Programme Area: Smart Systems and Heat

Project: EnergyPath

Title: Local Area Energy Planning Data Challenges Report

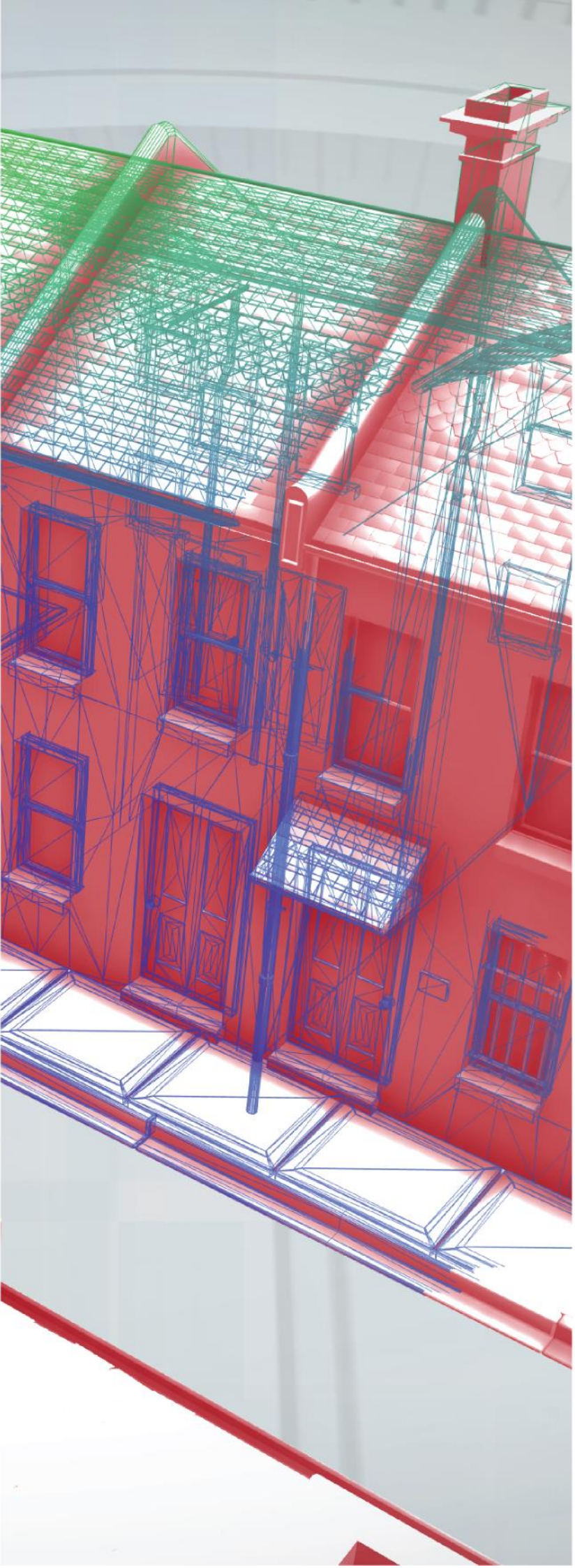
Abstract:

The purpose of this report is to provide an analysis of the data challenges involved in a whole system approach to local area energy planning informed by the piloting of EnergyPath Networks in three local areas. The report considers:

- The types of data that are needed, the sources from which it is available and any restrictions on use.
- The data processing required and associated costs.
- The assumptions, exclusions, limitations and value of data in supporting decision making.
- Other data that might be available and valuable for validation of modelling outputs and supporting local area energy planning in the future.
- Methods for gathering additional data and how these could be phased into local area planning.

Context:

Energy consultancy Baringa Partners were appointed to design and develop a software modelling tool to be used in the planning of cost-effective local energy systems. This software is called EnergyPath and will evolve to include a number of additional packages to inform planning, consumer insights and business metrics. Element Energy, Hitachi and University College London have worked with Baringa to develop the software with input from a range of local authorities, Western Power Distribution and Ramboll. EnergyPath will complement ETI's national strategic energy system tool ESME which links heat, power, transport and the infrastructure that connects them. EnergyPath is a registered trade mark of the Energy Technologies Institute LLP.



energy technologies institute

Smart Systems and Heat Phase 1

WP2 Bidders Pack

D3: Local Area Energy Planning Data Challenges Report

Delivered by CATAPULT Energy Systems

1 Introduction

1.1 Context

Local Area Energy Planning is seen by the Energy Technologies Institute (ETI) and the Energy Systems Catapult (ESC) as central to achieving national greenhouse gas emission reduction targets and the decarbonisation of heat[1]. Currently, there is no structured planning process to manage this transition at a local level across the UK. To decarbonise the UK's energy system efficiently, and at least cost, local area energy planning will be of integral importance in providing a long-term and whole system approach to planning local energy systems. .It will need to consider the unique characteristics of the local area and its existing energy system to guide the transition; aid decision making; prioritise resources; and support project and investment decisions.

As part of the Smart Systems and Heat (SSH) Phase 1 Programme the ETI has worked with the ESC on development and piloting of EnergyPath Networks with three different local areas in the UK. EnergyPath Networks is a local energy system optimisation analysis framework. It considers the whole energy system for a given local area of the UK, and can be used to investigate different options and choices to decarbonise buildings and the energy networks that support them. This can provide insights into cost effective decarbonisation pathways and inform future local energy system designs allowing evidence based decision making by key stakeholders including national and local government, network operators, communities and businesses to understand the costs and realise the benefits of smart low carbon heating solutions.

Producing a robust and valuable evidence base requires substantial and detailed data on both the current energy system and future options for change.

1.2 Purpose of this report

This report has been prepared as part of the Smart Systems and Heat (SSH) Phase 1 Programme with the objective of supporting more effective Local Area Energy Planning. The purpose of this report is to provide an analysis of the data challenges involved in a whole system approach to local area energy planning informed by the piloting of EnergyPath Networks in three local areas. The report considers:

- The types of data that are needed, the sources from which it is available and any restrictions on use.
- The data processing required and associated costs.
- The assumptions, exclusions, limitations and value of data in supporting decision making.
- Other data that might be available and valuable for validation of modelling outputs and supporting local area energy planning in the future.
- Methods for gathering additional data and how these could be phased into local area planning.

[1] http://www.eti.co.uk/library/an-eti-perspective-the-importance-of-local-area-planning-to-the-decarbonisation-of-heat

2 Data for local energy system modelling

In order to adopt a data driven and whole system approach to local area energy planning and the development and application of more advanced local energy system models a wide range of data is required. This is likely to come from a wide variety of sources and be held by a range of different public and private organisations.

This data includes:

- Spatial data:
  - Local topography (including building locations, sizes and heights)
  - Unique building identifiers and associated address information
  - Local road network
- Domestic buildings:
  - Locations and sizes
  - Types and ages
  - Current fabric / heating systems including any low carbon systems already installed locally
  - Future options for fabric retrofit and heating systems with associated cost and performance data
- Non-domestic buildings:
  - Locations and sizes
  - Current uses
  - Future options for fabric retrofit and heating systems with associated cost and performance data
  - Energy consumption profiles for different building use categories
- Network data for gas electricity and heat:
  - Current asset locations, capacities and connections
  - Future options for network build, reinforcement, re-purposing and decommissioning with associated costs
- Distributed generation (e.g. Combined Heat and Power, Solar PV) and storage for electricity and heat:
  - Current site locations with details of technologies installed
  - Future technology options with associated cost and performance data
- National information:
  - Current national energy prices and carbon contents
  - Projected future national energy prices and carbon contents
- Other local information:
  - Local carbon emissions inventory and future emissions target
  - Land boundaries

- Development information such as land allocations and associated build plans, local conservation areas and transport infrastructure

Some data might be considered commercially sensitive or be subject to data protection legislation. In addition, some data has been collated and managed commercially and can be subject to both licensing costs and restrictions.

