CREATING A THREE LEVEL BUILDING CLASIIFICATIONUSING TOPOGRAPHIC AND ADDRESS-BASED DATA FOR MANCHESTER

Masroor Hussain [a] , * and DongMei Chen [a]

[a] LaGISA, Department of Geography, Queen's University Kingston, ON, Canada – (hussainm, chendm)@queensu.ca

Technical Commission II

KEY WORDS: Building Classification, Urban land use, Learning Algorithms, Pattern Recognition, Cartrometry

ABSTRACT:

Buildings, the basic unit of an urban landscape, host most of its socio-economic activities and play an important role in the creation of urban land-use patterns. The spatial arrangement of different building types creates varied urban land-use clusters which can provide an insight to understand the relationships between social, economic, and living spaces. The classification of such urban clusters can help in policy-making and resource management. In many countries including the UK no national-level cadastral database containing information on individual building types exists in public domain. In this paper, we present a framework for inferring functional types of buildings based on the analysis of their form (e.g. geometrical properties, such as area and perimeter, layout) and spatial relationship from large topographic and address-based GIS database. Machine learning algorithms along with exploratory spatial analysis techniques are used to create the classification rules. The classification is extended to two further levels based on the functions (use) of buildings derived from address-based data. The developed methodology was applied to the Manchester metropolitan area using the Ordnance Survey's MasterMap®, a large-scale topographic and address-based data available for the UK.

1. INTRODUCTION

Buildings are one of the most basic elements in an urban layout and of great importance for variety of urban analysis and planning activities. Most socio-economic activities in urban areas are linked to building in one way or the other. These activities define the functional characteristics of buildings; which along with its form (e.g. size, shape, layout) defines the characteristics of a certain area. For example, the structure of buildings in a residential block is different than those in an urban commercial centre. By understanding building types and their associated function, one can have a base to launch complex spatial analysis to understand human to build environment interactions. In more densely populated countries such as the United Kingdom (UK), the identification of different types of urban areas (such as more compact, high-density dwellings, and mixed-uses) can help achieve sustainable policy development and resource management (Bramley and Power 2009).

In the neighbourhood model presented by Patricios (2002) there are three fundamental physical elements which shape urban land-use patterns at the atomic level: (a) buildings and their related open spaces, (b) plots or lots, and (c) streets (Levy 1999; Vanderhaegen and Canters 2010). Varied land-use patterns are formed because of different configurations of these three elements. Analyzing these patterns can help understand the underlying soci-economic activities that are shaping these patterns. This also opens up a possibility to understand urban land use patterns at different scales corresponding to a city's hierarchy.

Social scientists and policy makers, use various statistical measures based on dwelling types. For example, the dwelling classification based on house prices data from the Land Registry in the UK. Burdett et al. (2004) presented a model for measuring area homogeneity based on housing type. In the UK's 2001 census, dwelling type was used as a measure of homogeneity in constructing Output Areas (OAs) (Martin, Nolan et al. 2001). A significant interest in building -level and -type information can also be seen in those government departments that make policies regarding the sustainable development of cities (CLG 2011).

There is no national repository for building-type information for the UK. The Valuation Office Agency (VOA) maintained building type information until 1993 (Orford and Radcliffe 2007). The VOA principally collected building type information for calculating tax and property rates rather than for land-use inventory. Currently this information is not collected because of changes to the VOA's mandate. Another comprehensive source of such information is the UK's decennial census. There are, however, some critical issues related to the census information. First, it is only collected every ten years and, it becomes outdate quickly due to the continuous land use changes. Second, no mechanism is available for cross-validating the information provided in the census forms by individuals. Third, because of the privacy regulations, the individual building level information is not in the public domain, rather it is aggregated to the Output Area (OA) level and above geography units by the Office of National Statistics (ONS). This makes it impossible to link building information to the spatial counterpart (building footprints) available from other sources such as OS MasterMap®.

