Programme Area: Smart Systems and Heat

Project: EnergyPath

Title: EnergyPath Networks Modelling Local Energy System Designs Report

Abstract:

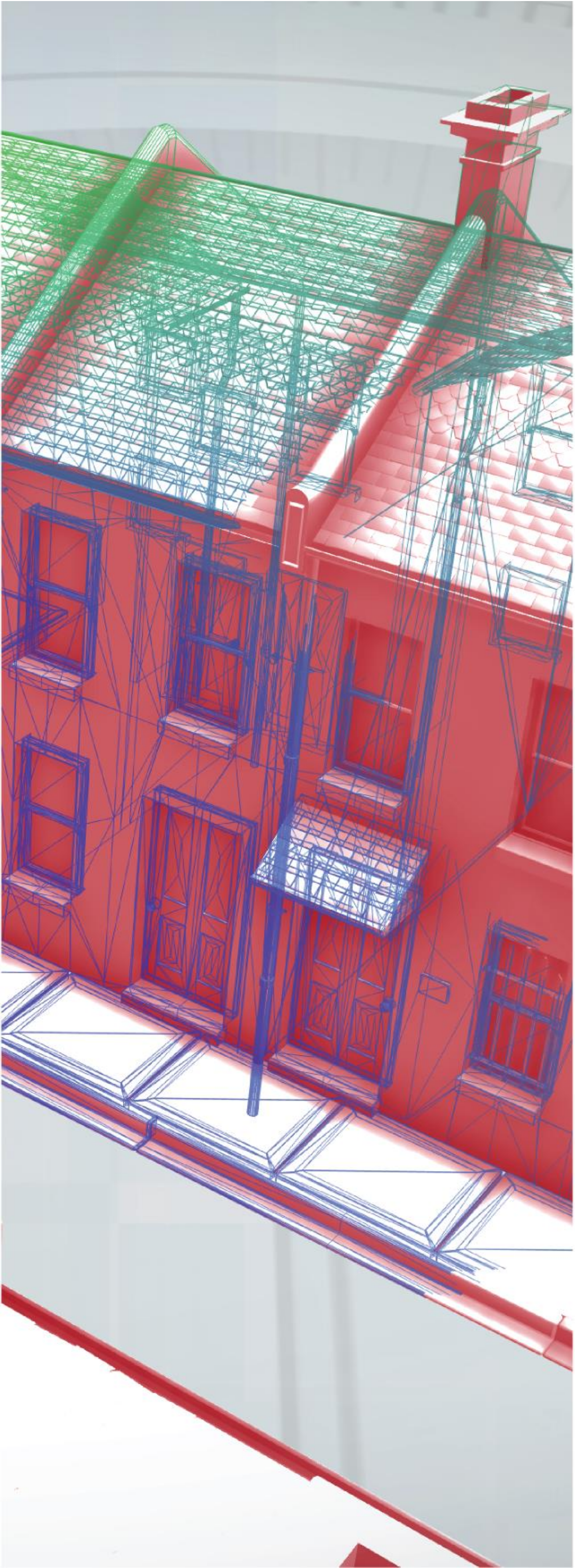
This deliverable package comprises:

- Bidders Pack deliverable D4, EnergyPath Networks Modelling Local Energy System Designs Report, incorporating:
- Bidders Pack deliverable D2 EPN Use-Cases Limitations and Uncertainty

This report is intended to provide an accessible summary of how EPN can be used to model local energy systems. It includes context on the development of EPN (including high-level interactions with other relevant tools such as ESME) and description of the scope of EPN modelling framework including Buildings, Technologies, Transmission/Distribution, Time, Cost, Carbon and Pathways.

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.



Smart Systems and Heat
Phase 1

EnergyPath™ Networks
Knowledge Transfer

Bidders Pack

D4 Modelling Local Energy
System Designs with
EnergyPath™ Networks

Final

Delivered by
CATAPULT
Energy Systems

Disclaimer

The information in this document is the property of the Energy Technologies Institute and may not be copied or communicated to a third party, or used for any purpose other than that for which it is supplied without the express written consent of Energy Technologies Institute LLP.

This information is given in good faith based upon the latest information available to the Energy Systems Catapult Limited and Energy Technologies Institute LLP, no warranty or representation is given concerning such information, which must not be taken as establishing any contractual or other commitment binding upon the Energy Systems Catapult Limited, Energy Technologies Institute LLP or any of its subsidiary or associated companies.

Contents

Acronyms

Table 1 Abbreviations and Acronyms

Acronym	Elaboration
BaU	Business as Usual
BCBC	Bridgend County Borough Council
BEIS	Department for Business, Energy & Industrial Strategy
BLPU	Basic Land and Property Unit
CAPEX	Capital Expenditure
CHP	Combined Heat and Power
DB	Database
DER	Distributed Energy Resources
DH	District Heating
DHN	District Heating Network
DNO	Distribution Network Operator
EHS	English Housing Survey
EPC	Energy Performance Certificate
EPN	EnergyPath™ Networks
ESC	Energy Systems Catapult
ESME	Energy System Modelling Environment
EST	Energy Savings Trust
ETI	Energy Technologies Institute
EV	Electric Vehicle
GIS	Geographical Information System
GOR	Government Office Region
GSHP	Ground Source Heat Pump
GUI	Graphical User Interface
HOM	Household Options Module
HSE	Health and Safety Executive
HV	High Voltage (11kV network)
LA	Local Authority
LiW	Living in Wales (survey)
LED	Light-Emitting Diode
LLP	Limited Liability Partnership
LP	Linear Program
LV	Low Voltage (400V network)
MCF	Master Control Framework
MIP	Mixed Integer Program
MPAN	Meter Point Administration Number
NAM	Network Analysis Module
NCC	Newcastle City Council
NDB	Non-Domestic Buildings
NPV	Net Present Value
NTS	National Transmission System
OS	Ordnance Survey

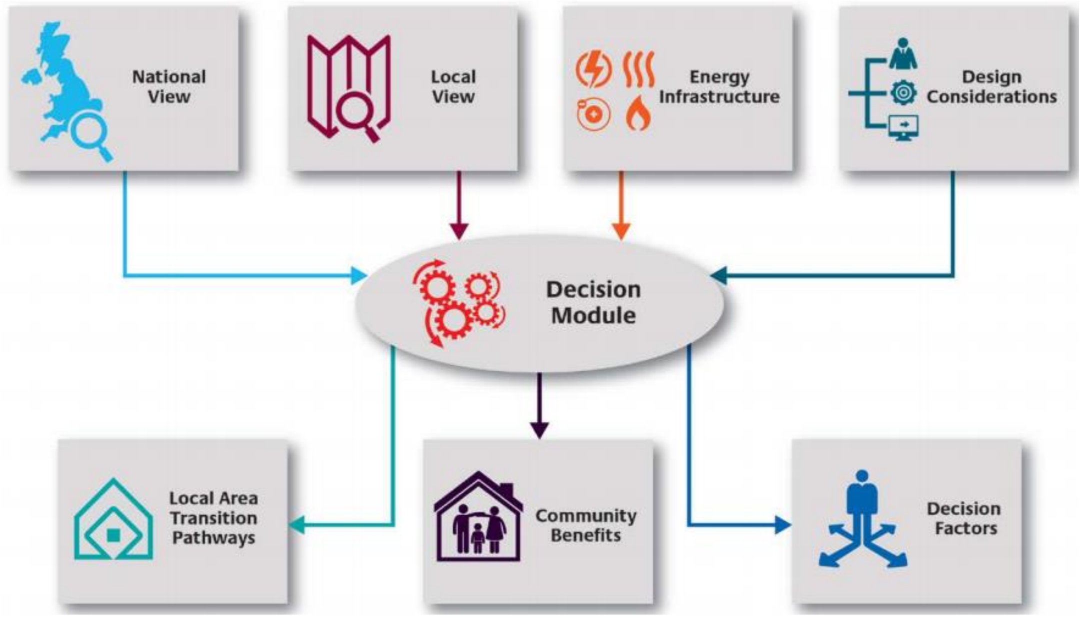
OTEoEH	Optimising Thermal Efficiency of Existing Housing (ETI project)
POM	Pathway Optimisation Module
PV	Photovoltaics
QA	Quality Assurance
SAM	Spatial Analysis Module
SAP	Standard Assessment Procedure
SHCS	Scottish House Condition Survey
SoS	Security of Supply
SQL	Structured Query Language
SSH	Smart Systems and Heat (programme)
UI	User Interface
UK	United Kingdom
UPRN	Unique Property Reference Number
TOID	Topographic Identifier
VOA	Valuation Office Agency
YHN	Your Homes Newcastle

Executive Summary

This document provides an overview of EnergyPath™ Networks (EPN) analysis framework, providing a summary of what EPN can be used for and its key design principles, components, input requirements and the main process steps involved in running EPN. This document is aimed at a technical audience familiar with energy system modelling and discusses the technical architecture and requirements of EPN. Three appendices are also provided outlining known limitations.

EPN is designed to help analyse multiple techno-economic pathways for decarbonising local area energy systems from now to 2050, by considering choices associated with buildings (heating systems and efficiency measures), networks (electricity, district heating and gas; including repurposing for hydrogen[1]), and distributed energy and storage. Figure 1 below provides a conceptual overview of the EPN analysis framework.

Figure 1 Conceptual overview of the EnergyPath™ Networks analysis framework



Using the EPN framework provides many benefits, the combination of which is believed to be unique, for example by:

- Taking a 'whole systems' view of a local energy system provides an independent and internally consistent framework for creating an evidence base to inform local area energy planning, whilst facilitating close collaboration with *all* of the key stakeholders (this can include the Local Authority, electricity and gas Distribution Network Operators (DNOs) and heat network developers).
- Basing the analysis on a detailed spatial representation of each local area, which helps to reflect the relationship between buildings and the networks that serve them and determine the associated costs and benefits of modelled pathways. Capturing the nuances of each real world

---

[1] Available with EPN R2.2.

local area also improves collaboration amongst stakeholders by giving them more confidence that their particular issues are well represented (e.g. the costs associated with network reinforcement in highly rural or urban areas).

- Using a least cost optimisation process to develop potential future decarbonisation pathways, which allows for a much wider set of options and trade-offs for the local area to be explored, compared to more manual scenario development. For example, the role of heat networks versus electrification of heat in multiple different configurations across any given local area.
- Ensuring the scope of the optimisation simultaneously covers buildings, distributed generation/storage and network options across multiple energy vectors and timescales (i.e. 10-year steps to 2050) helps to drive a more complete picture of how decarbonisation can be achieved whilst attempting to minimise total system costs, compared to analyses which only consider each vector in isolation.
- Incorporating a Monte Carlo mode of operation whereby a wide range of uncertainty in the value of future inputs (e.g. the costs of different heating systems or commodity prices) can be more easily explored, helping stakeholders to better understand which pathways and solutions are more robust to changing external conditions from now to 2050 and which cost inputs are more significant in terms of making decisions.

In addition, the detailed representation of the local area is beneficial for planning the future local energy system as it:

- Helps to provide a consistent and structured focal point to manage and audit available data on the local area energy system from a range of different stakeholders.
- Allows for area specific economic and wellbeing indicators to be derived by combining outputs from EPN's decarbonisation pathways with other socio-economic indicators to explore the impact on jobs, health, etc.

As part of a structured process of local area energy planning the outputs from EPN can be used to help answer questions such as:

- *What are the high-level features of the local area energy system pathways?* For example, how do total costs (capital and operational) vary for district-heating focused pathways versus those with electrified-heating over the horizon to 2050?
- *What are the key geographical and underlying features of different low carbon pathways?* For example, where should any decentralised energy centres supplying a heat network be sited? Or which areas and homes are most likely to be cost effectively served by electric heating?
- *What are the key points of uncertainty in the pathways and the value of reducing this?* For example, the sensitivity of delivered heat costs to key external factors such as long-term commodity or transmission level electricity prices.
- *What does the pathway imply for specific projects?* For example, the value in deploying an extensive building retrofit programme to improve the energy efficiency of homes in a specific part of a local area.

1 Introduction

This document provides an overview of the EnergyPath™ Networks (EPN) analysis framework. EPN has been produced with the objective of supporting the progression of Local Area Energy Planning in the UK. Local Area Energy Planning is seen by the ETI and the ESC as central to achieving national greenhouse gas emissions reduction targets and supporting the network choices needed to decarbonise heat. This document provides a detailed technical description of the EnergyPath™ Networks analysis framework, suitable for a technical audience, including the scope of the framework and associated limitations, describing:

- In Section 2 - What EnergyPath™ Networks can be used for, its key design principles followed by an overview of EPNs analysis framework; summarising the analysis framework's main modules and functions
- In section 3 – The key input data requirements and process steps when running the analysis framework
- In Section 4 - The technical architecture and requirements of the analysis framework

Three appendices are also provided outlining known limitations of the EPN analysis framework, describing the impact modelling choices could have on results, outlining how uncertainty related to data can influence results and highlighting the main challenges associated with using EPN.

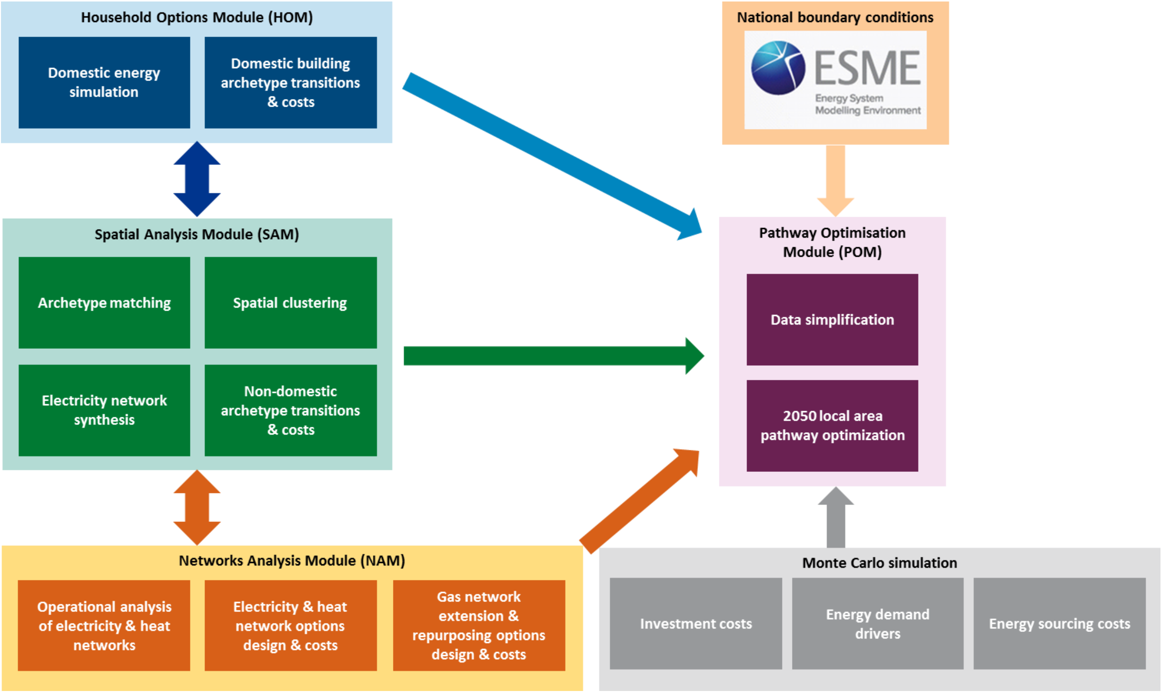
Background

EPN has been developed under the remit of the Energy Technologies Institute's (ETI) Smart System and Heat (SSH) programme. Phase One of the SSH programme was delivered by the Energy Systems Catapult (ESC) on behalf of the ETI, working with three Local Authority areas in Newcastle, Bridgend and Bury in Greater Manchester, to pilot a process of local energy systems planning in these areas. Scoping (of EPN) began in early 2014 with 'Release 2.0' deployed at the end of 2015 as part of work for Newcastle City Council (NCC).

The EPN Analysis Framework

EPN is an analysis framework that aims to support investigation of future local energy system pathways, initially focused on decarbonising the delivery of heat in a given local area, as well as providing evidence that can be used to inform long-term policy decisions on the impact of decarbonisation and energy network choices. For example, outputs can be used to assess the impact on energy prices of energy system changes, considering specific local energy infrastructure and plans whilst remaining consistent with a national decarbonisation pathway. A conceptual overview of this framework is depicted in Figure 2.

Figure 2 Technical overview of the EnergyPath™ Networks analysis framework



The core process within the EPN analysis framework, illustrated in the figure above, comprises two steps:

1. The first is to create a highly detailed and bottom-up representation of the current local area energy system and its future choices for buildings, networks and distributed energy options across the local area's spatial topography; via the *Household Options Module* (HOM), *Spatial Analysis Module* (SAM) and *Network Analysis Module* (NAM).

2. This representation is then simplified – based on user-defined levels of granularity – such that these future choices can be evaluated and traded-off against each other over the pathway to 2050 within a cost-optimisation engine; the *Pathways Optimisation Module* (POM). This minimises the total energy system costs (investment, operation, resource) across the whole pathway and local area simultaneously whilst ensuring all constraints are satisfied (e.g. that energy supply is sufficient to meet demand or CO_2 emissions are below target levels). The POM can also undertake Monte Carlo simulation of key inputs parameters (e.g. technology costs, commodity prices) to more easily explore uncertainty in the long-term pathway design.

EPN is a data intensive model, given the need to represent the existing local area as accurately as possible before simplifying this for the pathway optimisation. As part of this several commercial datasets are used (e.g. Ordnance Survey) along with information from DNOs and the local area to build a robust picture of current network topology, and where and what type of buildings currently exist in the local area, along with associated energy demands. EPN is also a complex and computationally intensive analysis framework requiring a powerful, dedicated workstation. Its design structure is highly modularised integrating several commercial 3rd-party tools[2] along with various bespoke components developed in Python[3].

---

[2] Including ArcGIS for spatial analysis, AIMMS for the POM optimisation, PSS Sincal for heat and electricity load flow modelling, EnergyPlus for detailed building energy simulation, @Risk for Monte Carlo simulation of inputs and MSSQL Commercial for data management.

[3] https://www.python.org/

The Objectives of EnergyPath^{TM} Networks

During its development, the overarching objectives for the EPN analysis framework were as follows:

- To create a strategic planning analysis framework to support decision-making on future local area energy systems to 2050, principally to help support relevant stakeholders such as Local Authorities (LAs) and Distribution Network Operators (DNOs);
- More specifically, for the analysis framework to provide evidence to help prioritise and plan interventions in the local area energy system, including generation, network, storage and buildings projects, aligned with a national 2050 decarbonisation pathway;
- Using the outputs of the analysis framework to focus on the development and application of a cost-effective Local Area Energy Strategy, which is subject to 'real world' constraints and uncertainty. This helps to inform, more subjective, strategic decision making on a transition pathway as part of wider business planning processes;
- To be able to account, at least indirectly, for the potential impact of consumers on local area energy system pathway design;
- To ensure that the analysis framework and associated databases are based around a scalable architecture, to set the ETI/ESC on the trajectory towards building full functionality and capability.

2 EnergyPath^{TM} Networks scope and approach

EPN is designed to help create and analyse techno-economic pathways for decarbonising local area energy systems from now to 2050, by considering choices associated with buildings (heating systems and efficiency measures), networks (electricity, district heating and gas), and distributed energy and storage.

