Programme Area: Smart Systems and Heat

Project: EnergyPath

Title: The EnergyPath ™ Design Tool Functional Specification

Abstract:
This document describes the functional specification of the EnergyPath™ tool that was developed at the beginning of the design phase of the project. The document does not describe the final specification of the tool or how the tool was implemented but is useful to provide context to the functionality of the final version.

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

Energy Technologies Institute
Smart Systems and Heat Programme
The EnergyPath™ Design Tool

(D1) Functional Specification

CLIENT: Energy Technologies Institute

DATE: 23/04/2014

Version History

Version	Date	Description	Prepared by	Approved by
V0_1	19/03/2014	Initial draft	JG, PT, LH, EP, FL, NAM, FT	OR
V1_0	23/4/2014	Final	JG, PT, LH, EP, FL, NAM, FT	OR

Copyright

Confidentiality and Limitation Statement

This document: (a) is proprietary and confidential to Baringa Partners LLP ("Baringa") and should not be disclosed without our consent; (b) is subject to contract and shall not form part of any contract nor constitute an offer capable of acceptance or an acceptance; (c) excludes all conditions and warranties whether express or implied by statute, law or otherwise; (d) places no responsibility on Baringa for any inaccuracy or error herein as a result of following instructions and information provided by the requesting party; (e) places no responsibility for accuracy and completeness on Baringa for any comments on, or opinions regarding the functional and technical capabilities of any software or other products mentioned where based on information provided by the product vendors; and (f) may be withdrawn by Baringa upon written notice. Where specific Baringa clients are mentioned by name, please do not contact them without our prior written approval.

TABLE OF CONTENTS

1. Introduction

1.1. Background and context

In 2012, the Energy Technologies Institute LLP (ETI) launched the Smart Systems and Heat (SSH) programme, aimed at developing a cost effective smart energy system supporting reduction in CO_2 emissions associated with energy use in buildings in the UK. This is undertaken in the context of the UK's commitment under the 2008 Climate Change Act to reduce emissions of greenhouse gasses by 80% versus 1990 levels by 2050. SSH comprises a number of Work Areas (WAs), including WA2 – EnergyPath™ design tools.

The objective of WA2 is to deliver a tool ("the EnergyPath™ Design Tool") that will enable the ETI to work with Local Authorities (LAs) and their key strategic stakeholders to answer two primary questions:

- For each area, what is the optimum contribution to the UK 2050 target?
- For a given area, what is the optimum plan for transitioning energy supply networks to meet local needs in a manner aligned with the 2050 target?

The ETI issued a Request for Proposals (RfP) for WA2 in June 2013, and subsequently selected a team led by Baringa Partners and including Element Energy, Hitachi Europe Limited and University College London (UCL), to undertake the work. The work is being undertaken in two stages:

- Stage 1 – detailed functional specification and development plan for Stage 2; and
- Stage 2 – delivery of a functional version of EnergyPath™.

1.1.1. Stage 1 deliverables

The Functional Specification (D1) is one of seven Stage 1 deliverables. It should be read in conjunction with the other deliverable documents:

- (D2) Design Architecture
- (D3) Data Architecture
- (D4) Data Acquisition Plan
- (D5) Stage 2 Proposal
- (D6) IP Statement
- (D7) QA processes and test plan

1.1.2. Proof-of-concept work and deliverables

During Stage 1 a number of more 'hands-on' proof of concept pieces were undertaken to help understand the technical challenges of some of the key logic steps being proposed in the functional specification, as well as to more directly assess a number of potential 3rd party software packages that could be integrated within the overall EnergyPath™ Design Tool. This work is summarized briefly within the main document and outlined in further detail in the relevant Appendices[1].

