

Modelling residential living space for individual properties using digital infrastructure and remote sensing data: a Cardiff case study

Scott Orford¹, Jonathan Radcliffe²

¹School of City and Regional Planning, University of Cardiff
Glamorgan Building
King Edward VII Avenue
Cardiff, Wales. CF10 3WA

Tel. +44 (0)29 20875272 Fax +44 (0)29 20874845

orfords@cardiff.ac.uk

<http://www.cf.ac.uk/cplan/contactsandpeople/stafflist/m-r/orford-scott-dr-overview.html>

radcliffej@cardiff.ac.uk

<http://www.cf.ac.uk/cplan/contactsandpeople/stafflist/m-r/radcliffe-jonathan-mr-overview.html>

KEYWORDS: Housing, Living space, OS Mastermap, LiDAR, urban modelling

1. Introduction

Planning is a key element to the success of sustainable urban environments and successful planning relies on access to good quality information. In strategic and spatial planning, this information is often derived from models and simulations of urban systems and hence there is a real need (and demand) for well-founded, data-rich environments in which to develop and test these models (Batty, 1995). With respect to models and simulation of housing markets, British cities are hindered by a lack of detailed structural attribute data at the level of the individual property. A key issue is the lack of data on living space as this is the single most important structural attribute in determining house price (Orford, 2000) and is the most important variable in any housing market model. The British Valuation Office Agency (VOA) holds the most extensive set of property attribute data in the UK, but the vast majority of these data are held on paper records (VOA, 2002) making them both difficult and costly to maintain and access. Researchers and policy makers often have to rely on a sample of properties for their models, typically without data on living space, and this can lead to problematic results – for example, the issues associated with using beacon properties in council tax valuations or the uncertain and frequently contradictory forecasts of the state of the national housing market.

This lack of property level structural attribute data may be resolvable, however. The past decade has seen the continuing development of geographic information handling technologies and the emergence of rich new sources of both GIS and remote sensing data. Commentators such as Longley (2002; 2003) have begun to examine how these developments are improving not only the digital data infrastructure of many cities but also the opportunities for undertaking urban modelling and urban analysis. The research agenda in urban GIS is now increasingly dominated by the development of integrated databases, methods for their analysis and models relevant to fine-scale urban geographies. Traditional sources of ancillary information such as the Census of Population and address and population registers are becoming enhanced by their integration with these data sets. The precision of these data and the frequency with which they are updated makes them increasingly useful as a source of micro-scale structure of the built environment in the absence of more detailed data (e.g. Longley and Mesev, 2000). This paper explores the opportunities provided by some of these new data sources for deriving property level structural attribute data. In particular, the paper

presents research on methods for deriving modelled estimates of living space for individual properties within a small area of Cardiff, Wales and analyses the success of these estimates for a sample of properties. The paper outlines how these estimates can be used in further research to improve housing market analysis.

2. Integrating digital infrastructure data with LiDAR data

Previous research by Orford and Radcliffe (2007) demonstrated how property level structural attribute data, in this case dwelling type, can be derived from Ordnance Survey's (OS) digital framework data product, OS Mastermap. It also highlighted errors in the modelling process that could be rectified by the integration of remotely sensed data relating to building height (Rottensteiner, 2003). This research on dwelling type has since been developed in order to provide a methodology for deriving modelled estimates of residential living space. Four principal data sources are used:

- OS Mastermap for Cardiff: topographic and address layers
- Environment Agency (EA) LiDAR data
- Estate agent data for a sample of properties in the study area
- 2001 Census data for Output Areas

In order to estimate living space within a particular property, it is necessary to have data on the building outline and the number of storeys in order to calculate an approximate measure of floor area. It is also important to capture any internal variation in building height across the building outline in order to estimate living space as accurately as possible. For instance, if a single storey extension on a property is treated as a two storey structure, then living space will be over-estimated. The research covers a small area in Cardiff that has been selected based on the findings in Orford and Radcliffe (2007). The OS Mastermap data provides the individual addresses of all the properties in this area, classified into residential and non-residential types, and also provides the building outlines for these addresses. The EA LiDAR data provides height data at a one metre resolution for the area— see Figure 1 which shows LiDAR and OS Mastermap data part of the study area.