2.1 Use cases

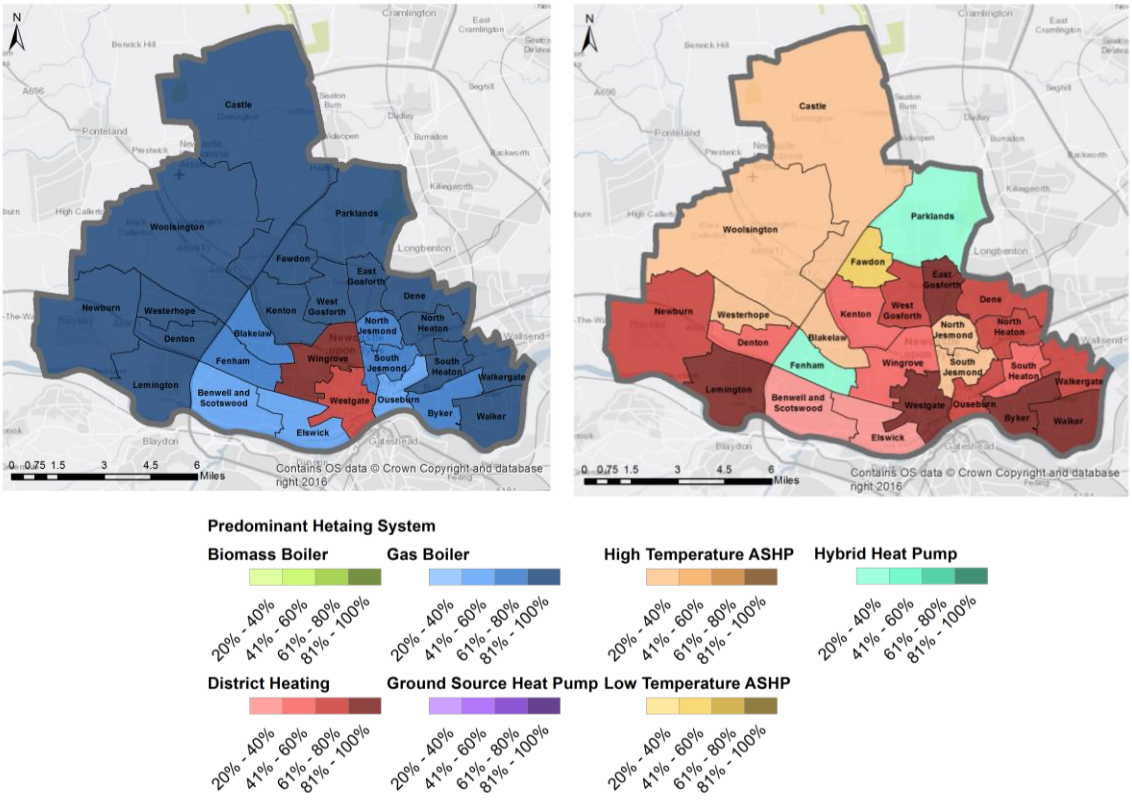
EPN is a flexible analysis framework that can be used in part or in full for several different use cases, including:

- **Local Area Energy Planning:** for organisations such as Local/Regional Authorities, Network Operators and Infrastructure/Housing Developers. This includes developing Local Area Energy Strategy for future energy networks; setting local energy objectives consistent with national targets (e.g. CO_2 emission reductions), or helping to specify local interventions, for example targeted retrofit or heat network expansion; in terms of identifying target areas and potential demonstration or deployment projects.
- **Local Area Existing Energy System Insights:** for Local Authorities. The EPN process is designed to first create a detailed picture of the local area's energy system, collating the best available data in a structured format, and applying techniques to fill data gaps. This may help identify early areas for intervention (e.g. target areas for building fabric retrofit programmes) ahead of creating a full Local Area Energy Strategy.
- **Pathways for Growth Towns and Cities:** for Local Authorities working with developers and network operators to better understand how major/accelerated growth of new homes and buildings near to existing towns and cities may alter the Local Area Energy Strategy (e.g. the potential for anchor deployment of district heat in the new development with extension to the wider existing local area).
- **Alternative Network Investment Pathway Optimisation:** for Network Operators, planning a multi-vector approach to assessing investment across an operator's licence area and to demonstrate the minimisation of cost for consumers as part of future price control regulation. For example, more explicit consideration of gas and electricity network interactions through the use of hybrid heating systems to manage peak load.

Test Bed for Innovations: for government institutions and new technology developers to help understand the potential role of future technologies that deliver whole system benefits and to inform the policy or market framework that would support deployment of the associated technology.For example, as part of a process of local area energy planning the outputs from EPN can be interrogated in various ways - and at various levels of granularity – to help answer questions such as:

- **What are the high-level features of future local energy scenarios?** For example, how total costs (capital and operational) vary for district-heating focused pathways versus those with electrified-heating over the horizon to 2050; how household energy costs are impacted[4]; and the technologies which are key in implementing the pathway. Figure 3 below shows predominant domestic heating system deployment in different areas of Newcastle in 2050 (the more deployment the darker the colour), comparing a Business as Usual scenario (left) to a future local energy scenario of deep decarbonisation (right).

Figure 3 Predominant domestic heating systems deployment in 2050



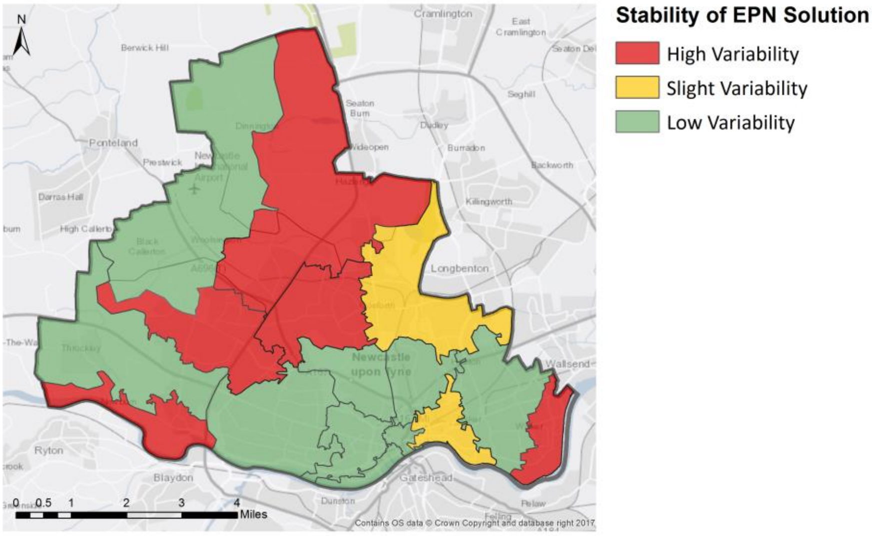
- **What are the key geographical and underlying features of different future local energy scenarios?** For example, how would a new heat network be laid out; what would be the location of plant supplying this network; and how would a map of cost of energy provided overlap with the map of fuel poverty?

- **What are the key points of uncertainty in the pathways and the value of reducing this?** For example, how results are distributed around the average and the risk of high cost outcomes; the reduction in cost uncertainty achieved from obtaining better data, for example on the installation cost of a particular measure in an area; and the key external factors to which a pathway is sensitive. Figure 4 illustrates the variability in deployment in district heating given Monte Carlo simulation of key uncertain inputs in EPN. This also serves to illustrate a level of confidence in deployment decisions, which could be translated into low-regret technology investment plans when all factors outside of the modelling framework are considered.

---

[4] As opposed to energy bills directly, as this is ultimately affected by the distributional considerations of various regulatory and market structures, for example whether network costs are incurred directly by the end-user (on a capacity or utilisation basis) or socialised across a wider set of end-users.

Figure 4 **Variability in deployment of domestic heating systems in 2050**



- **What does the pathway imply for specific projects?** For example, the roll out of a fabric retrofit programme focused on solid wall insulation in a specific part of the local area; informing the case for new distributed energy storage, or the viability of transitioning a small geographic location to electric heating and potentially decommissioning the gas grid.

2.2 Key design principles

The EPN Analysis framework has been built to the following design principles:

- **The nature of the EPN Analysis framework:** the analysis framework's overall objective when helping to construct a local area energy system pathway is to minimise total energy system costs whilst ensuring all system design standards are met (e.g. target comfort levels or emission limits), Total system costs are defined as cost to society associated with different energy system choices over a time horizon to 2050. They do not include taxes or subsidies as these are transfer payments. Future costs are discounted to 2015.
- **Co-optimisation of energy system choices:** building, technology and network choices are optimised together by considering trade-offs across multiple energy vectors and technology choices including buildings, networks and other energy system features such as distributed generation. This means that the whole system cost of energy choices is considered. For example, the unit cost of storage heaters is cheap whilst heat pumps are considerably more expensive for the same power output. However, a heat pump can require less than a third of the power required by a storage heater to deliver the same output power. From a whole system perspective the network reinforcement cost saved by installing heat pumps can more than overcome the additional heating unit cost compared to storage heaters. Co-optimisation of energy system choices allows these trade-offs to be understood and accounted for.
- **Analysis of uncertainty:** EnergyPath Networks is designed to help the user understand the resilience of particular low carbon pathways or deployment options. Exploration of the impact of uncertainty on future pathways can be explored through sensitivity analysis by changing individual input parameters. Alternatively, distributions can be defined for input parameters

which can then be sampled to form collections of different input parameters that are solved as a Monte Carlo analysis (see section 2.4).

- **Flexibility in required level of resolution:** the analysis framework is designed around an overarching vision of "data-driven" flexibility with respect to spatial resolution. This is based upon the use of a granular representation of local area building and networks and their future options, that can easily accommodate more real-world data as it becomes available (e.g. from building surveys). EPN accurately represents the costs and technical parameters of individual technology and network choices for the specific local area; as opposed to a more abstract representation for a region or country as whole. This can be subject to specific maximum limits in complexity within the networks and spatial modules. The analysis framework also supports optimisation at different resolutions to provide control over the computational effort required for solving. The user can aggregate up *time periods* or geographical coverage, or drill down to greater granularity, to the extent the data allows it.

- **Ability to apply real-world constraints:** for example, planning restrictions can be applied and physical constraints based on local knowledge and data. In addition, whilst consumers are not modelled as entities within the model itself, the indirect impact of consumer behaviour (on energy demand or technology uptake) on future local energy scenarios and energy system designs can be tested, examples could include assessing:

    - *Consumer technology preferences:* in the selection of building options such as insulation retrofits or new heating systems. Whilst the default approach assumes selection of building option based on cost, the implications of different consumer preferences on pathway designs can be tested. For example, this could be via adjusting the costs of discount rates applied to different technologies to monetise the 'hassle costs' or force in a given expected deployment of certain options.

    - *Behaviours that affect energy use:* as above the default assumption is one of economic rationality, for example the use of storage within a building to minimise overall energy use and costs whilst maintaining comfort levels, however, it is possible to test the impact of different consumer behaviours and subsequent energy demand profiles on the pathway design.

- **Linking to national level boundary conditions:** the primary links to the national level are through price of carbon and other nationally delivered resources, such as centrally generated electricity and natural gas. This retains flexibility in choices at the local level but considers the associated costs. The pathway optimiser design enables the user to impose additional constraints (such as a local area carbon emissions target) if required.

- **Flexibility in accommodating data:** the analysis framework maximises the use of existing data and a number of elements of anticipated future data, which are likely to be available in some, but not all LAs. For example, where detailed knowledge of individual buildings is available this can be used. This might come from social housing asset records or Energy Performance Certificates. For buildings where less information is available this is synthesised based on assuming that these buildings statistically match national data such as that available from the English Housing Survey.

- **Post processing integration of other data sets:** the analysis framework supports the layering on to the results of other data (usually though a GIS) to support end-user requirements, for example combining energy system cost with socio-economic datasets to explore the implications of a pathway on fuel poverty.

Importantly, the analysis framework is not intended to be used to identify a single fixed pathway (or scenario); EPN is not designed to forecast what will happen in a local area due to policy impacts, consumer behaviour, degree of market competition, etc. Instead, it is designed to provide an indication of potential future energy system options or choices, not predictions. For example, if EPN is being used for Local Area Energy System planning, it would be prudent to explore multiple possible scenarios, using combinations of assumptions and constraints to identify lowest-cost decarbonisation pathways. The modelled outcomes would not provide a definitive view of the future but represent an informed view of 'feasible futures'. They would provide an indication of which futures could be more or less costly and which potential decarbonisation pathways could be explored further.

In addition, when the analysis framework is used to produce a local area energy system pathway it assumes that achieving the pathway is possible; aspects such as the decision making needed by all actors to enable the pathway (e.g. government, regulators & consumers etc.) and real market behaviours are not modelled. However, constraints can be applied in the model as needed, for example, energy system choices can be restricted where they are deemed unsuitable for particular areas or buildings. Local knowledge and data can be applied to support this activity.

The high-level features of EPN are summarised in Table 2 below:

Table 2 Overview of key principles underpinning the EPN analysis framework

Category	Issue
Local energy systems coverage	Tight geographical scope (typically covering a single Local Authority area)
	Detailed representation within geography
	Multiple energy vectors: electricity, heat, gas, hydrogen
	All local heating energy demand, networks and supply infrastructure
Building centric	Energy demands anchored around actual buildings in an area, their underlying energy needs, and their ability to retrofit
Pathway optimisation	Assesses a range of potential investment options to give optimal investment pathway to reach a decarbonised state in 2050
	Optimises for lowest cost solution that meets all energy needs, SoS and decarbonisation targets
	Perfect foresight optimisation gives optimal pathway for deterministic cases, uncertainties are represented through scenarios or Monte Carlo simulations
Policy agnostic	Considers fundamental techno-economic information
	Does not consider real-world policy and regulation, nor specific businesscase constraints

2.2.1 Where EPN sits within the broader modelling landscape

Existing local area planning tools[5] can target one or more of the elements set out within the Key Design Principles section, but not all in combination. The closest to EPN is UrbEn[6] (integrated modelling

---

[5] A list of these tools has been compiled in the "Strategic Modelling Tools" document as part of this Bidders' Pack work stream.

[6] http://www3.imperial.ac.uk/urbanenergysystems

framework for urban energy systems – also known as "SynCity"), which was produced by Imperial College London as part of a BP funded research project. The question UrbEn is trying to tackle is, however, much broader than EPN and extends to transport planning and agent behaviour, which means there is correspondingly less detail in the pathway optimisation for the local energy system as a whole and more limited resolution of building energy use.

A core aspect of the EPN analysis framework is to bridge appropriately the various dimensions of spatial and temporal granularity and multiple trade-offs across different parts of the energy system - in particular the relationship between building energy demands and energy network build and reinforcement - over the full pathway to 2050. This is particularly important given the long-lead times for larger-scale network infrastructure upgrade and development. As it is likely to be intractable to hold significant detail on all of these aspects simultaneously, it is important that the end-user can easily flex the level of detail in different parts of the analysis framework and understand the impact on the pathway design.

Although the inputs to the pathway optimisation process are likely to be simplified across one or more of the dimensions it is important that the base input data (particularly for the understanding of building energy demands and networks) are as detailed as possible. This allows different aspects of the problem to be simplified to different degrees providing flexibility to analyse aspects of interest at different levels of detail. In contrast UrbEn and other models start with a simplified representation of the energy system to be studied which fixes the granularity and detail at which problems can be considered. The EPN approach inverts the problem, undertaking a detailed assessment of possible options *first* (e.g. via network flow modelling) and then parameterising this to inform the options in the pathway optimisation.

To enable the above, EPN has drawn on a number of well-established modelling practices in key areas such as network flow modelling, dynamic building energy simulation, and pathway optimisation. These have been used to increase both the sophistication of the analysis framework and make its development as cost-effective as possible.

2.3 EPN analysis framework overview

The EPN analysis framework can be divided conceptually into four main **modules** (see Figure 5 below), which include the:

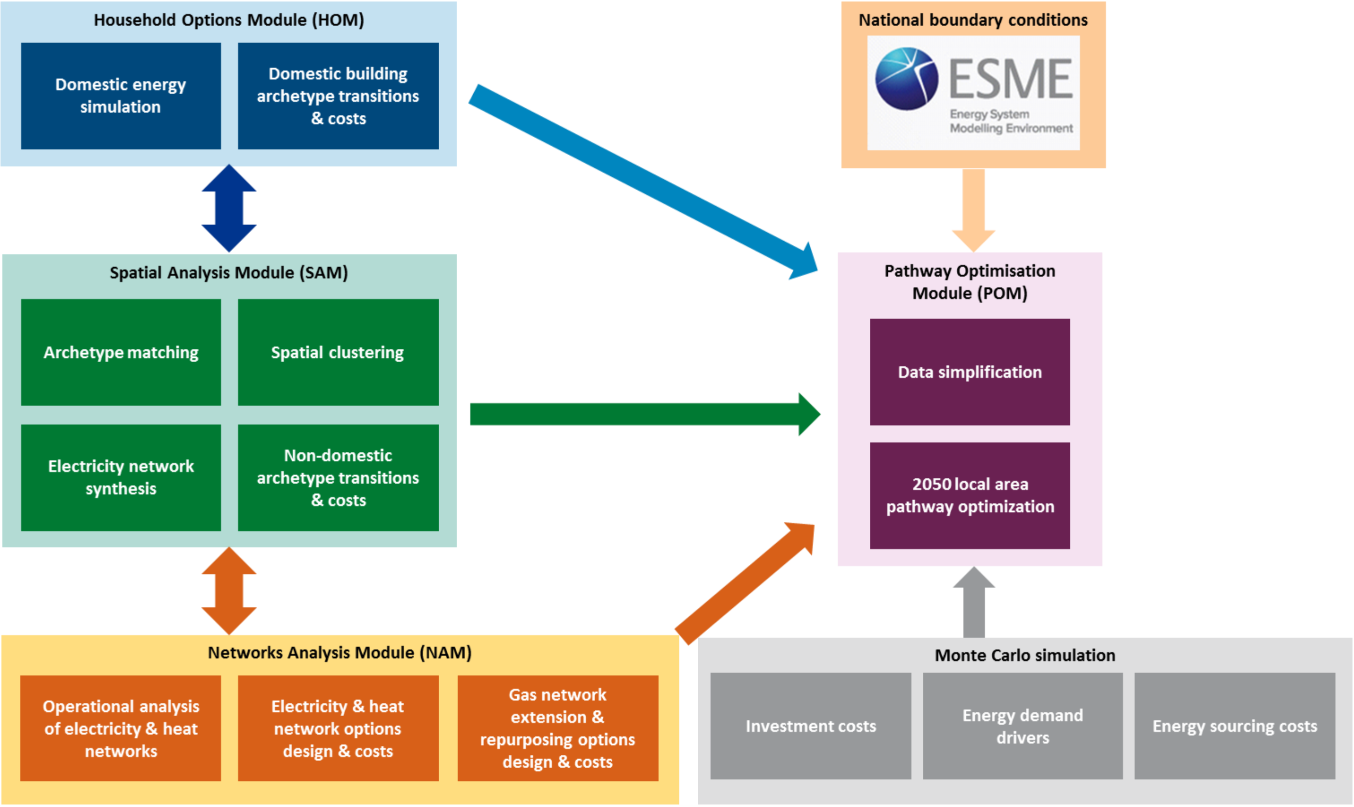
- **Household Options Module (HOM)**
- **Spatial Analysis Module (SAM)**
- **Networks Analysis Module (NAM)**
- **Pathway Optimisation Module (POM)**

As noted above, the core process within the EPN modelling framework comprises two steps:

- The first is to create a highly detailed and bottom-up representation of the current local area energy system and its future choices for buildings, networks and distributed energy options across the local area's spatial topography; via the **HOM**, **SAM and NAM modules**.

- This representation is then simplified – based on user-defined levels of granularity – such that future choices can be evaluated and traded-off against each other over the pathway to 2050 within a cost-optimisation engine; via the **POM module**. The **POM module** minimises the total energy system costs (investment, operation, resource) across the whole pathway and local area simultaneously whilst ensuring all constraints are satisfied (e.g. that energy supply is sufficient to meet demand or carbon emissions are below target levels[7]).