1.2. High-level objectives

1.2.1. Tool objectives

The key objectives with regard to the EnergyPath™ tool are as follows:

- To create a strategic planning tool to support decision-making on future local area energy systems to 2050, principally to help support Local Authorities (LAs) and their strategic stakeholders, such as Distribution Network Operators (DNOs)
- Specifically, for the tool 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
- For the tool to do this by focussing on the development and application of a cost-effective local plan, which is subject to 'real world' constraints and uncertainty; to help inform, more subjective, strategic decision making on a transition pathway as part of wider business planning processes;
- Be able to account, at least indirectly, for the potential impact of consumers on local area energy system pathway design; and
- To ensure that the tool and associated databases are based around a scalable architecture, in order to set the ETI on the trajectory towards building the full level of functionality and capability.

1.2.2. Key end-user requirements

The ETI is assumed to be the main 'model user', however, the outputs and insights generated by the EnergyPath™ Design Tool are likely to be used primarily by LAs. Although they currently have no explicit requirement to set emissions targets, a large proportion of emissions are within their influence (most particularly from residential buildings), and according to the Committee on Climate Change they have a vital role to play in decarbonising the economy. Their main means of intervention in local energy provision are summarised below.

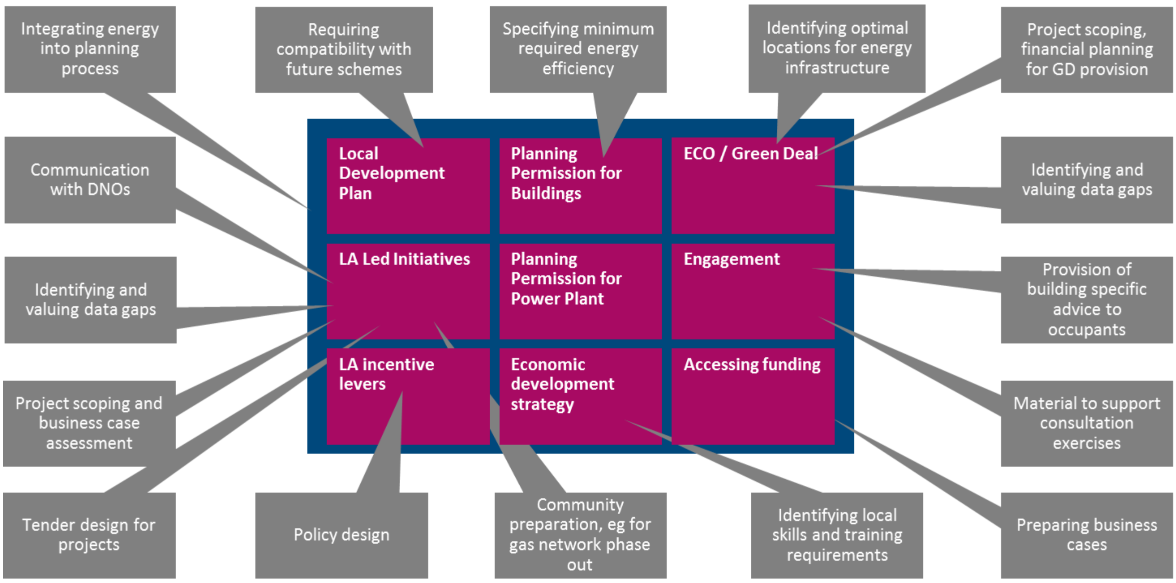
- **Local development plans:** used to identify sites for specific land uses and set out the criteria for approving planning applications.
- **Economic development strategy:** planning for the development of employment and skills in a local area.

[1] Supporting software files (Excel, GIS, etc) were also compiled and delivered to ETI.

