

Engineering Modelling of Building Energy Consumption in Cities: Identifying Key Variables and Their Interactions with the Built Environment

Javier Urquizo[1,3(✉)], Carlos Calderón[1], and Philip James[2]

[1] School of Architecture Planning and Landscape, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
carlos.calderon@newcastle.ac.uk
[2] School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
philip.james@newcastle.ac.uk
[3] Escuela Superior Politécnica del Litoral, ESPOL, FIEC, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador
jurquizo@espol.edu.ec
http://www.espol.edu.ec, https://www.ncl.ac.uk

Abstract. This paper focuses on a spatial domestic energy framework of sub-city areas in the United Kingdom using Newcastle upon Tyne as a case study. The framework estimates the energy end-use at the single dwelling level on three aggregate scales: district, neighbourhood and community. The framework uses two aggregation approaches, one using an approximated prototype-cluster of similar dwellings at the district scale and the other using a novel, sub-city DEM modelling of building and its micro-cohesive energy structures in neighbourhoods and communities. The validation strategy compare the modelled gas and electricity values in three representative districts against the DECC values in two aggregate hierarchical areas for electricity and gas: the DECC Medium Layer Super Output Area (MLSOA), and the Lower Layer Super Output Area (LLSOA). Our work discusses this framework in key areas: the availability of data, the number of surveyed variables and data processing methods for filling blanks in order to have an individual dwelling's complete energy Standard Assessment Procedure (SAP) profile, methods to provide evidence for finding patterns residing in the data set, methods for estimating the annual composite (gas and electricity) energy consumption for individual and aggregated dwellings, the validation strategy, the uncertainties associated in the data and model, and scaling and replication in other cities.

Keywords: Spatial interpolation · Record imputation · Distributed target scenarios · Domestic energy consumption · Structure detection · Urban energy modelling

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11624, pp. 3–25, 2019.
https://doi.org/10.1007/978-3-030-24311-1_1

1 Introduction

Swan and Ugursal [1, p. 1828] state that "the bottom-up engineering approach is the only method that can fully develop the energy consumption of the sector without any historical energy consumption information" and that "these techniques have the capability of determining the impact of new technologies" [1, p. 1833] which, in turn, makes the method ideally suited for an area-based delivery approach favoured by local authorities (LAs). This paper discusses a particular spatial bottom-up area-based energy framework, the Newcastle Energy Model Framework (NCRF) [2], using two modelling approaches: clusters in districts and individual Domestic Energy Modelling (DEM) in neighbourhoods and communities. Because of assumptions in the nature of physics-based energy modelling, it is inevitable perhaps that there will be significant discrepancies between energy estimates and actual published UK government consumption figures. However, the energy estimates may still provide useful insights, as they are largely based on the physical parameters of the dwellings. In our research, the results are encouraging in some ways as there is a degree of spatial consistency across the results.

Our work represents a framework to attempt to estimate spatial domestic energy consumption in individual dwellings and aggregated areas using a spatially enabled database to represent local area characteristics. We argue that, as the scale decreases, the local area characteristics are more important. Model estimates in the sub-city DEM are improved in neighbourhoods by taking into consideration more variables. However, assessing the individual dwelling's energy consumption in the sub-city DEM approach requires more than a simple counting of the individual dwellings. Specifically, what must be examined is the role of local structures like micro-cohesive energy supply structures (district and group heating), dwelling in electricity time-of-use tariff (in the UK the Economy 7 tariff E7 i.e. pay normal prices during the day, but cheaper rates for seven hours during the night) and planned permission granted for converting houses to Houses in Multiple Occupations (HMO) and houses and flats having sharing amenities, as a number of researchers have suggested, for instance, [3] and [4]. The reason for having an increased number of data set variables is that this may lead to a better insight into the effect of uncertainties in the aggregated process [5] by generating a more complete representation of the dwelling. In the case of dwellings using the Economy 7 tariff, an increase number of variables is needed to provide a better description of the dwelling (the hot water heating system and the tank insulation are the new variables).

Using the Unique Property Reference Number (UPRN) code of an individual dwelling, it is likely to link secondary sources, either by Topographic Identification (TOID), and/or address to produce a better representation of individual dwelling predictors, instead or adopting archetypes to simplify the complexity of housing stock, e.g. two archetypes in the [6] physical model, 1,000 archetypes in the [7] BREHOMES model, 47 archetypes [8] in the CDEM model, which are difficult to relate to the specific address of a dwelling.

2 Overview of the Newcastle Energy Framework Methodology NCRF

This section reviews the roadmap in terms of data flow, process and milestones to produce estimates of energy in a single dwelling and aggregates in sub-city areas. NCRF [9] is a spatially referenced parameterised per-dwelling domestic energy framework developed with the purpose of estimating the energy consumption in sub-city areas. The design of the framework is sufficiently flexible to scale-up in continuous geographic areas allowing energy estimates (electricity and gas). The roadmap described in Fig. 1 defines four processes that are completed sequentially: the spatially-referenced per-dwelling integration across multiple secondary data sets; the model refinement; the spatially-referenced per-dwelling and sub-city energy aggregates; and the validation strategy.

The per-dwelling integration process uses spatial interpolation and multiple imputation (see Fig. 1) procedures to replace an incomplete, missing, duplicated and inaccurate local 2009 convenience secondary data sources, the Newcastle Carbon Route Model (NCRM) [10] (see Fig. 2) with a set of plausible new data values. The new data values represent valid inferences that properly reflect the uncertainty due to missing values. At the end of this process, a complete set of valid records of the case study areas are ready for refinement.

The model refinement process provides precise data in cohesive structures by identifying at least one neighbour with a known value that can be used as a template for predicting the values in the whole urban structure. The refinement is for two cohesive structures: the first is those that affect the energy aggregation processes like district heating, group heating and Economy 7; and the second is those that provide a physical context to a cohesive community through planning regulations, in particular the change of use for dwelling houses due to the legislation that gives councils the freedom to choose areas where landlords must submit a planning application to rent their properties to unrelated tenants (houses in multiple occupation HMO). At the end of this process, a complete set of records is ready for input in a Building Research Environment Domestic Energy Model (BREDEM). Another refinement assigns a better estimate to floor area (one of the key energy drivers) to houses that are in shared occupancy.

The energy estimation process is the core process, as it estimates the annual electricity and gas consumption using three scales. The first is for individual dwellings (then groups of repeated dwellings to estimate the annual heating gas consumption of multiple repeated per-property types). The second and third are of annual electricity and gas consumption estimates in two levels of detail (sub-city aggregates): Neighbourhoods (Lower Layer Super Output Areas—LLSOAs) and Districts (Medium Layer Super Output Areas—MLSOAs). Because the DECC statistics display aggregates of annual electricity consumption of dwellings using the Economy 7 (E7) tariff, we also aggregate dwellings with households using E7.