* Dr. DongMei Chen
Email: chendm@queensu.ca
Tel: +1-613-5336045; Fax: +1-613-5336122

The research work carried out at the UK Government level has anticipated the potential of using large scale digital mapping for the purpose of urban analysis, characterising the built up environment, and especially developing National Land Use Database (NLUD) (Harrison 2000; Tompkinson, Morton et al. 2004). The recommendations included the formulization of automated or semi-automated methodology to classify the buildings types (Harrison 2002; Tompkinson, Morton et al. 2004; Wyatt 2004 ; ODPM 2006.).

The objective in this paper is to examine the possibility of using large scale digital framework the Ordnance Survey (OS) MasterMap® (OSMM), created and maintained by Ordnance Survey (OS), the national mapping agency for Great Britain, to model building types in the UK. A classification framework is presented to infer functional types of buildings based on the analysis of their form and spatial relationship. Although the modelling process presented here uses a specific digital data, the principle of classification developed here can be translated and/or modified to other urban areas where similar large scale cadastral datasets are available.

2. STUDY AREA AND DATA

Manchester metropolitan is selected as a test bench for this study. Manchester is situated in the northwest of England (geographical coordinates: 53° 30' 0" North, 2° 13' 0" West) and covers an area of around 116 km^2, with a population density of 4,313 people per km^2 (ONS 2012) shown in Figure 1.

Figure 1. Study area – Manchester Metropolitan UK

2.1 Data related to buildings

The OS captures and manages topographic features, both natural (e.g. rivers, forests, grass land) and man-made (e.g. buildings, roads), in a seamless and multi-layered geospatial database at large scale, typically 1:1250 in urban areas and 1:2500 in rural areas (Holland and Allen 2001; OS 2006; OS 2010). There are four separate data layers in the OSMM given in Table 1.

The features in the OSMM are georeferenced to the OS National Grid Reference (NGR) and have a unique identifier, known as Topographic Identifier (TOID, a 16 digit code). The topographic features are arranged into nine different themes one of which is 'buildings'. Themes are although not part of the formal feature classification, make it possible to select similar features. However, OSMM does not provide information on building types such as tenement, terraced, or semi-detached or detached buildings. This missing functional class information

limits the OMSS's utility. Methods for automatically enriching such databases are needed to be developed (Lüscher, Weibel et al. 2008). The building features for this study were extracted from the topographic layer.

Data Layer	Description
Topography layer	Polygonized representation of more than over 425 million geographic (surface) features.
Integrated Transport Network layer (ITN)	A detailed overview of GB's transport infrastructure. Routable link and node transportation network. ITN consists of the road network and road restriction information (RRI) themes
Address Layer	Approximately 28 million residential and commercial properties addresses (including coordinates).
Imagery layer	National colour aerial photography at 25cm resolution

Table 1. Layers in OS MasterMap®

Address Layer 2 (AL2) is address-based data maintained in OSMM, which it claims to be among the most comprehensive inventory of both residential and commercial addresses in the UK. The addresses in AL2 are represented as point features geo-referenced to NGR with a resolution of less than 1 meter (in most cases). AL2 has evolved from previous products at OS which were initially developed from the Royal Mail's postcode address file (PAF). However, Martin and Higgs (1997) and Smith and Crooks (2010) pointed out that PAF is not designed to cover the complex hierarchy of mail delivery to addresses such as flats inside a building that do not have unique delivery points and hence may be missing in the OS address database. In some cases a delivery point may be referenced to as a PO Box, a building under a railway arches, a temporary building or a houseboat which are not part of the OSMM framework (OS 2010). In this regard the positional metadata information of AL2 provides valuable and accurate (including positional) information for an address (OS 2011).

Two layers from OSMM, Topographic Layer and Address Layer 2 (Al2) were acquired for this study. Topographic layer is used to extract building features and addresses from AL2 are used.

3. METHODOLOGY

The UK has different building-type classifications including: the National Land Use Database (NLUD) classification schema, ONS's dwelling-type, Building Regulations 2000 classification, and the one by UK Census (CLG 2010; SI 2010). A hierarchy of urban residential house types[1] was earlier presented by Jones and Larkham (1991). Also, buildings have different definition in the UK such as dwelling types, accommodation types, house types, or property types, which include "maisonette", "semi-detached" and "terraced" (Orford and Radcliffe 2007; Orford 2010; Hussain, Barr et al. 2012).