- **Planning permission for buildings:** for individual new buildings or extensions, LAs can stipulate planning conditions.
- **Planning permission for power plant:** LAs have responsibility for authorising plant less than 50 MW.
- **LA led initiatives:** LAs may be best placed to initiate multi-party schemes such as district heating, possibly using Public Private Partnerships (PPP).
- **LA building stock management**: including both social housing and public buildings such as libraries, council offices, etc.
- **ECO / Green Deal:** LAs can be Green Deal providers or delivery partners.
- **LA incentive levers:** LAs have limited control over tax and spending policy (e.g., business rates) that could be used to accelerate and direct investment.
- **Community engagement:** advice from LAs, being responsible for residents' wellbeing, may be more accepted by the public as objective and impartial.
- **Accessing funding:** LAs play a key role in accessing funding for local energy schemes from devolved authorities, UK and EU level.

The diagram below identifies ways in which EnergyPath™ could assist LAs in performing their role in local energy provision.

Figure 1-1 The role of EnergyPath™ in assisting LAs plan local energy provision



LAs are likely to use the outputs of the tool to engage with other parties including network operators, property developers and landlords, energy companies, financiers, local companies, community energy groups and final consumers. This engagement will include investment planning in different elements of the local energy system, including buildings of different types, energy sources and networks. The actors in local energy and their roles in different elements of the local energy system are summarised in the table below.

1.2.3. Key end-uses

Across all applications and users, we have categorised the main contexts in which the tool's 'outputs' will be used under five headings, listed below with example uses.

- **Local planning:** developing local master plans; setting local energy objectives; or specifying local interventions

- **Network planning:** helping to support the strategic, as opposed to detailed, planning of the extension, reinforcement or retrenchment of energy networks; or promoting new generation in constrained areas to avoid network investment

- **Financial planning:** assessing project costs, revenues and risks for investment decisions; estimating investment requirements for regulated assets; or understanding the impact of transition on existing business models

- **Engagement:** informing decision makers on optimal choices; engaging with LA executives on energy plans; or engaging with local communities on energy plans

- **National planning:** establishing local targets consistent with national targets

It is envisaged that the direct outputs of the tool will need to be tailored (e.g. different graphical outputs or post-processing of results) for each of these uses, given the different stakeholders involved. In all these contexts, the tool must support and inform what is, ultimately, subjective decision strategic making. As such, it should:

- Allow the user to develop and analyse the timing and quantum of potential deployment options and associated costs that together create a viable pathway towards a desired outcome, and to be able to compare between alternatives.

- Enable comparison to outcomes with no interventions ("Business as Usual") to provide a counterfactual case without intervention.

- Indicate the uncertainty associated with the analysis, and allow for testing of different scenarios, to enable the robustness of a decision to be tested against different outcomes. This will be supported by providing transparency in outputs, such that they can be flexibly disaggregated and re-characterised so as to provide a more nuanced understanding.

- Once particular deployment options have been identified, the tool should also support the design of projects to implement these. As such it will output sufficient detail on the actions and measures employed for a particular option to allow scoping of projects.

1.2.4. Key end-user expertise

As outlined in section 3, the tool requires a complex mix of energy system modelling and subject matter areas. Whilst the user interface will be designed to minimise the interaction with some of the more complex components (such as network load flow modelling) they will still need a basic understanding of the underlying components and their concepts. For example, to be able to configure and apply the tool, undertake diagnostics of problems caused by input data (such as optimisation infeasibilities) and interpret the results.

As such, it is assumed that the model is designed only for an expert user, who has the following competencies:

- Strong quantitative skills, e.g. via a first degree in mathematics, engineering, economics, physics, etc.
- Basic understanding of key mathematical modelling techniques including optimisation and Monte Carlo sampling
- Basic understanding of key energy system modelling techniques such as
  - Electricity and heat load flow modelling
  - Whole energy systems modelling
  - Building energy simulation
- Basic understanding of key software packages including
  - GIS, to be able to interact with, add new or interpret key spatial input data and results
  - Databases, to be able to install and manage overarching sets of input data and results files across the tool
- Detailed energy systems understanding of
  - Domestic and non-domestic buildings
  - Energy networks - electricity, heat, gas and hydrogen
  - Other local area features – e.g. embedded generation
  - Other national level energy system features that impact on the solution – e.g. transmission connected electricity generation

In addition to the above, development skills will also require a core understanding of some of the key software languages proposed for the tool. These are outlined in more detail in the Design Architecture Deliverable (D2), but include SQL database programming, the AIMMS optimization language and Python.