In the final process, the per-dwelling property type and sub-city energy estimates (electricity, gas or E7) were validated (compared) against the 2009 Department of Energy and Climate Change (DECC) estimates in three different scales:

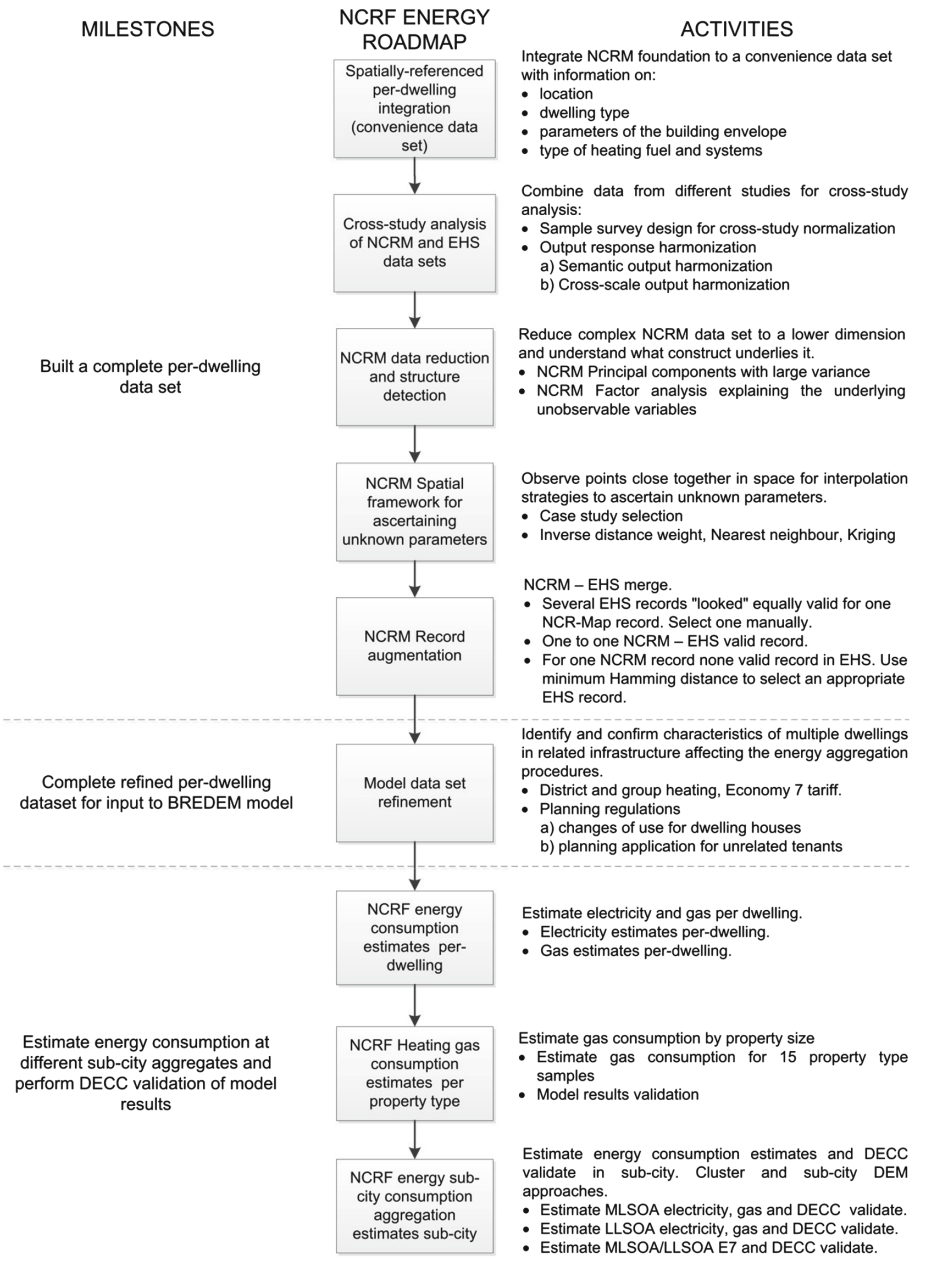

Fig. 1. NCRF energy estimates roadmap

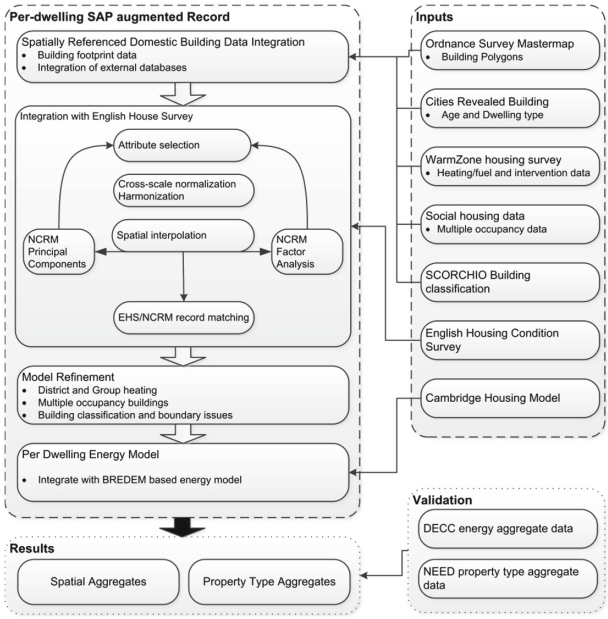

Fig. 2. NCRF data and processes

first the per-dwelling-type heating gas estimates set up by DECC with support from the Energy Saving Trust (EST) and energy suppliers for the National Energy Efficiency Data Framework (NEED) project [11]; and two aggregates of electricity, gas and E7 in the districts (MLSOA) and neighbourhoods (LLSOA) levels.

In the MLSOAs, the cluster model uses six dwelling archetype prototypes in every cluster to capture the mean effects in the interaction between the physical variables as the individual effect will cancel out at this scale. In every cluster prototype there are assumed regional weather characteristics and standard occupancy values. The cluster centre (mean) is the prototype of the cluster. In NCRF we want the centre (mean) in each cluster to be one particular dwelling. This is the reason why the six medoids were utilised to replace the six mean values in each cluster. The cluster energy method estimates the energy in every cluster by multiplying the Cambridge Housing Model (CHM) energy estimate [12] of the full SAP cluster prototype by the number of dwellings in every cluster. The process is repeated for every cluster in the MLSOA. In the LLSOA instead, the sub-city DEM model method uses city records to improve the data quality in the detailed model. The sub-city DEM uses secondary data sets from Arm's Length Management Organizations (ALMO) to gather rented and leased social housing characteristics and specific data on energy systems. The sub-city DEM also utilise the NCRM Registered Social Landlords housing information and HMO licensing from the local authority (LA), and shared housing data from Housing

Associations and the LA, to understand dwellings in group heating schemes, dwellings with an Economy 7 tariff, and the number and characteristics of residential dwellings sharing district heat schemes. The interpolated data are compared with accurate detailed city information, and in the case of discrepancy city records correct the interpolated values, i.e. the detailed model data set is refined. The sub-city DEM process finally aggregates individual dwelling energy estimation in the LLSOA and MLSOA areas.

3 Discussion

The discussion section is organised as follows. Section 3.1 discusses the cluster model against DECC statistics in MLSOA districts and Sect. 3.2 the sub-city DEM model against the DECC statistics in LLSOA neighbourhoods in the three selected case studies.

3.1 The Cluster Modelling Approach

The Newcastle Carbon Route Model (NCRM) data set components are generally available in some form to many, if not all, local authorities. The use of common values for UPRN and TOID (individual and aggregated) has made possible the use of spatial data from different resolutions and their combination with other types of spatial information.