3 Discussion of data sources

As discussed in Section 2 access to a wide range of data is needed to understand current local energy systems and the options for decarbonisation of energy in a given local area to inform local energy system designs.

This chapter provides details of the types of data that is required based on the work undertaken using EnergyPath Networks as part of the SSH Programme, the sources from which this data might be available, the types of processing that might be required and the possible costs.

Section	Data Description
3.1	Building and road locations, non-domestic building uses, building status and building heights.
3.2	Domestic building age and type
3.3	Local statistics for housing archetypes, heating systems and current insulation levels
3.4	Existing microgeneration
3.5	National statistics for housing archetypes, heating systems and current insulation levels
3.6	Profiles for domestic lighting and appliance use
3.7	Profiles for domestic hot water use.
3.8	Domestic heating system costs
3.9	Domestic building target temperature profiles
3.10	Domestic building thermal performance
3.11	Domestic fabric retrofit costs
3.12	Current and future weather
3.13	Non-domestic building uses and floor areas
3.14	Non-domestic building energy benchmarks
3.15	Non domestic building transition options
3.16	Current electricity network data
3.17	Location of high pressure gas entry points
3.18	Off-gas grid postcodes
3.19	Current and planned heat networks
3.20	Energy network options
3.21	Current and planned distributed generation and storage
3.22	Distributed energy generation and storage technology costs and efficiencies
3.23	Potential locations for distributed energy generation and storage
3.24	National fuel prices and carbon contents (boundary prices)
3.25	Land boundaries and development information

3.1 Building and road locations, non-domestic building uses, building status and building heights.

The primary spatial data on which analysis is based is the Ordnance Survey Mastermap and AddressBase Premium data. This is required to establish building locations and their spatial relationships to the road network which is generally assumed to be the location of energy networks. OS data provides some classification of domestic building types; the use category of non-domestic buildings and whether buildings are currently in use. In certain cases it can also provide mapping to other data sources such as the Valuation Office Agency data.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Ordnance Survey (OS): Topography (including bulding heights), AddressBase Premium, Integrated Transport Network (ITN) layer	The OS map is assumed to be up to date and a complete spatial representation of the study area. Height data is not included as standard by OS as part of the Topography data but can be provided at no extra cost if requested. It does not cover the whole of the UK and may be incomplete in the areas covered.	Essential	Data licensed from OS can only be used for the projects specified when the license is obtained. Some information is now available open source from OS (e.g. ITN layer) but AddressBase Premium and Mastermap Topography are not freely available.	If work is being done for a local authority then OS data can be made available under the Public Sector Mapping Agreement (PSMA) for the duration of the project. Otherwise this data must be licensed from OS. Cost depends on the map area required. Indicative cost is £6,000 for a local authority area for a 1 year license, with no cost if	In some cases this data can be incomplete and can need considerable processing to associate building points with building polygons and to establish which buildings are currently active. In other cases the processing required is less but it is unlikely that no processing will be required.	It is not possible to perform spatial planning without map data.	UKMap by the GeoInformation Group could be an alternative but is not available under PSMA and it only covers some areas of the country.	Processing of alternative sources of building height data can be required to supplement data available from OS.

				provided by an LA under PSMA.				

3.2 Domestic building age and type

This data is required is required to understand both current energy consumption and the heating system and thermal efficiency improvement options available to individual domestic buildings. Ideally this should be provided at individual Unique Property Reference Number (UPRN) level otherwise building characteristics must be modelled, introducing error and uncertainty. The more data that can be provided the better.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
GeoInformation. ESC processing of OS topographic layer for type.	This data is based on analysis of aerial photographs and then verified with some local surveying where possible. It should not be assumed to be 100% correct although quality is generally high. GeoInformation data is only available for urban areas. Data can be produced for areas not currently covered but this has an associated time and cost.	Required in as much detail as available.	Commercial data available under 1 year licenses.	Around £7,500 for a local authority area for domestic and non-domestic classifications. Depends on the number of buildings for which data is requested.	Little processing required if GeoInformation data is used. A mapping from GeoInformation classifications to those used in EnergyPath Networks must be produced.	In order to be able to estimate domestic building energy use the age and type of buildings must be known. The distribution of these in the local area will be assumed to match the national average if detailed data is not provided. For England this is based on the English housing survey. Similar data is available for Scotland but none is currently available for Wales such that other data must	Energy Performance Certificates can be used to understand individual building age and type. The OS topographic layer can be processed for type. Housing Association asset records and records from local retrofit schemes may also be sources of this data.	Data could be obtained by processing Energy Performance Certificate data where available and then inferring details for buildings with no EPC. Both commercial and academic organisations have developed methods to perform this process.

						be sourced for Welsh studies. In the case when building level data is not provided both current and future building information will be much less accurate leading to a significant increase in uncertainty in results for both buildings and networks.		

3.3 Local data for housing archetypes, heating systems and current insulation levels

Detailed information on some local buildings may be available. This can be used to improve understanding of the current housing stock to assist in estimating current energy demands and to allow better informed future pathways to be developed. This information is likely to be held by a variety of local organisations and may take considerable effort to gather and process.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local authority records for retrofit projects. Building control records for heating system and double glazing installations. Housing association asset records. Energy Performance Certificate data.	Assumed that detailed records are correct. These data sets are often incomplete or suffer from inconsistent classification (e.g. the heating system might be classified as gas boiler or just boiler). Energy Performance Certificate data is often inaccurate and should be treated with caution.	This data is valuable to improve understanding of local housing but is not essential to performing analyses.	Care must be taken if the original records contain personal information due to data protection legislation. The biggest challenge in accessing this type of data can be finding the people responsible for the individual data sets and getting access to the information required.	The underlying data is generally available at no cost but data processing costs are likely to be incurred in developing this data set.	Processing can be a complicated process as each data set will have been developed for another purpose, be held in a format suited to that purpose and will use a unique data schema. Matching of addresses between different data sources can be difficult if there are no unique references, such as a Unique Property Reference Number (UPRN).	Representation of some aspects of the local housing stock will have to be based on statistically matching national or regional data such that accuracy will be reduced.	The locations of data held locally will have to be identified by individual project teams with each location being unique. Examples of places where this data might be held are given in column 1 of this table.	A survey of local housing could be conducted to establish its make up but this could be extremely expensive.