1.2.5. Commercial use of the tool

The commercial model under which the EnergyPath™ Design Tool or its outputs will be made available to customers will be a decision by the ETI. This decision has not yet been taken but could include sale of the tool or the provision of products and services with use of the tool retained within the ETI as part of four possible approaches:

- Licensing to the ETI and its Members and potential also to their global affiliates
- Onwards assignment/sale of the tool to an entity
- Provision of commercial products and services by the ETI
- Sublicensing of the tool in order to provide commercial products and services by a third party (either alone or in conjunction with the ETI)

The design of the tool, particularly the incorporation of commercial 3rd party software tools and data will have implications for these possible approaches which are considered as part of Stage 1.

1.2.6. Outputs and their use

Ultimately, the tool should allow results to be interrogated and analysed in whatever manner and through whatever metrics are of value to the final users. Section 4 explores in more detail some anticipated means of presenting results along with the overarching use of the tool to deliver insights.

A typical interrogative process a user might be expected to go through, drilling down into the results from a high level overview down to specific projects, and some indication of outputs to facilitate each stage, are outlined below:

- **What are the high-level features of the overall area pathways?** For example, how total costs (capital and operational) vary for different pathways versus BaU over the horizon to 2050; how household energy costs are impacted[2]; the technologies which are key in implementing the pathway.
- **What are the key geographical and underlying features of different pathways?** For example, how would a new heat network will 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; the key external factors to which a pathway is sensitive.
- **What does the pathway imply for specific projects?** For example, the roll out of solid wall insulation in postcode XY12; the summary financials for a new CHP plant, determining the economic case for distributed energy storage, or the viability of transitioning a small geographic location to electric heating and potentially decommissioning the gas grid.

[2] 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.

Key end-users of the tool, such as the ETI are likely to be the most demanding, requiring the highest level of interaction with the assumptions and operation of the model. The ETI, or potential subcontractors, will work closely to support LAs as well as other *indirect* users. The latter will be focused primarily on the details of the results, as opposed to necessarily using the tool themselves. Indirect users will also have varying degrees of interest in the outputs.

The diagram below summarises how different types of user may attach different levels of importance to the level of their interaction with the model.

Table 1-2 Required level of model interaction by user

High level of interaction → Low level of interaction

User	Value of usage mode				
	Full interactive model use	Full local area model outputs	Results for local area consultation	Final local area pathway design(s)	Aggregated insights for the UK
ETI	High	High	High	High	High
ETI members	Low	Medium	Low	Low	High
Local Authorities	Medium	High	High	High	Medium
Electricity and gas Network operators	Medium	Medium	High	High	High
Heat network developers / operators	Medium	High	High	High	Medium
Academic / research users	High	High	Low	Low	High
Energy companies	Low	Low	High	High	High
Central Government	Low	Medium	Medium	Medium	High

1.2.7. Informing the national pathway

A key objective for the EnergyPath^{TM} Design tool is to examine local area pathways that are consistent with national decarbonisation pathways to 2050. However, by its nature the tool will provide a more realistic assessment of the costs and effective deployment options in a local area, by more accurately reflecting real-world spatial implications.

These insights may lead to a different aggregate solution at the national level, for example, the costs of large-scale district heat deployment may be higher than anticipated. This may focus national level solutions more towards electrification, but with commensurate impact on transmission connected electricity costs, which in turn affects the local area solution. Hence the interaction between the local area and national level pathway is ultimately a *2-way* iterative process.

It is clearly not pragmatic to wait for local area studies to be undertaken with every LA in the UK before updating the national pathway and hence there are two main routes for feeding back the insights gained at the local level:

- As early studies are undertaken with small numbers of LAs the key features of the pathway designs and the local area can be extrapolated, based on how representative they are as a proxy for other parts of the UK.