The objective of this paper is not to implement the full building classification schema given by NLUD or summarized by Orford and Radcliffe (2007). Nevertheless, it may be possible later to target the development of such a detailed classification system

---

[1] A detailed breakdown of different building types can be seen in Orford and Radcliffe (2007).

by integrating the one developed here with information from other sources. The methodology aims at using information contained in the topography and address layers and infer building classification for every building in the Manchester.

A three tier building classification schema was, hence, inferred for buildings given in Table 2. At the first level, buildings are classified into five classes purely based on their cartometric (morphological) properties and spatial relationships. The second and third levels of buildings classification (based on functions) are achieved by first linking address-based data sets to buildings feature and then by analyzing them.

Name	Class	Description
Level 1	Detached	Detached dwelling
	Semi-detached	Semi-detached dwelling
	End-Terrace	End -Terraced dwelling
	MID-Terrace	Mid -Terraced dwelling
	Complex Buildings	With much larger and complex structure to be considered such as offices, plazas, flats or offices.
Level 2	Residential	Solely residential usage
	Commercial	Solely commercial usage
	Mixed	Mixed residential and commercial usage
	Other	Not classified – missing flags
Level 3	1	One dwelling only
	2	Two dwellings
	3	Three or more dwellings

Table 2. A three levels classification hierarchy for buildings

3.1 Method for 1st level buildings classification

Pattern recognition based building classification was presented Earlier (Hussain, Davies et al. 2007; Hussain 2008; Hussain, Barr et al. 2012). It discussed using building morphological properties (i.e. area, perimeter, orientation) and topological spatial relationships (i.e. Adjacency, Neighbourhood) for classification. The concepts from cartometry were used to extract various quantitative variables from the map data (Maling 1989).

The overall process flow for the level building classification for level 1 is given in in Figure 2.
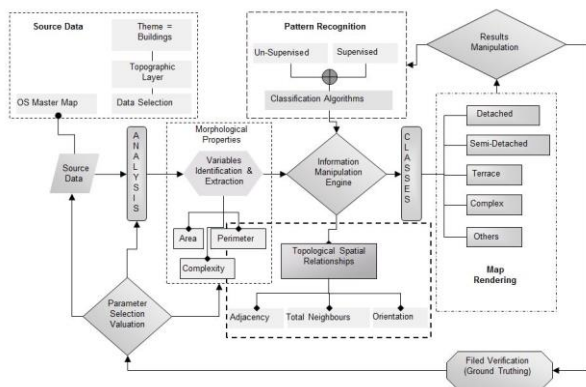

Figure 2. Level 1 Buildings classification model

Various variables extracted for each building is given in Table 3. Both unsupervised and supervised machine learning classification algorithms were considered. Initial exploratory analysis favoured using supervised classification algorithms. Two supervised decision tree algorithms (C5 and CART) were considered. Based on extensive evaluation of the accuracy produced by both the algorithms, it was found the C5 outperformed CART and hence was selected for building classification.

Variable	Variable
Building's area	Building's perimeter
Sides of a building	Orientation of a building
Elongation of a building	Concavity of a building
Building circularity	[2]Adjacency (see Table 4)
Total buildings in a local group of buildings	Areas of buildings in a local group
Average building size in a local Group	Smallest building's area within local group
Largest building's area within local group	Standard deviation of building area within local group

Table 3. Cartometric variables for buildings.

Case	Buildings adjacency count	Buildings in local group
Case A	= 0; no side is adjacent	One – Itself
Case B	= 1; Only one side is adjacent	Two – Itself and Second one
Case C	=> 1 ; One and More than One sides Adjacent	> 2 – Itself and other in the group

Table 4. Three cases of 'building adjacency' and 'buildings in a group'

3.2 Method for 2nd and 3rd Level buildings classification

The level 1 building classification results are used to extend the classification to next two levels. The OS AL2 contains entries for both the residential and commercial addresses. By linking addresses to buildings, it is possible to identify building type based on the type of associated address. OS has derived the functional information from VOA CTax and NNDR data sets for each addresses (corresponding to a building) in AL2 by using (OS 2007). There is, however, no classification flag to distinguish between residential, commercial and mix use address (buildings). Also no classification flag to distinguish a building based on the structural form, such as whether it is a detached, semi-detached or terraced building is available. Therefore, using OS AL2 in its current state it is difficult to: (a) identify addresses which correspond to residential, commercial or mix use buildings, and (b) identify addresses which correspond to detached, semi-detached or terraced buildings.

