Identifying local areas with higher than expected domestic gas use

Introduction

This article presents analysis undertaken in 2010 that has been carried out to improve the understanding of energy efficiency. The aim of this analysis was to produce a statistical model to predict expected gas use at a local area level using publicly available data. Modelled gas consumption data could then be compared with actual data produced by DECC at local levels (www.decc.gov.uk/en/content/cms/statistics/energy_stats/regional/electricity/mlsoa_llsoa/mlsoa_llsoa.aspx). Where actual gas use is higher than predicted gas use by more than, say, 10 per cent then this might imply a lower level of efficiency in homes (of course there could be other reasons why gas consumption may be higher) and hence may help direct efficiency investment in that area.

Methodology

The neighbourhood statistics[1] (NeSS) database is maintained by the Office for National Statistics and contains datasets that describe the characteristics of neighbourhoods in the UK. Data from the census, currently census 2001, makes up the main information in the database but other government departments have provided data to complement this, including DECC which has provided data on domestic gas consumption. At the time of the analysis, 2008 sub-regional gas consumption data was the latest available dataset, the 2009 data is now available with the 2010 data being released on 29th March 2012.

The data are available down to Lower layer Super Output Area (LSOA). There are over 34,000 LSOAs in England and Wales and each one has a minimum population of 1,000 (the average being 1,500). The analysis was carried out for England only, since some variables were not available for Wales on the database. The following variables in the neighbourhood statistics database were assumed as having the potential to influence gas consumption.

Variable	Groupings	Latest year at time of modelling
Index of Multiple Deprivation (IMD)	Scale taking value between 0-100[a]	2007
Age of population (%)	0-15, 16-59/64[b], 60/65+	2008
Dwelling stock by tenure (%)	Owner, Private rented Social rented	2001
Dwelling stock by type (%)	Detached, Semi –detached, Terrace, Purpose built flat, Converted flat	2001

[a] The higher the value the more deprived the area

[b] Includes females aged 59 and under and males aged 64 and under. It is also possible to disaggregate the working age population into smaller groups (e.g. 30-44)

The IMD[2] is a measure at LSOA level of multiple deprivation and is made up of seven domains indices (income deprivation, employment deprivation, health deprivation and disability, education deprivation, barriers to housing and services, living environment deprivation and crime). The overall IMD was picked as providing the most comprehensive picture of deprivation in the LSOA.

Data limitations

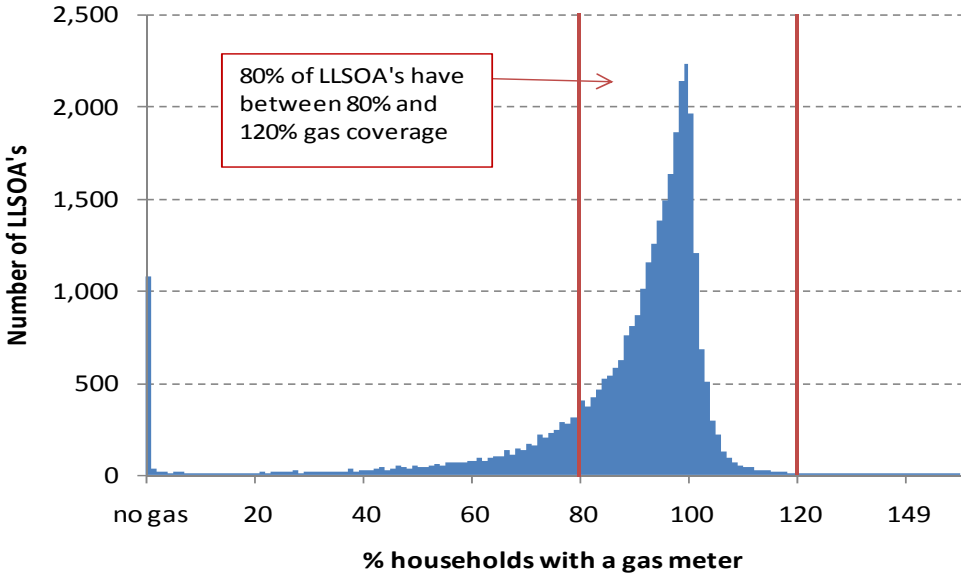
There were a number of data issues that were examined before undertaking the analysis. As the data are aggregated to LSOA level there is no way of knowing which houses use gas. As the focus of the work was to model gas consumption as the primary heating fuel, it was important to only consider local areas where this was the case. Therefore, an approximate 'gas coverage'