- 'Representative', hypothetical local areas (e.g. heavily urban, suburban, rural or mixed) could be generated for use in EnergyPath™, with insights from the resulting pathway designs used to inform the national level – i.e. under what key local area conditions are certain designs favoured over others[3].

1.3. The EnergyPath™ Design Tool within the modelling landscape

The EnergyPath™ Design Tool is unique in the problem that it is trying to tackle, which combines the assessment of what a 'good' pathway design for a local area energy system is whilst considering in parallel:

1. A granular representation of local area building and network level options, that can easily accommodate more real-world data as it becomes available (e.g. from building level surveys)

2. Comparing trade-offs across multiple energy vectors and technology choices, across buildings, networks and other energy system features

3. Consistency with relevant features of the 'National Pathway decarbonisation blueprint'

4. Evolution of the design over a time horizon to 2050

5. Consideration of real-world constraints (e.g. planning restrictions) and the indirect impact of constraints imposed by consumer behaviour (on energy demand behaviour or technology uptake) on the pathway designs

Existing tools target one or more of these elements, but not in combination. For example,

- National-level energy system optimisation models like ESME or UK MARKAL explore 2, 3 and 4, but have very limited or no representation of individual geographic areas

- A wide range of planning tools or stock models for buildings, or electricity and heat networks, exist individually to tackle 1. E.g. WPD's spatially detailed, full-load flow modelling Falcon SIM[4] project or the more aggregated Transform model[5] which only

---

[3] For the long-term development of the EnergyPath™ Design Tool it would be possible to design an area generator to quickly create a random set of local area conditions based on key input parameters such as the total number of buildings and % urbanity.
[4] http://www.westernpowerinnovation.co.uk/Falcon.aspx
[5] Created under the Auspices of the DECC / Ofgem Smart Grid Forum All GB DNOs, Ofgem and DECC have a

consider electricity. Or alternatively, CitySim[6] which provides a very detailed spatial representation of building energy simulation – at individual building level - across an area. However, all of these do not tackle 2 as they only consider a small set of energy system options or energy vectors; and as result do not attempt to trade-off the suitability of developing a heat network against longer-term electricity reinforcement, subject to wide-spread heat-electrification and insulation measures.

- In addition, many detailed planning tools focus on the relatively near or medium term rather than a longer term pathway (4) and are not understood within the context of wider national decarbonisation objectives (3)

The closest to the proposed EnergyPath™ Design Tool is UrbEn[7] (integrated modelling framework for urban energy systems – also known as SynCity) was produced by Imperial as part of a BP funded research project.   This is comprised of four inter-linked layers looking at i) the optimisation of overall city land use layout (primarily to minimise travel requirements), ii) agent based modelling of transport and energy demand behaviour, iii) optimisation of energy use and transport of resources, iv) detailed network operation. UrbEn passes information through layers i) to iv) and then iterates to explore solutions for urban city planning over time which help to minimise carbon emissions. The question UrbEn is trying to tackle is, however, much broader than the EnergyPath™ Design Tool and extends to transport planning and agent behaviour, which means there is correspondingly less detail in the pathway optimisation for the energy system aspects (only partial optimisation of some elements in layer iii) and more limited resolution of building energy use. For example, the more detailed engineering network configuration in iv) is estimated after the more aggregated flows have been established iii). This potentially runs the risk of specifying less appropriate upgrades by starting with the simplified optimisation earlier in the process. Although it would then be possible to iterate back to the simplified trade-offs and more detailed design, this loop makes it more difficult to explore wider uncertainty in the pathway itself.

