Annex A: What is NEED?

Contents

What is NEED? 2

1. Introduction 3
2. The framework 3
3. Scope 4
4. Data protection 5
5. Data in NEED 5
6. Creating an analytical dataset 7
7. Uses of NEED 8

Appendix 1: Data in NEED 9

National Land and Property Gazetteer 9

One Scotland Gazetteer 9

Office for National Statistics Postcode Directory 9

Energy consumption meter data 10

Homes Energy Efficiency Database (HEED) 11

Valuation Office Agency Council Tax data 11

Valuation Office Agency Non-domestic Rates 11

Experian – domestic data 11

Experian – non-domestic data 11

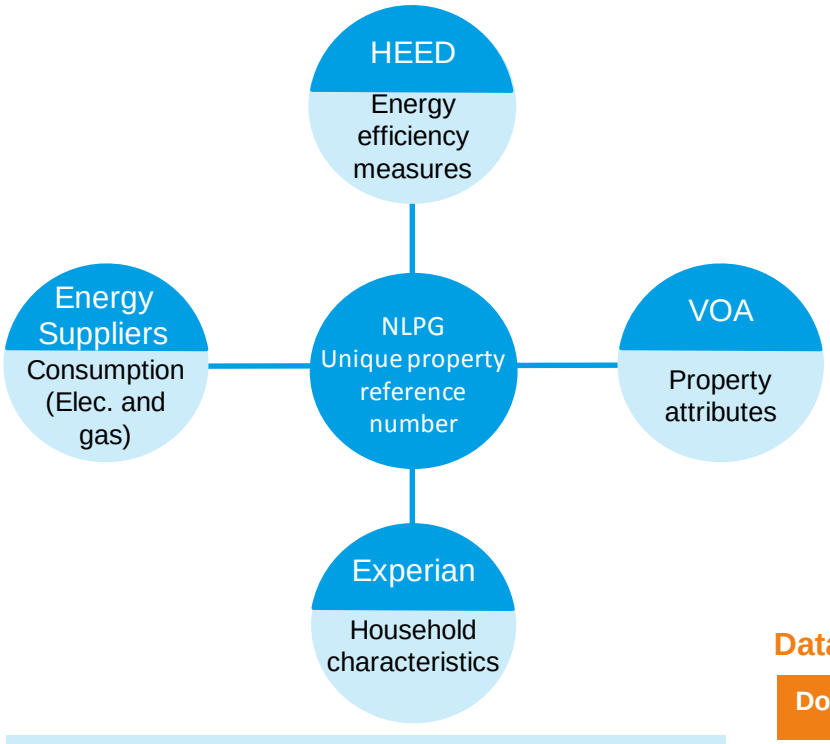
Output Area Classification 11

Display Energy Certificates 13

The National Energy Efficiency Data-Framework

What is NEED?

NEED is a framework for combining data from existing sources (administrative and commercial) to provide insights into how energy is used and what the impact of energy efficiency measures is, for different types of property and household.



The figure above shows how data are combined to form NEED. The address information in each dataset is used to assign a unique property reference number (UPRN) to each record within the dataset. Data from different sources can then be matched to each other via the UPRN. The principle is the same for the domestic and non-domestic sectors, though different data sources are used.

In Scope	Out of Scope
Domestic and non-domestic properties	Very large consumers e.g. power stations
Great Britain	Northern Ireland
Metered Gas and Electricity	Non-metered fuels e.g. oil, coal
Energy Efficiency Measures in HEED	DIY measures and others not recorded on HEED

Why is NEED important?

NEED provides the largest source of data available for analysis of energy consumption; previous evidence has been derived from surveys and small technical monitoring trials.

NEED forms an important element of DECC's evidence base and already plays a key role in development and evaluation of DECC policies, including the Green Deal.

Data in NEED includes:

Domestic	Non-domestic
Energy Suppliers	
• Electricity consumption 2004-2010 • Gas consumption 2004- 2010 • Electricity meter profile class	
VOA	
• Property size (m^2) • Number of bedrooms • Property type (e.g. terraced) • Property age	• Property description • Property area (m^2) • Car parking • Other additions
Experian	
• Income • Age of head of household • Length of residence • Tenure	• Turnover • Employment • Standard industrial classification (SIC) • Premises type
HEED	
• Energy efficiency measure installed (Yes/No) • Date measure installed	• None

1. Introduction

The UK has collected and published energy consumption data within the Digest of UK energy Statistics since 1948[1]. This has been produced and published at a national level and is based on aggregate information from energy suppliers. Data has relating to domestic energy use has also been published in Energy Consumption in the UK2. Whilst these data are still used, the development of UK energy policy has required more detailed data to help deliver and monitor reductions in energy use and emissions.