3.4 Existing microgeneration

To accurately model the existing local energy system it is necessary to understand how much micro generation (primarily solar PV) has been installed in the local area. Some data is available from national sources[2] broken down by local authority. Alternatively, the local electricity network operator may be prepared to provide their data on micro-generation installed in the local area.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Department of Business, Energy and Industrial Strategy (BEIS) data, Local Authority data, local Distribution Network Operator (DNO) data.	No data is publically available on the precise locations of all existing installations. Assumptions must be made as to how the total installed is spatially distributed across the local authority area. Published data is based on installations confirmed on the Central Feed-in Tariff Register. Any installations not on the register will not be included.	At current levels of microgeneration including this data is not essential in terms of obtaining reasonable estimates of current energy demands and network capacities as less than 3% of domestic buildings currently have solar PV installed.	BEIS data is publically available.	The underlying data is available at no cost but data processing costs are likely to be incurred in developing this data set for modelling processes.	Aggregated data must be disaggregated and distributed appropriately across the study area.	Estimates of imported electricity demand will be slightly too high.	The national data could be supplemented by information from local sources such as social housing stock records although the benefits of doing this are likely to be small compared to the effort involved.	None.

[2] https://www.gov.uk/government/statistical-data-sets/sub-regional-feed-in-tariffs-confirmed-on-the-cfr-statistics

3.5 National statistics for housing archetypes, heating systems and current insulation levels

Ideally full details of each house in the local area would be known. Whilst good quality data is often available for social housing this is not the case for owner-occupied and private rented houses. Where information is not available then it must be inferred from other data sources. This can be done by assuming that the local housing stock matches national statistics.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
English Housing Survey	It is assumed that the local housing stock matches the national data. For example, the proportion of cavity walls that has been filled in the local housing stock is the same as the national proportion. For England this is based on the English housing survey. Similar data is available for Scotland but none is currently available for Wales such that other data must be sourced for Welsh studies.	Essential. Without an understanding of the housing in a local area it is not possible to develop credible future pathway options.	English Housing Survey and Scottish House Condition Survey data are publically available. The first Welsh House Condition Survey is due to be published in autumn 2018. To understand how often combinations of different attributes are represented in the national stock the underlying data from surveys will be required. This will need to be specially requested.	The underlying data is available at no cost but data processing costs are likely to be incurred in developing this data set.	The data needs to be processed to establish the number of times each combination of different attributes is represented in the national stock.	In order to be able to estimate domestic building energy use the age and type must be known. These can be assumed to match the national average if detailed data is not provided. Without national or local data it will not be possible to understand the make-up and characteristics of the local stock.	It is possible to process the Energy Performance Certificates that exist for local buildings in order to understand the make-up of the local housing stock.	A survey of local housing could be conducted to establish its make up but this could be extremely expensive.

3.6 Profiles for domestic lighting and appliance use

Electricity demand for lighting and appliances must be included in analyses in order to correctly estimate current and future network loads. Within EnergyPath Networks there are no alternative options for these loads and the user-defined loads are assumed to be present regardless of the heating system installed.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Department of Energy and Climate Change (DECC) Household Electricity Survey[3]	The detailed monitoring data is assumed to represent after diversity demand for all homes. Future demand changes must be generated. These can be based on scenarios like National Grid's Future Energy Scenarios[4].	These demands are important in calculating total electricity demand on local networks in order to correctly understand current network capacity margins and future reinforcement requirements.	This is publically available data.	The underlying data is available at no cost but data processing costs are likely to be incurred in developing this data set.	Projections must be made for future demand changes. Data must be disaggregated to match building categories used in analysis.	Current and future electricity network options will be incorrectly sized.	None found.	It should be possible to use data from advanced Home Energy Management Systems / Home Energy Services Gateway systems to estimate these demands. This data would need to be anonymised and aggregated to get after diversity demand profiles.

[3] https://www.gov.uk/government/publications/household-electricity-survey--2
[4] http://fes.nationalgrid.com/

3.7 Profiles for domestic hot water use.

Estimates of domestic hot water demand are required in order to correctly size heating systems. Within EnergyPath Networks these demands are fixed by the user and must be met.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Energy Savings Trust (EST) Measurement of Domestic Hot Water Consumption in Dwellings[5]	The aggregated monitoring data is assumed to be representative of demand for all homes. Future demand changes must be generated if future demand is expected to be different to current demand. The detailed monitoring data has not been made publically available by EST. So demand profiles must be inferred from the published data.	These demands are important in order to understand the total heat demand in homes.	EST report is publically available but underlying data has not been published.	The underlying data is available at no cost but data processing costs are likely to be incurred in developing this data set.	Projections must be made for future demand changes. Data must be disaggregated to match building categories used in analysis.	Domestic heating systems are likely to be incorrectly sized if hot water demands are not included.	None found.	It should be possible to use data from advanced Hone Energy Management Systems / Home Energy Services Gateway systems to estimate these demands. This data would need to be anonymised and aggregated to get after diversity demand profiles.

[5] https://www.gov.uk/government/publications/measurement-of-domestic-hot-water-consumption-in-dwellings

3.8 Domestic heating system costs

The cost of fitting and maintaining heating systems must be included in any whole system analysis. Domestic heating systems have a typical lifetime of 15 to 20 years. When planning energy system change to 2050 it is expected that any individual system will be changed twice in that time frame. For the pilot projects the data was produced using evidence gathered by DECC to set the tariffs for the Renewable Heat Incentive (RHI).

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Data for RHI	Assumed that data gathered to create RHI tariffs is representative of current costs to install different heating systems. Assumptions must be made about how current costs will change in future. The RHI data only contains information for a limited number of different heating system technologies.	Estimating the future costs of different heating systems is essential to understanding the costs of future low carbon pathway options.	Publically available data.	The underlying data is available at no cost but data processing costs are likely to be incurred in developing this data set for modelling processes.	The data used for the pilot projects was for installation of heating systems with specific capacities. In order to allow estimation of installation costs into a wide variety of buildings this limited data set had to be processed to understand the fixed cost of installing each type of heating system (incurred regardless of the capacity of the system - £/installation) and the variable cost (£/kW capacity).	It is not possible to understand the costs of different future energy systems without knowledge of heating system costs.	Data from local schemes or trials could be used to supplement or modify the RHI data if it was available.	None.

3.9 Domestic building target temperature profiles

In order to perform modelling of domestic buildings an internal target temperature must be set for different times of day at different times of year. For the pilot projects the internal target temperature used in the Standard Assessment Procedure (SAP) calculations was used[6]. SAP is the methodology used by the Government to assess and compare the energy and environmental performance of dwellings. Its purpose is to provide accurate and reliable assessments of dwelling energy performances that are needed to underpin energy and environmental policy initiatives.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Standard Assessment Procedure (SAP)	It is assumed that all buildings are heated to the target temperatures used in the SAP calculation methodology. In reality different households will heat their homes to different temperatures at different times to those assumed in the SAP calculations.	Target temperatures are vital to understanding the heat demand of different buildings. This influences the building heating system power requirement, the annual and peak demand of the buildings and the networks required to deliver the energy.	Publically available.	None.	Target temperature profiles must be entered into the simulation tool used to perform building modelling. For the pilot projects this modelling was done using EnergyPlus[7].	N/A	Bespoke target temperature profiles could be easily produced if required. These could be based on survey or monitoring data..	None.