One possible solution is to infer these categories by analysing and combining information from classification fields, in AL2 database, describing different functional uses of corresponding buildings. The addresses in AL2 contain 'base function' information (for example dwelling, offices, and churches) and 'NNDR functional code' describing various non-residential uses of the corresponding building. It will make it possible to group buildings into three classes; (a) residential, (b) commercial, and (c) mixed use, by following the above described criteria.

---

[2] It is number of adjacent buildings to a building. Similar to 'meet' criteria explained by Egenhofer, M. J. (1994).

The building classification is then extended to 3rd level by analysing the count of addresses for each corresponding building. The number of addresses to each building potentially represents number of residences and/or businesses inside that building. The address to building relationship is also useful to classify buildings by understanding sub-building level uses. Buildings such as flats or multi-occupancy dwellings or offices can be identified by analysing the structure and context of and address. For example a residential building divided into flats will usually have more than one residential address associated with it and in some cases they are referred to as 'flat' in the text of the address. This involves the textual pattern analysis of the contents of an address.

4. RESULTS AND DISCUSSIOM

4.1 Level 1 buildings classification

The building features were extracted from OSMM and several pre-processing steps were taken to clean the data and isolate valid residential and commercial buildings. There is a set of 'sandwich' building polygons in topographic layer, which we refer to as non-trivial small buildings. Example of such building polygon is an entrance porch or a garage when two main buildings are connected at the upper stories or by a continuous roof line. Procedures were developed to handle these and much smaller buildings (non-residential) such as independent garages and sheds (Hussain 2008; Hussain, Barr et al. 2012). The number of buildings before and after the treatment is given in Table 5.

With non-trivial buildings	Count	Percentage
Valid buildings	170594	77.29
Small non-residential buildings	50134	22.71
Non-trivial buildings	6147	2.78
Without non-trivial buildings	Count	Percentage
Valid buildings	176741	80.07
Small non-residential buildings	43987	19.93

Table 5. Summary of the building distribution after processing

The C5 decision tree (DT) classification algorithm, belonging to the supervised-learning family, was applied for buildings classification. Classification rules were created by using the algorithms on the various variables calculated for building features. The level 1 buildings classification results are given in Table 6. The largest group of buildings (44.72%) in Manchester is semi-detached houses followed by terraced housing, 41.68% (end-terraced plus mid-terraced buildings). The smallest group, 5.72%, of residential buildings is detached houses; Group 5 (7.88%) comprises of complex buildings. A classification sample map is sown in Figure 3.

Class Name	Frequency	% age
Detached Buildings	9751	5.72
Semi Detached Buildings	76294	44.72
End Terrace Buildings	21327	12.50
Mid Terrace Buildings	49777	29.18
Complex Buildings	13445	7.88

Table 6. The summary of building classification based on selected decision tree model.

Figure 3. Level 1 Buildings classification map

4.2 Level 2 and 3 buildings classification

The 2nd and 3rd level building classification was created by linking and analysing AL2 addresses to corresponding buildings. The address analysis was based contents, context and address count per building. The AL2 contains entries for both the domestic and non-domestic addresses. The buildings which corresponded to the domestic addresses in AL2 were classified as "residential". Whereas the buildings corresponding to non-domestic addresses in AL2 were classified as "commercial" or "non-residential". In the cases when a single building acquired both the residential and non-residential addresses, that building was classified as "mixed use". Sub-building classification was carried out by understanding the address to building relationships. The addresses in AL2 contain various 'use codes' for the corresponding buildings and were used to identify and analyses commercial properties. The "Base Function" field in AL2 data was used to identify different address-types. The classification process consisted of following steps:

- Removal of unnecessary addresses
- Classification of residential and non-residential addresses
- Linking addresses to buildings
- Building classification – based on addresses

There are some non-geographic addresses in AL2 such as mail boxes in shops and offices [Post Office Box, (PO Box)] which do not link to any building seed. Also, AL2 provide positional flags (i.e. 'matched', 'unmatched', 'matched with discrepancy') which describe the accuracy of the position of an address on the NGR. There is another positional information flag which describe either the location of an address is final or it is provisional. This address accuracy information was used to acquire a clean cut of accurate addresses.