In 2004 DECC first started to collect individual meter point data primarily for the production of small area consumption data, but by working closely with the energy industry and other key energy efficiency stakeholders, plans were established for a future data architecture suitable for matching the consumption data with additional data sources. The National Energy Efficiency Data-Framework (NEED) is the means by which this has been achieved. Gas and electricity consumption data are matched, at an individual property level, with information about energy efficiency measures installed in households, property attributes and household characteristics.

The Framework was first announced in the Heat and Energy Saving Strategy[3] in 2009 and was developed by DECC, with support from the Energy Saving Trust (EST) and gas and electricity suppliers, in order to assist DECC in its business plan priority to "save energy with the Green Deal and support vulnerable consumers". It forms a key element of DECC's evidence base supporting DECC to:

- develop, monitor and evaluate key policies (including the Green Deal);
- identify energy efficiency potential which sits outside the current policy framework;
- develop a greater understanding of the drivers of energy consumption; and
- gain a deeper understanding of the impacts of energy efficiency measures for households and businesses.

The data framework provides the largest source of data available for analysis of consumption and impacts of energy efficiency measures. Previously DECC has relied on evidence from surveys and small technical monitoring trials. The first results from the framework were published in June 2011, as a pilot to test the framework approach worked. It demonstrated the value of NEED and its importance to DECC and a wider group of stakeholders, and as a result of this further work has been undertaken, enabling DECC to update and expand on previous work.

2. The framework

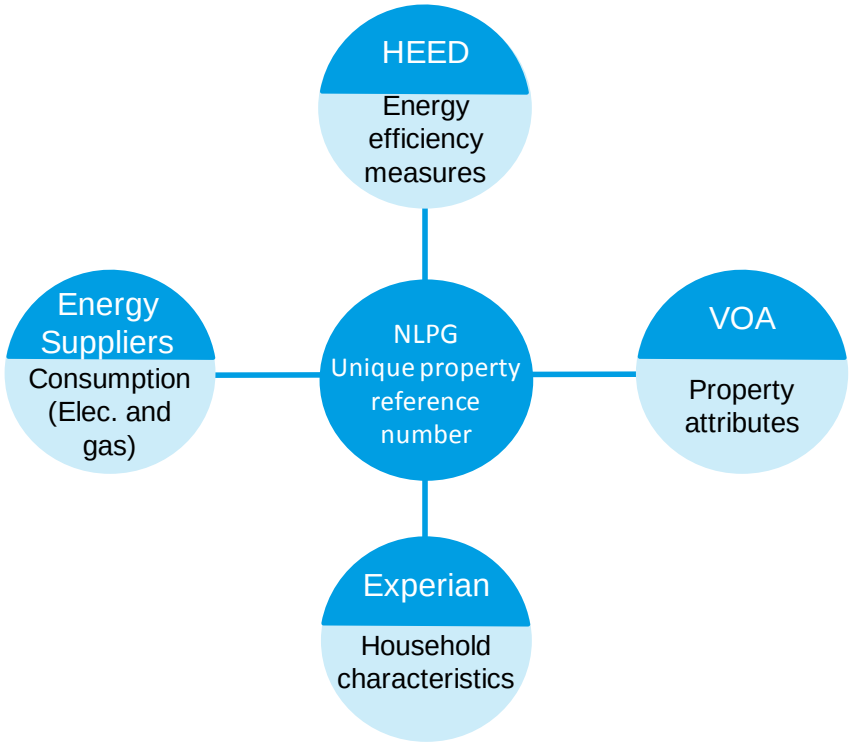
At the core of NEED is the National Land and Property Gazetteer (NLPG), the national standard for all buildings and addresses in England and Wales[4]. Each record within the NLPG has a Unique Property Reference Number (UPRN) which provides a reference key to join

[1] http://www.decc.gov.uk/en/content/cms/statistics/publications/dukes/dukes.aspx
[2] http://www.decc.gov.uk/en/content/cms/statistics/publications/ecuk/ecuk.aspx
[3] http://hes.decc.gov.uk/view_results/index.html
[4] The National Address Gazetteer is the equivalent in Scotland. Both of these, along with Ordnance Survey's Address Layer 2 (AL2), have now been replaced by AddressBase (http://www.ordnancesurvey.co.uk/oswebsite/products/addressbase/) which will be used for NEED in future.

related address records across different datasets. Even if a property is demolished, the UPRN can never be reused and retains its historical information.