[6] https://www.gov.uk/guidance/standard-assessment-procedure
[7] https://energyplus.net/

3.10 Domestic building thermal performance

The thermal performance of buildings has a direct influence on their energy demands. This in turn will influence local energy networks. Understanding thermal performance of domestic buildings is essential to developing decarbonisation pathways.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
SAP	The SAP calculation methodology includes details required to calculate the heat loss rate of buildings from walls, windows, roofs etc. It includes characteristics for different types of wall and window and for different levels of insulation. These are assumed to be valid for all buildings.	Essential.	Publically available.	Underlying data available at no cost. Data processing costs are likely to be incurred in developing this data set.	Processing required to get this data into the format required for modelling of domestic buildings – exact format will depend on the modelling tool used.	It will not be possible to estimate the energy demands of buildings in the local area.	Various academic papers are available which discuss the heat loss rates of different buildings. These could be used as an alternative to the SAP values.	None.

3.11 Domestic fabric retrofit costs

Fabric retrofit can have a direct influence on the energy demands of individual buildings. This in turn will influence local energy networks. Options to improve the thermal performance of domestic buildings should be assessed as part of developing decarbonisation pathways.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
ETI's Optimising Thermal Efficiency of Existing Housing project[8]. Element Energy report "Review of potential for carbon savings from residential energy efficiency Final report"[9].	Assumed that the costs of retrofit are controlled by the type of intervention and the building size. It can be hard to identify costs for hard to treat buildings and correctly include them. Data is often commercially sensitive and so can be hard to access.	Essential.	The full data set used for the pilot projects is not currently publically available. Alternative sources may need to be found.	Unknown. Data processing costs are likely to be incurred in developing this data set.	Significant effort is likely to be required to assemble the information required into a format that allows it to be used for modelling the options for building retrofit.	It will not be possible to consider fabric retrofit as part of the development of local decarbonisation pathways. The options considered in pathway modelling could be limited by data availability.	Innovate UK 'Retrofit for the Future project'[10]. Information from previous local projects could be used.	May require specialist consultant support to generate the information required.

[8] http://www.eti.co.uk/
[9] https://www.theccc.org.uk/wp-content/uploads/2013/12/Review-of-potential-for-carbon-savings-from-residential-energy-efficiency-Final-report-A-160114.pdf
[10] https://retrofit.innovateuk.org/

3.12 Current and future weather

Network capacity requirements are typically defined by the peak load on the network. In the UK this will normally occur on the coldest winter days but this is not always the case. Since the objective is to understand future options from a whole systems perspective it is necessary to estimate current and future heating demand for different days of the year for different heating systems in different houses. Within EnergyPath Networks this is done for domestic buildings by performing a series of simulations using EnergyPlus[11]. This requires information on the current weather in the study area and can include predictions of future weather to account for the influence of climate change. For the pilot studies EnergyPlus weather files published by Exeter University were used.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
University of Exeter[12]	Assumed that weather data published by Exeter University is correct. Exeter produce future weather files for defined climate change scenarios. It is also assumed that these scenarios are realistic.	In order to perform building modelling to produce detailed energy demand curves it is necessary to have a definition of the weather conditions for different days of the year.	Weather files from Exeter University are free to use but have certain license conditions attached. Users should ensure that they have read and understood these restrictions.	None.	None required in the case of the pilot studies.	It will not be possible to perform building simulations and therefore to understand both peak and annual network loads for different heating systems in different buildings.	None found.	None.

[11] https://energyplus.net/
[12] http://emps.exeter.ac.uk/engineering/research/cee/research/prometheus/downloads/

3.13 Non-domestic building uses and floor areas

Data on the age and construction method of non-domestic buildings is not widely available and so accurately modelling the demands of all buildings in a local area is not possible. Non-domestic buildings energy demand can be estimated using benchmarks which categorise demand per unit floor area based on the use of the building. For the pilot projects this information was gathered from a variety of sources.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
OS topographic data and building classification, GeoInformation Group, Valuation Office Agency (VOA).	Ordnance Survey use a detailed classification for non-domestic buildings with classifications provided for those buildings with an address point. The GeoInformation non-domestic building classification has 25 different classes. This limits the ability to map this classification to detailed energy benchmarks. However, GeoInformation classification includes buildings on large sites such as schools and factories without	Non-domestic building uses and floor areas are required in as much detail as possible. VOA data is valuable to improve classification of buildings - particularly when one building has several tenants each using their space for a different type of business.	Data licensed from OS can only be used for the projects specified when the license is obtained. GeoInformation is commercial data available under 1 year licenses. VOA is commercial data available for a one-off cost if updates are not required.	If work is being done for a local authority then OS data can be made available under the Public Sector Mapping Agreement (PSMA) for the duration of the project. Otherwise this data must be licensed from OS. Cost depends on the map area required. Indicative cost is £6,000 for a local authority area for a 1 year license, with no cost if provided by an LA under PSMA. GeoInformation costs around £7,500 for a local	Different data sources provide different classifications. These need to be mapped to appropriate energy benchmarks. Decisions need to be made on the priority order in which datasets will be used to classify buildings. VOA data must be mapped to individual buildings on the OS map.	Quality of non-domestic dataset will be less good if this data is not available. Non-domestic building energy demands may be incorrectly calculated and inappropriate decarbonisation pathway options may be considered for buildings.	OS Mastermap classification. Valuation Office Agency. Display Energy Certificates could be used to improve information for buildings where they are available.	Surveying of individual non domestic buildings would be extremely valuable to improve the quality of data available but would take significant resources to conduct comprehensively.

	an address point in the OS data. It is particularly useful to classify these. Mapping of VOA data onto individual buildings requires matching with the OS data. The OS data does not always include the information to directly map by the building's reference number in the VOA data. In this case address matching can be required. This is a difficult and unreliable process because address fields are often incomplete or erratically entered.			authority area for domestic and non-domestic classifications. Depends on the number of buildings for which data is requested. VOA is a national data set cost around £7,000. Updates £3,000 p.a.				

3.14 Non-domestic building energy benchmarks

The data available on non-domestic buildings is insufficient to allow accurate building models to be built. Within EnergyPath Networks a benchmark process is used where energy demand is calculated based on a half-hourly demand profile and building floor area. Individual demand profiles are provided for each non-domestic building use categories. For the pilot studies these were produced using annual demand data for gas and electricity from the CARB2 project run by University College London (UCL)[13] and standard Chartered Institution of Building Services Engineers (CIBSE) energy benchmarks[14]. These were disaggregated to hourly profiles using Elexon Class Profiles[15].

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
UCL CARB2 project, CIBSE Benchmarks	Assumed that all buildings which are used for the same purpose have the same energy use per m^2 of floor area. Accurate demand profiles are not available at a detailed level so judgment must be used when allocating demand profiles to different building use classes.	Estimation of non-domestic building energy demands is essential for two purposes: 1) To estimate current and future network demands and required capacities. 2) To understand the options for connection of different buildings to district heat networks.	Each data source used for the pilot projects has a separate license for use. Users should ensure that they have read and understood these restrictions.	Depends on data sources used.	Significant processing required to produce detailed demand profiles at ½ hour time periods for different days of the year for each building use category.	Energy demands of non-domestic buildings will not be able to be calculated and so their influence on energy networks will not be included in the analysis of future options leading to inappropriate decision making.	None found.	None.

[13] http://www.ucl.ac.uk/energy-models/models/carb2
[14] The Chartered Institution of Building Services Engineers*: Energy benchmarks (TM46: 2008)*
[15] https://www.elexon.co.uk/operations-settlement/profiling/

3.15 Non-domestic building transition options

Energy demand of non-domestic buildings can make up a significant proportion of total demand in some areas. In order to meet local carbon reduction targets it is likely that these buildings will need to move away from using gas as the primary fuel for space heating. Options might be use of large electric heat pumps or connection to a district heat network. Large buildings (such as offices, sports centres and hospitals) can provide a significant and predictable demand when connected to heat networks and are normally required as 'anchor' loads in order to make heat network schemes commercially viable. It is important that switching of non-domestic buildings to low carbon heating solutions is considered in any analysis as their influence on local energy networks is significant.