The crux modelling issue for the EnergyPath™ Design Tool is to create a framework that can bridge appropriately the various dimensions of spatial and temporal granularity and multiple trade-offs across different parts of the energy system (in particular network build and building reinforcement) over the full pathway to 2050. This is particularly important given the long-lead times for larger-scale infrastructure upgrade and development. As it is likely to be intractable to hold significant detail on all of these aspects simultaneously, it will be important that the end-user can easily flex the level of detail in different parts of the tool 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 demand and networks) are as detailed as possible. Hence in contrast to UrbEn and some other models the problem is inverted and a detailed assessment of possible options (e.g. via network flow modelling) is undertaken *first* and this is used to inform

---

Royalty free licence to use the software. Other users may access the model on a commercial basis. The model only assess example, rather than spatially specific networks, and does not undertake network flow modelling.

[6] http://citysim.epfl.ch/

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

the options in the pathway optimisation. The complexity in this approach is clearly in assessing a wide spectrum of appropriate network / building options in detail first. To enable the above, the tool should be able to draw on a number of well-established modelling practices in key areas such as network flow modelling, dynamic building energy simulation, and pathway optimisation. Hence it is important to understand how these could be integrated within the tool to increase both the sophistication of the tool and make its development as cost-effective as possible.

Finally, once the more simplified representations have been explored within the pathway under a wide range of conditions and their resilience explored, it is important to be able to loop back and explore some of the key features of the proposed pathway (such as the network development) in more detail; via the same tools and processes that created the detailed representation initially.

1.4. Purpose and structure of this document

The purpose of this functional specification is to document the required properties and behaviour of the EnergyPath™ Design Tool in sufficient detail that a software developer could develop the requisite code, and a user can understand what tasks the tool will perform. As such it details the data, calculation processes or logic, controls and interfaces the tool will need to employ to support the expected outcomes. It is structured as follows:

- Sections 2 and 3 provide an overview of the design requirements and the tool itself
- Section 4 discusses how the tool would be used as part of a wider set of business processes to support the objectives of the end-users
- Sections 4 to 8 outline in more detail the specification of the requirements, logic steps and implementation of the individual component modules of the overarching tool:
  - The Household Options Module (HOM)
  - The Spatial Analysis Module (SAM)
  - The Networks Analysis Module (NAM); and
  - The Pathway Optimiser Module (POM)
- Sections 9 consolidates the discussion with respect to the key areas of uncertainty across the tool (both data and modelling approach)
- Section 10 provides a summary of some of the open functional considerations that will be finalised early within Stage 2 of the project
- Section 11 provides a set of concluding remarks

1.5. Glossary of key terms

Table 1-3 Glossary of key terms

Term	Description
Area	The overall spatial area considered by the tool within its analysis, such as a Local Authority
Capital costs	The investment costs for new technology options, the lump sum investment is converted into a stream of annualized costs (at a given discount rate) over the economic lifetime of the investment
Cluster	The spatial building blocks seen by the POM, an individual cluster is comprised of a number of unique zones. All clusters together represent the whole area under consideration.
Constraint	A restriction on the values variables can take within the optimisation process
Correlation	The degree of dependence in the values of one factor due to another. Within the tool this focused on the degree of correlation in input factors simulated from different distributions.
Database	A comprehensive collection of related data organized for convenient access
Dynamic building simulation	A model of the operation an individual building across the day or multiple days to understand the dynamic evolution of heating supply and thermal loss / gain
Energy services	The requirement for space heating, hot water, lighting, transport, etc as opposed to the energy used to deliver these service
Energy vector	The energy type used to deliver the above services including coal, oil, gas, electricity, biomass, hydrogen, etc.
Load flow modelling	Modelling of the steady-state operation of energy networks to understand the need for control and/or expansion of the system. For example, for electricity it is used to calculate the voltage drop on each feeder, the voltage at each bus, and the power flow in all branch and feeder circuits.
Operating costs	The costs associated with the use of a technology over the year after it has been built. These can be divided into fixed operation costs (FOM), which must be incurred such as annual maintenance versus Variable Operating Costs (VOM) which are affected by the scale of operation, for example due to additional 'wear and tear' under very high operation.
Optimisation	This process aims to maximise or minimise some objective (e.g. profit or cost, respectively) given available variables.
Pathway	The combined timeperiods considered from now until the final modelled timeperiod (e.g. 2015, 2020 … 2050)
Probabilistic simulation	Repeatedly undertaking the same core modelling process (i.e. the optimisation) with different sets of input data that have been generated as part of a Monte Carlo simulation process
Timeperiod	A block of years considered within the tool, within which the timeperiod represents the mid-point. For example, if the pathway is represented as 5-yearly steps 2020 may represent the period from 2018-2022.
Timeslice	Within year disaggregation of time by characteristic days and within day period as part of supply/demand balancing and representation of peak demand requirements. The combination of the two allows for the representation of e.g. a typical winter overnight period or an extreme winter early evening period.
Vintage	The year within which a technology/network/building option has been constructed, which is separate from the year in which it is operating. For example, the timeperiod for 2030 may contain a mix of plant which has been built in 2030 itself, as well as those built in 2020 and 2025 which are still within their technical lifetime. The earlier vintages of the same plant type may have slightly different characteristics such as lower efficiency.
Variable	A choice within the optimisation process as part of maximising or minimising the objective function. The freedom to change the value of the variables can be restricted by constraints
Zone	The smallest set contiguous spatial building blocks within the tool (e.g. a street). Groups of contiguous zones form clusters and groups of contiguous clusters represent the area as a whole.

