon for Beaumont overlaid with the clipped GeoDirectory points that fell within its boundary.

Method 4: Population Density Estimation using GeoDirectory with MOLAND grids.

As a fourth approach and to attempt to more closely link the method to the final model, a decision was made to use the sample sites and ancillary data sets to try and model population densities for the specific 200 m × 200 m grid cells used in MOLAND. Given the size of the grid involved, it made little sense to try and link it to either ED or EA data. It made more technical sense, given the fine scales involved, to overlap the grid cells with the GeoDirectory data. In addition, the sample area chosen was North Wall to show the development from the previous model. The sample area was clipped against the grid cells that overlapped with its bounds using a spatial join process commonly found in GIS (Figure 4). Six individual 200 m × 200 m cells were used (and applied in all the subsequent samples bar one), and population densities (using the approach described in Method 3 linking household counts with average household size) were calculated as averages across all six cells. The final set of household counts from GeoDirectory that fell within the identified cells was then multiplied by average household size (for the sample area) to produce a population estimate. This was then divided by the size of the six cells, namely 24 hectares to provide a fourth estimate of population density per hectare.

Method 5: Population Estimation using Original CSO street-level data

This final approach was developed in conjunction with the 2000 land-use maps and involved requesting the original census returns from the CSO for a specific number of streets that fell within the sample polygons. In this case, very confidential data was made available under Officer of Statistics protocols, which provided a specific count of households and the numbers of residents in each household. Unlike the proxy count identified through GeoDirectory this was a real count of residents. But as noted previously, due to data constraints, it would be virtually impossible to reproduce this and here it functions solely as a validating count for the other approaches. For the different land-use classifications within MOLAND, a series of values for the same sample polygons were generated and they are reproduced in Table 2.

3 Results

The data on population densities were developed using the first four methods. There were eight different sample polygons reported in Table 2, representative of two different sample sites for each of the four classes identified from the MOLAND classification. One set, listed in the top part of the table are specifically found in the general vicinity of the Malahide Corridor and within the identified UEP

Table 2 – Population Density Estimates by Sample and LU Code

Land-Use Code	1.1.1.1 Continuous Dense	1.1.1.2 Continuous Medium Dense	1.1.2.1 Discontinuous	1.1.2.2 Discontinuous Sparse
Location	Ballybough	Phibsborough	Beaumont	North Wall
1. ED	97	113	44	66
2. EA	81	100	45	82
3. GeoDirectory	83	164	60	224
4. MOLAND Grid	117	121	58	194
5. CSO Data	138	195	56	131
Mean (all methods)	103	139	53	139
Location	Royal Exchange	South Dock	Palmerstown	Jobstown
1. ED	162	49	31	46
2. EA	108	34	30	18
3. GeoDirectory	496	155	76	126
4. MOLAND Grid	238	147	83	102
5. CSO Data	173	142	62	68
Mean (all methods)	235	105	56	72
All Sample Means	169	122	55	106
Working Mean	169	122	55	106

All values in persons per hectare (persons ha^{-1})

local study area on the north side of the city.

The second set of results, in the lower half of the table, are drawn from the south of the river to provide a balance and to test out the method in different city districts. The density figures were initially calculated as counted or estimated populations divided by the area of the sample polygons. In the case of Method 4, a set of MOLAND grid cells surrounding the sample polygons was used instead of the polygons themselves. While additional sample polygons for the same codes were also available in other parts of the city, the presented results based on two sample sites per classification were chosen as broadly representative. Broadening the sample and generating average population densities across the samples would provide a wider range of estimated values but in relation to the requirements of the model, it appears that a single validated number per code is sufficient for modelling purposes. In addition, given the considerable variation, even across the eight sampled polygons, the addition of more samples was likely to confuse, rather than clarify, the density estimates. The population densities could have been reported as population per square kilometre but given that MOLAND uses hectares for its cell measurements; this was chosen as a more applicable scale for the outputs.

From the data generated by the first four approaches and mindful of a number of caveats, some broad conclusions may be drawn. Firstly, it would appear that across the first four sample polygons in the Malahide Corridor there appeared to be a broadly similar estimate generated by Method 1 and Method 2. There seemed to be more variability in the samples south of the river. One possible explanation is that the first four sample were deliberately chosen to fit within as few EDs and EAs as possible. It was certainly the case that some of the sample polygons south of the river were much larger and were therefore drawing average data from, in some cases, up to five EDs and eight EAs. This was bound to introduce more variability. Using ED and EA population estimates, both seemed to provide similar results and given these were the quickest and easiest methods to apply, it was helpful to note this. It might be argued that there were subtle differences across the four categories. Both of the continuous categories seem to produce a more precise figure when estimated at EA level, which would make sense given the greater precision of the latter. This must be balanced by the use of multiple EAs to estimate a density for the sample polygons, whereas the ED estimates tend to draw from one, or at most two, areas. The corollary is the fact that for the two discontinuous areas on the Northside, the EA estimates are larger, especially in the case of the North Wall. On the Southside, the smaller estimate associated with the EA method can also be observed. There may be an argument to be made here in relation to the nature of the land use category and an expectation that EDs are more heterogeneous in the discontinuous urban areas. It should also be noted that prior to 2002, EA level data was only available for the larger Irish cities. From 2006 EA layer coverage has been extended to all towns with populations over 1,500. While this is an improvement in coverage, outside of urban areas EDs will still have to be the basis on which population density is calculated. Given the project's interest in a wider regional area, this full and consistent applicability must be borne in mind.