[1] http://neighbourhood.statistics.gov.uk/dissemination/LeadHome.do;jessionid=2yFyPJSpt2t6dGK1gtt0wJvHqYBjK35Lz2c6L2YXG69wKXrpGvcx!1519500766!1330958910756?m=0&s=1330958910756&enc=1&nsjs=true&nsck=true&nssvg=false&nswid=119

[2] www.communities.gov.uk/publications/communities/indiciesdeprivation07

indicator for each LSOA was created. This was calculated by dividing the number of domestic gas meters by the total number of electricity meters in each LSOA as, in theory, it is likely that all domestic houses will have an electricity meter. The spread is shown in chart 1 below.

Chart 1: Percentage of households with a gas meter in each LSOA



The distribution of the bars indicates that a large number of LSOAs have around 100 per cent of the households on gas. However, around 16 per cent of LSOAs had less than 80 per cent gas coverage (with 3 per cent having no gas). There are also 11 per cent of LSOAs with over 100 per cent gas coverage (the majority of these within 100-105 per cent), this could indicate that some electricity meters are missing or it could be due to the way 'domestic gas' is defined. For the gas consumption figures, a domestic meter is defined as one where less than 73,200 kWh is used in a year. However, the use of this definition is thought to categorise some small businesses as domestic, whereas for electricity the definition of 'domestic' is thought to be more robust as domestic dwellings have a different meter profile number.

LSOAs with a gas coverage indicator of over 80 per cent and under 120 per cent were included in the analysis. This means that around 80 per cent of LSOAs were used. Table 1 shows how this proportion varies by region. The results are in-line with expectations, for example, only 60 per cent of the households in the South West are included, where there are known to be areas in Cornwall off the gas network.

Table 1: LSOAs used by region

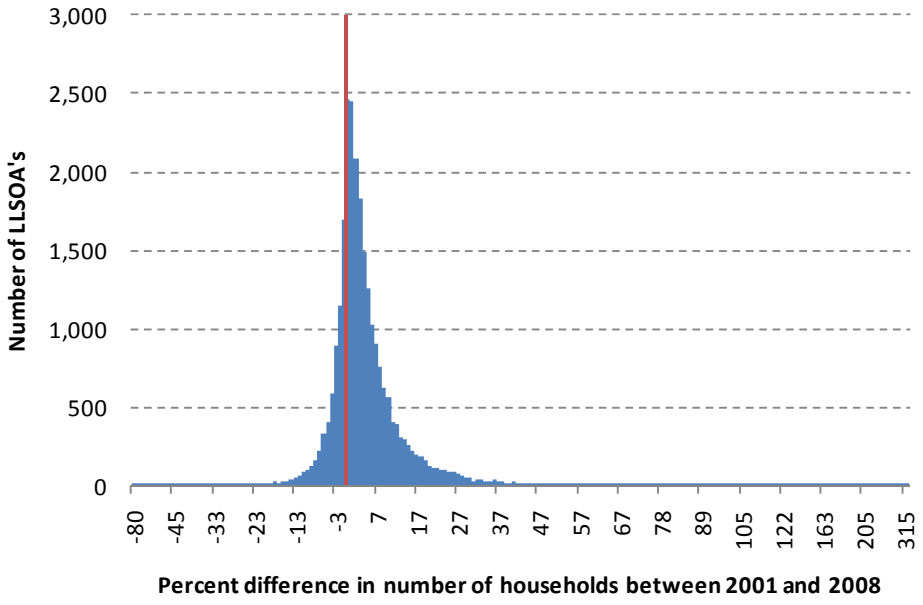
Region	Not used in analysis	Used in analysis	Total	% Used
East Midlands	492	2,240	2,732	82%
East of England	1,060	2,490	3,550	70%
London	743	4,022	4,765	84%
North East	190	1,466	1,656	89%
North West	501	3,958	4,459	89%
South East	1,252	4,067	5,319	76%
South West	1,273	1,953	3,226	61%
West Midlands	609	2,873	3,482	83%
Yorkshire and The Humber	488	2,805	3,293	85%
Total	**6,608**	**25,874**	**32,482**	**80%**

It should also be noted that although a household may have a gas meter, there is no way of knowing whether the gas is being used for heating only or also for cooking. If gas use for cooking

is of similar proportions between all LSOAs this will have little impact on the analysis but if there are some LSOAs where a large proportion of the households on gas only use it for heating and some LSOAs where the majority use gas for both heating and cooking then these LSOAs will appear to have higher than expected gas use. For this initial analysis, it has been assumed that gas use for cooking is broadly constant across LSOAs.