The main NCRM data sets are the Warm Zone Programme and the English Housing Survey (EHS). Warm Zone is a major initiative designed to systematically reduce fuel poverty on a local area basis in cities, and the English Housing Survey is a national initiative for providing accurate information on the type and condition of housing and the people living there. The spatial visualization of NCRM WarmZone records and their further integration with other data sets is achieved through two spatial elements: the national address Gazetteer unique property identifier, the UPRN (the dwelling code) and the Ordnance Survey (OS) MasterMap^{TM} unique topographic identifier, the TOID (the building code). All data sets used in our work (WarmZone, EHS, Gazetteer and Ordnance Survey) are generally available in some form to many if not all local authorities.

The modelling approach of the cluster model is top-down. The top-down method involves patterning dwellings with similar energy profiles in the city, and then disaggregating the patterns [13] based on the MLSOA's proportions. Our work has shown that six key surveyed variables (dwelling size, dwelling age, building form, wall construction, heating and number of floors) perform well in explaining the energy consumption of aggregated dwellings in districts, three of the six variables were composites (building form, wall construction and heating). One of the issues highlighted in integrating these spatial data sets is the definition of what constitutes the floor area. The NCRM dwellings are assigned their building foot print area [10] instead of an individual useable floor area which is required for CHM (and other BREDEM models). Therefore, the very first activity is to find the usable floor area of individual dwellings through

computation. In addition, the problem is further exacerbated where we have complex building with Homes in Multiple Occupations (HMOs) or individual rented rooms and mixed-use (commercial and residential) buildings.

Warm Zone Surveys do not have a value of usable floor area recorded for individual dwellings in complex buildings. Our work has provided a solution involving secondary sources that is time consuming and not ideal. It would be better if there were a straightforward method of calculating usable floor area or data were collected in a different manner. Usable floor area is a key variable in understanding energy consumption but, especially for complex or mixed-use buildings, a solution to estimate usable floor area is not obvious. In our work, a painstaking process using a variety of secondary data (described in Sect. 2) was developed. However, more robust solutions need to be developed to facilitate this type of modelling. These could be extensions of the NLPG/LLPG, energy audits or other means of better capturing the detailed usage of buildings from an energy perspective. An interesting development is the use of 3D structures that have both internal and external levels of detail providing a ready means to associate usage with spaces, such as the system developed in Berlin using a CityGML. The CityGML uses an application schema for the Geographic Markup Language (GML). Model for energy diagnostic and urban energy policy support [14,15]. In the UK, the Ordnance Survey is also investigating the addition of 3D products within their portfolio of survey products [19]. In 2014, Ordnance Survey began to release building height attribution with their 2D topography data [16] and additionally it has been enriching the 3D models with generic annotation e.g. from images of buildings [17].

Principal component analysis underpinning the cluster model (see Sect. 2) was used to provide both a method for explaining how many of the surveyed categorical variables are needed to explain the total variance within the reduced data set, and to detect the correlations between the variables. Besides these more obvious uses, our work has shown that there is an additional benefit as the principal component also provides group estimates of EHS 'donor' records before the NCRM nearest neighbour multiple imputations method actually selects the best 'donor' in an augmentation strategy. This shows that the principal component analysis is a worthwhile process; however, it remains to be seen whether the principal component results will be transferable to other cities because regional climates make for different insulation and heating supply requirements. However, other local authorities should consider this type of approach to collate data from existing sources with the requirement that there has to be a large sample so its application will be statistically valid.

One observation from the NCRM data set is that some of the variables are easy to collect, for example building types, because these are easy to see and also it is easy to add building data from other data sets such as the Ordnance Survey or buy it from commercial vendors such as Cities Revealed. However, integrating a number of sources from multiple collections throws up a number of semantic and cross-scale output harmonization issues but on the other hand also provides the ability to cross-check. An illustration of this in our case was the errors in the

Building Age in the NCRM Cities Revealed data set which were easily identified and rectified by NCRM YHN build dates (see Sect. 2). Building age is a difficult variable as it requires expert knowledge to age an individual building correctly. Other variables are difficult because they are hidden, such as wall insulation. Missing data was a real problem because for example 17% of the people surveyed in Warm Zone did not know the insulation type, especially in cavity walls. In these cases, the survey generates 'don't knows' and we decided to disregard the whole record. This will have implications when trying to replicate NCRM in other cities, because physically hidden variables are difficult for surveyors to identify, even to the point that there could be significant bias in the sample because the 'don't knows' are relevant, e.g. in rented properties. One way to avoid this problem is by having a much larger sample than normal, as is the case with NCRM.

Our work has shown that clustering provides evidence to find patterns residing in the NCRM data through partitioning of the six dimensional data vector into subsets based on the closeness or similarity among the data of the same kind. Another innovative approach is the use of the two-step sequential cluster to identify the cluster structure in the NCRM complete data set of observations (hierarchical cluster analysis) and the optimal cluster method to assign observations (k-means cluster analysis), i.e. both the hierarchical and the k-means cluster analysis techniques were used sequentially. Our work has shown that this hybrid approach gave good results, i.e. first, the Ward's method (hierarchical cluster) was used to obtain the possible number of clusters (eight for the NCRM data set) and the way they merge, as seen from the dendrogram, and then the k-means clustering was run with only the chosen optimum number (eight) of clusters in which to place all the observations.

The cluster method shows sensible energy estimate results both citywide and when disaggregating the cluster results over the MLSOAs. This approach, however, requires reasonably good data, in our case even though the NCRM WarmZone data have some variables that are difficult to collect; the sample size is citywide and in each MLSOA, is largely in excess of the minimum and preferable size. Formann recommends a minimum sample size of at least 2^m (where m=6 equals the number of clustering variables) and preferably $5 * 2^m$ [18, p. 4], i.e. in this study between 64 and 320 observations. If the minimum is not reached, the medoid shows weakness to outliers, but that is not the case for the Warm Zone Survey in Newcastle. Also, the high coverage guarantees that if the six variables were to be inputted in a different way (order) they would still produce the same cluster results. This shows the potential of integrating different data sets in energy modelling. However, access to this critical data set was only made possible by signing a non-disclosure agreement as the NCRM WarmZone data are commercially sensitive and potentially valuable. Because this type of data is so important to energy studies, it should be considered part of the open data strategy of the [20] so that other research studies can utilise the data easily.

Also critical in the open data strategy is a definitive single address register not only for local authorities but for the Royal Mail, Ordnance Survey and

DECC. Local authorities are working through the Improvement and Development Agency (I&DeA) for local government to establish the Local Government Information House (LGIH) by forming the GeoPlace Limited Liability Partnership (LLP) to produce a single definitive National Address Gazetteer (NAG). Government is also looking at options to ensure that the Postcode Address File (produced by the Royal Mail and regulated by Ofcom) continue to be a part of the single definitive address (the NAG).

However, the NEED framework data sets are combined using the NLPG unique property reference number (UPRN) as the integrating code in a twofold process. First, the address data from each of the data sets included in NEED is used to assign a UPRN to each record within that data set and then the UPRN is used to link records from one data set to the corresponding record in each of the other data sets. In Newcastle, as in other cities, the address matching in mixed-use buildings is more complicated for two reasons: firstly, the mixed-use building is less likely to be primarily addressed by a number, as is the case with domestic homes, and, secondly, several businesses may occupy the same buildings and it is difficult to identify the domestic dwellings arrangement within the building. Additionally, address information in HEED (which is part of NEED) did not allow a reliable match for individual flats within a building. One possible solution for NEED is replacing the NLPG UPRN with the AddressBase^{TM} UPRN; this should allow better analysis of the impact of measures for flats and increase confidence in the address information. AddressBase^{TM} is actually three data sets from GeoPlace and the Ordnance Survey: Royal Mail Postcode Address File (PAF), NLPG and OS AddressLayer^{TM} with further information from the Valuation Office. A final comment is that all these data sets are well trusted but an open access policy needs to be effective.