Figure 5 Technical overview of the EnergyPath™ Networks analysis framework



Each module operates independently, with key data that is shared between the modules. Modules are further subdivided into a number of discrete steps. A single end-to-end "run" of EPN consists of the sequential completion of a series of steps across all modules. A high-level overview of each module and their interactions across the analysis framework is shown in above. This also highlights the analysis frameworks interaction with ESME and the Monte Carlo Simulation (discussed in section 2.4).

ESME

In the context of decarbonisation of local energy systems it is important to have a view of how national energy systems will decarbonise and the influence this will have on local options. For example, the cost and carbon content of nationally generated electricity imported into the local area will have a fundamental impact on the cost and carbon emissions of that local area. In EnergyPath networks this view of the future national energy system is provided through use of a national pathway produced using ESME.

---

[7] Decarbonisation pathways can also be explored by adding a carbon price (i.e. as cost that the optimiser must minimise) rather than an explicit emissions cap.

ESME (Energy System Modelling Environment) was developed by the Energy Technologies Institute. It is a least-cost optimisation model designed to explore technology options for a carbon-constrained UK energy system, subject to additional constraints around energy security, peak energy demand and more. ESME covers the power, transport, buildings and industry sectors, and the infrastructure that underpins them, in five year time-steps from 2010 to 2050. It provides least cost decarbonisation pathways for the whole UK energy system in a similar way to how EnergyPath Networks does for local energy systems.

2.3.1 Household Options Module (HOM)

The primary role of the **HOM** is to generate a database of domestic building archetypes that can be used to represent the existing building stock in an area and the options for evolving this stock over time (referred to as "pathways"). Specifically, the HOM provides other modules with data on energy consumption, peak demand and the costs of retrofit interventions (including insulation, heating systems, etc.) in different building types as they transition over time to 2050.

Domestic building "archetypes"

The characteristics of buildings are represented in the HOM using a so-called "archetype" approach. This is similar to the creation of a large catalogue or a reference library of pre-defined building types against which data on real world buildings can be matched. This library is refined by information from the SAM (see 2.3.2) reflecting which buildings have been matched to which archetypes across the different parts of the local area (i.e. if certain types of buildings do not exist locally they do not need to be included in the final catalogue[8]).

A challenge for the HOM is refining the many possible characteristics used to identify and represent real world domestic buildings to only those which are most material for understanding the key outputs of energy consumption, peak demand and retrofit costs in a particular local area. This is achieved using mathematical clustering (k-means [9]) techniques to reduce the number of building types into a manageable number; where it is possible to aggregate similar building types into one archetype with limited loss of information. To support this selection process the HOM uses a fast-parameterised SAP (Standard Assessment Procedure) model to simulate the typical annual levels of electricity and heat provision required to maintain comfort conditions in each dwelling type. The approach used is aligned with the method used for assessing building regulations compliance in the UK[10].

Domestic energy simulations

Once the short-list of the most relevant (for the local area) domestic building archetypes has been created, bottom-up dynamic thermal modelling of space heating and hot water, using the EnergyPlus[11] software, is then used to generate possible demand profiles for a number of 'characteristics days[12]' at 30-minute granularity. Electricity profiles for lighting and other appliances, which may indirectly affect energy demand through internal gains, but are not affected by the weather are also included. The EnergyPlus simulations are undertaken for a range of different heating technology options for each

---

[8] This is a deliberate data architecture design decision to help simplify the final pathway optimization undertaken within the POM.

[9] Note that clustering here refers to a specific set of mathematical techniques and is separate to the broader, model specific concept of spatial clustering in EPN.

[10] For these calculations, EnergyPath™ employs an ISO 13790 model based on the UK Standard Assessment Procedure (SAP) methodology

[11] https://energyplus.net/

[12] As part of this assessment it is necessary to simulate a period of time around the day of interest to account for e.g. the slower variation in thermal mass

building archetype, including gas boilers, heat pumps, hybrids, district heating, etc., as well simulations for the technology options when combined with features such as storage (including various size of store). EnergyPlus also allows for modelling solar hot water as well as solar photovoltaic (PV) output profiles based on the same weather file used for temperature.

The number and definition of characteristic days that can be considered is flexible and can cover typical average seasonal days (winter, spring, summer, autumn) as well as more extreme cases to test the resilience of the pathway or to design to a specific standard, for example, to meet a 1-in-20 or a 1-in-50 cold winter day.

Domestic building retrofitting

The HOM also generates domestic building transitions for both heating system combinations (including storage and heat control) as well as insulation measures. The availability of transitions is controlled by factors to minimise the maximum number of combinations, such as:

- Heating systems are only replaced at the end of their technical life;
- Once a heating system transitions to a low carbon option it cannot switch to a higher carbon option later in the pathway;
- Insulation should always be replaced by equivalent or stronger insulation measures;
- Enhanced insulation measures are installed in the same calendar year as heating system retrofits to minimise disruption.

Once the transitions are generated, the HOM calculates their costs based on the building stock underpinning the “pathway”, based on the number of buildings assigned to each archetype; this information can then be assessed in the POM.

Non-Domestic Buildings

The additional complexity and variation of Non-Domestic Buildings (NDB) means that their energy requirements are treated in a more simplistic manner, with user-defined energy benchmarks applied to a relatively small number of non-domestic “activity classes” as opposed to fundamental modelling of energy demands in EnergyPlus.

In addition, the user must adjust future energy benchmarks as an *indirect* way of representing efficiency improvements or heating system changes. These pathways are passed to the POM as investment decisions to be considered in the cost minimization exercise.

2.3.2 Spatial Analysis Module (SAM)

The role of the SAM is to create a detailed local area representation that contains sufficient spatial data to enable the:

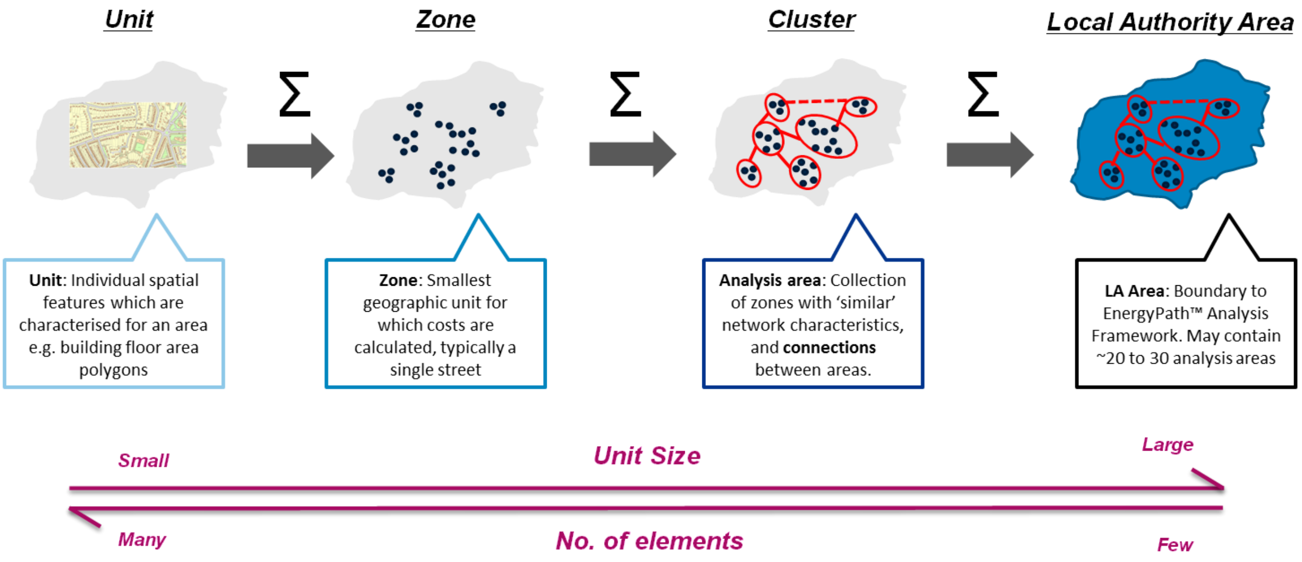
- HOM (see section 2.3.1) to focus its archetypal grouping analysis on the relevant domestic buildings for the considered local area;
- NAM (see section 2.3.3) to assess the costs and feasibility of potential network reinforcements and new build;

- POM (see section 2.3.4) to generate evidence based energy system pathway choices for the area and elements within it (e.g. buildings, networks upgrades, etc.).

Spatial clustering

As part of its internal processes the SAM needs to be able to simplify the spatial representation of a real local area by aggregating it into larger regions to make the optimisation process tractable, as shown in Figure 6. The spatial information associated with each of these analysis areas and the number of each building archetype contained within them are then passed to the POM for whole system energy cost optimisation. The user can define the analysis areas manually (with semi-automated mechanisms to help streamline this process) or via an auto-cluster algorithm, which can suggest a set of starting areas for the user (see section 2.3.3).

Figure 6 Overview of spatial aggregation simplification process



Representing bespoke spatial features

The SAM is designed to develop geographic information system (GIS)-based layers containing detailed spatial and topographical information about the local area, for example containing the specific location of each building and the current road network.

In addition to existing spatial data, the SAM also provides an interface for the user to enter future spatial data (i.e. that assumed to occur as part of a scenario) or future options that could be utilised or impact on the pathway design via the integration of ArcGIS into EPN. This can include, but is not restricted to: geothermal or biomass availability, possible sites for large heat sources or embedded generation for Distributed Energy Resources (DER), sites for potential new build housing, existing PV uptake, Electric Vehicle (EV) uptake projections, etc. This information is stored in the GIS layers and passed to the POM for the pathway analysis. Other existing data sets can also be used such as the boundary for a conservation area restricting archetype upgrades.

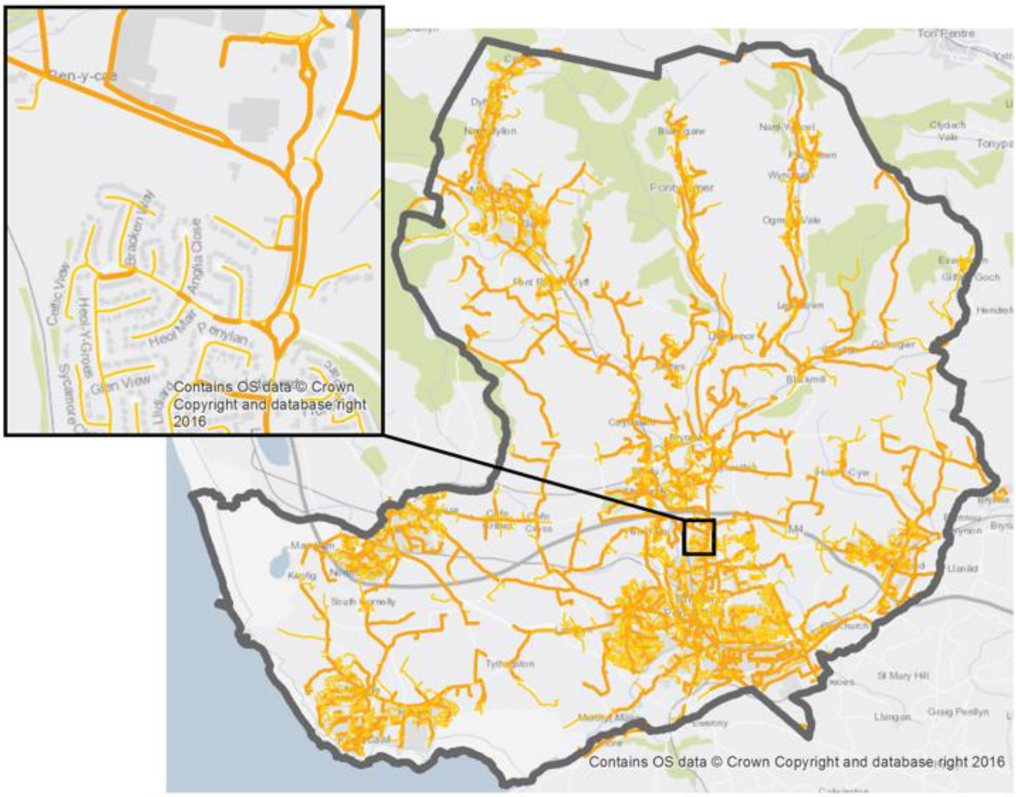
Buildings characterisation and matching

The SAM matches the most detailed set of domestic building archetypes (defined in the HOM) and non-domestic activity classes to each individual building address in the area. The matched building archetypes are also passed back to the HOM to help better define the potential upgrade pathways for these buildings as they are aggregated as part of the simplification process, grouping similar buildings together for the purposes of decision making in the POM.

Electricity and heat network synthesis

The SAM synthesises a representation of the electricity distribution network serving the area by connecting DNO-specified sub-station locations to individual buildings (subject to shortest path routing and other factors). The user can also incorporate further DNO data about the typology of the network where it is available. An example output from the detailed EPN electricity network synthesis is shown in Figure 7.

Figure 7 Example of detailed EPN electricity network synthesis[13]


Contains OS data © Crown Copyright and database right 2016

Contains OS data © Crown Copyright and database right 2016

The topology of existing, as well as potential, heat networks can be represented in the SAM. While representing the span of existing heat networks requires user input, the NAM can generate heat network topologies consistent with the spatial clustering and the electricity network topology. Potential heat network topology is anchored around a 'load weighted centre' within each spatial analysis area. This helps to define a point where a heat generation/source provides heat to pipes laid in a radial manner along the shortest path to serve building demands.

Additionally, the user can configure possible inter-area transmission line connections for gas and heat manually. A functionality to represent meshed electricity networks at High Voltage (HV) feeder level is also available.

---

[13] © Local Government Information House Limited copyright and database rights 2016, licence number 100057254

The topological network information and building archetype information is passed to the NAM, such that the NAM can calculate electricity and heat network upgrade cost curves, and gas network extension or hydrogen repurposing[14] costs for individual network elements (e.g. pipes, feeders, substations etc.).

2.3.3 Network Analysis Module (NAM)

The NAM provides the detailed analysis for quantifying potential network upgrade and new build options (their technical characteristics and costs) associated with an increase in peak energy demand for electricity, heat, gas and hydrogen. These options are then simplified and passed to the POM (see section 2.3.4) as part of the pathway analysis.

The NAM contains three main key components, each with a distinct set of requirements:

- The ‘operational analysis’ sub-module, which is responsible for performing load flow studies using PSS®Sincal[15] to simulate the steady-state operation of electricity and heat networks (existing and potential) in the area under investigation at different levels of load. For heat networks, the user can define different combinations of operating pressure and temperature to test, which may lead to different combinations of heat losses and electricity requirements for pumps.
- The ‘network options design and cost’ sub-module, which is responsible for generating costed options for all zones, analysis areas and their connections. These cost functions are based on network upgrade or new-build costs depending on the number of user-defined options that are available (e.g. multiple sizes of heat pipe or feeder cable).
- The ‘Gas and H_2 Options’ sub-module generates a simple representation of options and costs for the gas network, whereby the optimiser is given the choice between keeping the existing gas network, extending it to cover off-gas grid buildings, decommissioning it or repurposing it for use with hydrogen for each analysis area. Costs of extending the gas network are estimated in relation to the length of the road network from existing gas infrastructure to new connection points.

NAM and SAM interactions

The NAM interacts closely with the SAM and together they help to define the final spatial analysis area boundaries, which are used to simplify the geographical representation of the area, whilst trying to ensure that:

- These boundaries do not cut across infrastructure that should be treated as a single item for the purposes of pathway decision making in the POM. For example, assigning half of the buildings on a single electricity feeder to one analysis area and the other half to another.
- The costs of network options are as representative as possible across all vectors within an area, for example, trying to avoid combining areas with very high and low-cost upgrade options so that the low-cost option is ‘lost’ as it cannot be chosen in the optimisation process without also

---

[14] Available with EPN R2.2

[15] http://w3.siemens.com/smartgrid/global/en/products-systems-solutions/software-solutions/planning-data-management-software/Pages/overview.aspx

choosing the high cost option. This is done to avoid unfairly biasing the choices given to the POM in a particular area (e.g. towards district heat ahead of electricity or hydrogen)

- The total number of analysis areas leads to a tractable optimisation problem in the POM.

2.3.4 Pathway Optimization Module (POM)

As described in the previous sections the HOM / SAM / NAM modules help to create a detailed spatial picture of the existing local energy system, along with the ***options*** available for evolving the system over time, such as building retrofits or installing a new district heat network. The modules also create the necessary data to evaluate these options including cost, performance and various other characteristics both now and in the future. These detailed understandings are then reduced to a more simplified form so that the **POM** can compare them against each other in a manageable manner – i.e. the optimisation problem can be solved in a ‘reasonable’ time. For example a detailed cost-capacity curve for electricity network reinforcement that contains 200 individual cost-capacity points might be reduced to a simplified curve with 7 points that are sufficient to capture the main characteristics of the detailed, underlying curve. These include capacity points where a large investment is required to achieve a small step in network capacity or parts of the curve where large capacity increases can be achieved through small investments.

The **POM** is the engine that constructs the local energy system pathway from the available options by deciding:

- **What *options*** could be deployed, **where** and **when** (e.g. network development/reinforcement, building energy/efficiency and distributed energy)?
- For some ***options*** *(*e.g. energy centres or storage*)*, **how** should these be utilised once deployed?

The process needs to consider whether the options that are chosen as part of an energy system pathway are collectively appropriate (i.e. are delivered at lowest total system cost) and satisfy various other goals (e.g. relevant for building type, ensure households are provided with sufficient comfort, is consistent with meeting security of supply constraints and climate change targets).

The national pathway produced by the ETI’s ESME model is also used to inform some of the **boundary conditions** within the POM that are not modelled endogenously such as carbon and fossil fuel prices, and the availability, price and carbon intensity of transmission-connected electricity.

The POM also allows the user to more efficiently investigate key areas of uncertainty, such as the future cost and performance of different technologies or the future cost of energy resources, within the development of a local area pathway design, via the use of a Monte Carlo simulation process (see section 2.4).

Temporal representation

There are two distinct aspects to the representation of time in the POM:

- **Time periods**: which reflect the years across the pathway, from the starting year 2014 to 2050, within which the POM is able to make a decision around when energy system changes occurs[16].

---

[16] Except for building archetype groups and their pathways, where the decision is made for the pathway as a whole, but each pathway choice implicitly reflects a number of transition points for the buildings in the group over time.

Data is provided by the other modules at an annual level but is averaged across time periods to simplify the problem for the POM. While the averaging is flexible, and user defined, in practice 10-year steps have been used. The underlying modules calculate all data for individual years. For each parameter the average value over each 10-year step is calculated for input to the POM.

- **Time slices:** these represent periods within the year for which the POM must balance the operational supply and demand of energy (accounting for factors such as losses) across the local area system. They are further categorised into characteristic seasons and within-day periods for each season. As per *time periods*, the *time slice* data is produced at a higher level of granularity in the other modules (e.g. building energy demand profiles) by peak[17], winter, spring, summer, autumn season and at half-hourly level within day, but these can be aggregated to simplify the problem that must be solved in the POM.