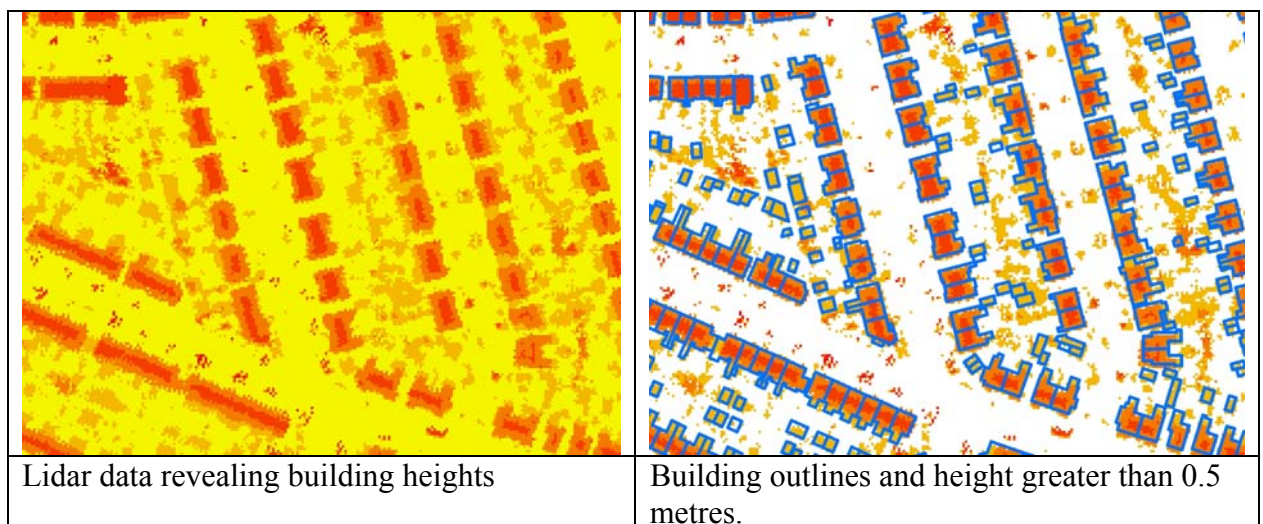


Figure 1. EA LiDAR and OS Mastermap data for part of the study area

In order to capture internal variations in building height within individual properties, the OS Mastermap building outline was sub-divided into rectilinear polygons using a custom written GIS script. Building heights were then calculated for each polygon within a building outline

using a weighted average of the LiDAR data and then this was converted into a measure of the number of storeys – see Figure 2. In some cases, the LiDAR data also identified extensions that had not been captured by the OS Mastermap building outlines. In such cases the building outlines were re-drawn to include these extra structures.