The energy model used to estimate the energy end-use of every cluster prototype was the Cambridge Housing Model (CHM). CHM is a national model requiring a full SAP input (115 variables). This study has shown that by combining data from different geographies and sources it is possible to augment limited NCRM WarmZone data with EHS data to create a full SAP input for use in energy models (see Sect. 2). This approach saves a great deal of time and also repurposes the NCRM WarmZone data, which were difficult and expensive to collect, for energy modelling. Moreover, validating (with respect to DECC consumption values) the modelled data at the MLSOA level needs full coverage of the whole district, but cluster data did not have a full coverage in the MLSOAs and did not have the complete SAP input records. To fulfil these requirements, NCRF has proven that record generation (for missing records to achieve complete coverage) and augmentation (for missing fields in every record) algorithms lead to the desired full coverage SAP input in the MLSOAs (see Sect. 2). Spatial interpolation algorithms were used as a record generation strategy and a nearest neighbour multiple imputations algorithm was used as a record augmentation strategy both with a reasonable success given that performing a physical survey for a full SAP data on each and every dwelling in Newcastle is not possible. Spatial interpolation and multiple imputation algorithms are not straightforward

implementations because we have to consider the practical problems associated with cohesive structures e.g. shared houses, communal heating structures and district heating as well as regional weather variations, local built types (e.g. the Tyneside flat) and non-standard occupancy [21]. The sources for information about cohesive structures are available in a limited selection of useful websites containing local sources (e.g. HMO licensing) but they require patience and time to link the data to the building databases.

However, urban spaces are not only two dimensional areas; they also have the third dimension that defines the volumetric characteristic of the buildings, i.e. there is a need for a three dimensional (volumetric) patterns to define the physical urban space. [22] argues that the built environment can be seen as a setting for human activities and the volumetric patterns of the residential areas are in three different densities, low, medium and high rise.

NCRF has shown that the volumetric patterns and initial coverage of the survey data could be used to assign an appropriate point-based method for spatial interpolation in the three selected case study areas. For a cluster with low density building patterns, the Inverse Distance Weighting (IDW) deterministic method is commonly used, whereas for a compact building distribution with high coverage the nearest-neighbour (NN) is appropriate. For complex high rise structures of different building and uses-classes, a stochastic method is a sensible choice. The IDW algorithm was applied in Castle because it shows a cluster distribution; the NN was applied to South Heaton because it shows a compact distribution. Westgate is complex because the building classes (including tower block buildings) are diverse in the area; this made appropriate the use of the Kriging stochastic method. Kriging not only considers the distance between surrounding points but also the spatial autocorrelation (dwellings being in the same building) among the observations. In summary, the volumetric patterns and initial coverage of the survey data simple analyses could be used to inform the best strategy for filling-in the gaps, although further work would be needed to fully understand the boundary conditions. Although this seems to be a promising approach to completing the gaps, the link between urban building density patterns and spatial interpolation needs to be fully investigated. However, understanding urban form in this way has the potential to inform LAs and the like about the areas where there is a need to improve the data collection strategy.

To cope with the fields missing in the NCRM record (to obtain a full SAP record); NCRF has shown that an imputed method is an algorithm for record augmentation (see Sect. 2). In our augmentation strategy the two data sets are from different geographies. NCRM cluster prototype (the acceptor) is a local data set and EHS (the donor) is a national data set. Predicted values were then generated for all NCRM cluster prototype records using the following two-step procedure: first, 'donor' records were sorted by the 'acceptor' common fusion variables (CFV) (see Sect. 2), and then the missing values were given by the 'donor' record nearest to the acceptor. Imputed values were used only after establishing the matches, as this preserves the variability in the data as well as the correlations across the NCRM survey [23].

The fact that in the record augmentation strategy (see Sect. 2 for details of the strategy), the 'perfect outcome', there is a single best suitable EHS record with all CFV being equal in both, the NCRM and the EHS data sets—leading to a single EHS 'donor'—is only possible on 11% in South Heaton and 1% in Westgate suggests that one possible solution is to apply the nearest neighbour imputation as a method for handling missing values to acquire a full SAP record input to the energy model. To improve the search effectiveness and reduce the computing time, nearest neighbour multiple imputations require a previous pre-filter stage in which the principal component analysis gives the appropriate (reduced) number of EHS records. This two-stage process suggests that missing values of NCRM were estimated even though EHS covers only 0.073% of the English housing stock and the Warm Zone programme covers approximately 63.69% of Newcastle residential housing stock.

The reader should also consider that these processes can be misleading and cause the user to assume that the NCRM data has more data observations than actually surveyed and may magnify the confidence in the energy estimation results. Realistically, the goal of spatial interpolation, multiple imputation and aggregation strategy are initially to provide completed SAP records for areas where there are DECC statistics (i.e. MLSOA and LLSOA) and to make the validation process possible. Interpolation and imputation are techniques that produce values respectively within the range of NCRM and EHS data values and are not meant to recreate individual missing values by optimal point prediction.

The three case study selection is adequate because they cover a wide range of volumetric patterns in the city. Furthermore, Castle and South Heaton show high concentrations (above city average) of certain energy clusters, making these districts more homogeneous. Westgate, having close to average and low concentration of energy clusters is more diverse. From a policy perspective, it is reasonable to suggest that an increase in the accuracy of the data in places where there are complex mixed-use properties rather than homogeneous areas would be useful to provide better inputs to energy models, and this also has implications for reducing the cost of data collection.

The cluster energy method estimates the energy in every cluster by multiplying the CHM energy estimate of the full SAP cluster prototype by the number of dwellings in every cluster. The process is repeated for every cluster in the MLSOA. The aggregated cluster result suggests that the cluster size and composition not only reflect the energy efficiency of the Newcastle stock, but what was encouraging, the potential impact of applying certain retrofitting measures is possible. What was good from the cluster model district results is that they enable us to model aggregate energy consumption using a reduced number of variables.

NCRF has shown that clustering (or classification of houses into groups) represents a measure of homogeneity in energy consumption. The areas with high concentration of specific energy clusters show a close approximation with the DECC values, whereas areas having close to average and low concentration of energy clusters show a more distant approximation with DECC values. Castle

and South Heaton have concentrations of particular clusters which show close approximation to DECC values of 5% and 3%, respectively. The Westgate result shows less close approximation to DECC value (27%). The Westgate result may suggest that the micro-level energy structures in Westgate, planned permission granted for converting houses to HMOs and houses and flats sharing amenities in neighbourhoods play an important role in the overall energy consumption in the case study areas, particularly in complex areas, which has implications for better data collection methods.

The scaling of the cluster model to a city-level is not problematic as the clusters were produced at the city-level, and the number of record guarantees robustness to outliers in all MLSOAs. The main challenge to replicate the cluster model in other cities is bringing the information together in a way that it will be useful for meaningful analysis. Currently, the Warm Zone programme throughout England is grouped into larger area hubs, covering Hull, London/South East, Midlands, North East, and Nottingham/the East Midlands. For other regions, similar surveys have to be implemented to collect the physical characteristics of dwellings.