1.6. Acronyms

Table 1-4 Acronyms and associated elaborations

Acronym	Elaboration
API	Application Programming Interface
ASHP	Air Source Heat Pump
BaU	Business as Usual
CAPEX	Capital EXpenditure
CHP	Combined Heat and Power
DCLG	Department of Communities and Local Government
DEC	Display Energy Certificate
DECC	Department of Energy and Climate Change
DH	District Heat
DHN	District Heat Network
DNO	Distribution Network Operator
EHS	English Housing Survey
EPC	Energy Performance Certificate
ESME	Energy System Modelling Environment
EST	Energy Savings Trust
EV	Electric Vehicle
FEED	Front End Engineering Design
FOM	Fixed Operating and Maintenance (costs)
GIS	Geographical Information System
GOR	Government Office Region
GSHP	Ground Source Heat Pump
GUI	Graphical User Interface
HLR	Heat Loss Rate
HOM	Household Options Module
HV	High Voltage (network)
LA	Local Authority
LiW	Living in Wales (survey)
LLSOA	Lower Layer Super Output Area
LP	Linear Program
LV	Low Voltage (network)
MIP	Mixed Integer Program
MLSOA	Middle Layer Super Output Area
NAM	Network Analysis Module
NPV	Net Present Value
OS	Ordnance Survey

OTEoEH	Optimising Thermal Efficiency of Existing Housing (ETI project)
POM	Pathway Optimisation Module
PV	PhotoVoltaics
SAM	Spatial Optimisation Module
SAP	Standard Assessment Procedure
SBEM	Simplified Building Energy Model
SHCS	Scottish House Condition Survey
SWI	Solid Wall Insulation
TM	Thermal Mass
UPRN	Unique Property Reference Number
VOA	Valuation Office Agency
VOM	Variable Operating and Maintenance (costs)

2. Overarching design requirements

2.1. Key design principles

Key design principles for the EnergyPath™ Design Tool were tested and refined with the ETI and key stakeholders through a series of workshops and discussions. They have guided the definition of the structure and proposed operation of the tool, and are listed below.

- **The nature of the EnergyPath™ Design Tool:** the tool's overall objective when helping to construct a local area energy system pathway is to minimise energy system 'resource costs[8]', not to maximise social welfare or profit.

- **Co-optimisation of energy system choices:** building, technology and network choices will be co-optimised. However, it is important that the user has full flexibility over the choices that are controllable versus uncontrollable such that tool be applied to tackle a wide variety of possible questions.