POM model ontology

The key entities within optimisation, created primarily via the data provided by other modules, are as follows:

- **Products[18]**: are produced, transmitted and consumed within the energy system. One product can be converted to different products using an appropriate technology. For example, a gas boiler converts the gas product into two products, heat and CO_2. Products cover *energy resources* (e.g. biomass, natural gas), *energy carriers* produced by technologies in the model (e.g. a CHP plant producing heat and electricity) which are sub-divided by grid level[19], *CO_2 emissions* and '*tradable products*[20]' at the boundary of the local area such as transmission-level electricity.
- **Archetype groups:** within each analysis area domestic (and non-domestic) buildings with similar characteristics (e.g. starting energy demand, heating system, building shell features) are grouped together for the purpose of decision making. Each group has a number of possible pathway options with different costs and energy demands, from which the POM can choose. The pathways reflect the transition of the buildings in that group (from the 2014 base year to 2050) to different heating systems and/or building shell efficiency measures. This pathway representation helps to simplify the complexity processed by the optimiser and as a result improves processing time.
- **Energy supply technologies**: which represent a process that converts one or more input products into one or more output products. These are used primarily to represent local distributed energy resource such as small-scale solar or wind farms or energy supply centres (e.g. CHP, boilers, large scale heat pumps) that provide energy into local networks.
- **Storage technologies**: are capable of storing energy in one *time slice* and releasing it in another. These are used primarily to represent electricity distribution network storage such as batteries

---

[17] This can be defined by the user but is broadly used to represent an e.g. 1-in-10 or 1-in-20 cold weather period.

[18] Unlike other energy system models, such as ESME, end-use services for heat, transport, etc. are not present in the POM as these have already been accounted for and translated into energy requirements in previous modules. For example, the HOM calculates the final energy demand for a given building archetype to meet a target temperature level.

[19] For example, High-Voltage 11kV and Low Voltage 400kV products.

[20] Such that EPN can choose to 'import' or 'export' into/out of the local area, respectively.

or large-scale district heat network storage, but can also be used to represent small-scale building battery storage[21].

- **Network technologies**: used to manage the flow of products *within* a spatial analysis area for electricity, gas and heat. These cover transformers and feeders at different grid levels for electricity, and heat exchangers and pipes for heat networks. For example, an LV substation would take in the HV electricity product and convert this into an LV product for use in buildings (or vice versa to export surplus electricity to higher grid levels).
- **Transmission technologies:** these represent the networks which manage the flow of electricity, heat and gas or hydrogen products *between* the spatial analysis areas defined by the user.

Key optimisation components

The POM objective function represents the total local area system costs that must be minimised as part of finding an optimal least-cost solution. This being the net present value[22] of the sum of all costs across the pathway from 2014 to 2050 including:

- All resource and tradable product costs (these are a negative cost where energy is exported from the local area).
- The sum of all annualised[23] investment, fixed and variable operating costs for building archetype groups (reflecting heating system and efficiency improvements), energy supply and storage technologies, and network and transmission technologies.

As part of finding an optimal least-cost solution, the POM can change the value of key ***decision variables***, including:

- What quantity of resource / tradable products are sourced by analysis area (where they are available), by *time period* by *time slice*.
- The choice of domestic/non-domestic building archetype pathway by archetype group by analysis area. As noted previously, there is no direct time dependence in the pathway choices as they embed information about the possible transitions over time (and e.g. the impact on costs and energy demands). This pathway structure helps to reduce complexity and improve solving time in the optimisation.
- The choice of energy supply and storage technologies, where these need to reflect lumpy investments and mutually exclusive choices by analysis area by build year vintage[24]. Lumpy investments are those that have fixed capital costs associated with individual choices. For example, CHP plant is manufactured at particular capacities where each capacity option has its own cost. It is not possible to buy a CHP plant whose capacity is mid-way between two available options and pay the average price of the two available options. The options available are discrete choices which give a lumpy set of cost-capacity investment choices rather than a

---

[21] Note that the impact of building heat storage is represented implicitly in the final energy demand profiles produced for different building archetypes by EnergyPlus in the HOM.

[22] Subject to a user defined discount rate, usually aligned to the Treasury's Green Book rate of 3.5%

[23] These are annualised to reflect user defined assumptions on the cost of capital by technology and to account for only the costs incurred within to 2050. E.g. the objective function will only see 10 years' worth of costs for an investment with an economic life of 25 years built in 2040.

[24] I.e. the year the technology is first commissioned, this technology will then be active in future *time periods* until it reaches the end of its technical life. For example, a CHP built in 2020 will still be active in 2030 and 2040.

sloping cost-capacity curve where any point on the curve can be chosen. The capacity of other technologies (e.g. building level solar PV) can be represented appropriately by an incremental expansion, by analysis area by build year vintage.

- The operation – or activity – of technologies within each analysis area in each *time period* in each *time slice*. In some cases, these technologies may also have multiple modes of operation, for example a CHP plant with a number of different heat to power ratios. Analogous decision variables manage the operation of storage in terms of its injection and withdrawal.
- The choice of network build options (new build or reinforcement) by network type, by analysis area by build year (of the network). Analogous decision variables also exist for new transmission network capacity installed by area pairs (i.e. between areas), by build year.
- Network activity variables representing the flow on each network technology by analysis area, by *time period*, by *time slice* within year and by mode of operation. The modes reflect the ability to flow both up and down different network levels in the case of electricity. For heat networks these modes represent different combinations of temperature and pressure operating regimes where these have been parameterised for testing in the NAM. Analogous activity variables also exist to represent the operation of network transmission technologies flowing products between areas.

The freedom to change these decision variables is subject to ***constraints***, these split into mandatory constraints that are always active and user-defined constraints for fixing aspects of the solution or testing a particular scenario. Key mandatory constraints include those which:

- Ensure the **balancing of supply and demand** (accounting for losses, spill, etc.) by product by analysis area by *time period* by *time slice.* These constraints are supplemented by security of supply constraints to ensure supply capacity can meet technical design requirements (e.g. capacity margin);
- **Enforce a single, mutually exclusive choice** for building archetype pathways, network, energy supply technology or storage options, by analysis area by build year vintage;
- **Constrain the maximum operation** of a network, energy supply or storage technology to **within technical availability** by technology, by analysis area, by *time period*, by *time slice*.

Key optional constraints include:

- Limiting the maximum, or forcing a minimum, build of network capacity by analysis area by *time period*;
- Forcing specific energy supply or storage technology options to be built, by analysis area by *time period*. Note that restricting the availability of these options is undertaken by restricting these in the pre-processing stage of the POM rather than an explicit constraint.
- Restricting the maximum quantity of CO_2 emissions that can be produced in the local area by *time period*. The user can define whether, for example, the CO_2 emissions associated with imported electricity at the boundary of the local area should be included in this total. The user can also drive decarbonisation via use of a carbon price instead of a quantity restriction.

- Restrictions on the maximum resource that is available by analysis area by *time period*, for example, to reflect restrictions on local waste availability for incineration.

It should be noted that for performance reasons some constraints (e.g. forcing building transitions) are handled by passing only the relevant options to the POM[25], rather than via explicit constraints. These building-related constraints can be imposed for the analysis area as a whole or for individual building archetype groups within an analysis area (e.g. to target efficiency measures to a subset of buildings) and by timing of intervention across the pathway[26].

The POM has been designed to reflect both "lumpy" investment choices for technologies and networks[27] as well as mutually exclusive choices, for example, to reflect two or more choices for an energy centre sited on the same land area in a city centre[28]. This makes the POM optimisation a Mixed Integer Problem where discrete choices must be made between individual options as opposed to a Linear Problem where any value can be chosen for each option (as ESME for instance).

Representation of peak demand

Instantaneous peak demand is the key driver of network reinforcement, but it is challenging to represent this within the POM due to computational limitations. Although the temporal simplification process is flexible, dividing a peak day into three diurnal *time slices* still makes it difficult to calculate both the underlying demand and network headroom at the point of greatest stress. As a result, the POM represents this indirectly via two main mechanisms:

- An estimate of **peak demand diversity -** i.e. reflecting the fact that the highest point of demand across many buildings connected to the same part of the network is unlikely to be fully coincident. Demand diversity can be implemented in EPN via a pre-processing adjustment to the POM input data or via a set of dynamic variables and constraints that link the outturn level of demand diversity to the choices made within the POM (i.e. the numbers of buildings shifting to district heat or electrified heating). The dynamic approach is the more accurate representation but has significant performance implications for the optimiser.
- **Peak security of supply margin** constraints. For each network type (electricity, heat, gas) and grid level, this constraint tries to ensure there is sufficient headroom available on the network to meet a proxy of actual peak demand. Where this is not sufficient, the optimiser has either to reduce demand or install more network capacity. The input to the constraint adjusts the peak demand, by adding a mark-up to the highest demand calculated in the main supply/demand balance constraints. In addition, available network headroom is also adjusted, by applying a de-rating factor to the installed network capacity. This could act to reduce the available headroom, for example by reflecting N-1 contingency requirements, or conversely these could be offset by user-defined estimates of the benefits of direct smart network operation[29].

---

[25] I.e. in the case of forcing a particular pathway by offering it as the only available option.

[26] Although each pathway is only a single decision, defining a constraint which says all buildings after 2030 must be connected to a heat network will only retain available pathways where this condition holds.

[27] I.e. to reflect that it is not always possible to incrementally expand their capacity over time at the same per unit cost.

[28] These can be extended to compare two 'packages' of options each of which may contain multiple technologies and/or storage options, but where the packages as a whole are still mutually exclusive.

[29] Such as Active Network Management, or Real Time Thermal Ratings on transformers, which are difficult to model in EPN due to the limited time granularity that can be accommodated in the POM and its use of perfect foresight optimisation.

Problem simplification in the POM

To make the problem of comparing multiple options more tractable, their representation in the POM must be **simplified relative to the other modules**. Simplification naturally involves a loss of detail and may then affect the options chosen. To counter this, in-built flexibility allows the user to easily vary the level of aggregation/simplification and understand how the solution changes as you move along this spectrum. Key simplifications include the number of:

- *Time periods* on the pathway to 2050 (e.g. annual to 10-year steps);
- Number of *time slices* within year for which supply-demand balances are resolved (e.g. half-hourly daily demand profiles by season to a peak day and an annual average day, each with three within day *time slices*[30]);
- Number of spatial analysis areas in the SAM (e.g. 20 compared to thousands of underlying zones or streets);
- Number of building pathways (i.e. the set of future choices for building efficiency and heating system upgrades) and analogous set of user-defined non-domestic pathways in the SAM. This is driven by a variety of factors e.g. the available heating and insulation options and heuristics guiding the transitions as described in section 2.3.1;
- Number of discrete network investment options for electricity and heat as defined in the NAM per component type (e.g. heat pipe, HV feeder), analysis area and *time period*. For gas and hydrogen, the number of options only vary based on the number of analysis areas[31].

EPN has been designed so that the detailed, bottom-up representation of the local area buildings, network options, etc. can be simplified before being sent to the pathway optimiser. This is particularly important as the optimiser is a Mixed Integer Program, where run-times can increase dramatically as a function of increasing problem complexity (e.g. number of variables and constraints).

The main features where granularity can be adjusted before passing information to the optimiser are represented in Table 3 below.

---

[30] Overnight, Midday, Evening

[31] Maintain existing gas network, decommission, extend network to off gas-grid, repurpose gas network to hydrogen, extend and repurpose to hydrogen.

Table 3 Summary of data simplifications for the pathway optimiser

Category	Data simplification
Time and space	The number of discrete spatial analysis areas used to sub-divide a local area, as opposed to representing each individual building or electricity feeder as a discrete object in the pathway decision-making process. Figure 6 represents the various stages of spatial aggregation (from individual buildings to local area-wide) used in EPN. The typical number of spatial analysis areas used to-date has been in the order of 10-20, focused around buildings connected HV sub-stations.
	The number of time periods used to represent the choices on the pathway from now to 2050, e.g. 10-yearly steps as opposed to annual.
	The number of within year *time slices* used to model the operational flows associated with balancing of supply and demand across the local energy system.
Building archetypes design	The maximum number of possible options for converting a building's heating system or improving its energy efficiency over the pathway to 2050. For example, analysis to-date has used 3 possible insulation retrofit packages and 21 heating system combinations, with two opportunities for change (due to the assumed lifetime of a heating system) across the pathway to 2050.
	The extent to which buildings that are in close proximity and which share similar energy characteristics (e.g. in terms of their size and heat loss rated) can be treated as the same for making decisions about their conversion.
	Non-domestic buildings are grouped according to user-defined energy demand and retrofitting cost levels. These are in turn determined based on building characteristics e.g. activity class, size, etc.
Network options design	Number of final network reinforcement options by year, component type and energy vector.

2.4 Uncertainty modelling

2.4.1 Systematic exploration of uncertainty

Analysis of pathways for decarbonising local energy systems is subject to multiple areas of uncertainty, many of which become increasingly uncertain as a pathway moves further into the future. For example, future technology costs and commodity prices etc. EPN implements two approaches for exploring some aspects of uncertainty in designing local energy system pathways:

- **Scenario/sensitivity analysis**, to generate pathways consistent with particular scenarios e.g. high prices, deployment of selected energy centres.
- **Monte Carlo simulations**, to explore the influence of input parameters on solution variability (or stability/robustness).

These are explained briefly below, further details on uncertainty handing within EPN can be found in the "*EnergyPath™ Networks Functional Specifications*" document[32].

---

[32] The document is available here: http://www.eti.co.uk/programmes/smart-systems-heat/energypath

2.4.2 Scenario analysis

EPN is designed to allow users to more easily explore the implications of different drivers on least cost decarbonisation pathways. This might include the relative cost of building heat networks compared to reinforcing the existing electricity network or re-purposing the gas network to use hydrogen or the influence of particular energy or technology costs. Whilst all of EPN's exogenous inputs can be re-configured as part of testing a new scenario, EPN also includes a scenario analysis feature providing timesaving facilities for e.g. running the model in batches as well as storing input / results data sets related to several scenarios in the same database. This feature can handle parallel data sets for:

- **Domestic archetype demands**: these parameters can impact building archetype feasibility as far as meeting energy demands
    - **Target comfort temperature profile** defining heating requirements (with cooling to be included in EPN v2.2)
    - **Weather file scenario** including simulated weather for future years, driving domestic demand as well as solar PV output
- **Non-domestic archetype demands**: several scenarios can be parameterized off-model
- **Spatial clustering**: several sets of spatial analysis areas and their connections can be defined using the built-in GIS user interface e.g. to focus on areas for likely district heat network deployment.

2.4.3 Monte Carlo simulation

EPN also includes a Monte Carlo simulation function to explore the local energy pathway variation subject to analysing sets of randomly sampled and potentially correlated input parameters. Each pathway simulation – corresponding to one set of simulated inputs – is still solved deterministically with perfect foresight (the process assumes that the least cost pathway can be perfectly planned for and achieved), as per the core EPN running mode. However, this functionality allows users to more easily assess which input parameters have the largest impact on the chosen solution e.g. commodity prices, heating system costs or performance of efficiency measures. Rather than providing a single point solution, users can interrogate distributions of outputs (e.g. the number of heat pumps installed), which allows a greater understanding of the robustness of different pathways. For example, where significant deployment of a technology occurs across most of the simulations the user has more confidence that this is a 'low regret' measure.

Monte Carlo simulation requires significant computation time: experience shows that on a high-performance machine (see section 4), exploring 100 simulations can take several days. To maintain an appropriate balance between performance and accuracy, only key parameters (based on estimated impact on solution) should be simulated as part of the Monte Carlo mode of operation as shown in Table 4.

Table 4 Overview of simulated parameters

Category	Simulated parameter
Energy demand drivers	Non-domestic building demand profiles e.g. heating demand
	Efficiency of domestic heating systems e.g. gas boilers
	Efficiency of insulation measures e.g. u-values of cavity wall insulation
Investment costs	Domestic heating system costs e.g. heat pumps, hot water storage tank
	Distributed energy generation costs e.g. district heat biomass boiler, solar PV panels
	Retrofitting cost for non-domestic buildings, e.g. transition to district heating
	Network retrofitting options costs e.g. gas network extension, district heating deployment
Energy sourcing costs	From local e.g. heat, regional e.g. hydrogen, national e.g. electricity, or global e.g. biomass, oil sources or markets

3 Key input data

3.1 Principles

EPN is a complex analysis framework with extensive data requirements and the availability and management of input data has been a key consideration since the initial scoping of the model. Key principles mean that the analysis framework is designed:

- **To be data driven** as far as possible, for example in terms of adding new network or technology supply objects through data tables. Some degree of hard coding is necessary particularly within EnergyPlus (used in the HOM) where this has been configured with a wide variety of heating systems that can be tailored to simulate energy demands in different types of buildings (e.g. by size or level of insulation). However, adding new types of heating system object (e.g. different forms of micro-CHP), would need to be configured via coding changes first, before object specific data can be added.
- **With *de Minimis* data requirements** in mind. For example, understanding the topology of the local area energy system (e.g. where the buildings, roads and substations are located) is important to understanding the cost of electricity network and the analysis framework has been designed around availability of these sources of data.
- **To accommodate better local area information**, the SAM module can use information about the topology of the existing electricity network from DNOs. However, if this is not available, EPN contains an underlying process to synthesise the network topology based on the *de Minimis* data above. In a similar manner, the user can specify known information about the existing building stock at a building-by-building level (e.g. efficiency retrofits in LA owned properties) which is factored into the SAM building archetype matching process. Where this is not available, the model uses less spatially granular datasets (e.g. English Housing Survey in England[33]) and the matching is undertaken on a statistical basis.

3.2 Sources of data

The EPN analysis framework draws on a range of data sources. The availability, costs and licencing conditions for these were considered carefully as part of the scoping process:

- **Large data sources** a small number of large (typically UK-wide) datasets which are generally provided on a commercial (i.e. paid for) basis with associated licencing conditions. These are generally required for representing local area topology and matching building archetypes and currently include:
    - *Ordnance Survey data* (mandatory) provides GIS-layers for a local area's road network and building location[34], footprint size and height. In addition, it provides some basic

---

[33] Other data sets may be available in other regions e.g. Welsh Housing Conditions Survey (WHCS).

[34] And Unique Property Reference Number (UPRN)

categorisation of building features for use in domestic and non-domestic building matching. It is covered by the Public-Sector Mapping Agreement licence.

  - *English Housing Survey* (mandatory) provides a high-level regional view of the split of building archetypes. This free data set is used to inform the building matching where no other data is available, a similar dataset is available for Scotland and Northern Ireland[35], with an equivalent for Wales due to be published in autumn 2018[36]. For the pilot study in Bridgend data from UNO Energy was used[37].
  - *Valuation Office Agency* (mandatory) data contains further categorisation of non-domestic buildings at an individual building level, as well as an understanding of multiple occupancy within a single building footprint.
  - *The GeoInformation Group* (optional) provides further data on building characteristics at an individual building level to improve the archetype matching process.
  - *EnergyPlus Weather files* (mandatory) EPN currently uses the freely available PROMETHEUS[38] project UK location specific weather files from the University of Exeter.
  - *Xoserve* (optional) provides data on which postcodes are connected to the gas grid. This is used to estimate gas network extension costs and for defining current heating systems.

- **ETI data sources** describing archetype and network upgrade costs were used for the pilot projects with electricity network reinforcement costs adjusted using data from the local distribution network operators. However, other data sources could be substituted for future use where relevant.
- **Primary data sources** typically real world local data, generally provided by the LA and network operators, or other parties owning local infrastructure on the location and existing specification of key components (e.g. the rating and grid coordinates of each HV and LV substation).
- **Other freely available data sources** different from those identified above, such as large publicly owned datasets, specific reports, industry standards, etc.

3.3 Key exogenous inputs

The tables below summarise the key exogenous input data across the different modules necessary to run the EPN analysis framework; as distinct from the user definable constraints or data aggregation settings, which are used to configure scenarios based on this input data such as the carbon target that must be met.

---

[35] Scottish and Northern Irish House Condition Surveys could be substituted to the English Housing Survey, although it is important to note that EPN has not been deployed in Scotland or Northern Ireland as of yet.