For the pilot projects the only option considered for non-domestic buildings was connection to heat networks. In future projects the inclusion of electric heat and fabric retrofit options should be considered. In all cases the cost of options must be known alongside their influence on energy demand in order to perform the analysis.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
ESC internal data.	Assumed that: 1) Buildings that use gas for purposes other than space heating (such as industrial process heat) cannot switch to alternative fuel supplies. 2) Detailed knowledge of local buildings will not be available and representative options will have to be considered	Essential if consideration is to be given to tight carbon reduction targets as these will be unachievable without changes to how non-domestic buildings use energy.	Data availability will be dependent on the individual sources used for any particular study.	Unknown. Data processing costs could be significant.	Any available data is unlikely to be in a format required for analysis and will require processing to match the assumptions and calculation methods used in the analysis.	Any analysis will not be able to consider changes to non-domestic buildings and these aspects of local area energy planning will need to be considered in further work.	Some data is available from the Building Energy Efficiency Survey conducted by DECC / BEIS[16]. Information from previous local projects or studies could be used.	It is likely that a wider variety of sources will need to be consulted in order to develop a comprehensive catalogue of technology options. Surveying of individual non domestic buildings would be extremely valuable to improve the quality of data available but

[16] https://www.gov.uk/government/publications/building-energy-efficiency-survey-bees

	based on building use category.							would take significant resources to conduct comprehensively.

3.16 Current electricity network data

In order to understand the current electricity network some information is required on the location and capacity of existing network assets. EnergyPath Networks includes functionality to synthesise the feeder routings, connections and capacities if the substation locations and capacities are provided for HV (33kV/11kV) and LV (11kV/400V) substations. Network synthesis is then done by assuming that buildings are connected to their nearest substation as measured along the road network. Feeder connection data can be used if provided.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local electricity network operator	It is assumed that substation locations and capacities will be available. If the network is synthesised then it is assumed that buildings are connected to their nearest substation as measured along the road network, as long as the substation capacity is not compromised.	Electricity network data is essential for whole systems network analysis considering the influences of changes to buildings on network demands.	This data will need to be requested from the network operator and may only be provided under a Non-Disclosure Agreement. Network operators can be reluctant to provide customer connection data due to data protection concerns. If connection data is provided the primary key is meter number. It can be difficult to match meter numbers to addresses.	The underlying data is generally available at no cost but data processing costs are likely to be incurred in developing this data set.	If only substation location and capacity data is provided then little processing is required. If connection data is provided this can require significant processing to correctly identify which buildings are connected to which substations.	Without electricity network data (substation locations and capacities as a minimum) it is not possible to perform a whole systems analysis.	None.	In some areas there can be inset networks operated by a different network operator. In these cases data will have to be sourced from these other operators if inset networks are to be included in the analysis.

3.17 Location of high pressure gas entry points

As part of local area energy planning it is important to understand the level of local gas demand and have an idea of the gas flows in the local area. Within EnergyPath networks the points at which the gas network enters the study area and connections between different areas are used to represent the gas network in a simplified way.

Primary Data Source used in Pilot Studies	Assumptions / exclusions / limitations	Required / Optional	Availability / restrictions of use	Cost	Processing Required	Implications if not available	Alternative Sources	Additional data gathering procedures
Local gas network operators	All gas used in the study area is assumed to be imported through the gas entry points unless technology options are provided to produce gas within the area.	Some form of representation of the gas network is required to enable modelling of current and future gas demand.	Some network data might be publically available such as the network operator's Long Term Development Statement. Otherwise, this data will need to be requested from the network operator and may only be provided under a Non-Disclosure Agreement.	None.	Network data will be typically provided in Geographic Information System (GIS) format. This can then be used to define the gas entry points to the local area and to build a simplified gas network representation for modelling purposes.	If data is not available from the gas network operator then dummy data could be used but it would not, then, be possible to estimate gas flows around the area.	None.	None.

3.18 Off-gas grid postcodes

It is useful to understand which properties in the study area are not connected to the gas network. This helps to define both their current heating system and the cost of future options. If off-gas grid properties are to have heating systems that use gas then the cost of extending the gas network to them must be considered. Xoserve is the Central Data Service Provider for Britain's gas market. It produces a spreadsheet of postcodes that are off the gas grid[17].

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Xoserve	Assumed that the Xoserve data is correct as it is not easy to identify errors. For example, a group of electrically heated flats can appear as an off-grid island in the middle of an area of housing that is on-grid. It is not easy to establish if areas like this are erroneous. Note that Xoserve compile this data from other sources and do not verify it.	Essential to correctly understanding the extent of the current gas grid and the number of properties that are connected.	Publically available data. Updates are irregular so this data could be out of date in some areas.	None.	Local postcodes must be extracted from a national list.	If this data becomes unavailable either an alternative source will need to be found (possibly the local gas network operators) or distribution of off-gas grid properties will need to be inferred from sources like the English Housing Survey.	None found.	None.

[17] http://www.xoserve.com/wp-content/uploads/Off-Gas-Postcodes-V2.xlsx

3.19 Current and planned heat networks

It is important to correctly represent any existing, or planned heat networks when building a representation of the study area. This should include both identifying which buildings are connected to each network as well as the location of the energy centre and the heating systems it contains. For planned networks, costs should be provided if available. Otherwise the cost data listed in Section 3.20 will be used.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local Authority / Network Operator	Assumed that network operators can provide this data. Planned heat networks can be represented in two ways within EnergyPath Networks. They can be 'forced' such that they are built under all future scenarios or offered as options so that the deployment depends on the wider cost-optimal choices to be made.	Essential if the current situation is to be correctly represented	This data may be considered commercially sensitive by the heat network operator. Records may also include personal data which will need to be handled correctly with regard to data protection legislation.	The underlying data is generally available at no cost but data processing costs are likely to be incurred in developing this data set.	Considerable effort might be required to identify exactly which buildings on the OS map are connected to individual heat networks and to process the provided data into the correct format to allow analyses to be completed.	Without this data it must be assumed that there are no local heat networks and that all buildings currently use gas or electric heat solutions.	None	None

3.20 Energy network options

The options to build, reinforce and maintain gas, electricity and heat networks must be understood in order to correctly understand the implications of different choices. In addition the costs associated with decommissioning gas networks or re-purposing them to distribute hydrogen must also be considered. For the pilot projects this data was gathered from a variety of sources. Ideally costs should reflect the differences between, for example, hard and easy digs and between rural, suburban and urban working.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Electricity network operator reinforcement cost data. ETI Infrastructure Cost Calculator[18] for all networks and other ETI data for heat networks.	It is hard to identify the true costs of many network projects as they are commercially sensitive. Assumptions have to made that publically available data is representative of local costs. The ETI data used in the pilot projects is not currently publically available. Within EnergyPath Networks it is assumed that gas demand will	Essential	There is limited data that is currently publically available. It is likely that any energy planning project will have to work with local stakeholders to develop a good quality data set that is representative of the local area.	Unknown. Data processing costs are likely to be incurred in developing this data set.	Any data gathered is likely to require considerable processing to get it into a common format where assumptions are aligned so that it can be used for energy network modelling purposes.	Without good quality data on energy network costs whole systems analysis of decarbonisation pathways is not possible meaning that inappropriate decisions are likely to be made if choices for buildings do not consider their impact on the networks that supply them.	Previous local energy planning exercises may provide appropriate data to help with some aspects of assembling this data. Network operators are obliged to publish some data as part of their license agreements but does not contain all the information required. H21 Leeds Citigate[19] for gas network	May require specialist consultant support to generate the information required.