Datasets are combined within the framework using the NLPG Unique Property Reference Number (UPRN) as a spine. Address data from each of the datasets included in NEED is matched with the address information from the NLPG and used to assign the UPRN to each record within that dataset. The UPRN is then used to link records from one dataset to the corresponding record in each of the other datasets. The diagram below shows how this works for the core data used for domestic analysis. The principle is the same for the domestic and non-domestic sectors though different data are used.

Figure 1: How NEED works



By using this approach to bring together existing data, it offers a cost effective solution enabling DECC to undertake detailed analysis of the impact of energy efficiency measures and gain a better understanding of how energy is consumed.

3. Scope

Data in NEED cover the domestic (or residential) and non-domestic (commercial/industrial) sectors across the whole of Great Britain. It includes information on metered gas and electricity consumption for all properties in Great Britain as well as information on energy efficiency measures installed in the domestic sector. Table 1 summarizes the scope of data in NEED, at the high level. Appendix 1 includes more detail on the scope of each of the individual datasets in NEED and highlights where there are differences in scope, for example between countries within Great Britain or between the domestic and non-domestic sector.

Table 1: Scope of NEED

In Scope	Out of Scope
Domestic and non-domestic properties	Very large consumers e.g. power stations
Great Britain[5]	Northern Ireland
Metered Gas and Electricity	Non-metered fuels e.g. oil, coal
Energy Efficiency Measures in HEED	DIY measures and others not recorded on HEED

Work to date has focused on the domestic sector in England. An initial assessment of the quality of non-domestic data for England and Wales is included in Section 6 of the NEED report.

4. Data protection

NEED is constructed to ensure it is compliant with the Data Protection Act. Data in NEED are gathered from a variety of sources including publically available data and through commercial licences, voluntary agreements and service level agreements with owners of datasets. As part of these agreements and to ensure no individual property or energy supplier can be identified only aggregate results are published. In line with this, all outputs of analysis from NEED are based on a minimum of 30 observations for each result quoted; any results based on fewer than 30 properties are suppressed.

5. Data in NEED

Table 2 below provides a summary of data used to form NEED. Further details of each of the datasets is provided at Annex 1, with information on the quality of the data set out in Annex B of the NEED report.

Table 2: Data in NEED

Category	Source	Description
Premises	National Land and Property Gazetteer (NLPG)	Contains a unique identifier for each address in England and Wales which acts as the spine of NEED.
	One Scotland Gazetteer (OSG)	As above, for Scotland, provides a unique identifier for every address in Scotland.
Geography	Office for National Statistics Postcode Directory (ONSPD)	Provides information on postal and administrative boundaries (including Local Authority).

[5] Only England is included within the scope of analysis of the domestic sector in this report (England and Wales for the non-domestic sector) though data for all of Great Britain are available for a number of variables.

Category	Source	Description
Energy consumption	Energy suppliers and Xoserve/Gemserv	Gas and electricity meter consumption data for all domestic and non-domestic meters in GB, 2004 – 2010 and meter profile number for electricity meters. Gas data are weather corrected.
Measures installed	Homes Energy Efficiency Database (HEED)	Information on energy efficiency measures installed through government schemes (including EEC, CERT and CESP[6]).
Property attributes	Valuation Office Agency (VOA)	VOA is an executive agency of HM Revenue and Customs responsible for business rates and council tax. They collect property detail data to inform this function. The data include floor area (domestic and non-domestic), property type and property age (domestic only).
	Experian	Modelled data for domestic property attributes such as number of bedrooms, property age and type. Purchased for 10% of households in England.
	HEED	Includes wall construction and property type, but accuracy and coverage vary.
Households characteristics	Experian	Modelled data for household characteristics such as income and tenure. Purchased for 10% of households in England.
	Output Area Classification (OAC)	Categorises geographic areas each containing approximately 125 households into 21 socio-economic groups, based on the 2001 census.
Business characteristics	Experian	Information for non-domestic properties, including; turnover, employment and standard industrial classification (SIC) code. Quality and coverage varies.