[36] http://gov.wales/statistics-and-research/welsh-housing-conditions-survey/?lang=en

[37] http://www.unoenergy.co.uk/UNO-2010-Information/

[38] Licence terms are available here: http://emps.exeter.ac.uk/engineering/research/cee/research/prometheus/termsandconditions/

Table 5 Key HOM inputs

Type	Item	Purpose
Heat demands	Target temperature levels	Design standard for building comfort
	EnergyPlus weather files	Background data for heating system operation (also informs building solar PV output)
	Hot water demand profiles	Design standard for building comfort
Building options	Primary / secondary / heat storage / controls, heating system cost and performance data (e.g. efficiency)	Used to characterise buildings according to SAP methodology, to simulate energy demands and to cost future transition pathways. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.
	Building efficiency measures and technical data (e.g. U-values)	
Other demands	Other appliance and lighting electricity demand profiles	Exogenous demands that must also be accounted for in terms of network reinforcement (and building incidental gains where relevant)
	Electric Vehicle demand profiles	
Microgeneration	Solar PV cost and performance data	Used for costing of option and to simulate potential output (in EnergyPlus) for this option
NDB options	Non-Domestic Building (NDB) energy benchmark and cost pathways	Used to specify future NDB transition pathways e.g. to district heating. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.

Table 6 Key SAM inputs

Type	Item	Purpose
Existing buildings	Domestic building local area characteristics (various)	Improve understanding of existing buildings as part of building archetype matching process e.g. insulation, gas network connection, heating, etc.
	NDB local area characteristics (various)	
Existing networks	HV / LV electricity substation locations	Used to help define known topology characteristics of existing networks
	Additional gas and electricity network data (connection topology / current ratings)	
	Existing heat network and energy centre location	
Future DER options	Distributed generation and storage potential locations	Used to define location of potential DER options

Table 7 Key NAM inputs

Type	Item	Purpose
Future network options	Electricity network components (feeders, transformers) cost and performance data	Component data used as part of testing electricity reinforcement options. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.
	Heat network components (pipes, heat exchangers) cost and performance data	Component data used as part of testing heat new/build reinforcement options, both intra-area network and inter-areas transmission. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.
	Heat network operating modes	Used to define pressure and temperature modes as part of testing heat new/build reinforcement options.
	Gas network extension / hydrogen repurposing costs	Used to create simple costed options for evolving the gas grid by analysis area

Table 8 below describes the additional exogenous inputs required by the POM module – i.e. beyond those produced indirectly by the other modules.

Table 8 Key POM inputs

Type	Item	Purpose
Network demands	Building demand diversity functions	Used to adjust building level demands for network reinforcement (electricity / heat)
Resource prices	Resource prices (biomass, waste, coal, oil)	Used to set prices of resources consumed within the area or imported at the boundary (as opposed to those produced by technologies within the local area)[39]. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.
Boundary prices	Import prices (e.g. gas and electricity transmission, large-scale waste heat and carbon emissions)	
DER options	Distributed generation and storage options (costs, performance)	Used to create the costs of the potential DER packages in the local area. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.
Security of supply	Security of supply design standards and network de-rating factors	Used to proxy peak operating conditions as part of network reinforcement. Triangular distribution and correlation parameters are also required if these parameters are simulated as part of the Monte Carlo process.

In addition, bespoke constraints can be configured in the POM by defining caps on, for example, CO_2 emissions and biomass sourcing; forcing build of networks and energy centres; or restricting building transitions.

---

[39] In EPN, these boundary prices are calculated from ESME model results for consistency with national-level decarbonisation scenarios. These prices include the cost of infrastructure e.g. Carbon Capture and Storage, generation and storage asset construction, transmission networks, etc. Carbon accounting can be either physical (kg of CO2 per kWh of electricity or financial through a carbon price adder into the electricity price).

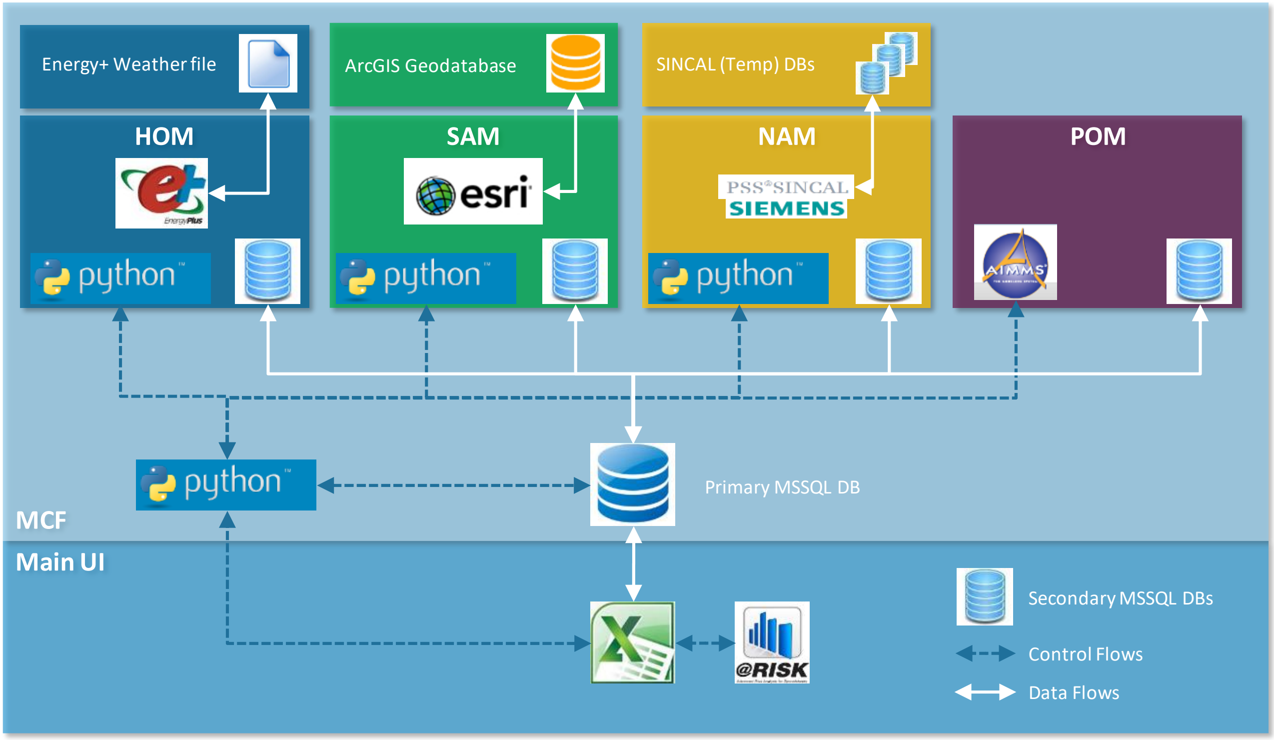
4 Technical architecture

4.1 Design architecture

An overview of the technical design architecture is shown in Figure 8 below. This highlights the four main sub-modules, the Master Control Framework (MCF) and main User Interface (UI), location of third party tools and primary / secondary and in some cases tertiary databases. It also illustrates the high-level control and data flow interactions between them.

An Excel User Interface (UI) is used to enter and interrogate data and results as well as indirectly execute the process steps within the analysis framework. The ArcGIS UI is also used to support the input of spatial data and visualisation of interim and final spatial results.

Figure 8 Overview of technical design architecture



The architecture reflects a modular structure with a hybrid centralised / decentralised DB configuration. All the tools are tightly integrated – or hard-linked – via MCF orchestrator, which manages the execution process and data flow across all the modules. This modular design:

- Enables flexible, parallel development across the analysis framework,
- Isolates 3rd-party components and the main UI so that they are easier to update or replace in future if required,
- Centralises management of primary input and results data to facilitate an audit trail.

4.2 Requirements

4.2.1 Hardware

EPN has been designed to work on a dedicated, high performance workstation. The specification of the computer will affect performance. Table 9 below shows a typical specification that allows for acceptable run times. A full end-to-end run of the analysis framework on a machine of this specification across all modules would typically take 5-7 days of computation time. Many process steps have already been designed to take advantage of multiple cores and are parallelised; hence if more cores and RAM are available this will reduce the computational run time of the analysis framework.

EPN is approximately 30GB in size before being run, and approximately 40GB – 70GB when completed due to the size of the primary database. However, this is heavily dependent on the size of the area being covered, number of analysis areas, and whether it is in deterministic or Monte-Carlo mode.

Table 9 Example specification for EPN dedicated model machine

Part	Specification
Processor	32 cores @ 3.1GHz
General machine Memory	256 GB
Dedicated GPU memory (primarily to speed ArcGIS visualisation)	6GB

4.2.2 Software

A 64-bit Windows operating system architecture is required for compatibility with all 3rd party software within the analysis framework and because 32-bit operating systems limit memory to 4GB for a single process, a single POM optimization problem typically requires far more than this.

The following software components comprise the overall analysis framework:

- Python 2.7 (64-bit) – has been used to code key functional requirements across the HOM, SAM, NAM and MCF as well as ‘wrap’ key 3rd party software[40].
- EnergyPlus v8.1 (64-bit) - used within the HOM for dynamic building energy simulation[41].
- ArcGIS 10.2.2 (64-bit) – used within the SAM to support the spatial analysis requirements and spatial elements of the GUI. This include the Network Analyst and Spatial Analyst extensions along with the Productivity Suite.
- PSS®Sincal 10.5 (32-bit) – used within the NAM for steady state analysis of electricity and heat distribution networks.
- AIMMS 4.23 (64-bit) – used within the POM for generating the optimisation problem. This includes the CPLEX 12.6.3 solver.
- @Risk 6.2 – used via the Excel GUI for generation the Monte Carlo simulated inputs.

---

[40] This uses the following python libraries: numpy, pandas (includes scipy), SQLAlchemy, pyodbc, pywin32, matplotlib, sphinx, numpydoc, scikit-learn

[41] In addition, the ETI’s Optimising Thermal Efficiency SAP model (modified version) – is used as part of the simplification of the building archetype representation for one of the sub-modules within the HOM (not shown in Figure 10).

- Microsoft Excel 2010 (32-bit) or above which is used as the main GUI (along with ArcGIS). It is also used to manage @Risk (via VBA) as this is an Excel plug-in).
- Microsoft SQL Server *Commercial* 2008 R2 (64-bit) and management studio. For all databases except for ArcGIS, which uses its own geodatabase format to store spatial information[42].

Further information discussing the rationale for selecting each of these and for the overall analysis framework design architecture is outlined in more detail as part of the documentation from the scoping phase of EPN[43].

4.2.3 Parallel computing for Monte Carlo simulation

As noted above EPN's processes have been designed, as far as is practical, to use multiple cores in parallel on a single machine. The Monte Carlo functionality of the POM – optimising 100 or more deterministic pathway optimisations – is computationally intensive and each individual optimisation tends to use the full resource of the single machine. EPN has been designed so that individual Monte Carlo POM simulations can be run on multiple machines in parallel.

One machine acts as the 'master' with the full EPN analysis framework and manages the parallel execution of simulations on the other machines. The other machines contain only a copy of the POM to execute each simulation, with input data being read and output results being written back to the database on the 'master' machine.

---

[42] Relevant attribute data required from this is still stored within the MSSQL databases.

[43] EnergyPath™ Networks – Deliverable (S1D2) Design Architecture – April 2014 available here: http://www.eti.co.uk/programmes/smart-systems-heat/energypath

Appendix A : Scope and limitations

A.1 Overview

As described in Section 2, the EPN analysis framework has been designed for a particular purpose (understanding local area energy systems) and is a based on a specific modelling paradigm (techno-economic analysis facilitated by least-cost optimisation). As a result, there are types of questions that EPN is not well placed to answer, but these are not considered a limitation of the model *per se*.

This section focuses on some of the challenges associated with the approach and scope of the EPN analysis framework, given the significant complexity of the issue it has been designed to explore.

A.2 Representation of buildings

A.2.1 Archetypal buildings

As described in section 2.3.1, the representation of buildings is simplified in EPN by using the concept of *archetypes* i.e. EPN groups buildings together based on similar:

- Building shell characteristics[44], which drive comparable energy demands, and
- Insulation and heating systems retrofitting options and costs.

Determining archetypes of local domestic buildings

Research conducted during the initial design of EPN led to a list of 12 key building attributes driving energy demand and retrofitting costs; several millions of different building archetypes can be derived by combining these attributes. For instance, buildings are classified by age band as described in section 2.3.1, which results into the following simplifications:

- A single age band covers the period from 1980 to 2015. Due to changes in building regulations, 1980s properties could have very different thermal efficiency to properties built in 2015. There is a risk that heat demand is over-estimated with this approach. Data is available to incorporate more age bands but the magnitude of this influence has not been established.
- The representation of new buildings constructed after 2016 is simpler as they are assumed to have standardized high-performance attributes, with limited variability compared to the existing stock.

Similarly, EPN defines a unique U-value for each type of window and wall[45]. In reality, it is known that:

- Older double-glazed windows are less thermally efficient than modern ones.
- The quality of installation of cavity wall insulation can vary considerably. In the worst cases records may show that cavities have been filled but an inspection will show that the work has never been carried out.

---

[44] For instance, wall types, windows, floor area, etc.

[45] Note that the Monte-Carlo mode allows for simulating U-values to as to assess the impact of uncertainty of this parameter on the decarbonisation pathways generated by the optimiser.

- Some walls that are classified as solid actually contain a small air gap that can significantly improve their thermal performance.
- Thermal bridging between the inside and outside surfaces of a wall can have a significant influence on thermal efficiency but the locations and scales of this problem are unique to individual buildings.
- The air tightness of individual buildings is unique and hard to establish without individual building testing (and is also a function of occupier behaviour – see A.2.2).

A list of rules is applied to exclude improbable combinations of features e.g. no ground source heat pumps in flats[46]. Despite simplifying the representation of the existing building stock into a more manageable set of characteristics, mapping each building in the area to a base archetype - which still number in the order of 10^5 - is challenging due to limitations in available data[47]. This is described further in section B.2.

In addition, managing unfettered transitions for all these archetypes would be intractable[48] in the final pathway optimisation process so that two further sets of simplifications are required:

- Undertaking detailed half-hourly simulation of domestic building energy performance and heating system and storage operation in the EnergyPlus component of EPN (described further in section A.2.2) is important to help understand demand profiles, in particular the implications for peak network demands as the key driver of reinforcement[49]. But this is time consuming so EPN only runs these simulations after grouping archetypes that are likely to be similar in energy terms.
    - The user can define the number of such archetype groups by floor area band but it is important to note that a large amount of data points is required for results from the grouping process to be stable: this is in general true for small buildings, but large buildings are rare in most local areas.
    - For performance reasons, the number of archetype groups is in practice limited to around 10 so that buildings with very different characteristics can sometimes be grouped together and be assigned the same demand profiles.
- In an analogous manner, groups of buildings that are similar from both an energy and cost perspective - considering energy demands from the above along with property type and size, wall type, starting heating system, etc. - are treated as one for the purpose of determining the choices for retrofitting the buildings over the pathway to 2050.

The above process aims to balance the need to represent the current building stock and its choices in sufficient detail, whilst maintaining computational tractability. The simplification process, moving from base archetypes to their pathway representation in the optimiser, can be flexed to add more or less detail, but some insight is naturally lost as detail is removed.

---

[46] We can also note that EPN will assume a property is off-gas if the user maps it to a non-gas heating system.

[47] In principle, the EPN framework allows the user to specify all known information about a building (as defined by the base archetype characteristics) on a building-by-building level.

[48] Assuming 2 unconstrained transitions to 2050 and a stock of ~10^5 different archetypes, ~10^{15} possible domestic pathways would be generated.

[49] In a more robust manner than, for example, using a simple building Standard Assessment Procedure (SAP) model.

Buildings in the local area are mapped to the representative archetypes defined above, where aggregate characteristics match the EHS[50]' at a Government Office Region (GOR) level. This statistical matching is a good representation for non-location specific features such as wall type but:

- Not for district heat networks which are usually spatially close together whereas EPN would tend to scatter buildings connected to DHN across the map;
- It can overlook local characteristics e.g. if terraced buildings in the local area have very different characteristics than the average in the GOR; and
- This process can also produce variable results: it has been shown that manually updating characteristics of a few dozen buildings (maybe as part of a sensitivity analysis) can lead to a very different mapping of buildings across the whole local area. This can make comparison of results between the different analyses difficult. However, allowing the user to update known information about individual buildings in the area is a fundamental EPN functionality: over time it is expected that this becomes the primary source of data, with decreasing reliance on statistical matching from more spatially aggregated datasets.

Treatment of non-domestic buildings

The heterogeneity of the non-domestic building stock means that creating an equivalent archetypal representation to domestic buildings is much harder. Instead, high-level activity classes such as shops, warehouses, schools, etc. are used to map and assign simple benchmark data (on energy demands and costs of retrofitting per m^2) for the non-domestic buildings in the local area. Multiple benchmarks can be defined to provide optionality in the pathway optimisation, for example, a lower demand level at higher cost as a proxy for energy efficiency improvements. However, the representation of non-domestic buildings is still relatively crude and hindered by the paucity of good data (see section B.2.1).

It is therefore important to consider this aspect of the local area energy system as part of sensitivity analysis or focus attention on gathering better primary data at an individual building level (e.g. energy data for Local Authority owned buildings such as schools).

Some LAs contain industrial areas that have high networks loads. Much of the gas demand in these areas could be for process load rather than space heat and transition of these demands is out of scope for studies, which are intended to focus on delivery of space heating. One approach adopted was to identify the industrial area and give it a benchmark that could not transition. This gave an estimate of the energy demand to be included in network loads, but these buildings were considered out of scope in terms of decarbonisation, leading to direct impact on the level of decarbonisation that could be achieved. In some cases, buildings were removed from the analysis if they had their own HV substation i.e. an electricity feed at above 11kV. These limitations to what is modelled, and how that is done, need to be clearly understood to correctly interpret the model outputs.

A.2.2 Domestic energy demand

Running detailed domestic energy demand simulations in EPN as described in section 2.3.1 and parameterising the results for use in the final pathway optimisation leads to a number of potential

---

[50] Or equivalent data set for Scotland, Wales or Northern Ireland.

simplifications and associated limitations, in particular as this must be undertaken ahead of knowing the future outturn system choices, which are only resolved in the final pathway optimisation step:

- *Consumer behaviour*: is treated as a typical design standard, such as a target temperature level or hot water demand profile, that must be achieved and aspects of consumer response are not captured dynamically; for example, preference for lower target temperatures, manual ventilation after cooking, etc. Variations in behaviour must then be tested indirectly via scenario analysis.
- *Heating system sizing*: EPN determines the appropriate size of the heating system for an archetypal building by iteratively increasing the system's size until it is able to meet the design standards. However, for systems with secondary heating systems and/or storage the user must define in advance the ratio of primary to secondary heating system size and storage size, respectively. This is because EnergyPlus lacks all of the necessary information to size of all components of the heating system effectively (e.g. implied energy prices that are only present in the pathway optimisation process).
    - It is therefore necessary for the user to define an appropriate spread of heating system configurations to ensure that issues such as the appropriate mix of heat pump to gas boiler in a hybrid system are tested sufficiently. This will increase run times for domestic energy simulation as well as all steps downstream e.g. domestic retrofitting, and pathway optimisation.
    - Similarly, the ex-ante sizing and operation of the heating systems means that the operational paradigm of storage must be defined ex-ante. For example, the user can specify how storage is used to minimise peak demand across the entire day or at specific points in the day (e.g. during early evening peak). However, at this stage it does not contain the information about the best way to shift peak load for the system as a whole, as this is an outturn decision only considered in the final pathway optimisation step and is dependent on a wider set of factors, such as the operation of grid level electricity storage.
- *Demand diversity*: EPN embeds a simplified representation of domestic buildings' heating demand[51] diversity to reflect the non-coincidence of demand across a group of buildings. The methodology flattens and widens the daily demand profile driven by the number of buildings connected to the same local infrastructure, typically a high voltage substation.
    - The more buildings, the lower the aggregate peak demand due to diversity. Note that this approach works best for large clusters with diverse buildings.
    - The scalars used are based on historic data and are not necessarily representative of new features such as new forms of demand side response, increased cooling demand, etc. In practice, this approximation works well when the spatial representation of the local area is defined as large analysis areas covering many buildings (e.g. at HV substation level) but would potentially overestimate peak demand if analysis areas are defined at a lower level (e.g. LV substation and below).