A second linked conclusion is the fact that the densities for the two continuous areas are significantly higher at ED, and in most cases at EA level, when compared with the discontinuous sample areas. This again makes sense given the background categories for which the densities are being estimated. The fact that densities in continuous urban areas are higher than those in discontinuous areas confirms expected findings. However, there are some counter-intuitive results when comparing the estimated densities for discontinuous and discontinuous sparse, with the latter unexpectedly exceeding the former in most cases. This suggests some issues with both the classification of, and sampling from the discontinuous sparse areas which might bear further investigation. The third and last comment is in relation to the fact that for the latter two, Methods 3 and 4, the GeoDirectory modelling approach throws up significantly higher figures in all cases except Ballybough. There are a number of issues with the process, not least of which are problems in relation to the use of residential address points as proxies for households and the subsequent use of the average ED household size as a numerator to calculate population density. Using the cell based approach in Method 4 provides broadly similar density estimates as for Method 3, though in the majority of the cases they are smaller, and arguably more accurate. This is countered by the fact that Method 4 ultimately takes its sample from a wider area than the initial sample polygons and this has a flattening effect on the count. While it would appear to be potentially a more problematic process, there was an initial assumption that this should provide a more accurate estimate than the 'area' based Methods 1 and 2. When the 'real data' (Method 5) was introduced, it did intriguingly, provide higher density values in the cases of the two continuous classes than any of the estimation methods used. This did suggest that the higher values generated by the GeoDirectory based methods may be likely to provide better estimates for the two continuous classes. For the discontinuous class, the Beaumont sample was consistent across all methods, whereas in general, the 'real data' seemed to provide values between the area-based approaches (Methods 1 and 2) and the point-based approaches (Methods 3 and 4).

In this sense it would appear that of the four trialled methods, it is quite difficult to identify which are likely to be the most fruitful ones to use in estimating population densities for individual urban and other land use codes within MOLAND. All are flawed in some form or other. It might be considered that averaging out the density estimates across the ED and EA methods would be effective but this adds complexity and it might be more statistically consistent to take the sampled average of the single preferred method, Method 2 (EA) instead. In addition, the validated counts provided by Method 5 suggest that the higher values generated by the point-based GeoDirectory approach may be closer to the true figure. In truth, an average value for each land use class, based on just two of the sampled areas averaged across all methods may provide as good a value as any. This suggests that the following densities, despite some counter-intuitive findings discussed below, be taken from Table 2 and applied to urban areas in the model;

Continuous Dense:	169 persons ha^{-1}
Continuous Medium Dense:	122 persons ha^{-1}
Discontinuous:	55 persons ha^{-1}
Discontinuous Sparse:	106 persons ha^{-1}

One final comment on the population density estimates was that it was considered advisable to check in broader terms how these validation estimates compared with any estimates of population density used in the original calibration of the model. The original figures for population used in the calibration were based on taking the polygon of the original single category of RDCU (see Table 1) and cookie-cutting that against background ED level population counts. By taking the single area and population measurement of that area it was possible to calculate an implied population density as embedded in the calibration of the model. The figure identified for residential continuous dense land use was 110 which could be compared to a figure of 146 (based on the average of continuous and medium dense above) identified above. This suggests that there may be a disparity between the validated and calibrated population densities, though given the issues with sampling discussed below, this may represent an acceptable variation. One possible explanation may be that, since all the population in each county must be attributed to residential cells, people in fact living on cells shown as non-residential are attributed to residential cells. Thus mean residential cell densities will be higher than actual mean densities.

The above technical test has identified a number of methods to identify population densities for use with the model. Given the requirements to provide a broad estimate for the four different land use classes for use within MOLAND, the process has arguably been successful, especially given the modest requirements of the model itself. However, a number of important caveats must be noted for both the estimation and application of these modelled population densities within MOLAND. While greater levels of precision may be possible, the requirements of the model mean that taking average densities from representative EDs will suffice. However those population densities, have, in this worked example, been taken from just two sampled polygons, one to the north and one to the south of the River Liffey. The addition of other sampled values is possible, drawn from other parts of the city. This would average the density value out from different sampled locations, which statistically would provide a better estimate (Wilson and Fotheringham, 2007). However, it is difficult to see what major differences a single figure averaged from a number of samples will provide. Essentially all the model needs is a single value for each of the land-uses that it uses. This throws up a number of issues around the model itself, which deserve further consideration.