The addresses were categorized into residential and commercial classes by using this criterion discussed in section 3.2. The NNDR codes for all non-residential addresses were analysed and addresses were grouped into seven classes given in Table 7.

Class	Description
C	Commercial
E	Educational, Training and Cultural
I	Industrial
L	Leisure
M	Miscellaneous
N	Non Formula Assessed Public and Other Utilities
T	Treasury (Crown)
Empty	Not Dwellings - VO NDR Code Missing

Table 7. A summary for classification of addresses in AL2

The address to building analysis (point in polygon) resulted in a lookup table which allowed to link the results back to building polygons using TOIDs. A large percentage (94.4%) of addresses did not have NNDR code. Only 2.73% of all addresses got one of the commercial flags. Comparatively, a small percentage (5.59%) of non-residential addresses was found in the results. The addresses which got a NNDR code were used to classify corresponding buildings as non-residential. The seven NNDR classes were not used to classify corresponding buildings into those seven classes.

The classified addresses were used to classify corresponding buildings. The results are shown in Table 8. The addresses which did not link to any building feature were not considered for second and third level building classification and were labelled as 'others'.

Case A – Address to Buildings	%age
Number of residential address inside buildings	99.96
Number of residential address outside buildings	00.04
Number of non-residential address inside buildings	98.79
Number of non-residential address outside buildings	01.21
Case B – Buildings to Addresses	**%age**
Number of buildings contained addresses (both residential and non-residential)	93.50
Number of buildings which did not contain addresses (both residential and non-Residential)	06.50

Table 9. Results of addresses and building matching for AL2

It can be seen from the results that a large percentage of addresses were linked to buildings. From residential addresses 99.96% and from non-residential addresses 98.79% addresses were linked to buildings. Similarly, large percentage (93.50%) of buildings got addresses linked to them, however only 6.50% buildings did not get any address and hence did not get 2nd level classification. The buildings with residential addresses were classified as residential and buildings with non-residential addresses were classified as commercial. The cases where a building was found having both the residential and commercial flags attached to it was classified as "mixed uses". The buildings which did not get any address, residential or non-residential, were classified as "Other". The 2nd level building classification result is given in Table 9.

Class Level: Second	Percentage
Residential	88.76
Commercial	04.04
Mixed Use	00.70
Others	06.50

Table 9: Results for second level building classification

It was possible to create first level classification for all the buildings in the study area. A large proportion 93.50% of buildings got second level classification. However 6.50% buildings were unable to get second level classification. This is because of the reason that no address information was found for these buildings.

There can be 1:1 (one to one) and/or n: 1 (many to one) linking between addresses and buildings. Also, the structure of addresses can also be different. A simplest case is when one address correspond to one building and it has only building and street level information. A more complex situation is when more than one addresses are linked to a single building the address structure is different. In such cases address may have sub-building level (inside a building) information.

The third level building classification was created by interpreting the address structure and number of addresses linked to each building. In many cases, it was found that addresses with sub-building level information contained some key words such as "Flat", "Court", "Apartment", "Room", "Floor", "Suite" or 'Caravan Site'. These key words were used to select and classify corresponding buildings to as "Flats and Apartments" and 'Caravan Site'. The addresses were aggregated to building level and number of addresses for corresponding building was calculated.

Buildings were classified based on three configurations of addresses: the cases where there is only one address, the cases where there are two addresses and cases where there are three or more addresses attached to buildings. A subset (as the table was too big) of the results based on 1st, 2nd and 3rd level are shown in Table 10.