---

[51] Diversity scaling does not apply to NDB demands in EPN v2.2.

  - A further complexity is that the outturn level of diversity is a function of future choices in the pathway optimiser – i.e. whether more or fewer buildings choose to base their heating system on electric or district heat options. This effect can be represented dynamically in the pathway optimisation[52], but only with a significant performance penalty. An alternative option is for the user to pre-process an assumed level of diversity benefit ex-ante, however, this may not align sufficiently with the outturn pathway choices.

A.2.3 Domestic building retrofitting

Building retrofitting is a key element to achieving a decarbonized energy system at local level. EPN represents simplified building fabric and heating system transitions for domestic (and non-domestic) buildings and assesses their total lifetime costs as part of preparing pathway options for the optimization engine.

The retrofitting pathways are first costed at the most detailed archetype level before being aggregated into groups (as described in section A.2.1) where buildings in close proximity in the area are deemed similar enough in terms of energy demands and retrofitting costs. The user can control the final number of retrofitting pathway options by:

- Altering the number of heating system configurations and energy efficiency packages (e.g. cavity wall insulation with draught-proofing and double glazing) being considered;
- Altering the number of transition points (e.g. 2-3) that a retrofit can be undertaken across the pathway;
- Specifying invalid combinations of transition – e.g. once a building has transitioned to a low carbon option it is not allowed to transition back to a ‘high’ carbon option. However, within the pathway costing process it is important to note a number of fundamental restrictions, which the user cannot change directly:
  - Retrofitting transitions can only be triggered by a heating system change (or replacement of the current system) and hence energy efficiency only transitions cannot be considered separately. The costing of efficiency measures does correctly account for lifetimes and replacements, which may span these transition points.
  - EPN does not currently allow for restricting deployment of some heating systems to meet engineering constraints, e.g. deploying GSHP in more than 30% of properties in a street could cause the ground to freeze.

The costs of retrofitting for heating and efficiency measures considers both fixed and marginal cost components (e.g. that change per kW or per m^2), however, this may not fully capture the variation across buildings, for example, in the case of hard to treat homes.

In other cases, modelling different retrofitting costs would require creating separate instances of the same technology. For instance, the cost of heat network connection to flats with electric storage heaters would be more than for flats with a wet heating system. Correctly modelling this would require setting up two district heat options (with or without wet heating system installation) and defining the right

---

[52] The maximum diurnal temporal granularity is half hourly, but due to performance limitations in the pathway optimisation only more limited diurnal time granularity is generally possibly, which may not represent the full value of SMART control e.g. turning down heat pumps.

transitions. Similarly, it would theoretically be possible to represent different remaining technical lives for existing heating systems by creating several copies of these heating systems. This of course would increase runtime.

A.3 Representation of networks

Electricity, gas and heat networks are a central part of the EPN model. The simplifications and assumptions made in designing the representation of these network types in EPN are described in the three sub-sections below.

A.3.1 Electricity networks

In order to synthesize the existing electricity network topology, EPN uses the locations of existing high (33kV/11kV) and low (11kV/400V) voltage substations[53] and the road networks. The user can also define directly the connections between high and low voltage substations and even connections between low voltage substations and their buildings to better represent the DNO's network. Where this data is not available, EPN will synthesize an approximation of the existing electricity network. User-defined or automated network syntheses follow these principles:

- *The network is assumed to be radial (with no tapering of cables along feeders)*, which may not represent existing networks in some urban areas where there is a high degree of meshing.
- *The network follows the road network*, which is in general true for urban networks but may not hold for some rural networks[54]; note that this assumption must still hold even where DNO data is available to specify LV substation to building connections.
    - If two substations are located close to each other (i.e. on the same road node) EPN must effectively treat them as one site serving the surrounding buildings.
    - In reality, some road nodes may be served by multiple LV feeders, whereas EPN assigns each road node to a distinct LV feeder.
- *The network minimizes connection lengths*, this assumes buildings are connected to the closest LV substation following the road network. This implies:
    - Areas within networks have to be spatially contiguous. EPN cannot model "doughnut-shaped networks" where an island in an area is connected to a different substation to the buildings surrounding it.
    - Networks which are routed across fields in rural areas will be modelled as following the road network and may be modelled significantly longer than they actually are.
- *The network synthesis process estimates current network capacity* (transformer ratings, cable sizes) based on its own approximation of existing peak load, unless DNO data is available to specify the capacity of the network directly.

---

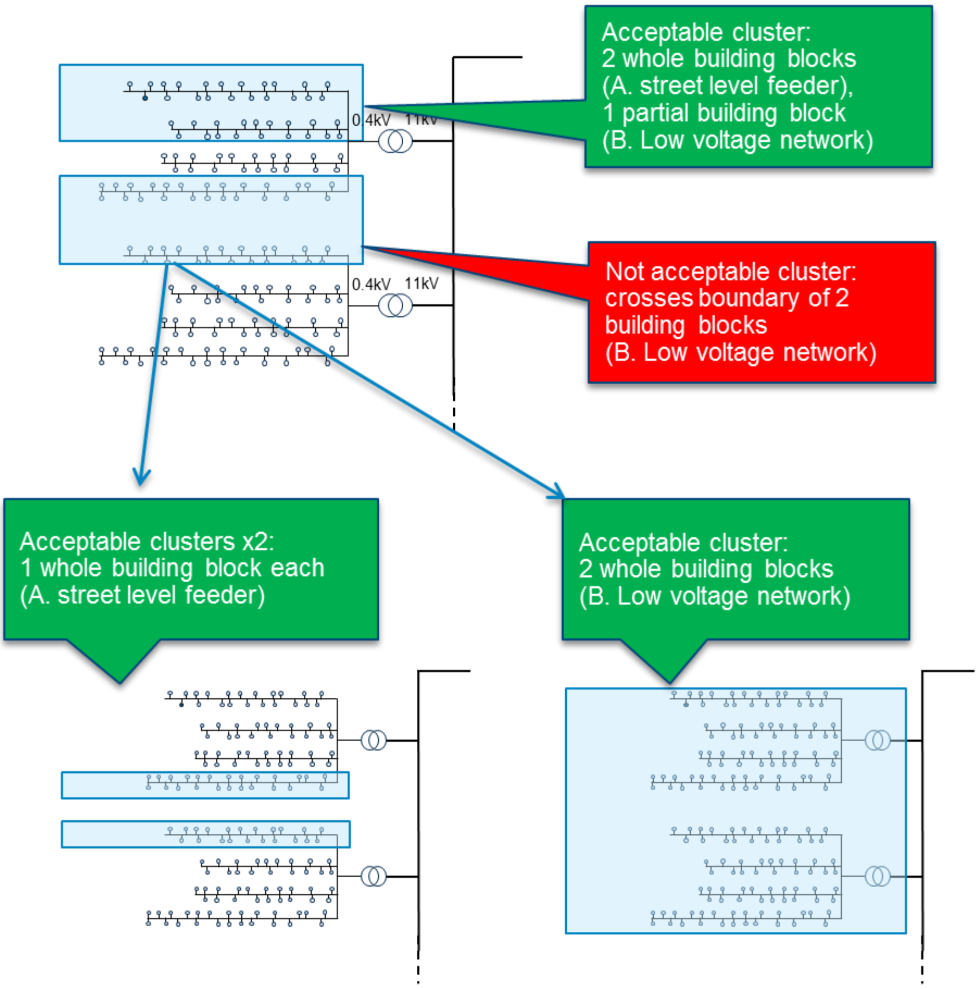
[53] Network reinforcements above a 33kV/11kV transformer are not considered in EPN but could represent a small additional network cost.

[54] This would mean that costs of rural networks could be lower than calculated in EPN because of their possible shorter length and cheaper excavation costs.

  - Some data calibration may be necessary if EPN estimated peak demands for the existing area exceed the known network capacity – i.e. implying reinforcement is already needed.
  - Note that where there are only a small number of buildings connected to a specific part of the network (e.g. LV feeder) this makes it harder to approximate existing peak load, given the more limited diversity effects.
- *The user can define whether the network is urban/rural or underground/overhead* – with appropriate variations in network cost – at a spatial cluster level, however, it is not possible to mix this representation within a cluster (e.g. 50/50 underground/overhead).

For performance reasons in the optimiser, the local area is represented by a set of discrete, contiguous spatial analysis areas. The user can define analysis areas flexibly (e.g. all buildings connected to the same HV substation, the same LV feeder, etc.) as long as the fundamental structure of the electricity network topology is respected, as represented in Figure 9 below[55]. For instance, EPN will not allow the user to define analysis areas containing two LV substations that do not connect to the same HV substation.

Figure 9 Defining spatial analysis areas to whilst reflecting electricity network topology



---

[55] Note that experience from the three deployment projects shows that this data manipulation can be quite time consuming to configure.

Future network options

EPN uses an integrated load flow simulation tool (PSS®Sincal[56]) to simulate thousands of possible reinforcement configurations (considering thermal and voltage limits). EPN then parameterises this detailed information into sets of simpler, aggregated network cost options.

These simpler options will provide analysis-area and component-level lumpy network reinforcement cost curves to the pathway optimisation process. The user can define the final number of network options, but information is naturally lost as the outputs of the more detailed load flow analysis are aggregated.

In addition, this parameterisation process necessarily introduces a number of implicit assumptions, in particular as the network testing and parameterisation must be undertaken ahead of knowing the future outturn loads, which are only resolved in the final pathway optimisation step. These include:

- An assumption that load grows spatially uniformly across different parts of the network. This is needed to keep the testing process tractable (i.e. thousands of tests in PSS®Sincal as opposed to hundreds of thousands). For example, when testing increases in aggregate load on an HV feeder it assumes that the share of total load at each LV substation offtake point stays the same.
    - This simplifying assumption works relatively well where the characteristics of the connected buildings, and their propensity to electrify in future, are fairly homogenous (e.g. area of terraced houses with gas boilers). Where this is not the case, it can be mitigated by increasing the level of spatial granularity seen in the final pathway optimisation process (by using more spatial analysis areas) at the expense of software run time. For example, splitting a single HV substation spatial analysis area into separate sub-areas for each connected HV feeder.
    - Alternative assumptions for how load growth is spatially distributed are being reviewed as part of future EPN enhancements.
- The testing and parameterisation process is focused on conventional network reinforcements. ‘Smart’ network options that involve load shifting on the wider energy system to manage constraints, such as distribution-level storage, can be represented directly in the pathway optimisation process as distributed energy assets.
    - However, other ‘smart’ options (and conversely the need to maintain N-1 contingency in parts of the network) need to be parameterised off-model and fed in as exogenous assumptions.
    - Meshing is a special case within EPN; the impact of HV feeder to HV feeder meshing on network reinforcement costs can be parameterised and tested via PSS®Sincal. However, this is currently limited to where this occurs in the *same* spatial analysis area when passed to the pathway optimisation, for performance reasons.

---

[56] http://w3.siemens.com/smartgrid/global/en/products-systems-solutions/software-solutions/planning-data-management-software/planning-simulation/pss-sincal/pages/pss-sincal.aspx

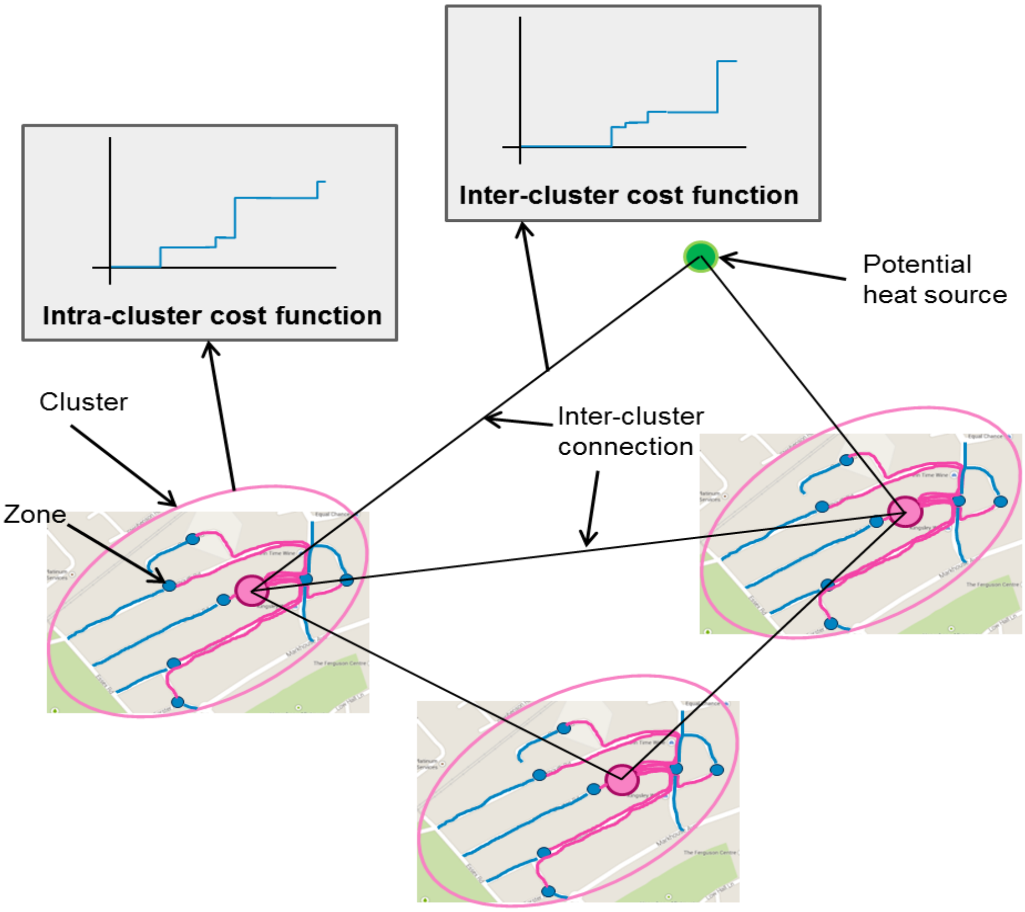
A.3.2 Heat networks

Heat networks follow the same basic process as for electricity, in terms of detailed load flow testing of reinforcement configurations (also in PSS®Sincal) followed by parameterisation of this analysis into a set of simpler, aggregate cost options for use in the pathway optimisation.

This parameterisation process introduces a number of other assumptions and limitations, which are specific to heat networks:

- **Representation of altitude differences** in the district heat network is simplified in EPN. Only the difference in head between the pumps and the average height of the network is captured, which will not fully reflect the true influence of the local topography. There is a risk that cost of heat network capital and operational costs are underestimated due to a need for larger pumps and more energy input to operate a heat network layout that crosses hills. Having said that, this is expected to be a relatively small proportion of the total system cost.
- Heat networks can operate under **different conditions of temperature and pressure**, with different implications for operating costs and losses. The user can define how many of these should be tested as part of the load flow analysis, but in practice they are limited to avoid creating too many additional options for the pathway optimiser to resolve.

Figure 10 Overview of process for developing potential heat network options



- Unlike electricity, where reinforcements take place based on a known starting network topology[57], the final topology of the heat network is not known in advance. To avoid dramatically adding to the number of load flow tests it is assumed that heat networks are deployed as part of an '*all or nothing*' decision within each of the spatial analysis areas, as defined by the EPN user. The load flow testing is undertaken for each analysis area independently and then separately for the heat 'transmission' options, which are used to connect analysis areas together. In the pathway optimisation EPN still decides which analysis areas heat networks are developed in across the local area (and the level of heat demand they can accommodate) and how these analysis areas are connected (e.g. to share heat supply from a single energy centre across multiple analysis areas). Figure 10 above illustrates the heat network design process in the NAM.
    - However, the deployment of a heat network across a single large analysis area may be a poor approximation if in reality only part of the analysis area provides a suitable high load density: this could cause the pathway optimiser to underestimate the deployment of district heating. This is because a small heat network might be the least cost solution when the optimiser choice is limited to building a larger, more expensive network which leads to higher total system cost.
    - Modelling existing or planned heat networks by defining the buildings it serves, or will serve, will force the build of a heat network across the whole analysis area(s) at the capacity required to supply the heat needed for the defined buildings. This can lead to EPN underestimating the cost of extending the existing network as buildings connected to existing DHN can straddle several analysis areas, so EPN will deploy the existing DHN for too many buildings.

As noted previously, the EPN framework was designed to be flexible and easily alter the granularity of a number of parts of the analysis framework. In this case sub-dividing a large spatial analysis area into multiple smaller analysis areas (whilst still respecting the electricity network topology) would provide a better representation of heat network choices and more accurately represent existing heat networks, but at the expense of performance in the pathway optimisation[58].

A.3.3 Gas and hydrogen networks

Modelling of gas networks in EPN relies on the assumption that where they are present, there is enough capacity to service current demand and that decarbonisation will result in reduced gas demand. Therefore, it is highly unlikely that the gas network will require further capacity. As a result, the current EPN framework does not include steady state modelling for gas networks.

As per heat networks, **the decision to extend, decommission or repurpose the gas grid can only be taken for the spatial analysis area as a whole**.

- Decommissioning the gas network e.g. to avoid ongoing fixed costs can only take place where:

---

[57] Topology extensions for new housing estates or non-domestic buildings are specified with perfect foresight across the pathway and are therefore known for the purposes of load flow testing.

[58] Alternatives to allow representation of a partial roll-out out of a heat network within a single specified analysis area are being explored as part of EPN v2.2

  - *All* gas demand (including from non-domestic buildings) is removed from the analysis area, which is in practice rarely the case. Future updates to EPN may consider proportional decommissioning based on e.g. a peak demand requirement within the spatial analysis area.
  - The analysis area does not provide intermediate transit between the local area's gas entry point and a subsequent analysis area where gas is still required.
- The decision to repurpose the gas network to hydrogen[59] is subject to the same conditions as gas network decommissioning and requires an appropriate source of hydrogen supply being available (e.g. a gasification plant in or connected to the analysis area).
- The limitations stated above can be worked around by defining more and smaller interconnected analysis areas[60], at the cost of increased runtime.

EPN does not currently represent **greenhouse gas emissions from natural gas leaks**, thus underestimating local area emissions[61]. The carbon target can be adjusted off-model to approximately account for these extra emissions.

A more detailed representation of the **costs of repurposing the gas network for hydrogen** is being considered as part of future EPN releases. For example, driving costs based on both the length of the network (already considered) and number of connections and allowing the user to specific the extent to which iron gas mains have already been replaced within each spatial analysis area.

A.4 Representation of other local area features

A.4.1 Boundaries to national energy system

EPN's geographical scope is limited to the modelled local area, but the choices it makes are dependent on factors (and assumptions about these) that are outside of the boundary of the model, including:

- EPN implicitly assumes the **Local Area is a price-taker** with respect to commodity prices that are at the boundary of the local area (e.g. transmission-connected electricity and gas, along with biomass, hydrogen and transmission scale waste heat). The maximum availability of these commodities also needs to be limited appropriately.
- In a similar manner, EPN assumes **exogenous technology learning curves** for all technologies and is not able to represent dynamic interactions between technology deployment and market price. However, given the small size of the geographic area under analysis this assumption is appropriate.
- EPN does not model **reinforcement for electricity distribution networks above 33/11kV transformers** and hence may be missing some additional costs for pathways with heavy electrification. Transmission system costs are included in the cost of importing electricity at the

---

[59] Available with EPN R2.2.