However, when the cluster model is downscaled to LLSOA (neighbourhood) level the limits of this approach start to appear. Some homogenous areas still showed reasonable comparison with the DECC figures (e.g. Castle LLSOAs 8308 and 8311) but there was little consistency with other values being widely off (e.g. Westgate LLSOAs 8394 and 8399). In Westgate, the increased local building, socio-economic and physical characteristics of the urban area play a more important part in the estimation of dwelling energy consumption, and that is the potential reason for having increased differences.

In the policy arena, the government is leading the spatial information publication and reuse of location-specific data through the [24] UK Location Programme. The programme ensures UK compliance with the EU INSPIRE Directive, of the 900+ data sets published, around 200 are mandated by INSPIRE; however, data quality is still an issue and so is harmonization and scale. NCRF has proven that spatial enabling sub-city scale energy modelling is a potential way forward, and within this scope, the emphasis of the Location Data Programme should be reconsidered and replaced with a more rigorous analysis in which the building is the unit of analysis. If that is possible, then all the methods described in this section can be replicated in every city in the UK to compute district level energy aggregates.

3.2 Sub-city DEM Modelling Approach in Neighbourhoods

The main motivation of creating a sub-city DEM model that allows the estimation of individual dwellings annual energy consumption is that modelling the physical processes in individual dwellings with standard occupancy has the potential to provide LAs with realistic, simulated estimated energy values for buildings that can be aggregated at many scales to provide a baseline for use in retrofit or micro-generation or other energy related planning. The initial variables of the sub-city DEM energy profile are: usable floor area, dwelling type,

construction date, number of floors above ground, predominant type of wall structure, cavity wall insulation, main heating fuel, primary heating system, boiler group and tenure. As with the cluster model, the usable floor area of all individual dwellings in the sub-city DEM model were corrected from the initial building footprint area value using secondary sources in the same three-step process detailed earlier in Sect. 3.1.

NCRF uses record generation algorithms with the sub-city DEM NCRM complete ten-variable records data set in the three case study areas (Castle, South Heaton and Westgate) to obtain complete coverage in the corresponding MLSOAs areas. As with the cluster model, the record generation algorithms prove to be a useful tool to obtain full coverage in geographic areas. The Inverse Distance Weighting (IDW) algorithm was used in Castle because dwellings show a cluster distribution, while the nearest neighbour (NN) was used in South Heaton because dwellings show a uniform distribution. Stochastic Kriging was used in Westgate because it has one of the most diverse collections of building classes (including tower blocks buildings) in the Newcastle area.

NCRF has shown that city records can be used to improve the data quality in the sub-city DEM model because of several reasons, first local authorities have extra information to help with the validation process, as an example, YHN council homes have an accurate build date taken from the deeds. This study uses secondary data sets from Arm's Length Management Organizations (ALMO) to gather rented and leased social housing characteristics and specific data on energy systems. The research utilised the NCRM Registered Social Landlords housing information and HMO licensing from the LA, and shared housing data from Housing Associations and the LA, to understand dwellings in group heating schemes, dwellings with an Economy 7 tariff, and the number and characteristics of residential dwellings sharing district heat schemes. The interpolated data are compared with accurate detailed city information, and in the case of discrepancy city records correct the interpolated values, i.e. the sub-city DEM model data set is refined.

Additional procedural refinement is needed when computing the residential consumption load-share in mixed residential/commercial district heating schemes, because is difficult to model and disaggregate each component in the mix. Mixed residential/commercial load sharing in existing district heating data is also problematic because of different time-of-use. There is no individual residential meter (so it is problematic to model individual behaviour and practices) and no charge for the actual domestic energy use and residential energy prices are subsidised. The need for understanding and modelling the load share in shared district heating is becoming increasingly important. The current model for the Newcastle district heating is to focus initially on connecting existing district heating, larger commercial loads, and new commercial and residential developments, and then to extend the network into the existing domestic sector [25, p. 33]. Connecting Byker's existing district heating is the most problematic task because it is by far the biggest (almost 2,000 households) and most complex district heating site in Newcastle. One possible solution is to introduce and manage

a transparent charging system and install controls and meters in homes [26]. Better models for Newcastle's future district heating will become a bigger issue because we need better methods to disaggregate and assign domestic energy consumption to individual residential and commercial loads because of different time-of-use; we also need to model individual household behaviour. One further area of procedural refinement concerns interpolating data close to the boundary. Spatial modellers should be aware of the limitations of interpolation boundaries and adjust results accordingly.

NCRF has shown that there are a higher number of data quality refinements in areas with low Warm Zone Survey penetration, complex urban energy supply structures, complex building types and a large number of non-self-contained household properties. Westgate and South Heaton required more adjustments due to the relatively low penetration of survey data in tower blocks, adjustment to the individual dwelling's usable floor area, a significant number of complex load shared energy structures, and rented on room-by-room properties. The refinement was made, respectively, in 69.46% and 24.53% of the complete ten-variable records for a complete coverage. Castle's simple building types (mostly semi-detached and detached houses) and high penetration survey required refinement in only 4% of the complete ten-variable records for a complete coverage. Urban complexity in terms of the built form and usage patterns potentially signpost data collection priority areas [27]. The sub-city DEM model uses the same augmentation algorithm that the cluster model uses. What is different is that the common fusion variables are individual (rather than composite), have increased to ten, and include tenure. The sub-city DEM model also uses the CHM to estimate the energy consumption in individual dwellings.

Tenure is an interesting variable in the case study areas because they show different composition. There are major contributions from owner occupied and council rented properties in Castle, private rented student accommodation, council rented properties and owner occupied in South Heaton, and social housing in Westgate. Private rented properties are inefficient; this perhaps increases the amount of households in fuel poverty [28]. Difficulties arise if the private tenant is poor, because due to the recent government reform [29], the cheapest 30% of private rental properties should be available for those in receipt of housing benefit (HB). The maximum payable HB is also limited to the maximum four-bedroom rate, and there is an age cap for singles living in a shared apartment; this will probably have the end result of encouraging private rented householders to behave in a 'cost-conscious' manner and move to cheaper accommodation if they cannot afford to meet the gap between their Local Housing Allowance (LHA) entitlement and the rent charged by the landlord [29]. However, social housing differs from housing in other tenures in many ways, as social housing occupants are more likely to live in areas of high unemployment [30, p. 8]; also, social tenants enjoy less space per person than other tenures, and because of living in small size dwellings they are more likely to be dissatisfied than others. Moreover, new build social housing is still disproportionately in the most deprived neighbourhoods [31]. However, YHN housing is efficient housing due

to the following four reasons: 70% of the houses meet and maintain the decent homes standard [32]; 45% of their housing has efficient heating system installed; 60% has double glazing; and 40% has insulation in loft and walls [33]. However, social tenants are more likely to be in fuel poverty because their income is low, even considering the Housing Benefit due to fuel price rises [34, p. 6], despite the small size and good condition of the social housing. One possible solution is to have a policy that monitors: the housing condition to all tenures, the impact of HB cuts in the residual incomes, their job opportunities (relative to the location) and perhaps supports and increases HB among older home-owners.

The NCRM RSL tenure has a significant share in Westgate. NCRM RSL tenure shows households in a non-standard occupancy and dwellings in shared-ownership. Shared-ownerships are difficult to model due to two reasons: they are both a rented and an owned property, and the household size is small, so occupancy may not be standard. This is indicative of the complexity of the energy modelling when trying to model at the level of the individual dwelling.