[6] http://www.decc.gov.uk/en/content/cms/funding/funding_ops/funding_ops.aspx

Category	Source	Description
	Display Energy Certificates (DECs)	Includes information about the amount of energy used in a property, size of property and type of heating and cooling systems. DECs are required for buildings with a useful floor area of over 1,000m^2 which are occupied by public authorities or institutions providing public services. Other building can have a DEC on a voluntary basis.

6. Creating an analytical dataset

Matching

The first step in creating an analytical dataset was to ensure that each component dataset had been matched to the National Address Gazetteer Address UPRN. GB Group carried out the matching, first ensuring they had Royal Mail Postcode Address File (PAF) formatted addresses on each dataset. These conditioned data were then matched with the NLPG reference set to assign a UPRN to each record. A quality assurance of this matching process was then carried out by Katalysis. The outputs of this showed that the match rate achieved had exceeded expectations. The table below shows the match rates for England and Wales for each of the datasets.

Table 3: Matching statistics at building level (sub-building[7] match rates in brackets)

Data source	Match rate
Electricity consumption[8]	94% (87%)
Gas consumption[9]	97% (93%)
HEED	99% (98%)
Experian	82% (69%)
Non-domestic rates	73% (54%)
VOA[10]	100%

The match rates set out in the table are calculated based on the number of records for each relevant source, not the number of UPRNs in England and Wales. For example, DECC only leased Experian data for a representative sample of about 3 million properties; the match rate shows how many of these 3 million records could be matched to the NLPG. The figure quoted

[7] Sub-buildings are properties within another building such as flats within a converted house, or individual units within shopping centre.
[8] Note that the match rate quoted for electricity and gas consumption include domestic and non-domestic properties.
[9] Note that the match rate quoted for electricity and gas consumption include domestic and non-domestic properties.
[10] The match rate for VOA data is 100 per cent as only VOA records that could be matched to the NLPG were included in the sample.

for HEED excludes flats which were excluded from the analysis of impacts of energy efficiency measures due to difficulties with matching.

Selecting a sample

In order to help data confidentiality, increase processing speeds and reduce cost it was decided that a sample would be used for analysis of the domestic sector. The aim was to create a representative sample which would be suitable for analysis. Data on the Valuation Office Agency Council Tax database was used to draw a representative sample.

Four million records (a 17.8 percent sample) were drawn from the VOA database; this number were selected to allow for attrition and still have a sample of at least ten percent of properties in England. Four variables were used to ensure the sample was representative; region, property type, number of bedrooms and age of property. A uniform random sample was drawn from each category. More detail about the distribution and representativeness of the sample is provided in Annex B.

The final step in creating an analytical file was to match the data from all sources together using the NLPG UPRN. The One Scotland Gazetteer was used as the spine for Scotland; where only HEED and consumption data are currently available to DECC.

7. Uses of NEED

NEED provides the largest source of data available for analysis of energy consumption. DECC had previously relied on evidence derived from surveys and small technical monitoring trials. This new source of data forms an important element of DECC's evidence base and already plays a key role in development and evaluation of DECC policies. For example it has been used to inform the savings expected from measures installed in households under the Green Deal.

It is intended that NEED will continue to be developed in future, including through the move to AddressBase as the spine for NEED which should improve the quality of the data and allow more analysis of flats. There are also plans to include more data sources in the Framework, including energy performance certificates data. Results will be published on the NEED webpage as they become available:
http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/en_effic_stats/need/need.aspx

DECC welcomes feedback on these data and outputs, including the usefulness of this publication and future work it would be useful to undertake. Please send any comments to: EnergyEfficiency.Stats@decc.gsi.gov.uk.

Appendix 1: Data in NEED

National Land and Property Gazetteer[11]

The National Land and Property Gazetteer (NLPG) is the brand name for the compilation of local authority Addressing Datasets. Local authorities in England and Wales have a statutory responsibility for street naming and numbering. They update the NLPG and the National Address Gazetteer Database on a continual basis providing comprehensive coverage of all addresses in England and Wales.

The NLPG was initiated in 1999 to become the master address dataset for England and Wales. It is the central hub for the 348 address creating local authorities' Local Land and Property Gazetteers (LLPGs), which are also knows as Authority Addressing Datasets. The creation and maintenance processes are tested, combining local knowledge with central validation.