[60] It is important to ensure all analysis areas with gas demand are connected to the NTS supply point either directly or indirectly through other analysis areas – i.e. given implicit meshing of some parts of the gas network.

[61] Total loses from gas network are approximately 0.6% of throughput across the medium and low pressure systems including theft, venting, leakage and that used by the network operators themselves.

boundary. Similarly, EPN does not model **reinforcements for the gas network** (upstream of the NTS supply point) although these are considered unlikely in GB given current gas penetration.

- Buildings near the edges of the geographic study area might be fed by electricity networks that are outside of the study area. For example, an LV feeder may cross the study area boundary to supply them. The user will need to decide whether to connect these buildings to network infrastructure that is in the study area, define an additional analysis area to allow their networks to be modelled separately or to exclude these buildings from the analysis. None of which will accurately reflect the true network topography and load flows as:
    - Connecting these buildings to an alternative substation within the study area so that the load on this substation will be modelled as larger than it should be; or
    - Adding the substation which is outside the study area and connecting the buildings to it will result in the load on this additional substation being underestimated. This can also involve additional data gathering and integration e.g. road network, electricity substations, etc.
    - Removing these buildings will underestimate total load.
    - In general, the first approach has been preferred in the pilot studies as this reduces the number of analysis areas required and so reduces the problem complexity in the optimiser.
- **Decarbonisation incentives** can be applied as an absolute limit on emissions or indirectly via a carbon price assumption. In the case of an absolute quantity constraint care needs to be taken with respect to the scope of emissions covered by the boundary of the EPN model, for example, whether transmission connected electricity is assumed to be zero carbon or at the national electricity generation carbon intensity.

A.4.2 Distributed energy and transport

In practice, solar PV output depends on panel orientation (North, South-facing) and pitch (roof incline compared to horizontal) for each building in the local area. However, to keep the problem tractable in EPN and because gathering data on individual buildings is impractical, a weighted average mix of these aspects is used when modelling solar PV output across the local area for building and ground-mounted solar PV panels.

Currently, Monte Carlo simulations cover technology costs but not performance (i.e. efficiency of large scale heat pumps), which limits the ability to understand the influence of improving technology performance on future outcomes. This must instead be explored via a series of manual sensitivities.

The focus of EnergyPath™ Networks is in delivering low carbon heat and as a result has a simplified representation of local transportation, in particular:

- Energy demand of public transportation can be represented with an exogenous load profile, but EPN will not make any recommendation on transportation infrastructure i.e. scenarios of public transport development in the local area e.g. subway, tram, buses would need to be parameterized off-model, and
- Electric Vehicle (EV) deployment and its associated electricity demand can be added as an input assumption at an individual building level (including public charge points), however, this is an

exogenous assumption (with pre-processed diversity adjustments) and EPN does not make the decision to 'purchase' an EV itself. In addition, it does not currently represent detailed operational issues such as smart EV charging management.

A.5 Representation of pathways

The following section outlines the key modelling simplifications and assumptions underpinning the 'local energy system pathway representation' in EPN. These aim to provide a modelling framework that achieves a good balance between model accuracy, ease of use and performance.

A.5.1 Summary of data simplification impacts

Table 10 below recaps how data simplification will affect the local area energy system representation in the optimiser and options for mitigating these in EPN.

Table 10 Overview of impacts of data simplifications in optimiser

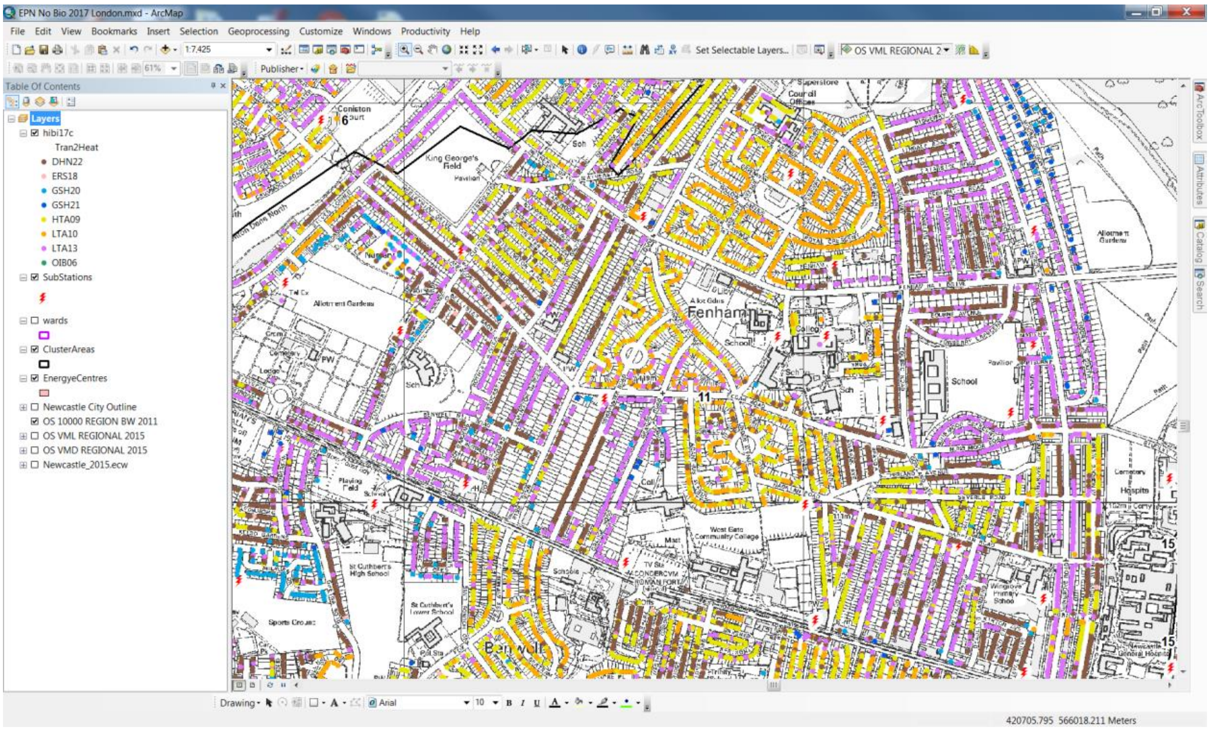
Topic	Impact of simplification	How EPN compensates for it
Security of Supply	Optimiser receives average demand across diurnal period of several hours, and therefore does not model energy dispatch for the "true" 30min peak demand period.	An off-model calculation is required to estimate "true" 30min peak demand above and beyond the average value over an aggregated diurnal time slice passed to the optimiser. This is added to a further security margin e.g. 5% for electricity to size heat and electricity distribution networks.
Assessment of storage assets	Limited opportunities for electricity storage optionality as diurnal periods are very blocky	It is possible to increase the intra-day resolution with an impact on performance.
Analysis area-level energy demand	Energy demand balancing is done at spatial analysis area level i.e. the optimiser does not represent energy flow on each individual LV feeder (unless the tool is specifically setup to reflect this level of granularity, which is likely to be computationally intractable). Similarly, energy flows between areas are considered as aggregated quantities, which may not be the case in reality e.g. several gas pipes connecting two areas are represented by a single flow.	It is possible to increase the analysis area granularity with an impact on performance e.g. from HV substation downstream to LV substation downstream.
Archetype transitions in groups	Buildings retrofitting to heating systems, insulation measures as well as solar thermal is modelled by group of domestic archetypes. Domestic heating systems and insulation measures are grouped into retrofit packages.	It is possible to increase the number of building groups e.g. by increasing the number of energy states with an impact on performance.
Simplified network options	Network options available for build in the optimiser are averaged from detailed load flow analysis. This uses a limited set of network components.	Network options by energy vector and component type can be made more granular with an impact on performance.

A.5.2 Limitations of coarse spatial resolution

The pilot projects have defined large analysis areas for performance reasons, comprising one HV substation and all the buildings that are connected to it. These analysis areas can cover a geographic area that has quite varied characteristics such as going from a dense town centre to surrounding rural areas. This coarse level of granularity could result in different decisions to those that might emerge with a more granular definition of analysis areas. For instance,

- As **buildings** are grouped based on their thermal characteristics rather than their spatial closeness within the analysis area, and transitions are selected individually for each group of buildings, identifying the transitions generated by EPN on a map can show that the buildings selected for heat networks are interspersed with those selected for electric heat solutions, as represented in Figure 11 below. However, stakeholders might require outputs to be produced at a building level and DHN operators want to see where the proposed heat scheme would be laid on a street by street level. EPN can produce some outputs at building level but aggregation (spatial and building archetypes) limits insights that can be gathered. Technically this could be resolved with more spatial and domestic building granularity, albeit with a cost on performance. As an intermediate step, the Catapult have a developed a post-processing methodology for re-allocation of buildings within an analysis area so that buildings selected for heat network connections are more spatially clustered. This methodology works on the basis that the network capacities defined in the optimiser outputs are still respected.

Figure 11 Example of allocation of heating systems to UPRN showing 'pepper pot' nature of the results



- Similarly, **network** characteristics like capacity requirements and associated reinforcement costs for LV feeders are aggregated at analysis area level. Some stakeholders have considered this to be of limited value as it does not identify the particular feeders (and hence streets) that will require work[62]. It has been suggested that this more granular level of output is required to hold meaningful discussions with both residents and local network operators[63]. It should be noted, however, that EPN is a strategic planning tool. It should not be expected that any project that might be proposed as a result of this type of analysis would be initiated without a significant amount of more detailed analysis and planning which would enable, and require, these types of discussions.

A.5.3 Parameterising building transitions

Step-changes in retrofitting buildings

EPN models building transitions as happening across the whole stock in the same calendar years, linked to heating systems technical lifetimes. This requires aligning remaining technical lifetimes of the technologies comprising the existing heating system and insulation measures[64] and therefore approximating 'real world' asset replacement cycles with standardized cycles e.g. window replacement times need to be aligned with boiler replacement so that retrofit packages can be deployed at the same time as heating systems change. Increasing the granularity of either time or heating system representation would improve this point.

The system value of heat pump / gas boiler hybrids as a stepping stone[65] on the path to full electrification of heat could also be better represented with increased granularity, as:

- Prior to grid decarbonisation hybrids are an expensive technology option which give limited CO_2 reduction benefits and which must be replaced due to being out of life by 2050.
- Post electricity grid decarbonisation it is not possible to have a large number of hybrids if a low 2050 emissions target is set.
- Network reinforcement is often required for both full electric and hybrid deployment. The incremental increase required from not deploying hybrids is not significant when compared to the costs of hybrid deployment.

Together these mean that a cheaper option is often to remain on a gas boiler in early years and replace this with a heating system powered only be electricity in later years, depending on the rate of electricity grid decarbonisation and the local carbon target used.

Right-sizing energy centres

Within the optimiser the decision to build any particular energy centre is a discrete choice in early time periods[66]. Energy centre options must be defined as particular technology choices with associated capacities and costs to best reflect the costs of what is inherently a large "lumpy" investment. This must be done before the optimiser has been run which is difficult without knowing what scale of heat network might be selected in different parts of the study area. There is a risk of under-sizing the options such that heat network deployment is curtailed due to a lack of heat capacity or only having options that

---

[62] Increasing spatial granularity within EPN to model each LV feeder individually could theoretically resolve these issues but the expected performance impact could make the model intractable.

[63] Local DNOs have dedicated tools and processes in place for these purposes already.

[64] This is also true of the modelling of non-domestic buildings in EPN, where heating system change dates are aligned across different systems.

[65] Hybrid systems could give consumers confidence in new technologies before they are rolled out at a larger scale. EPN does not however attempt to model this behavioural factor.

[66] Note that this linear relaxation based on investment decision time period is user-drive.

are too large and expensive for smaller networks to be priced appropriately. Similarly, the choice of available heat centre sizes will interact with the decision to deploy a small number of larger energy centres with heat transmission pipes to supply heat to other areas, versus a larger number of small energy centres - one for each spatial analysis area.

Whilst this could be overcome by defining a large range of different energy centre options for every analysis area this process is time consuming and would increase problem complexity and runtime. This could be partially mitigated by linearizing the investment variable for energy centres.

A.5.4 Least-cost optimisation framework

The least-cost optimisation framework underpinning the design of future pathways can sometimes lead to:

- ‘Penny-switching’ as small differences in costs can drive very different investments[67]. For instance, the balance between electric and district heat solution selection is heavily influenced by their relative costs, so that there is a risk of one solution dominating the results when the balance could switch with a small change in cost of either. However, understanding how sensitive the proposed ‘least-cost’ solution is to small changes in inputs is itself valuable.
- Decisions that assume perfectly rational behaviour, but which may not take into account other factors e.g. the full distributional value of social schemes to reduce fuel poverty. Hence broader off-model analysis of benefits from these policy options is required to understand how they change between different model solutions.

---

[67] Note that this variability can be explored using the built-in Monte Carlo functionality.

Appendix B : Input data uncertainty

B.1 Overview

This appendix describes the uncertainties in data used to populate EPN inputs. Table 11 below presents a qualitative overview of estimated uncertainty and impact on results. These issues are discussed in more detail in the following sections. In the absence of obtaining better data, the impact of input data uncertainty can be partially mitigated through careful sensitivity testing (or use of the Monte Carlo functionality where available).

Table 11 Impact of input data uncertainty in EPN

Data item	Uncertainty	Impact on results
Non-domestic buildings heating systems mapping (B.2.1)	High	Medium
Non-domestic buildings heating systems costs (B.2.1)	High	Medium
Non-domestic buildings demand (B.2.1)	High	Medium
Domestic buildings fabric and heating system mapping (B.2.2)	High	High
Domestic buildings insulation and heating system costs (B.2.3)	High	High
Domestic buildings demands (B.2.3)	High	High
Network characteristics (B.3.1)	Medium	Medium
Network costs (B.3.2)	High	High
Service demands (B.4.1)	Medium	Medium
Spatial mapping of distributed energy (B.4.2)	Medium	Medium

The main challenges faced with validating data quality are:

- Data gathered sometimes comes with embedded assumptions which may not be appropriate for the considered LA. For example, future cost reductions for technologies may assume large scale installation with accompanying economies of scale, whereas a particular local area may only have a limited number of properties where the technology is suitable.
- Data gathered is sometimes of poor quality (e.g. missing or out-of-date data, errors), which can be partially resolved by cross-checking one source against another where alternatives are available, although data are often not presented on the same basis.

Challenges with bespoke local area data

In all 3 pilot studies local data was used to improve the local area representation. This included items such as:

- Social housing stock records.
- Records of local energy efficiency schemes.
- Buildings connected to existing heat networks.
- Existing local energy generation such as solar PV, Gas CHP and wind turbines.
- Planned development (new buildings).

This data was made available in a variety of formats which used different classifications. In all cases local data needed to be cleaned to remove inconsistent information and had to be mapped to the options recognised by EPN. For example, local data may contain different age and type classifications to those used in EPN.

B.2 Representation of buildings

B.2.1 Non-domestic buildings heating systems and demands

The representation of non-domestic buildings (NDB) currently includes:

- Two different non-domestic gas heating systems to map to the current stock based on their use classification, using engineering advice from Arup. However, actual knowledge of individual NDBs and their actual heating systems is extremely limited such that these choices might be inappropriate.
- Transition options for district heating, as there was little robust data available to easily include options for insulation and electric heating at the time of the pilot projects.

Currently non-domestic building demand is calculated using the Carb2 dataset[68], which estimates energy usage per m^2 based on building activity class. However, in reality, there will be variations in demands between buildings, which is difficult to assess precisely given the lack of data. This simple approach can result in inaccurate demand estimates which themselves result in a requirement for excessive network capacity. It can also cause inappropriate future NDB pathways to be selected.

Data checking for non-domestic buildings is complex and time consuming due to the need to cross-reference multiple data sources, which all contain different building activity classifications (Ordnance Survey (OS), Valuation Office Agency (VOA) and GeoInformation were used for the pilot projects).

B.2.2 Domestic buildings features identification

At the time of running the pilot studies, there was no central source of energy-related data for buildings in the UK at an individual building level; however, Energy Performance Certificates are now becoming more widely available. Within EPN the main characteristics of domestic buildings (e.g. wall type, property age and type, heating system, etc.) are imputed from several data sources (primarily OS, English Housing Survey (EHS) and GeoInformation in the pilot projects).

General mapping

The building points (and associated Topographic Identifiers or TOIDs) present in the OS dataset are used as the basis for identifying the numbers, locations and types of buildings in the study area. As such, the quality of OS building data will have a strong influence on final outputs of the EPN study. In one of the three study areas there were several aspects of the building data that were of poor quality:

- The Basic Land and Property Unit (BLPU) code should define whether a building is currently in use – a primary factor when considering energy use. In 90,408 cases this field was set to “unknown” with only 10,560 buildings classified as “in use”.

---

[68] http://www.ucl.ac.uk/energy-models/models/carb2

- 3,086 buildings were classified as under construction but this field had not been updated for over 10 years.
- Nearly 10,000 building points did not have a corresponding building TOID.
- For 1,913 non-domestic buildings (out of around 8,000) it was not possible to identify their exact location, or whether they existed at all.

Some detailed data on local buildings is often available, particularly for social housing. However, it is generally very time consuming to perform data cleansing on this data as it is often not to the standards required. For example, names of the same item within a given field may vary (‘gas boiler’ or ‘boiler, gas’). Mapping of this data to specific UPRNS can also be time consuming and inaccurate if the records are stored by address rather than UPRN. This is partly because address fields are often defined inconsistently between data sources and may be inconsistently completed within any one data source.

An indication of the level of uncertainty in building classification can be gained by comparing the same attribute from different data sets. The OS classification of mid and end terrace properties in Bridgend was found to match 77% of the GeoInformation classification but only 35% of the EPC derived classification. Overall, GeoInformation provided data for 93% of buildings with a UPRN.

There are over 120,000 domestic buildings in Newcastle. Data was provided from a variety of sources which were cleaned, combined and incorporated into the local area representation. The data included:

- Detailed asset records for 29,000 homes operated by Your Homes Newcastle (YHN).
- Records for 83,000 surveys from the Warm Zone scheme including information on installs for 44,000 properties. This gave information on 61,500 homes which were not included in the YHN data although data quality was poor due to how the surveys had been conducted.
- Measures installed in 8,700 properties as part of the Warm Front scheme.

In all these records updated at least one field for 91,751 properties.

Within EPN it is possible to define areas of new build properties with planned construction dates and the expected housing mix. This data is often not available in a single file from Local Authorities. Sometimes map shape files are available but on other occasions it might be necessary to reconstruct these from maps in pdf documents. In addition, the planned housing mix may not be defined in detail, or might be included in separate documents or spreadsheets. This meant that data preparation and entry for new domestic buildings was very time consuming in all three pilot projects, taking up to a month of effort for one person. This may be disproportionate given that new development is normally only a small proportion (less than 10%) of predicted future energy demands. This is due to the relatively small number of new buildings when compared to the current stock and the expected high efficiency of any new builds.

Thermal efficiency mapping

There is no equivalent to the English Housing Survey in Wales. The available Energy Performance Certificates (EPC) for Bridgend, along with information from the local social housing providers, and building control data were used to both define individual building characteristics and the make-up of the local housing stock as a whole. It is generally accepted that the quality of some EPCs is very poor. Whilst efforts were made to filter out those which were clearly wrong it is likely that this dataset is less robust than the English Housing Survey, which is based on high quality detailed building surveys. In addition, it is likely that the social housing stock has better thermal efficiency than average since these buildings are professionally managed. This means that the average characteristics applied within the modelling work were probably worse than is really the case.