NCRF has shown that the validation of NCRF with respect to comparable DECC statistics is problematic. The validation cannot compare the number of meters because DECC statistics use aggregated consumption by Meter Point Administration Number (MPAN) to ensure energy consumption is accurately recorded, whereas in the case of YHN district heating schemes there are no individual heat meters. YHN argues economical and operational reasons for not having individual heat meters. The economic reasons are the cost of billing and renewal of meters and the operational reason is that district heating was originally designed to provide a constant and steady supply of hot water, then metering is likely to change people's use of the system and create high peaks in demand, which could make the heating system less reliable [35, p. 5]. [36, p. 198] argue that as for heating use, there is consistent evidence that metering in combination with consumption-dependent charging can act as an incentive for more sustainable heating behaviours in district heated dwellings, and [37] argue that indoor climate habits relate to experience with valves and meters gained at their workplace. The validation is also problematic because YHN and Newcastle City Council with other councils in the North East England purchase the fuel for district and group heating in bulk, and DECC statistics are on individual MPAN.

NCRF has shown that Economy 7(E7) (the off-peak electricity tariff) mostly uses electric storage heaters although there are also using warm air heating systems. If the heat in the dwelling and water is by electricity, but don't have storage heaters or a hot water tank, E7 probably will not be so cost effective. If the heat in the dwelling is with gas, E7 could still be a good fit, provided the bulk of the electricity use is at night. The E7 model needs two additional variables for the hot water in the energy profile: the hot water heating system and the tank insulation. E7 is used in a variety of tenures and building types. It is prevalent in owner occupied small area properties in the Castle area, while the city owned tower block in South Heaton uses storage heaters, and the housing association properties in Westgate use storage heaters. This also potentially has an impact

on fuel poverty policy and strategy given the amount of E7 heating existing in social housing tenures. This again is indicative of the complexity of the model and the difficulty of national data sets and models in properly addressing this type of fuel tariff. These three micro energy structures (HMOs, District Heating and E7 tariff) are difficult to model both from existing data and with current energy models. As an example, [26] argues that "approximately 2/3 of all tenants in the community heating schemes receive housing benefits. Currently housing benefits can be claimed for infrastructure and communal heating charges, but not for energy charges. Therefore any changes in energy charges have a profound effect on all tenants, while a modification in infrastructure or communal heating charges will only benefit or disadvantage tenants who do not receive housing benefits." Understanding their spatial distribution and potential impact is a first step to address shortcomings in current national energy models and demonstrates the importance of a detailed understanding of local physical characteristics of the building stock.

NCRF has shown that the sub-city DEM energy model allows estimation of energy consumption aggregates in a bottom-up approach. The aggregating strategy is as follows: by building, by block, by building classes and by LLSOAs; a further aggregation was made to the MLSOA level for comparison purposes with the cluster approach. Bottom-up is a sensible approach because it allows for ad-hoc energy retrofit planning scenarios and seeing patterns of base line energy consumption. NCRF are energy estimates using the physical properties of the building which would still be useful for retrofit evaluation or evidence based comparison of different strategies at an area level. It would even provide the means of implementing different energy efficient strategies with a great impact at the local level, rather than the more general approach made at citywide levels.

In summary, NCRF has integrated a number of data sets in a way that we were able to integrate with well-respected national level energy models such as the CHM, so every individual building has their own energy SAP profile. There are many uncertainties in extending the coverage and getting full SAP records to compare to DECC statistics in aggregated areas. However, this approach highlights an interesting way to move forward in energy modelling and it has set-out different ways of reducing the uncertainties using secondary sources. There are also issues in the way the data are being collected and restrictions on the distribution.

Given the uncertainties, it is not surprising perhaps that aggregated sub-city DEM model building totals are some way from DECC aggregates. Perhaps the biggest influence here is the lack of real occupancy data which becomes more important in a per-building analysis. This highlights a potential research avenue for a per-building approach to area energy modelling through the integration of better socio-economic data. In addition, the issues facing the city with respect to the physical characteristics of the building stock and current energy efficiency measures are highlighted through the detailed analysis, both spatial and aspatial, of individual buildings. Even the collation and integration of data sets without an

accompanying energy model has potential to provide insights into the make-up of a city's residential building stock.

NCRF (and therefore CHM) does not take into account the microclimate, however, at the scale of the individual buildings detailed models exist, such as EnergyPlus^{TM}. Urban microclimate effects on energy demand were analysed by [38], who used an urban microclimate model and the building energy software EnergyPlus^{TM} [39]. These models have to be supplied with suitable boundary conditions, which represent the urban microclimate. However, in order to consider interactions between energy demand and the urban microclimate, more complex tools are needed. The interactions between buildings and the landscape in low density Castle presumably create a real increase in energy consumption because there is an increase in the mean daily heat output from the heating system due to smaller increases in the outdoor air temperature due to heat island effects. One possible way to improve the weather data would be to spatially merge NCRM data with detailed weather data that is readily available. Urban microclimate is a key element during the design stages of sustainable and comfortable urban spaces, although the physics underlying the interaction of urban microclimate with buildings is complex to model.

The NCRF and DECC statistics show different aggregated counts in the same geographic areas because DECC statistics shows unallocated loads (where the address information within the Gemserv database is incomplete, invalid or missing); therefore, this underestimates the aggregated energy in neighbourhoods. However, it is unclear whether such incidental reductions (or implicit bias) of energy consumption in some neighbourhoods will last over a period of time or whether DECC is intentionally implementing strategies in a long-term goal to reduce this implicit bias.

The NCRF number of thermal zones is not well defined in several building types: rented on room-by-room accommodation, linear terraces and tower blocks. For example, a separate thermal zone for each heated room in rented on room-by-room properties would require information (e.g. which windows affect the space) that is difficult if not impossible to automatically extract, but important for estimating the internal temperatures for each room or space. The NCRF probably underestimates the energy consumption because a reduced number of thermal zones in areas of South Heaton with linear terraces, tower blocks in Westgate, HMOs and social housing shared accommodation. A way forward is to have statistics for space per person in the EPC and SAP. However, there are several practical and policy issues to address. First, HMOs with shared essential facilities have no individual tenancy agreement and there is no need to provide an EPC in the housing market in the UK. Secondly, the EPC methodology considers the number of habitable rooms but what constitutes a room is not well defined (the energy rating is adjusted for the floor area of a building). Meanwhile, [40, pp. 90–121] does consider the number of habitable rooms and the number heated habitable rooms, but as all SAP and BREDEM models do not consider occupancy in the heated rooms, this made it less possible to compute the space per person.

DECC underestimate the energy consumption in LLSOAs with district heating because it only considers individualized energy consumption and YHN purchases energy in blocks for households in District Heating. These bulk consumptions do not appear in individualized DECC statistics. More complete statistics on district heating networks are presumably about to be published in accordance with the Energy Efficiency [41, Article 11].

There is no clear cut-off for residential and small commercial premises so DECC underestimates the energy consumption due to the fact that the ground floor part of a building has, for example, been converted to a small shop. Domestic cut-off is based on arbitrary consumption figures used by industry; DECC uses the gas industry cut off threshold of 73,200 kWh. This means a number of smaller commercial/industrial consumers were wrongly allocated as domestic, i.e. DECC underestimates the energy consumption in this LLSOA. This is an issue in South Heaton; however, the use and integration of secondary sources into sub-city DEM energy modelling databases as demonstrated in this study does enable this level of fine-grained building-use modelling.