1st Level	2nd Level	3rd Level	%age
MID-Terrace	Residential	1 Dwelling	31.98
Semi-detached	Residential	1 Dwelling	30.72
End-Terrace	Residential	1 Dwelling	14.87
Detached	Residential	1 Dwelling	3.25
Semi-detached	Residential	2 Dwellings	1.23
MID-Terrace	Commercial	1 Dwelling	0.89
Complex Buildings	Residential	Flats and Apartments: 3 or More Dwellings	0.69
Detached	Commercial	1 Dwelling	0.17
Complex Buildings	Residential	2 Dwellings	0.16
Detached	Residential	3 or More Dwellings	0.12
Complex Buildings	Mixed Use	2 Dwellings	0.12
--------	--------	--------	--------
--------	--------	--------	--------

Table 10: Results for 1st, 2nd and 3rd level building classification

Thematic maps were produced using GIS to render results (an example is Figure 4). The accuracy of the classification was evaluated by ground truth exercise. Samples for various sites in the town were selected and classification results were manually cross checked. A comprehensive statistical accuracy assessment was out of the scope. The filed survey indicated the accuracy to be more that 92% and most of the errors were tracked back to errors in data. There were data issues such as missing addresses in AL2 and missing building or wrongly classified buildings in the topographic layer. The errors related to the classification algorithms were found to be comparatively fewer. Example of site used for field survey is given in Figure 5.

Figure 4. An example for Level 3 building classification

Figure 5. An example for filed verification

5. CONCLUSION

The objective in this paper was to evaluate the potential of using structured large scale topographic and address data for the UK to infer building classification. There are several land use classification schema exist in the UK. Beside the long history of interest in the land information and surveys, there is unfortunately no single data set in the public domain providing building type and use information in the UK. A classification encompassing built type and the functional use can provide a base to various socio-economic analyses and also can help sustainable management in densely populated countries.

A three level building classification schema was inferred and methodologies designed to achieve the classification. The results showed the potential of using existing large scale and the address based data to achieve the desired building classification. The methods used allowed exploiting the cartometric properties and spatial relationships to infer the building type classification at level 1. Also, the integration of address based data to building enabled achieving level 2 and 3 classifications.

A limited ground survey was carried out to validate the results. However, a comprehensive statistical analysis using existing classified data would help better evaluate the results and methodology. The evaluation of the results indicated that most of the times the errors were related to the errors in the source datasets. The completeness of both the building and address database is a benchmark for the percentage of the accuracy. In the current methodology it was not possible to cross check the completeness of both the products by OS. Other address database such as The National Land & Property Gazetteer (NLPG) in the UK can be used to verify the accuracy and completeness of AL2. The topographic layer also does not provide vertical (height) information regarding buildings. One

solution would be to use the 3D data from LiDAR or introduce attribute based measurements of premise size from sources such as VOA surveys.

The framework was tested in Manchester, UK, however, it is recommended that it be re-run and tested in other areas of the UK, using the same sources of data, to evaluate its generality and accuracy. While the model is currently based on specific UK data sets, all the concepts can be applied to any cadastral data which include building footprints and geo-referenced attribute data. Such applications could either be area specific, where particular data sets are available for the whole of the area, or could be adapted to allow land use classes to be fully generalized between countries.

REFERENCES

Bramley, G. and S. Power (2009). "Urban form and social sustainability: the role of density and housing type." Environment and Planning B: Planning and Design **36**(1): 30-48.

Burdett, R., T. Travers, et al. (2004). Density and Urban Neighbourhoods in London. London, Minerva plc**:** 82.

CLG (2010). English Housing Survey: Housing stock report 2008. London, Communities and Local Government**:** 174.

CLG (2011). Planning Policy Statement 3 (PPS3): Housing. London, Communities and Local Government**:** 29.

Harrison, A. (2002). Extending the dimensionality of OS MasterMap™: land use and land cover. the agi conference at GIS.

Harrison, A. R. (2000). The National Land Use Database: developing a framework for spatial referencing and classification of land use features. AGI Conference at GIS, Olympia.