Data on thermal efficiency improvements to owner-occupied and privately rented buildings is often very hard to find. Any improvements completed on these buildings will have been arranged by the building owners and no central record exists of the improvements that have been carried out.

Validating modelled building demands

It is useful to benchmark EnergyPath™ Networks energy demand estimates against alternative sources. BEIS publish domestic gas and electricity demand at postcode level. This is based on metered consumption. Several features of this data make it difficult to match exactly with EPN data:

- The BEIS data is aggregated when there are 5 or fewer buildings in a postcode to avoid the risk of disclosing individual building energy use. Even allowing for this not all postcodes appear to be included in the BEIS data e.g. in Bury 948 of 4,333 postcodes in the OS Mastermap did not have a match in the BEIS electricity consumption data and 629 of 4,190 did not match for gas.
- The BEIS data is adjusted to account for missing information, however, the exact process used to do this is not publicly available.
- In order to perform a meaningful comparison, both the BEIS and EPN data need to be aggregated to some level. Within the pilot projects this has been done by Postal Sector. Postal Sector boundaries do not coincide with the local authority boundaries used to define the EPN study areas. This can make comparison difficult.

Challenges have been experienced in validating against other data such as consumption data from Display Energy Certificates. These alternative data sets are often themselves modelled to some degree and can be incomplete. Problems have also been encountered where these alternative data sets are, themselves, internally inconsistent. Examples have been found where a single building has several Display Energy Certificates all produced at different times which contain significantly different energy consumption values.

Similarly, it can be hard to understand the exact nature of any data provided by network operators but this is essential to ensure valid comparisons are made. Demand data from network operators is typically measured at particular points in the network and is a measure of total demand at that point. It is not possible to separate domestic from non-domestic demand so comparison to EPN demand estimates is only possible at a high level of aggregation. Due to the cost of installing monitoring equipment energy networks are typically only monitored at points higher up the network. For example, electricity transformer data is unlikely to be available at less than 11kV. This limits the options for using this data.

In one case peak demand data was provided to the project team for the local gas network. However, this data proved unusable as it was not possible to derive anything that could be compared to EPN gas demand estimates from it as:

- It gave no indication of the direction of gas flow in the network at the measurement point.

- The data was for the absolute peak flow shown at each measurement location. These peak flows were all at different times and dates so it was not possible to establish the actual demand on any particular part of the network at a single instant.

B.2.3 Future heating systems and insulation costs

The scale of future deployment of insulation and heating systems[69] could have a strong impact on their costs:

- Single building retrofit insulation measure costs for the pilot studies were derived from the ETI's 'Optimising Thermal Efficiency of Housing' project. It is hard to quantify the influence of large scale, area wide schemes on cost reductions. It is also difficult to predict whether retrofit will occur piecemeal based on individual building owner decisions (especially in areas which are predominantly owner-occupied) or would be driven by large scale schemes.

- Heating system costs for the pilot studies were derived from the underlying data used to set the Renewable Heat Incentive subsidies for different technologies. These costs probably do not represent future costs with high uptake and it is hard to quantify what these might be.

- Current Heat Interface unit costs are high in the UK as many are bespoke units built in small numbers for individual schemes. There is potential for significant cost reductions with large scale manufacture and increased use of common parts. However, it is hard to predict the scale or timing of any future cost reductions.

B.3 Representation of networks

B.3.1 Mapping existing networks' characteristics

There are many areas of uncertainty associated with energy network connection and capacity data.

Connection points

EnergyPath™ Networks lets the user define network connection data flexibly where available. In practice gathering consistent network topology data is challenging since:

- Network operators have typically provided information from a variety of sources for the pilot projects and it is not unusual for these sources to be inconsistent. For example, substation names are often different which can lead to uncertainty when assembling a set of input data that contains substation locations, capacities and connections.

- Existing DNO data sets may not be complete and imputing missing values can be very difficult. For example, whilst capacity data was generally made available for electricity substations there were occasional instances when capacity data was missing for particular substations. There are also privacy issues with household-level data such that 2 out of 3 network operators in the pilot projects were unwilling to share customer level connection data.

- Gas and electricity network operators use the Meter Point Administration Number (MPAN) as their primary reference for building connection information whereas EPN uses the OS Unique Property Reference Number (UPRN) as the primary building identifier. This can lead to

---

[69] This is hard to estimate in EPN alone as technology learning curves can rely on national or even global supply chains and industries.

significant problems when attempting to map network data to buildings. In the case where this was attempted for electricity networks over 18% of the mappings from MPAN to UPRN were not available. An assumption had to be made that MPANs will generally have similar numbers in a particular geographic area. There was no data available to test this assumption so it is unknown what level of error it might introduce to the results.

- If building data is provided using addresses rather than a UPRN it can be extremely difficult to match the data to buildings. This is because address fields may not have a consistent naming convention between data sets, are often filled in inconsistently, or are likely to contain spelling mistakes and have missing fields. The result is that, even when additional building information is available, it can difficult to match and use it reliably.
- There are inaccuracies in the gas network connection data based on a list of off gas grid postcodes issued by Xoserve. As an example, a single postcode in an off-grid area can be missing from the list. There is no alternative publicly available data source that can be used to quantify the magnitude of these errors.

Capacity

Whilst capacity data has generally been made available for electricity substations there have been occasional instances when capacity data has been missing for particular substations. In general, this is not expected to significantly influence EPN results as missing capacities can be sized on estimates of current network load with headroom added. The capacity increase required for transitions will be similar provided the initial network capacity estimate is approximately correct.

When electricity networks are built in EPN their capacities are modified if the estimate of current load produced from building modelling is larger than that which can be met using the defined asset capacities. Errors in building modelling (normally related to the difficulties associated with obtaining good quality data on non-domestic buildings) can result in estimates of network loads on individual network assets being significantly different from actual network loads with the result that network capacities are incorrectly set in the capacity allocation process. These excess demands and corresponding network capacities are then carried forward and may result in inappropriate optimisation decisions.

B.3.2 Reinforcement and operational costs

Data is sparse on the capital costs of heat networks as operators do not have the same reporting requirements as the gas and electricity network operators and so do not generally make this commercially sensitive data publicly available. In addition, since there are far fewer heat networks in the UK the number of examples that could be used to estimate typical costs is far smaller.

The biggest factor in determining the cost of installing or reinforcing buried network assets will be the local ground conditions. In general the assumption has been made that ground conditions are such that average network installation costs apply. However, it is known that actual installation costs can increase by over 70% if dig conditions are particularly difficult. Whilst the EPN input data has scalars to account for the differences between rural, suburban and urban network installation costs and can apply additional scalars to allow for the difference between easy and hard dig conditions the data is not available to robustly map these scalars to specific areas. Similarly, data on which parts of the local electricity network are pole mounted and which are buried were not always available but can be expected to influence network costing.

When specifying network assemblies assumptions are made regarding the precise make-up of the underlying components. Examples include the number of pumps required per meter of heat network or the spacing of electricity network pylons. In the case of electricity network assets these assumptions have been agreed with the local DNO. For heat networks the assumptions have been based on advice from Arup. In both cases they are unlikely to correctly reflect specific locations.

Within the EPN modelling framework **operation and maintenance costs** are calculated based on costs per unit length of network and vary by capacity. However, network operators do not gather or report their data based on these metrics. The data is currently reported across the whole of the network that they operate rather than being broken down to more local areas. In addition, it reflects the fact that most of the current networks are in the middle of their technical life such that annual asset failure rates are only around 1% of the installed asset base. This means that the reported figures may be inappropriate when considering longer term maintenance costs to 2050.

B.4 Representation of other local area features

B.4.1 Service demands

Little data is available for **domestic hot water demand** profiles. The best UK data appears to be from a monitoring exercise conducted by the Energy Savings Trust (EST) which measured hot water use in 120 houses for a year[70]. Whilst this data is valuable, the sample size is small so applying these demand profiles to the whole of the housing stock might be inappropriate.

Household size will have an influence on hot water use. The EST data gives some information on how hot water demand varied by the number of household occupants although this shows a wide variation. In EPN hot water demand profiles are defined by floor area, which carries an assumption of occupancy (e.g. small properties will have only one occupant). The demand profiles used in EPN were approximated from the data relating to household size but there is no data available to validate the assumptions used in this process.

Lighting and appliance demands in EPN are based on historic data from the DECC Household Electricity Survey of 250 homes for one year. This data can be broken down by house type and size. However, the sample size is not large when this is done to a very granular level and so it is not clear that it is truly representative of typical UK demands. In addition, predictions for future demands are hard to define due to the large number of factors that could influence them such as use of more efficient devices (e.g. LED lightbulbs), the increased uptake of home electronics and the possible penetration of air conditioning systems into domestic buildings. For the pilot studies future domestic lighting and appliance trajectories were based on National Grid's Future Energy Scenarios.

Electric vehicle demand profiles used in the pilot projects were based on National Travel Survey data. However, there are many assumptions that need to be made to define future EV charging demands, all of which lead to a level of uncertainty in the predictions. These include:

- Vehicle uptake, in particular where in the LA uptake happens first.
- Electric range for plug-in hybrids.

---

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

- Rate of technology improvement.
- Charging time of day (e.g. arrive and charge vs. overnight charging).

Little data is available to correlate the charge demand estimates as available trial data is for small sample sizes and the people involved tend to be early adopters whose behaviour may not be representative of ‘typical’ users.

B.4.2 Distributed energy

Data on distributed energy is not available at a household level.

- **Feed-in Tariff data** is available from BEIS at a Local Authority level and was used in the pilot studies to establish the current amount of PV in each study area. However, the distribution of these installations across each analysis area is not known and was assumed to be spread evenly across the housing stock. It is known that, in reality, installations are often clustered such that the influence on local electricity networks (e.g. reverse flows) will be much higher in these areas than that which has been modelled using an even distribution.
- In one study area data was provided by the local DNO which gave the total PV capacity installed by substation. This allowed a better allocation of the installed capacity onto the electricity network. It should be noted that the total capacity reported by the DNO closely matched the BEIS data.
- Less data is available on solar hot water installations than solar PV installations. The only data that could be used in the pilot studies for solar hot water was that provided by social housing providers. This means that any private installations have not been included in the analysis although their number is small so the influence on overall energy demand is negligible.

Appendix C : EPN usability limitations

C.1 Overview

This section presents the main challenges in running EPN LA energy strategy studies, including:

- **Running the analysis framework and validating results**
- **Interpreting and post-processing EPN results** for use by stakeholders e.g. crafting messages and visualisations

C.2 Running EnergyPath™ Networks

C.2.1 Built-in data validation and error reporting

Automated data validation checking is used across EnergyPath™ Networks, but has developed organically in response to ongoing use of the tool. However, at present:

- In some places, data validation errors stop execution. This can make it difficult for users to understand exactly why the data validation step has failed.
- Not all data errors are captured, with some processes continuing without generating a warning. This can lead to incorrectly populating the outputs from that process step and the user may not readily be able to identify that a failure has occurred. For instance, Problems have been experienced with the steps in the Network Analysis Module that calculate network cost-capacity curves. This process uses multiple processor cores to run individual analyses in parallel. In some cases, the analyses run on one particular core will fail to be written to the results tables. This error appears to be random and is not flagged to the user.
- In some cases, an automated data validation will cause a process step to fail when the problem has actually been caused in an earlier process step. This can lead to lengthy data debugging exercises and the need to re-run several process steps. For instance, some data errors are only apparent in the domestic costing step of the Household Options Module when they are caused by data errors that could have been identified and fixed earlier in the process.

Future updates to EPN will further refine how and where the existing data validation checks are undertaken in response to recent user experience. This reflects a balance between improving usability and excessive model development time, in terms of create a large number of potentially redundant validation checks to capture as many potential issues as possible.

C.2.2 Performance

Assuming a fully configured input dataset it can take around 2-3 days of computational time to run all of the EnergyPath™ Networks modules from end to end given the complexity of the model (circa. 2 weeks if running 100 simulations in Monte Carlo mode[71]). Particularly time-consuming steps in the process are recapped in Table 12 below:

---

[71] Noting that each simulation can be run in parallel on multiple machines (see section 4.2.3).

Table 12 Most computation-intensive EPN steps

Module	Process step	Time taken (estimate)
HOM	Heating system sizing and demand profiles from running EnergyPlus	~2 hours
	Domestic building archetype pathway generation and costing	~2 hours
NAM	Electricity network cost-capacity curve calculation using PSS©Sincal	~2 hours
	Heat network cost-capacity curve development using PSS©Sincal	6h to 3 days
POM	Deterministic run (one sim)	30min to 12h
	Monte Carlo run (100 sims)	~two weeks

These practical timescales limit the ability to explore options and perform sensitivities. Sensitivities that only require changes to the POM input data are often quick to perform. If the sensitivity requires any of the SAM, NAM or HOM modules to be re-run then the time required for completion can be significantly longer.

A large number of performance improvements have been undertaken across the tool already, including parallelisation of many processes to take advantage of multi-core machines. Whilst further significant improvements are potentially possible, these are likely to require more substantial refactoring of code and/or representation of the optimisation problem[72], as most of the 'quick wins' have already been undertaken. Performance improvements can also be obtained through faster hardware and potentially via updates to the integrated commercial software (e.g. improvements to the optimisation solvers).

C.3 Results Interpretation and Visualisation

C.3.1 Output Processing

The direct outputs from EPN are contained within database tables. It is normally the case that these tables must be post-processed to produce maps, graphs or tables which can be used to communicate the results to stakeholders. This post-processing must be done with care to ensure that the different outputs are joined correctly and that they form a self-consistent set of results. The production of outputs for stakeholders is a largely manual process and can be time consuming.

The ESC has identified that the process could be streamlined by including some of the post-processing software as views in the EnergyPath™ Networks databases or as an add-on to the Python framework.

C.3.2 Output Presentation

Different types of stakeholders have different priorities and levels of understanding associated with Local Area energy strategies, both of which need to be considered when presenting EPN outputs. This is partly related to the complexity of the outputs with multiple factors (technology types, technology capacities, different archetypes, different networks etc.) all changing over space and time.

The ESC has developed an output dashboard to help with results presentation. It allows users to slice the data in different ways to see what they are interested in. It works at a ward or analysis area level of disaggregation. Analysis areas are considered to be the level at which outputs are most robust as this is

---

[72] For example, converting network reinforcement choices to a series of "pathway" choices, analogous to that for buildings could significantly reduce the problem complexity by removing less or irrelevant choices (e.g. never choosing to reduce capacity of heat or electricity networks after reinforcement). However, this would require a substantial re-working of how data is passed from the other modules to the POM.

the scale at which decisions are made within the optimiser. Feedback from Local Authority users included:

- Local Officers and Councillors are likely to want to be able to get maps showing which homes and streets move to particular energy networks or heating technologies. They see this as very important in terms of planning long term action and associated engagement with both residents and their political representatives.
- It would be valuable to be able to undertake simple “what-if analysis”, although this requires the relevant “what-if” runs to have been undertaken in EPN to allow them to be incorporated in the dashboard.

This gives a good indication of the challenges involved in communicating both the strengths and limitations of the modelling process and the validity of the modelling outputs to non-expert stakeholders. The ESC is investigating potential solutions to these issues at the time of writing.

Whilst councillors are interested in the pathways for individual buildings in the city, this level of disaggregation is not appropriate for the optimiser outputs. Although the functionality technically exists to spatially disaggregate the analysis at close to street level, this is currently computationally intractable and more importantly suggests a level of certainty and knowledge from the results that is inappropriate.

ESC’s preferred approach would be to use the EPN outputs to understand which heating systems have been selected for particular housing types in each analysis area and to use this to identify the preferred network solution for different areas. A manual operation would be undertaken to identify target areas for deployment of particular solutions. This would allow consideration of important aspects that are not included in the modelling work such as fuel poverty or other schemes already planned in particular areas that might influence decision making.

Document Control

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

Programme	Smart Systems and Heat Phase 1
Project	WP2 Bidders Pack
Title	EPN Knowledge Transfer Bidders Pack D4 Modelling Local Energy System Designs with EnergyPath™ Networks
ETI Project Number:	SS9016
ESC Project Number:	ESC00066
Version:	Version 2.3
Status*:	Final
Restrictions**:	Confidential
Completion Date:	27/02/2018
Author	Baringa
Reviewer	Grant Tuff
Reviewer	Richard Halsey
Approver:	Emma Harrison

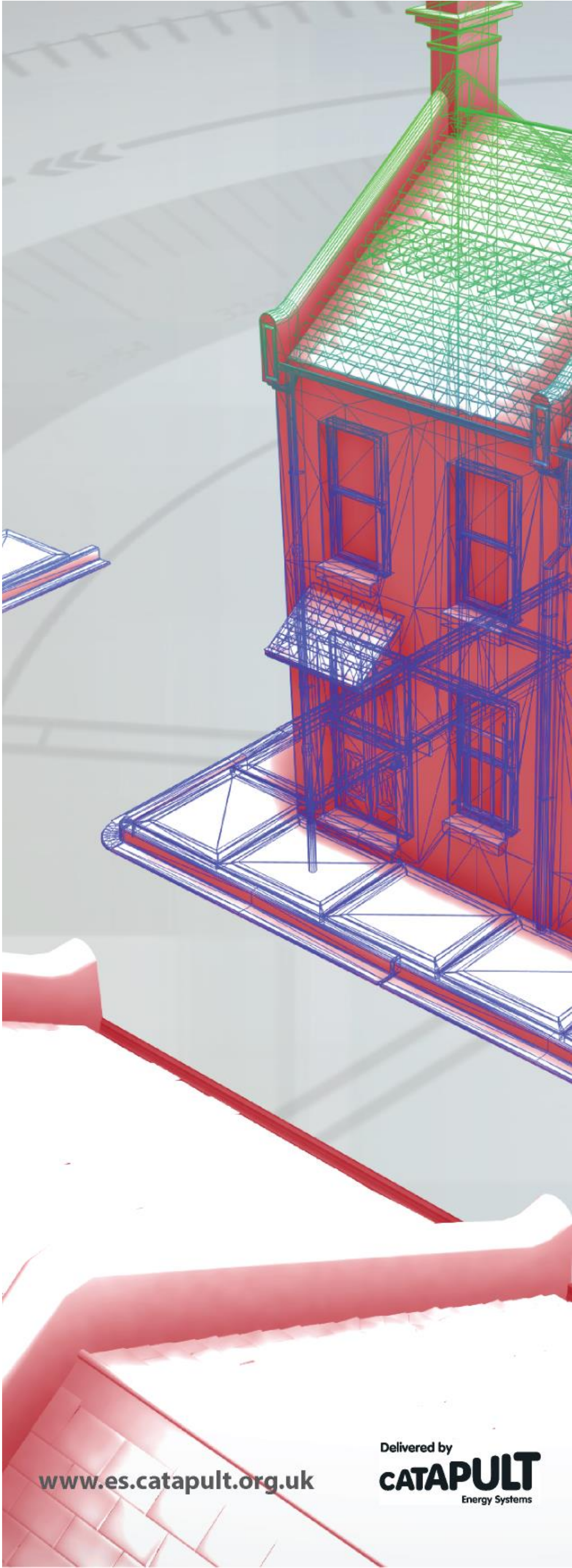
* Status defined as follows – Draft: Contains preliminary information only. Released: Contains reviewed and approved content.

** 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 Energy Systems Catapult Ltd 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.

Revision History

Date	Version	Comments
11 October 2017	V0.1	Initial Draft of skeleton structure
31 October 2017	V0.3	Full draft
27 November 2017	V1.0	Full draft including integration with EPN Limitations and Uncertainty document
27 February 2018	V2.0	Final version incorporating ESC comments
16 March 2018	V2.2	Review for approval
10 August 2018	V3.0	Amended following ETI feedback



www.es.catapult.org.uk

Delivered by
CATAPULT
Energy Systems