There is no definitive source for the number of households 'not on the gas grid.' The energy use of individual households in multi-tenanted commercial/residential buildings is difficult to estimate if there is no agreed measure or sub-metering system. With this type of building, it can also be difficult to build a clear picture of energy use in the building because landlords use energy to provide services in the shared parts of the building (e.g. lifts and pumps), plus the use in the tenanted areas. Additional issues are when some tenants pay for their own metered electricity supply or purchase other fuel directly. One possible solution for delivering sub-metering is through the incentives on the Carbon Reduction Commitment (CRC), announced in the Energy White Paper 2007.

NCRF does not consider the inter-building effect created by surrounding buildings. The energy consumption is underestimated in aggregated buildings because of the reduced solar radiance. However, this effect is complex and requires the modelling of a network of buildings. The South Heaton case study shows that the LLSOAs energy consumption may have modelling inaccuracies created by the nearby buildings. Physical building models would need detailed topological information (as provided by NCRM) in order to model inter-building effects effectively. For densifying urban environments, this is likely to be a relatively significant effect. Specifically, the role of inter-building effects must be examined as a number of researchers have suggested (e.g. [38,42,43]). Also, urban form from the point of view of environmental performance in cities as addressed in [44]; and, energy use and density as the results of [45] seem to suggest.

Weather correction methodologies in DECC/NEED and CHM lead to discrepancies, the main reason being that the exact methodology for weather correction for NEED/DECC is not fully disclosed. In order to be confident in the use of DECC/NEED statistics for validation purposes, clear and unambiguous information needs to be provided on weather corrections as these can have significant impact on estimated values.

In summary, there are some discrepancies in the sub-city DEM model results partly because of the uncertainties in the way national models model individual occupancy, but also because there are no individual consumption data to calibrate with the dwelling energy estimations. However, having this sub-city DEM model opens-up the development of low-carbon energy technologies at the community and neighbourhood levels through providing a physically-based energy consumption estimate at the building level that provides realistic energy consumption patterns and detailed spatial information based on the physical characteristics of the building and the outputs from a national energy model.

4 Conclusions

The aim of the research was fulfilled as the spatially enabled model and rich thematic database facilitates, first and foremost, integrated mapping practices, i.e. a NCRM data set accessible by a digital computer including a set of maps in the grid reference system used in Great Britain; a NCRF framework of varying and selectable scale from an individual dwelling to streets, communities, neighbourhoods and districts, and capable of presentation individual and aggregated domestic energy consumption on a computer display; tools for calculating, delineating, and displaying building dimensions in an orthographic projection; and means for coupling the computer to other computers for data exchange.

Previous modelling approaches adopted a number of house archetypes which together represent all dwellings in the stock. This study identifies individual dwellings' energy profiles to estimate individual and aggregated energy consumption. The cluster model uses six categorical variables (dwelling age, wall construction, building form, dwelling size, heating and number of storeys) to capture the mean effects in the interaction between the physical variables in districts as the individual effect will cancel out at this scale. The framework estimates in the sub-city DEM model are improved in neighbourhoods by taking into consideration ten variables (usable floor area, dwelling type, construction date, number of floors above ground, predominant type of wall structure, cavity wall insulation, main heating fuel, primary heating system, boiler group and tenure) and local micro-cohesive structures. The reason for the increased number of variables is to obtain a better insight into the effect of uncertainties of the aggregated process by generating a more complete representation of the dwelling.

Acknowledgement. We would also like to acknowledge the help we received from Newcastle City Council, who permitted me to use their data sets (especially Adrian McLoughlin). This proved to be unique and invaluable for this paper.

References

1. Swan, L.G., Ugursal, V.I.: Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renew. Sustain. Energy Rev. **13**(8), 1819–1835 (2009)
2. Urquizo, J., Calderon, C., James, P.: Understanding the complexities of domestic energy reductions in cities: integrating data sets generally available in the United Kingdom's local authorities. Cities **74**, 292–309 (2018a)
3. Weber, C., Keirstead, J., Samsatli, N., Shah, N., Fisk, D.: Trade-offs between layout of cities and design of district energy systems. In: Efficiency, Cost, Optimization, Simulation and Environmental Impact, ECOS 2010 (2010). https://workspace.imperial.ac.uk/urbanenergysystems/Public/ecos2010_weber.pdf
4. Cheng, V., Steemers, K.: Modelling domestic energy consumption at district scale: a tool to support national and local energy policies. Environ. Model. Softw. **26**(10), 1186–1198 (2011)
5. Urquizo, J., Calderon, C., James, P.: Using a local framework combining principal component regression and monte carlo simulation for uncertainty and sensitivity analysis of a domestic energy model in sub-city areas. Energies **10**(12), 1–22 (2017b)
6. Johnston, D.: A physically-based energy and carbon dioxide emission model of the UK housing stock. Ph.D. thesis, Leeds Metropolitan University (2003). http://www.leedsmet.ac.uk/as/cebe/team/johnston_thesis.pdf
7. Shorrock, L., Dunster, J., (1997). The physically-based model brehomes and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock. Energy Policy **25**(12), 1027–1037 (2003). http://www.sciencedirect.com/science/article/pii/S0301421597001304
8. Firth, S., Lomas, K., Wright, A.: Targeting household energy-efficiency measures using sensitivity analysis. Build. Res. Inf. **38**, 25–41 (2010)
9. Calderon, C., James, P., Urquizo, J., McLoughlin, A.: A GIS domestic building framework to estimate energy end-use demand in UK sub-city areas. Energy Build. **96**, 236–250 (2015). URL http://www.sciencedirect.com/science/article/pii/S0378778815002212
10. Calderon, C., James, P., Alderson, D., McLoughlin, A., Wagner, T.: Data availability and repeatability for urban carbon modelling: a CarbonRouteMap for Newcastle upon Tyne. In: Retrofit 2012, Manchester England (2012). http://www.energy.salford.ac.uk/page/day_2_papers
11. DECC: National energy efficient data need framework - report on the development of the data-framework and initial analysis. Technical report, Department of Energy and Climate Change (2011). http://www.cewales.org.uk/cew/wp-content/uploads/National-Energy-Efficiency-Data-Framework.pdf
12. Hughes, M.: A guide to the Cambridge housing model. Technical report, Department of Energy and Climate Change (2011). https://www.gov.uk/government/statistics/cambridge-housing-model-and-user-guide
13. Urquizo, J., Calderon, C., James, P.: Modelling household spatial energy intensity consumption patterns for building envelopes, heating systems and temperature controls in cities. Appl. Energy **226**, 670–681 (2018b)
14. Kolbe, T.: Representing and exchanging 3D city models with CityGML. In: Lee, J., Zlatanova, S. (eds.) 3D Geo-Information Sciences. LNGC, pp. 15–31. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-87395-2_2