The only variables on the neighbourhood statistics website which describe the housing stock in an area are based on 2001 census data, i.e. it describes the housing stock 7 years prior to the consumption data. To get an indication of which areas are particularly affected by the time lag, the number of electricity meters in each LSOA was compared to the number of households given in the census. Chart 2 shows the distribution of the percentage difference, 24 per cent of LSOAs had fewer electricity meters in 2008 compared to census dwellings listed and 66 per cent had more electricity meters (10 per cent showed no change in housing stock between 2001 and 2008). Over 80 per cent of the LSOAs were within +/- 10 per cent of the number of census dwellings.

Chart 2: Percentage difference between number of households in 2001 census and the number of electricity meters in 2008[3]



In order to avoid reducing the LSOAs included in the analysis further, it was decided that all the remaining LSOAs would be included even if it appeared that the structure of the area may have changed since the census was taken. The percentage change will be reported for each LSOA so users can decide if a particular area has been particularly affected by housing change. Newer houses will be built in line with current building regulations and these will have higher levels of efficiency, so higher concentration of new houses may explain lower actual gas use.

Groupings within the 'dwelling stock by type' category are related, e.g. if the percentage of detached properties in an LSOA is 90 per cent no other category can be more than 10 per cent. One option would be to just put the strongest category into the model (e.g. per cent detached). However, average gas consumption for an LSOA that is 80 per cent detached and 20 per cent purpose built flat will be different to an LSOA with 80 per cent detached and 20 per cent semi-detached. Therefore it was decided that a weighting system would be used to produce one overall figure to capture the housing stock in the area. One of the main reasons that detached houses have higher gas consumption is because they are likely to be bigger than other house types (e.g. semi-detached houses). Data from the English Housing Survey 2007 was used to provide an average floor area for each type of dwelling. The floor area was then multiplied by the proportion

[3] Only looks at LSOAs selected for analysis (i.e. those with a 'gas coverage' indicator of 80-120%).

of houses of that type in the LSOA (i.e. the dwelling type information from the census was used to get a proxy for the average size of dwelling for each LSOA). The floor areas used are given in table 2 followed by an example calculation.

Table 2: Dwelling type and corresponding average floor area[4]

Dwelling type	Average floor area (m^2)
Detached	146
Semi-detached	93
Terrace	82
Converted flat	66
Purpose built flat	56

In an LSOA with 80 per cent detached and 20 per cent purpose built flat the average floor area would be 128m2 ((146 x 0.8) + (56 x 0.2)) compared to 135.4m2 for an LSOA with 80 per cent detached and 20 per cent semi-detached ((148 x 0.8) + (93 x 0.2)).

For the remainder of this document this variable will be referred to as 'average floor area'. It is however, only a proxy measure and assumes that each house type has the same floor area in all areas of England, in reality it is also possible that some flats may have a bigger floor area than a terraced house, for example.

The variable for 'dwelling stock by tenure' and age has the same issue of local variation as described for 'dwelling stock by type'. For these variables the percentage of owner/occupiers and percentage of pensionable age were used as these showed a positive correlation with gas consumption, with a correlation coefficient of 0.42 and 0.27 respectively.

The gas consumption data used in this analysis is temperature corrected, therefore different temperature patterns across the UK have already been taken into account.

Descriptive analysis

To get a clearer idea of the relationship between average household gas consumption and the five variables of interest, each variable was grouped into quintiles and box and whisker plots were drawn to illustrate the variation in gas consumption within each quintile.

Chart 3 looks at the relationship between IMD and gas consumption. The line in the middle of the box indicates the median average gas consumption for LSOAs in each quintile. The top and bottom of the box indicates the upper quartile and the lower quartile respectively and the line the 5th and 95th percentile (i.e. the range excluding the most extreme LSOAs). Overall, the median average gas consumption decreases as deprivation increases. It can also be seen that there are a number of LSOAs in the most deprived quintile that have an average gas consumption that is higher than gas consumption of some LSOAs in the least deprived quintile. The width of the lines also shows that variation between LSOAs is higher in the least deprived areas compared to the most deprived. In the most deprived areas 90 per cent of the LSOAs have an average gas consumption between 10,800 kWh and 18,400 kWh. In the least deprived areas the range is 14,500 kWh to 28,100 kWh.