Within geocomputation and spatial analysis, there has been a general concern with the rationale behind using global or local statistics to represent area-based statistics (Fotheringham et al., 2002). This has some relevance for MOLAND given that within the model, a specific land-use category such as continuous dense urban, runs across the whole of the central city area of Dublin. By taking a global approach one assumes that the population density within the model is the same across that whole area. Again this is all that the model, being global in its basic foundation, actually requires. However, clearly there are variations in population density across the central city area, which would be and indeed are clear from all the sampling used in this paper. Applying an approach that tries to represent this local variation might be something for the model to consider in the future. There are two possible ways it could do this, though each would require a considerable amount of further thinking. One would be to take a preferred method such as Method 2, the use of EA population densities. By having a differential EA density for all the different units that overlay the model, a local statistic could be applied within each of the EAs and applied to the cells within the model. A similar but alternative approach might be to go back to the grid cells from the MOLAND model (as used in Method 4 but without using GeoDirectory) and use these as the sample polygons to extract differential population densities from the background EA level data that could then provide local estimates to the modelling. While both are problematic to develop within the constraints of the model, they do identify a potential finer level of precision within the model and do bear consideration in any refinements made to the model in the future.

Within the study area for Ireland, clearly the spatial scale of EDs and even EAs continue to affect their applicability. Though less of a problem within the city, the size and heterogeneity of ED populations in the wider GDA study area is problematic with ranges from around 150 to over 32,000 in 2006. This means that using the ED for population density estimation in rural areas, or as is likely to be the case, taking a standard population density for all peripheral or non-urban areas, will again lead to representational difficulties. One possible future solution will be the introduction in 2011 of a set of atomic small areas (SAs) across the country. These are currently being developed in conjunction with the CSO by the National Centre for GeoComputation at NUI Maynooth (Foley et al., 2005). These new units will have three specific advantages for the type of modelling envisaged through MOLAND. Firstly the size of the new SA units will be smaller, at around 125 households on average, than the current EDs and EAs. Secondly they are likely to be much more homogenous across the range and therefore any application of locally derived indicators at a global scale will be more statistically sound. Thirdly, the size of unit will be of a much more similar standard between urban and rural areas. This means again that the application of density indicators will be more consistent for the two modelling scales within MOLAND.

All of the above discussions are based on the premise that there are wider issues around how to set densities within a cellular model. A number of the suggestions put forward are likely to be difficult to implement but could perhaps be taken on board for the developing use of MOLAND in urban land use mapping. Within the constraints of that model, this technical paper had as its aim the need to come up with a working set of densities for each urban fabric land use within MOLAND. The methods identified are ones that have been investigated primarily for replicability and ease of use within the model and the resultant values can be used in the subsequent modelling. While not the brief of this paper, similar methods can clearly be applied to most of the other land-use categories used in the model as well with a probable dependence on a mix of the methods discussed here. The specific method to be used may vary depending on the land-use but this is a quite separate discussion from the outcome of this piece of research.

Acknowledgements

The authors would like to acknowledge the support of the Environmental Protection Agency (EPA) through the National Development Plan (NDP) under contract 2005-CD-U1-M1. All work undertaken with the MOLAND model for the Greater Dublin Region is subject to the licence conditions of the software developers, Research Institute for Knowledge Systems b.v. (RIKS b.v.), and the dataset owners, The Joint Research Centre (DG JRC) under licence number JRC.BWL.30715.

References

Charlton, M. (2007). *Encyclopaedia of GIS*, chapter Modifiable Arial Unit Problem, pages 288–290. Sage, London.

Engelen, G., White, R., and Uljee, I. (2004). *The MURBANDY and MOLAND models for Dublin.* Research Institute for Knowledge Systems (RIKS), Maastricht.

Foley, R., Brady, G., Fahey, D., Haase, T., McCormack, B., Morrin, P., Murphy, K., and O'Shea, C. (2005). *Final Report of the Small Areas Project II: Proofs of Concept. The Value and Potential of Small Area Spatial Data in Ireland in the 21st Century.* Report to the Information Society, Dublin.

Fotheringham, A., Brunsdon, C., and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* Wiley, Chichester.

Lavalle, C. and Barredo, I. (2004). *The MOLAND model for urban and regional growth forecast: A tool for the development of sustainable development paths.* European Commission Joint Research Centre, Ispra.

Longley, P., Goodchild, M., Maguire, D., and Rhind, D. (2005). *Geographic Information Systems and Science.* Wiley, Chichester, 2 edition.

White, R. and Engelen, G. (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land use patterns. *Environment and Planning A*, 25:1175–1199.

Wilson, J. and Fotheringham, A., editors (2007). *Handbook of Geographic Information Science.* Blackwell, Oxford.