15. Nouvel, R., Schulte, C., Eicker, U., Pietruschka, D., Coors, V.: Urban energy analysis based on 3D city model for national scale. In: International Building Performance Simulation Association IBPSA (2013). http://www.ibpsa.org/proceedings/BS2013/p_989.pdf
16. Sargent, I., Harding, J., Freeman, M., Holland, D.: The building blocks of user-focused 3D building data. In: Efficient Capturing of 3D Objects at a National Level: With A Focus on Buildings and Infrastructure, EUROSDR/ISPRS Workshop (2014). http://www.eurosdr.net/workshops/eurosdrisprs-workshop-efficient-capturing-3d-objects-national-level-focus-buildings-and
17. Jones, C., Rosin, P., Slate, J.: Semantic and geometric enrichment of 3D geo-spatial models with captioned photos and labelled illustrations. In: 25th International Conference on Computational Linguistics, Coling 2014, Workshop on Vision and Language (2014). https://vision.cs.bath.ac.uk/VL_2014/programme.html
18. Dolnicar, S.: A review of unquestioned standards in using cluster analysis for data-driven market segmentation. In: CD Conference Proceedings of the Australian and New Zealand Marketing Academy Conference 2002 (ANZMAC 2002). Deakin University, Melbourne (2002). http://ro.uow.edu.au/commpapers/273/
19. Capstick, D., Heathcote, G.: Moving towards 3D from a national mapping agency perspective. In: Abdul-Rahman, A., Zlatanova, S., Coors, V. (eds.) Innovations in 3D GeoInformation Systems. LNGC, pp. 491–500. Springer, Heidelberg (2006). https://doi.org/10.1007/978-3-540-36998-1_38
20. HM Government: Open Data White Paper: Unleashing the Potential (CM 8353). TSO@Blackwell and other accredited agents, London (2012). ISBN 9780101835329. https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential
21. Urquizo, J., Calderon, C., James, P.: 2016 IEEE Ecuador Technical Chapters Meeting (ETCM), 12–14 October 2016 (2016)
22. Rapoport, A.: CHAPTER 1 - Urban design as the organization of space, time, meaning and communication. In: Rapoport, A. (ed.) Human Aspects of Urban Form, Pergamon, pp. 8–47 (1977)
23. Rubin, D.B.: Multiple imputation after 18+ years. J. Am. Stat. Assoc. **91**, 473–489 (1996). http://www.tandfonline.com/doi/abs/10.1080/01621459.1996.10476908
24. DEFRA: Defra open data strategy. Technical report (PB13785), Department for Environment, Food and Rural Affairs, London (2012). https://www.gov.uk/government/publications/defra-open-data-strategy Directive 2012/27/EU, 2012. of the European Parliament and of the Council of 25 October 2012 on Energy Efficiency, Amending Directives 2009/125/ec and 2010/30/eu and repealing directives 2004/8/ec and 2006/32/ec (text with EEA relevance). Technical report (L 315/1-56), Off. J. Eur. Union Brussels. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:315:0001:0056:EN:PDF
25. NAREC: Newcastle City Council energy masterplan. Technical report, NAREC Distributed Energy, Blyth Northumberland (2012). https://www.newcastle.gov.uk/sites/drupalncc.newcastle.gov.uk/files/wwwfileroot/environment/energy/ncc_energy_masterplan_may_2012.pdf
26. French, C., Joyce, S., Gorton, I., Heppenstall, T.: Critical friend report on district & group heating. Technical report, Newcastle upon Tyne: YHN and IRES: Newcastle University, Newcastle upon Tyne (2007). http://www.yhn.org.uk/pdf/MainBoard22May07DistrictandGroupHeatingTLCItem11.pdf
27. Urquizo, J., Calderon, C., James, P.: Metrics of urban morphology and their impact on energy consumption: a case study in the United Kingdom. Energy Res. Soc. Sci. **32**, 196–212 (2017a)

28. DECC: Private rented sector energy efficiency regulations (domestic) (England and Wales): consultation on implementation of the energy act 2011 provision for energy efficiency regulation of the domestic private rented sector. Technical report (URN 14D/228), Department of Energy and Climate Change, London (2014). https://www.gov.uk/government/ploads/system/uploads/attachment_data/file/346767/Domestic_PRS_Regulations_Consultation_Draft_v1_6_No_tracks_final_version.pdf
29. DWP: The impact of recent reforms to local housing allowances: summary of key findings. Technical report, Department for Work and Pensions, London (2014). ISBN 978-1-910219-31-7. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/329902/rr874-lha-impact-of-recent-reforms-summary.pdf
30. Lupton, R., et al.: Growing up in social housing in Britain. A profile of four generations, 1946 to the present day. Technical report, Tenant Service Authority, London (2009). http://www.jrf.org.uk/sites/files/jrf/social-housing-britain-FULL.pdf
31. Hills, J.: Ends and means: the future roles of social housing in England (casereport 34). Technical report, The London School of Economics and Political Science, London (2007). ISSN 1465-3001. http://eprints.lse.ac.uk/5568/1/Ends_and_Means_The_future_roles_of_social_housing_in_England_1.pdf
32. DCLG: A decent home: definition and guidance for implementation. Technical report (06HC03962), Department for Communities and Local Government, London (2006). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/7812/138355.pdf
33. Gallagher, I., Langhorne, D.: Environmental sustainability strategy. Technical report, Your Homes Newcastle, Newcastle Upon Tyne (2011). http://www.yhn.org.uk/pdf/EnvironmentalSustainabilityStrategy.pdf
34. Palmer, G., MacInnes, T., Kenway, P.: Cold and poor: an analysis of the link between fuel poverty and low income. Technical report, New Policy Institute, London (2008). ISBN 978-1-902080-24-6. http://www.poverty.org.uk/reports/fuelpoverty.pdf
35. YHN: District and group heating - a guide for tenants. Technical report, Your Homes Newcastle, Newcastle upon Tyne (2014). http://www.yhn.org.uk/pdf/Districtandgroupheatingtenants.PDF
36. Morgenstern, P., Lowe, R., Chiu, L.: Heat metering: sociotechnical challenges in district-heated social housing. Build. Res. Inf. **43**(2), 197–209 (2015)
37. Gram-Hanssen, K.: Residential heat comfort practices: understanding users. Build. Res. Inf. **38**(2), 175–186 (2010)
38. Yang, X., Zhao, L., Bruse, M., Meng, Q.: An integrated simulation method for building energy performance assessment in urban environments. Energy Build. **54**, 243–251 (2012)
39. Crawley, D., et al.: EnergyPlus: new, capable, and linked. J. Arch. Plan. Res. **21**(4), 292–302 (2004)
40. BRE: The governments standard assessment procedure for energy rating of dwellings. Technical report, Building Research Establishment (2009). http://www.bre.co.uk/filelibrary/SAP/2009/SAP-2009_9-90.pdf
41. Directive 2012/27/EU (2012) of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC (Text with EEA relevance) (L 315/1-56). Brussels: OJ. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:315:0001:0056:EN:PDF. Accessed 12 Aug 2014

42. Pisello, A., Taylor, J., Xu, X., Cotana, F.: Inter-building effect: simulating the impact of a network of buildings on the accuracy of building energy performance predictions. Build. Environ. **58**, 37–45 (2012)
43. Bueno, B., Norford, L., Pigeon, G., Britter, R.: A resistance capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate. Build. Environ. **54**, 116–125 (2012)
44. Adolphe, L.: A simplified model of urban morphology: application to an analysis of the environmental performance of cities. Environ. Plan. B: Plan. Des. **28**(2), 183–200 (2001)
45. Steemers, K.: Energy and the city: density, buildings and transport. Energy Build. **35**(1), 3–14 (2003)