[4] Source: Table SS2.0, Stock profile 2007, English Housing Survey
www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousecondition/ehcsdatasupporting/ehcsstandardtables/summarystatistics/

Chart 3: Average household consumption by IMD quintile

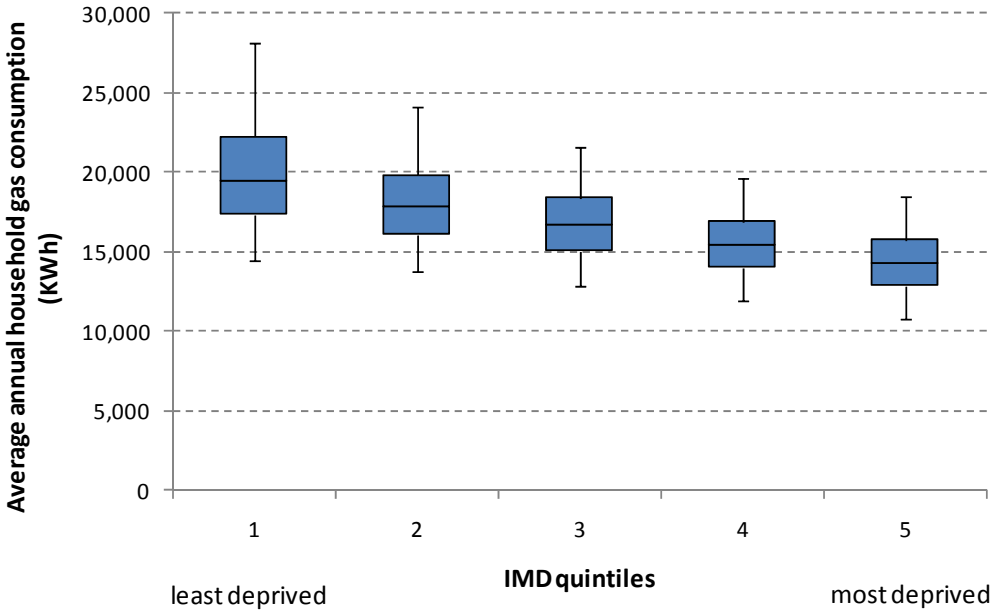
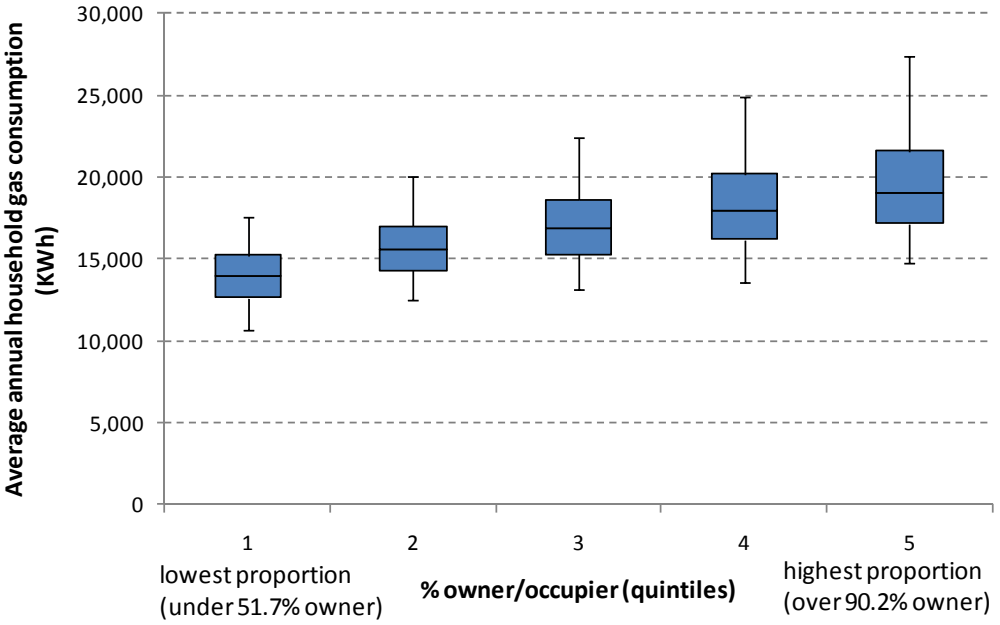