[18] http://www.eti.co.uk/programmes/energy-storage-distribution/infrastructure-cost-calculator
[19] https://www.northerngasnetworks.co.uk/wp-content/uploads/2017/04/H21-Report-Interactive-PDF-July-2016.compressed.pdf

	reduce due to decarbonisation and that current network capacity is sufficient to meet current demand. As such, there is no option to increase gas network capacity beyond extension of existing networks.						repurposing to hydrogen.	

3.21 Current and planned distributed generation and storage

It is important to correctly represent any existing or planned energy distribution and storage when building a representation of the study area. This should include both identifying the locations of existing sites and the technology types and capacities installed.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local Authority / Network Operator	Assumed that network operators can provide this data. Planned distributed energy schemes can be represented in two ways within EnergyPath Networks. They can be 'forced' such that they are built under all future scenarios or offered as options so that the deployment depends on the wider cost-optimal choices to be made.	Essential if the current situation is to be correctly represented	This data may be considered commercially sensitive by the site operators.	The underlying data is generally available at no cost but data processing costs are likely to be incurred in developing this data set.	Processing is likely to be required to either map technologies to those that already exist in the modelling technology catalogue or to create new technology definitions if the local technology is not already represented.	Without this data it must be assumed that there are no generation and storage assets. Analyses will not correctly reflect the current situation.	Site operators may be prepared to provide the information required.	None

3.22 Distributed energy generation and storage technology costs and efficiencies

Local energy generation and storage is expected to be a significant part of future energy systems. This might include facilities to provide heat for heat networks (such as gas boilers, gas combined heat and power plants, energy from waste plants and large scale heat pumps); local electricity generation using renewable technologies; and heat and battery storage. In order to understand the options for using these technologies the costs of building and operating them must be understood.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
The ETI's Energy System Modelling Environment national energy system model[20]. The ETI's Macro Distributed Energy project[21]	Assumptions will need to be made about technology development rates and how their cost and performance will change over time. In order to correctly understand how distributed energy technologies might fit into future energy systems it is necessary to understand how they can be expected to perform at different times of day and times of	Valuable in order to understand the costs and benefits of different future options and to aid decision making.	Data on the cost and performance of many technologies is commercially sensitive and can be hard to access.	Unknown. Data processing costs are likely to be incurred in developing this data set.	It is unlikely that all the information required to fully define any individual technology will be available from a single source. Definitions are likely to require that information from several sources is collated.	If data is not available for a particular technology then it cannot be considered in any technical analysis of energy system options.	Information from previous or planned local projects or studies could be used. ESC are unaware of any publically available data that provides all the information required to perform detailed analyses.	It is likely that a wide variety of sources will need to be consulted in order to develop a comprehensive catalogue of technology options.

[20] http://www.eti.co.uk/programmes/strategy/esme
[21] http://www.eti.co.uk/programmes/distributed-energy/macro-de/

	year. This is particularly true for technologies such as solar and wind.							

3.23 Potential locations for distributed energy generation and storage

In order to understand how new distributed energy generation and storage assets will influence local energy systems it is important to identify possible installation sites. These might include publically owned land or be incorporated into wider development planning exercises. Experience in the pilot projects suggests that most local areas have not completed this type of study across the whole of their local area.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local Authority data.	Once the technology options for different possible sites have been identified further work will be required to confirm that those technologies are suitable for those sites. Considerations might include air quality, noise, flood risk, conservation areas and existing infrastructure (for energy and transport)	Identification of these sites prior to completing an analysis will reduce uncertainty in the planning process.	Since this data must be generated locally there are unlikely to be restrictions on its use for planning purposes. However, there may be local concerns around particular locations so decisions will need to be made about the best time to make this information public.	Initial site identification might require input from engineering consultants to perform a high level screening process on all the possible sites that have been identified. The cost of this will vary depending on the scale of the task.	Once sites have been screened and confirmed as realistic options little further processing is required.	If sites are not identified then either distributed generation and storage cannot be considered in the analysis or possibly inappropriate sites will have to be used in the analysis such that its influence on energy networks may be incorrectly accounted for in the analysis.	None.	Initial site identification might require input from engineering consultants to perform a high level screening process on all the possible sites that have been identified.

3.24 National fuel prices and carbon contents, (boundary prices)

In order to understand the costs and benefits of different local decarbonisation pathways it is necessary to model the future costs of importing different types of energy into the local area. Alongside this the carbon content of those different fuels is likely to change over time and this projection must also be estimated. In many cases the cost of a particular energy source will be closely related to its carbon content. For example, the cheapest methods currently available to generate large quantities of electricity involve burning fossil fuels. If traditional gas and coal fired powered stations are to be replaced by low carbon generation alternatives then the cost of the electricity produced is likely to rise as its carbon content declines. Similarly, low carbon gas production and introduction to the gas network is likely to be more expensive than continuing to use natural gas.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
The ETI's Energy System Modelling Environment (ESME) national energy system model[22]	It is required that the costs of energy used in any modelling process correctly reflect their carbon contents. This requires some method of modelling the whole energy system in order to understand the interactions between the cost and performance of different options and their future mix.	Essential.	The full data used for the pilot projects is not currently publicly available. Other organisations that conduct the types of whole energy system analyses required to produce this data may be able to provide the information required.	Unknown. Data processing costs are likely to be incurred in developing this data set.	This data must be processed such that it is consistent with all other assumptions associated with analysing the local energy system. For example, the inclusion, or exclusion of tax, and the costs of energy transmission to the local area.	Any whole system analysis which considers changes to national energy supplies requires this information. Inappropriate local decisions will be made if they are not set in an appropriate national context.	Unknown. It may be possible to obtain this data from sources such as the Committee on Climate Change, BEIS or academic institutions such as UCL which conduct whole energy system analyses but using ESME data ensures consistency between carbon contents, energy prices and resource availability.	None

[22] http://www.eti.co.uk/programmes/strategy/esme

	It should be noted that any projections used will have considerable uncertainty associated with them and that variations in national solutions can have a profound influence on local options and choices. In all studies a range of national scenarios should be explored to aid understanding of these interactions.							

3.25 Land boundaries and development information

It is useful to understand wider spatial planning when conducting energy planning. Items of interest might include ward or other political boundaries, conservation areas, transport infrastructure or environmentally sensitive areas. Of particular importance is understanding areas where new development is planned. To correctly understand how energy demands are expected to change in the future the scale and timing of new development must be considered. This includes residential, commercial and industrial development.