Holland, D. A. and L. E. Allen (2001). "The Digital National Framework and Digital Photogrammetry at Ordnance Survey." Photogrammetric Record **17**(98): 291 - 301.

Hussain, M. (2008). Automated Urban Land Use Classification and Change Monitoring. Geography. Manchester, The University of Manchester. **Doctoral**.

Hussain, M., R. Barr, et al. (2012). Building-based urban land use classification from vector databases in Manchester, UK. The 20th International Conference on Geoinformatics (GEOINFORMATICS), Hong Kong.

Hussain, M., C. Davies, et al. (2007). Classifying Buildings Automatically: A Methodology. GISRUK 2007, National University of Ireland Maynooth, NUI Maynooth, Ireland.

Jones, A. and P. J. Larkham (1991). Glossary of Urban Form, Historical Geography Research Group.

Levy, A. (1999). "Urban morphology and the problem of the modern urban fabric: some questions for research." Urban Morphology **3.2**.

Lüscher, P., R. Weibel, et al. (2008). Where is the Terraced House? On the Use of Ontologies for Recognition of Urban Concepts in Cartographic Databases Headway in Spatial Data Handling. A. Ruas and C. Gold, Springer Berlin Heidelberg**:** 449-466.

Maling, D. H. (1989). Measurements from maps: principles and methods of cartometry. Oxford, England ; New York, Pergamon Press.

Martin, D. and G. Higgs (1997). "Population georeferencing in England and Wales: basic spatial units reconsidered." Environment and Planning A **29**(2): 333-347.

Martin, D., A. Nolan, et al. (2001). "The application of zone-design methodology in the 2001 UK Census." Environment and Planning A **33**(11): 1949-1962.

ODPM (2006.). National Land Use Database: Land Use and Land Cover Classification. DCLG.

ONS (2012). Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2010, Office for National Statistics UK.

Orford, S. (2010). "Towards a data-rich infrastructure for housing-market research: deriving floor-area estimates for individual properties from secondary data sources." Environment and Planning B: Planning and Design **37**(2): 248-264.

Orford, S. and J. Radcliffe (2007). "Modelling UK residential dwelling types using OS Mastermap data: A comparison to the 2001 census." Computers, Environment and Urban Systems **31**(2): 206-227.

OS (2006). The OS MasterMap® Part 1: user guide. Southampton, United Kingdom, Ordnance Survey. **v6.1.1:** 205.

OS (2007). "OS MasterMap Address Layer 2: Full Release." Retrieved September, 2007, 2007, from http://www.ordnancesurvey.co.uk/oswebsite/products/osmaster map/layers/addresslayer2/index.html.

OS (2010). ADDRESS-POINT User guide and technical specification. Southampton UK, Ordnance Survey. **v5.2:** 36.

OS (2010). OS MasterMap Topography Layer User guide and technical specification. Southampton, UK, Ordnance Survey. **v1.9:** 142.

OS (2011). OS MasterMap Address Layer and Address Layer 2: User guide Southampton, Ordnance Survey. **v1.4:** 62.

Patricios, N. N. (2002). "Urban design principles of the original neighbourhood concepts." Urban Morphology **6**(1): 21-32.

SI (2010). Building and buildings, England and Wales: The Building Regulations 2010 H. M. s. S. O. a. Q. s. P. o. A. o. Parliament. London, The Stationery Office Limited**:** 56.

Smith, D. A. and A. T. Crooks (2010). From Buildings to Cities: Techniques for the Multi-Scale Analysis of Urban Form and Function London, CASA University College London**:** 64.

Tompkinson, W., D. Morton, et al. (2004). An optimised semi-automated methodology for populating a. national land use dataset. GISRUK. University of East Anglia.

Vanderhaegen, S. and F. Canters (2010). Using Remote Sensing Data for Improving the Distinction Between Distinct Types of Urban Land Use and Form Using Spatial Metrics. 30th EARSeL Symposium: Remote Sensing for Science, Education and Natural and Cultural Heritage, Paris, EARSeL.

Wyatt, P. (2004). "Constructing a land-use data set from public domain information in England." Planning Practice and Research **19**(2): 147 - 171.