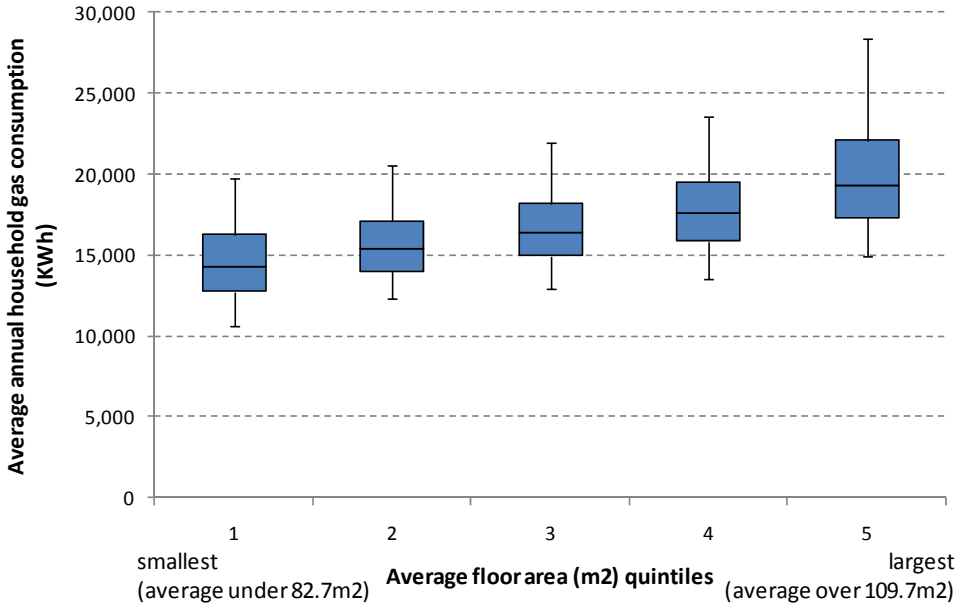
Chart 4 looks at the relationship between the proportion of owner/occupiers in the area and gas consumption. As the proportion of owner/occupiers increases so does the median average gas consumption. The median LSOA in the highest quintile (where over 90 per cent of the householders are owner/occupiers) has an average annual gas consumption of 19,000 kWh compared to 13,900 kWh for the median LSOA in the lowest quintile (where under 50 per cent of the householders are owner/occupiers).

Chart 4: Average household gas consumption by proportion of owner/occupiers



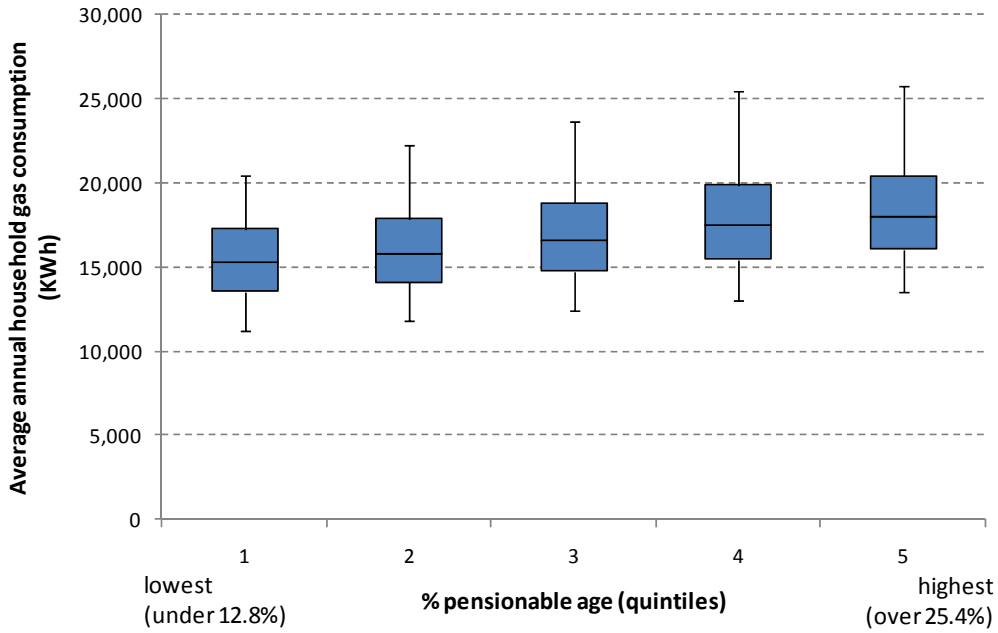
As expected, as the average floor area increases, gas consumption also increases (see chart 5), although again there is an overlap between the quintiles. For example, 25 per cent of the LSOAs with smallest average floor area (lowest quintile) have an average annual gas consumption of over 16,300 kWh a year and 25 per cent of the LSOAs in the (highest quintile) have an average annual gas consumption of under 17,300 kWh.