Primary data source used in pilot studies	Assumptions / exclusions / limitations	Value of data in decision making	Availability / restrictions of use	Cost	Processing required	Implications if not available	Alternative sources	Additional data gathering procedures
Local Authority data. Ordnance Survey Boundary Line data.	When assessing future pathways, it will be assumed that planned development will come to fruition. If development does not occur, or occurs at different locations or different times to that planned then future energy demand estimates (and associated network capacities) will be incorrect. Assumptions must be made around the building standards and heating systems that will be built.	Planned development data is essential if future energy demands are to be correctly represented. Useful to ensure that energy projects fit within wider plans and respect existing constraints. Political boundaries can also be required for results visualisation, e.g. if stakeholders wanted to see the modelled heating system transitions related to different wards or communities.	This information should be readily available within all local authorities or even publicly available.	None.	Processing is required to convert information on planned development into information on exact types and sizes of buildings to be modelled. Little processing is required for Boundary Line data. Some of this data will be used as an overlay on energy system data to understand spatial relationships. This might just be for information or could require processing if particular spatial	Future energy demands will be under-estimated if new development is not considered. Proposed energy projects may be undeliverable if certain restrictions are not considered during their development.	None.	None.

					definitions are being used to restrict or force different options in energy system modelling (e.g. removing external insulation as an option in a conservation area).			

4 Validation data

Chapter 3 discussed the data required for whole systems analyses of local energy systems in order to understand the options for decarbonisation in a given local area.

This chapter discusses other data which was not used for local energy system modelling as part of the SSH Phase 1 programme but which has been utilised for validation of modelling outputs and to inform the local area energy planning process with local stakeholders.

Section	Data Description
4.1	BEIS domestic demand data at postcode level
4.2	BEIS non-domestic demand data at Middle Layer Super Output Area (MSOA) level
4.3	BEIS National Energy Efficiency Data-Framework (NEED) data
4.4	Network operator monitoring data for peak and annual demand
4.5	Local authority data on public building energy use
4.6	Energy Performance Certificates
4.7	Display Energy Certificates

Data Set	Validation Use	Strengths and Weaknesses
4.1 BEIS domestic demand data at postcode level[23]	This data can be used to check current gas and electricity demand estimates from building modelling.	Building demand estimates must be aggregated to postcode level to enable comparison to the BEIS data. There are several characteristics of the BEIS data that must be considered when performing this comparison. 1) These are estimates of energy demand at a postcode level. 2) The data is aggregated from building level data which is based on estimates as well as actual meter readings. 3) It does not include postcodes where there are less than 6 connections in the postcode. 4) The dataset includes some meter points that have low or no consumption which can have an impact on the median and average consumption values for that postcode. 5) Some meters and consumptions cannot be allocated to a postcode due to insufficient or incomplete address information. 6) If a meter's consumption has been overestimated in the past, the meter reading for subsequent years will be provided as negative consumption to compensate for the overestimation. In reality, these considerations mean that comparisons need to be performed at a level of aggregation that is larger than individual postcodes (e.g. at postal sector level). However, this data is still valuable to confirm that modelling assumptions are correct and to conduct more detailed correlations at lower levels of aggregation to investigate the energy demand of particular groups of buildings if required.
4.2 BEIS non-domestic demand data at MSOA level[24]	This data can be used to check current gas and electricity demand estimates for non-domestic buildings.	MSOAs cover an area comprising 5,000 people or 2,000 households. An individual MSOA can contain close to 1000 non-domestic buildings. There are 7,201 middle layer super output areas (MSOA) in England and Wales meaning this data is at quite a high level of aggregation. 1)Meters that have not successfully been assigned to a geography due to insufficient address information are not included in MSOA level data. 2)If the number of energy meters in an individual MSOA is too small this data is considered to be "disclosive" and the data for that MSOA is not included. When looking on a map the number of non-domestic buildings within any individual MSOA is generally larger than the number of meters specified in the BEIS data for that MSOA. This is true

[23] https://www.gov.uk/government/statistics/postcode-level-electricity-estimates-2015-experimental
https://www.gov.uk/government/statistics/postcode-level-gas-estimates-2015-experimental
[24] https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-electricity-consumption
https://www.gov.uk/government/statistics/lower-and-middle-super-output-areas-gas-consumption

		for both gas and electricity meters. Some discrepancy is expected as all the buildings on a single site might be fed through a single meter so the meter numbers in the BEIS data should be less than the number of buildings. However, the differences can be as large as 50% suggesting that the BEIS demand data does not include all demand within most MSOAs. This limits the ability to correlate this data with model outputs. Comparing energy use per meter from the BEIS data with energy use per building from model results can produce good correlation but if differences occur it can be hard to identify whether this is due to modelling problems or energy consumption that is missing from the BEIS data. Problems can also occur as MSOA boundaries do not necessarily align with administrative boundaries such that data is hard to correlate for MSOAs whose area lies partly outside the local authority area.
4.3 BEIS NEED data[25]	The NEED data framework matches gas and electricity consumption data, collected for BEIS sub-national energy consumption statistics, with information on energy efficiency measures installed in homes. It also includes data about property attributes and household characteristics, obtained from a range of sources. The data provides electricity and gas consumption for different house age and type broken down by floor area. This can be used to compare with energy consumption estimates for different buildings derived from modelling.	If the floor areas used to categorise buildings in the modelling do not match those in the NEED data then comparisons can be difficult. In some cases, NEED data could be based on a small number of properties. BEIS advise that figures based on fewer than 2,000 observations should be treated with caution.
4.4 Network operator monitoring data for peak and annual demand	This can be used to compare with modelled predictions of annual and peak demand on network assets.	Much of the electricity network is not routinely monitored. Distribution Network Operators publish data at HV substation level (33kV/11kV). Each primary substation will typically serve an area that is larger than a single local authority and so is of little use for correlation with model results at local authority scale.

[25] https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

		Some network operators can provide monitoring data at lower voltages which can be used to correlate both annual and peak demands. Correlation is difficult for those network assets that serve buildings both inside and outside the study area. Due to the meshed nature of gas networks they often cross study area boundaries in multiple places at several different pressures. This can make isolating gas network consumption and flow data that is linked to buildings within the study area particularly challenging. In one case a gas network operator could provide modelled values for peak gas flow across their gas distribution network. Some measured data was also available. However, this data was not time based such that the modelled peak flows covered a variety of different demand scenarios and did not represent a particular 'peak demand' event. Similarly, the measured values all occurred at different times and so could not be added up to give the total network demand. In addition, gas networks often have a grid configuration and the flow data provided gave no indication of the direction of gas flow at the time of peak flow. These two factors combined made this data unusable for correlation purposes.
4.5 Local authority data on public building energy use	Local authorities regularly collect data on the energy consumption of buildings in their estate. This can be used to correlate with the modelled energy demands for those buildings.	Where data was provided for the pilot projects this was often incomplete and did not include all buildings owned or run by the local authority. In addition, it was based on billed energy demand for a particular period which may not represent a 'typical' year as modelled. This data can be useful, however, to correlate non-domestic energy benchmarks and to check that assumptions around building use and building floor area in the model are appropriate.
4.6 Energy Performance Certificates	Energy Performance certificates are now available for a large number of buildings with some buildings having several certificates available if the building has been sold or rented several times in recent years. They include building characteristics that can be compared to those assigned in a modelling process. They also contain estimates of annual gas and electricity demands that can be compared to model estimates.	The quality of EPCs varies widely with some containing significant errors and omissions. In cases where building characteristics or energy consumptions given in an EPC do not match modelled values it can be hard to understand whether the EPC or the model is incorrect. The energy consumptions given in any EPC are modelled and heavily reliant on the quality of survey such that they are unreliable as a prediction of energy consumption. Any use of EPC data requires that the raw data is checked to attempt to identify and remove particularly bad examples. However, this can be hard in cases where errors and incorrect assumptions are not obvious.
4.7 Display Energy Certificates	Display Energy Certificates must be produced by owners and occupiers of	DECs only cover a limited range of buildings and so have limited value for model correlation purposes. The values displayed on a DEC are based on actual energy consumption but require