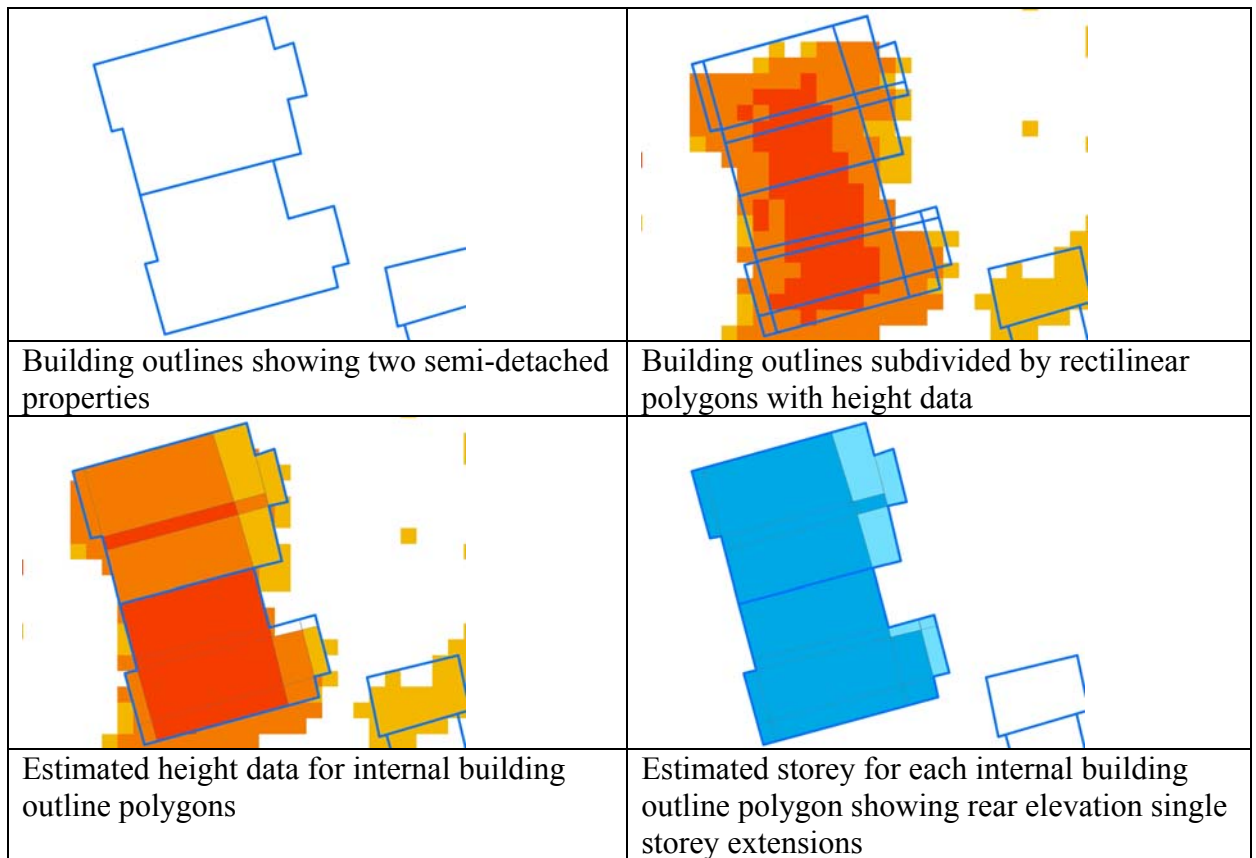


Figure 2. Calculating building height and number of stories for internal polygons within building outlines

3. Deriving estimates of residential living space

In order to derive estimates of living space based on building outlines and height data, Estate Agent data for a sample of properties in the study area was integrated into the GIS. These data included information on dwelling type, number of storeys and the total floor area of rooms. A regression model was estimated for the Estate Agent sample of properties which specified living space (as measured by floor area) as a function of structural attributes and a measure of location (see Equation 1):

$$LS = f(DT, NS, ABO, LN) \quad (1)$$

LS = Living space as measured by floor area (Estate Agent)

DT = Dwelling type (Estate Agent)

NS = Number of stories (Estate Agent)

ABO = Area of building outline (OS Mastermap)

LN = Local neighbourhood (OS Mastermap/Census)

The model was then used to estimate living space (LS) for every single property in the study area using the measures of DT, NS, ABO and LN derived from OS Mastermap and EA LiDAR data. These modelled living space estimates were validated against floor area calculations from a sample of Estate Agent properties which had not been included in the calibration of the regression model. By categorising the types of properties and locations in which modelled living space estimates were similar to those calculated using Estate Agent floor area data and those in which there were substantial differences, the utility of the methodology in estimating living space in different urban contexts is explored. These include problems with the digital framework data such as the building outlines being incorrect (e.g. missing extensions); EA LiDAR data not being sensitive enough to capture the internal variation in building heights within a property and both data sources unable to capture particular types of internal spaces such as attics, basements, cellars, integral garages and certain layouts that have an important effect upon living space. There are also issues with the accuracy of Estate Agent data and the measures of floor area and living space. However, the research suggests that these problems have particular spatial configurations within the built environment and the paper provides some means of resolving them.

4. Conclusion

In this paper we presented a method for deriving living space estimates for a sample of individual properties in Cardiff. By doing so, we have moved towards developing a data-rich environment for modelling micro-scale structure in the built environment. The aim of future research is to scale up the study area to the whole of Cardiff. With the integration of sales price data from the Land Registry, the GIS will provide a laboratory for the development and testing of housing market theories through models and computer simulations of the Cardiff housing market.

References

- Batty M (1995) Planning support systems and the new logic of computation, *Regional Development Dialogue* **16(1)** pp1-17
- Longley PA (2003) Geographic Information Systems: developments in socio-economic data infrastructures. *Progress in Human Geography* **27** pp114-121
- Longley PA (2002) Geographical Information Systems: will developments in urban remote sensing and GIS lead to 'better' urban geography? *Progress in Human Geography* **26** pp231–239
- Longley PA & Mesev V (2000) On the measurement and generalisation of urban form, *Environment and Planning A* **32(3)** pp473-488
- Orford S & Radcliffe J (2007) Modelling UK residential dwelling types using OS Mastermap data: a comparison with the 2001 census, *Computers, Environment and Urban Systems* **31(2)** pp206-227
- Orford S (2000) Modelling spatial structures in local housing market dynamics: A multi-level perspective. *Urban Studies* **13(9)** pp1643-1671
- Rottensteiner, F. (2003) Automatic generation of high-quality building models from Lidar data, *IEEE Computer Graphics and Applications* **23(6)** pp 42-50
- VOA (2002) *CTR(E) IA 111202 Data quantity / quality audit – property attributes data*, at <http://www.voa.gov.uk/publications/CouncilTaxIas/021211-ctre-ia.htm>

Acknowledgements

This research has been supported by the Royal Geographical Society (with The Institute of British Geographers) with an EPSRC research grant. The authors would like to thank the Ordnance Survey and the Environment Agency for providing OS Mastermap and LiDAR data under license.

Biography

Dr Scott Orford is a lecture in GIS and Spatial Analysis in the School of City and Regional Planning, Cardiff University. His research interests include housing market analysis and modelling socio-economic processes.

Mr Jonathan Radcliffe obtained an MSc GIS with distinction from University of Glamorgan. He is currently a research associate for the Wales Rural Observatory at Cardiff. He has recently started a part time PhD investigating the use of visualization tools in planning and policy in Wales.