Chart 5: Average household gas consumption by average floor area



The proportion of the population over pensionable age had the weakest association with average gas consumption, with a positive correlation coefficient of 0.27. This is also evident in chart 6, where although the median average gas consumption increased as the proportion of pensionable age increased, the overlap between the quintiles was greater than that observed for the other variables. It should of course be noted that correlations exist between the factors described in charts 3 - 5.

Chart 6: Average household gas consumption by proportion of pensionable age



Regression analysis

As the four variables above are correlated (e.g. the proportion of owner/occupiers is higher in areas with low deprivation), stepwise regression analysis was carried out to try and unpick the effect of each individual variable. This is an iterative process whereby the most significant variable is entered at each step and any variables already in the model are re-tested to determine whether they are still statistically significant. The first variable to enter was '% owner/occupier', this variable alone explained 31.6 per cent of the variation in average gas consumption between

LSOAs (indicated by the R2 term). Adding in the average floor area increased the R2 to 36.6 per cent. Following this, the 'IMD' variable was entered, increasing the R2 by only 0.5 per cent. The percentage of pensionable age was the last variable to enter the model and although statistically significant the R2 did not increase noticeably when given to 3 decimal places.

The final model explains 37.2 per cent of the variation in average gas consumption between LSOAs. As seen in charts 3-6, the quintiles containing what are likely to be the most affluent LSOAs (the ones with the largest floor area, the highest proportion of owner/occupiers and the lowest IMD score) tend to have the biggest variation in average household gas consumption. This suggests that any predictive regression models that do not account for individual actions or choices, are likely to be less accurate for more affluent LSOAs.

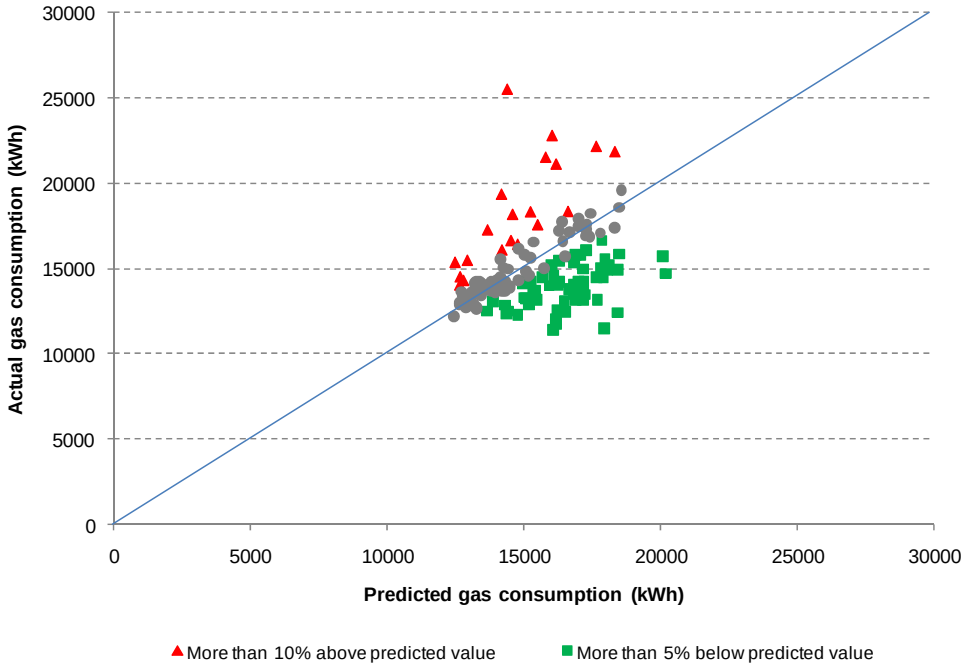
Using the coefficients computed by the regression model, a predicted gas consumption figure can be calculated for each LSOA using the equation below:

Predicted gas consumption in LSOA (kWh) = **8034.9** + **43.3** * % owner/occupier + **67.2** * Average floor area - **28.3** * IMD+**8.3** * % pensionable age

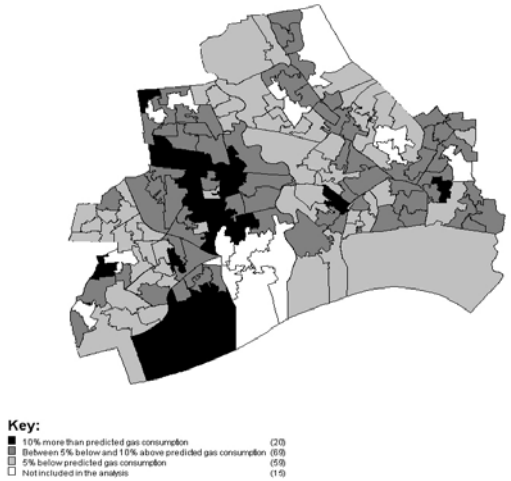
Selecting areas to target

One method for selecting LSOAs would be to set a criteria above which the actual value should lie in relation to the predicted value (e.g. if actual gas consumption is more than 10% above the predicted value then this LSOA will be targeted). This also helps to allow for the fact that the model has a lot of unexplained variance so the fact that an LSOA is slightly above the predicted value may not necessarily mean that gas use is higher than you would expect given the characteristics. The following example provides an illustration of the output that can be produced using this methodology.

Chart 7: Kingston upon Hull - Predicted average annual gas consumption compared to actual, 2008

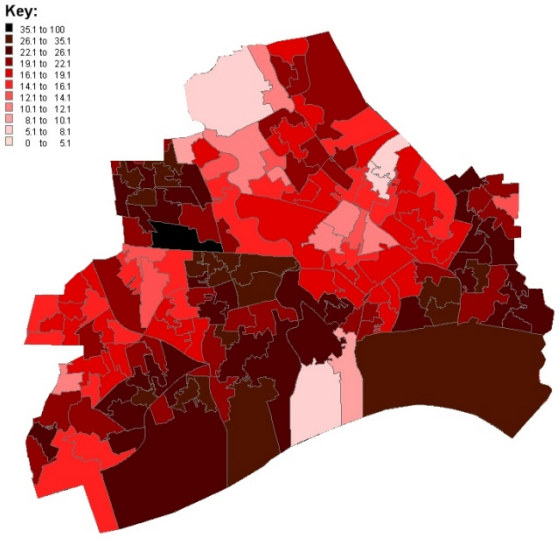


Kingston upon Hull: Actual gas consumption compared to predicted, 2008



Source: ONS, Super Output Area Boundaries. Crown copyright 2004. Crown copyright material is reproduced with the permission of the Controller of HMSO

Kingston upon Hull: Percentage of households in fuel poverty in each lower layer super output area, 2008



Source: ONS, Super Output Area Boundaries. Crown copyright 2004.

Both chart 7 and the map on the left-hand side show that, 14 per cent of LSOAs in Kingston upon Hull (20 out of 148) included in the analysis have actual gas consumption 10 per cent above predicted gas consumption, i.e. potential areas to target for energy efficiency measures. These areas are highlighted in black. In addition, the map on the right-hand side shows the fuel poverty ratios for each LSOA in this local authority, where the darker areas represent LSOAs that are more fuel poor. Ninety per cent of those LSOAs identified in Kingston upon Hull (18 out of 20) also have higher than average fuel poverty levels (the national fuel poverty ratio in 2008 was 15.6 per cent). As discussed, high gas use is often associated with low deprivation/high owner occupancy which are generally not characteristics of fuel poverty. Therefore, this high gas use is potentially being driven by energy inefficient housing. As such, these LSOAs could potentially be areas where efficiency investment could be directed to first.

User engagement

There is still some work to be done to refine the model but in the meantime DECC would like some feedback from users on this piece of analysis. If you would like to see the dataset for your local authority, please contact Mita Kerai using the details below. We are particularly keen to work with local authorities and housing associations who have a good level of knowledge of local housing, which could contribute either knowledge or inferences (such as high proportion of private rented). DECC statisticians will continue to work on the model, but working with others will help us to develop the work.

Mita Kerai
DECC Energy Statistics
Tel: 0300 068 5044
E-mail: Mita.Kerai@decc.gsi.gov.uk