	buildings occupied by a public authority that are larger than 250m² and frequently visited by the public. They can be used to check modelled annual energy demands. They also provide floor areas which can be compared to those calculated from OS map data.	other building information (such as dimensions) which are heavily reliant on the quality of survey. It can often be found that the details given on two consecutive DECs for the same building are different. In addition critical information such as energy use or emissions can be missing[26]. For sites with multiple buildings such as schools and hospitals an aggregate DEC will sometimes be produced that covers some, or all, of the buildings on the site whilst other buildings on the same site will have individual DECs. It can be hard to understand which buildings on the site are covered by which certificate in these cases.

[26] https://www.cse.org.uk/projects/view/1259

5 Future data improvements

Data is considered to be a critical foundation of digitally enabled and whole system approaches to local area energy planning. The previous chapters outlined the main data sources that have been used for modelling and correlation purposes in the ESC's pilot projects with three local areas as well as alternative sources where these have been identified.

This chapter considers what data would be valuable in supporting local area energy planning in the future but which isn't currently available and how support for improving local energy data management could provide a number of benefits.

5.1 Valuable new data

The following data sets are considered to be those that would be most valuable and beneficial in supporting future local area Energy planning

5.1.1 Improved non-domestic building data

Due to the wide range of construction ages and types of non-domestic buildings it is not possible to classify them into archetypes that have similar energy performance as can be done for domestic buildings. To develop a credible strategy to decarbonise the non-domestic building stock a detailed knowledge of the both the buildings and the uses to which they are currently put is required. This would enable an analysis of the available options for different buildings to be performed and detailed decarbonisation strategies to be developed. Some work has been done at UCL to start to improve the understanding of the non-domestic building stock and the options for decarbonisation.[27]

5.1.2 Sources of heat for heat networks.

Whilst some information is available on heat available from water sources[28], research to establish the locations and magnitudes of other heat sources in local areas that might be upgradeable for heat networks would be valuable for energy planning purposes. For the most promising heat sources further analysis should be completed to identify the options for exploitation, any limitations or constraints on exploitation and the associated indicative costs. Heat sources that could be researched include, but are not limited to: waste water / sewage, heat in tunnels of underground railways, coastal heat, ground heat, industrial heat sources, ground or mine water and electricity substation heat.

5.1.3 Domestic hot water use demand profiles

There is little publically available data on domestic hot water use demand profiles in the UK. This would be valuable to improve understanding of the timing and level of demand in different domestic building types. This information could be developed from data gathered by Home Energy Management Systems such as that developed as part of the ETI's Smart Systems and Heat Programme[29].

[27] https://www.ucl.ac.uk/bartlett/energy/news/jul/2014/new-database-and-3d-model-energy-use-national-non-domestic-building-stock
[28] https://www.gov.uk/government/publications/water-source-heat-map-layer
[29] http://www.eti.co.uk/programmes/smart-systems-heat/home-energy-management-system-hems/

5.1.4 Hydrogen data

When the pilot projects were initiated little data was available on the production of low carbon hydrogen and the costs of re-purposing the gas network to provide hydrogen as a heating fuel. It was considered that sufficient data was not available to enable a robust analysis of this option. Since that time some information has been made available through the H21 project[30] whilst BEIS is also collecting further data to better understand this option.[31]

5.1.5 Low carbon / green gas

In all study areas the local gas network operator was interested in the options for use of low carbon gas. To understand and assess this several pieces of information are required.

1) The volume of low carbon gas that could be available to a local area.
2) The timescales on which different volumes of gas will be available.
3) The cost of producing low carbon gas.
4) The carbon content of the gas.
5) Whether the gas is mixed with natural gas or substitutes it completely.

A reliable, evidence based analysis that provided this detail would be valuable for local area energy planning.

5.1.6 Data on future installations

Local authorities currently collect information on installation of condensing boilers and double glazing through their building control departments. In future it would be extremely useful for this data collection to be expanded to include installations of low carbon heating systems and fabric retrofit at individual building level. This would allow both better local energy planning and tracking of actions against existing plans.

5.2 Local energy data management

Local areas face a major challenge in efficient access and effective management of data. Ad hoc projects in local areas proliferate many data fragments which quickly get lost or become outdated. In addition, data collected by individual departments within a local authority is often not readily available to other departments who may not even know of its existence.

Our work has identified that there could be a significant value in local energy data management as a foundation of local area energy planning in bringing together energy data for local areas. This would then have several benefits:

- It would enable a better understanding of current buildings and energy networks.
- It could help to identify areas with poor quality housing which may suffer from both fuel poverty and health problems.

[30] https://www.northerngasnetworks.co.uk/wp-content/uploads/2017/04/H21-Report-Interactive-PDF-July-2016.compressed.pdf
[31] http://www.government-online.net/hydrogen-supply-chain-technical-evidence-and-modelling-tool/

- It could help to identify areas where existing energy networks are constrained or have spare capacity.
- It could provide insight into a wide range of different activities in local areas including building retrofit, decarbonisation of heat, electrification of vehicles and deployment of low carbon and renewable energy.

Document Control

This Document

This document has been prepared by the Energy Systems Catapult on behalf of the Energy Technologies Institute LLP.

Relevant document release information is shown below. For further details of distribution restrictions, please refer to the "License/Disclaimer" provided, or contact the ETI (www.eti.co.uk).

Type:	Project Report
Subject:	Data for Local Area Energy System Modelling
Title:	Bidders Pack D3: Local Area Energy Planning - Data Challenges Report
ETI Project Number:	SS9016
ESC Project Number:	ESC00066
Version:	V 1.1
Status*:	Final
Restrictions**:	Confidential
Completion Date:	03-May-2018
Prepared by	Grant Tuff
Reviewed by	Natalie Lowery
Reviewed by	Richard Halsey
Approved by:: (Approval Denoted by Signature)	Chris Open

* Status defined as follows – **Draft**: Contains preliminary information only. **Released**: Contains reviewed and approved content.

** Restrictions defined as follows:
Public: Regarded as "within the public domain".
Confidential: Contains confidential information of the ETI and comprises intellectual property rights, including copyright, belonging to or licensed to the ETI.
Confidential (R): As Confidential, however certain information or data has been removed due to confidentiality, commercial, or license requirements. To request access to the full (Restricted) version, please refer to the document provider and/or contact the ETI (www.eti.co.uk).
Restricted: As Confidential, however additional restrictions apply (as detailed in this chapter) due to confidentiality, commercial, or license requirements.

Note that for all documents, copyright, trademark, license, and disclaimer restrictions apply as detailed within this chapter.

Revision History

Date	Version	Comments
03 May 2018	V1.1	First draft.
09 May 2018	V1.2	Initial Review and comments
31-May 2018	V2.0	Final, approved.