The data is created and maintained at local level to an agreed methodology under the LLPG data entry conventions document (DEC-NLPG), and passed to the hub which tests its structural conformance to the agreed implementation of BS7666 (2006) Parts 1 & 2. The hub also checks the quality through a regular data audit against third party national address datasets such as the Valuation Office Agency's Council Tax and Non Domestic Rates lists of addresses.

Each record has a Unique Property Reference Number (UPRN) which provides a reference key to join related address records across different datasets. Even if a property is demolished, the UPRN can never be reused and retains its historical information.

The NLPG has recently been superseded by AddressBase[12], which will be used as the spine for NEED in future.

One Scotland Gazetteer[13]

The One Scotland Gazetteer is an address database made up of all 32 individual local authority gazetteers in Scotland. All addresses are created in accordance with the national standard for addressing, BS7666:2006 and the Scottish Gazetteer Conventions.

AddressBase includes information for Scotland and will be used as the spine for NEED for all of Great Britain in future.

Office for National Statistics Postcode Directory[14]

The ONS Postcode Directory (ONSPD) links all current and terminated UK postcodes to the administrative, health and other geographic areas in which each postcode falls. The ONSPD is a Gridlink® branded product that pulls together data from members of the Gridlink® Consortium (Royal Mail, Ordnance Survey, National Records of Scotland, Land & Property Services (Northern Ireland) and ONS).

---

[11] www.nlpg.org.uk
[12] http://www.ordnancesurvey.co.uk/oswebsite/products/addressbase/
[13] http://www.onescotlandgazetteer.org.uk/
[14] http://www.ons.gov.uk/ons/guide-method/geography/products/postcode-directories/-nspp-/index.html

Postcodes that straddle two geographic areas will be assigned to the area where the mean grid reference of all the addresses within the postcode falls.

Energy consumption meter data

DECC has collated and analysed property level electricity and gas consumption data since 2004 for the purpose of producing aggregate statistics at sub-national level. DECC publish these data at the Middle Layer Super Output Area (MLSOA) and Lower Layer Super Output Area (LLSOA) [15]. These are both UK Census based areas, covering around 5,000 and 1,000 people (2,000 and 400 homes) respectively[16]. DECC also has a agreements in place with suppliers to use meter point gas and electricity consumption data in NEED.

The analysis in this report uses gas and electricity meter consumption data from 2004 to 2010. Analysis primarily focuses on the domestic sector in England, but consumption data are for all domestic and non-domestic meters in Great Britain are supplied to DECC - around 30 million electricity meters and 25 million gas meters.

Data are obtained from the existing administrative systems of the energy companies and provided by the data aggregators. Electricity data are provided by data aggregators who provide meter point consumption information. Gemserv provide the addresses for each meter point administration number (MPAN). Gas data are aggregated by Xoserve and a number of independent gas transporters for each meter point reference number or gas meter (MPRN).

Xoserve provide annualised estimates of consumption for all the MPRN's based on an Annual Quantity (AQ). An AQ is an estimate of annualised consumption using consumption recorded between two meter readings at least six months apart. The estimate is then adjusted to reflect a 17 year weather correction factor. The AQ for each MPRN represents consumption relating to the gas year – the period covering 1 October through to the following 30 September . The gas data provided to DECC are weather corrected before DECC receive the data.

For electricity, annualised estimates are based on either an annualised advance (AA) or estimated annual consumption (EAC). The AA is an estimate of annualised consumption based on consumption recorded between two meter readings. In comparison an EAC is used where two meter readings are not available and an estimate of annualised consumption is produced by the energy company using historical information and the profile information relating to the meter. These annualised estimates provide a good approximation of consumption, but do not cover exactly the calendar year. For example, 2010 annualised consumption estimates cover the 365 days up to 30 January 2011.

Data are currently collected on an annual basis, with the introduction of smart metering DECC will be reviewing the frequency of the data collection and subsequent analysis.

Further information on these data including the data production methodology and the quality of the data are available in Annex B.

[15] http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/regional/regional.aspx.
[16] http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas--soas-/index.html

Homes Energy Efficiency Database (HEED)[17]

HEED is a national database developed by the Energy Saving Trust. It was set up to help monitor and target carbon reduction and fuel poverty work. It contains details of energy efficiency and micro-generation installations such as cavity wall insulation and solar hot water, including the date of installation. Data have been recorded in HEED since 1995 including activity reported from Government programmes, such as the Energy Efficiency Commitment (EEC) and Carbon Emissions Reduction Target (CERT), and activity reported by trade associations such as Gas Safe and FENSA. Approximately 50 per cent of UK homes have a record in HEED[18].

Valuation Office Agency Council Tax data

The Valuation Office Agency (VOA) are responsible for allocating homes in England and Wales to the appropriate Council Tax band. In order to do this it maintains a property database covering all properties in England and Wales. It includes information on the age of dwelling, dwelling type, number of bedrooms and floor area. Further information about the data and statistical publications relating to these data are available on the VOA website: http://www.voa.gov.uk/corporate/Publications/statistics.html.

Valuation Office Agency Non-domestic Rates

The VOA are also responsible for valuing business premises for the purpose of charging Business Rates in England and Wales. Information available from this database includes type of premises (e.g. office, shop, place of worship) and floor area.

Experian – domestic data

Experian is a commercial organisation which produces modelled data of household characteristics at address level. Variables include income group, number of adults, age of occupant and length of residence. The Experian model derives these variables for each address using data from a range of sources including 2001 Census outputs and Experian's consumer survey. DECC purchased a sample of these data for 10 percent of households in England. A quality assessment of these data are provided in Annex B.

Experian – non-domestic data

The Experian non-domestic file is compiled from a variety of business registers and provides address level information about the occupying business. Variables include turnover, employment and business sector (e.g. Standard Industrial Classification (SIC) code and Thomson Directory code).

Output Area Classification[19]

Output Area Classifications (OAC) were created to enable an understanding of the character of local areas at low level of geography.

Output Areas are the lowest level of Census geography each containing approximately 125 households and are based on socio-economic data from the 2001 Census[20]. Output Area

[17] http://www.energysavingtrust.org.uk/Organisations/Local-delivery/Free-resources-for-local-authorities/Homes-Energy-Efficiency-Database/Introduction-to-HEED

[18] There are records for approximately 50 per cent of homes in the UK, however there may not be full information for each of these records. There is no information on measures that a household has installed itself (DIY) or measures installed in the home when they were built.

[19] http://areaclassification.org.uk/getting-started/getting-started-what-is-the-output-area-classification/

[20] A project is underway to create a new Output Area Classification using 2011 Census. Once this is available it will

Classifications are a three tier hierarchy consisting of 7 (Super-groups), 21 (Groups) and 52 (Sub-groups). The classification was created from 41 census variables and classifies every output area in the UK based on its value for those variables.

It was created in a collaboration between the ONS and the University of Leeds using the same methods as the related classifications of local authorities and wards and is freely available from ONS. OAC are used for a range of purposes including profiling of populations, structuring other data, and the targeting of resources.

Results in NEED are shown for the 21 groups, set out in Table 4.

Table 4: Output Area Classifications

Super-group	Group
Blue collar communities	Terraced
	Younger
	Older
City Living	Transient
	Settled
Constrained by circumstances	Older workers
	Public housing
	Senior communities
Countryside	Village life
	Agricultural
	Accessible
Prospering suburbs	Younger families
	Older families
	Thriving
	Semis
Typical traits	Settled Households
	Least divergent

be used in NEED.

Super-group	Group
	Younger families in terraced houses
	Aspiring households
Multicultural	Asian
	Afro-Caribbean

Display Energy Certificates[21]

Display Energy Certificates (DECs) are required for buildings with a total useful floor area over 1,000m^2 that are occupied in whole or part by public authorities and by institutions providing public services to a large number of persons and therefore frequently visited by those persons. Other buildings can choose to have a DEC on a voluntary basis.

The purpose of the DEC is to raise awareness of energy use and to inform visitors to public buildings about energy use of a building. DECs provide an energy rating of the building from A to G, where A is very efficient and G is the least efficient. The ratings are based on actual amount of metered energy used by the building over a period of 12 months. DECs also include information on the main heating fuel, the amount of energy used, size of the property and type of heating, cooling and ventilation systems. The DEC Operational Rating compares the actual energy consumption with a benchmark set for typical energy use of such a building.

[21] http://www.communities.gov.uk/publications/planningandbuilding/displayenergycertificates