- **Analysis of uncertainty:** the model will capture uncertainty in inputs probabilistically, and facilitate exploration of their impact, but not attempt to produce optimal hedged positions via formal stochastic optimization. The tool will be designed to help the user understand the resilience of particular pathway or deployment options.

- **Required level of resolution:** the tool should be designed around an overarching vision of "data-driven" flexibility with respect to spatial and temporal resolution, subject to specific maximum limits in complexity within the networks and spatial modules. The user should have the ability to aggregate up time periods or geographical coverage, or drill down to greater granularity, to extent the data allow it.

- **Accounting for consumer behaviour:** consumers will not be modelled as entities within the model itself. Rather, the flexibility to test the impact of consumer preferences and behaviours through exogenous input assumptions or constraints will be included. This can be divided into two main areas

  - *Consumer technology preferences* in the selection of building options such as insulation retrofits or new heating systems. Whilst the default approach assumes rationale economic decision making by consumers it will be important for the tool to be able to test the implications of different consumer preferences on pathway designs. For example, this could adjust the costs of discount rates applied to different technologies to monetise the 'hassle costs' or force in a given expected deployment of certain options.

  - *Energy using behaviour* as above whilst the default assumption is economically rational, for example the use of storage within in a building to minimse overall

---

[8] With the ability to include other monetised 'social costs' where appropriate such as air pollutant damages

energy use/costs whilst maintaining comfort levels, it should be possible for the tool to test the impact on different demand profiles on pathway design.

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

- **Flexibility in optimisation problem granularity:** the tool will support optimisation at different temporal and locational resolutions to provide control over the computational effort required for solving.

- **Flexibility in defining cluster boundaries:** the tool should support manual intervention to override clustering decisions or set thresholds or parameters used in the clustering algorithms.

- **Flexibility in accommodating data:** the tool should allow integration and make best use of existing data as well as anticipated future data to be gathered from LA surveys or other channels.

- **Post processing integration of other data sets:** the tool should support the layering on to the results of other data (potentially in a GIS) to support end-user requirements, for example combining energy system cost with a socioeconomic datasets to explore the implications of a blueprint design on fuel poverty

Strategic and 'free market' investments

As mentioned above the overall tool framework is designed to be highly flexible, so that the user can apply it to explore a wide variety of questions. This means that automated parts of the tool, such as the optimisation processes, should be easy to configure such that the user can fix or control these aspects or leave them 'uncontrolled'.

For example, as an extension of accounting for consumer behaviour in the previous section, it is important within the analysis undertaken by the EnergyPath™ Design Tool to be able to understand the impact of wider 'free market' investments on more strategic investments, such as network infrastructure development. For example, consumer preferences may mean that the uptake of new heating systems does not follow an economically rational deployment pathway. A much longer continuation of gas boiler use may be followed by a rapid uptake of alternatives, or conversely the alternatives may take market share more rapidly than expected.

Given that key local area energy system stakeholders such as LA and DNOs have more limited control over these wider 'free market' investments it is important to understand how sensitive or resilient their strategic investments are to unexpected investments. For example, how is the decision to build a heat network affected if its cost-effective deployment requires a large number of buildings to connect outside of the scope of those managed or owned by the LA themselves.

From a practical user perspective within the pathway analysis functionality (see section 8 for further details), this means that instead of letting the tool cost-optimise *all* available choices this should be limited to network decisions only. As part of the analysis process the user could define a number of fixed levels of building upgrades for different heating systems (i.e. pre-specifying the consumer uptake preferences to be tested rather than letting them be cost-optimised) and explore to what extent the network upgrades change under these sensitivities.

2.2. Other general requirements

The design and development of the tool should also take into account a range of broader requirements, noting that in some cases a balance will need to be struck given competing objectives.

- **Retention of IP by ETI:** in general, it would be preferable for all IP to be retained by ETI so that it has full capability to maintain and modify the tool
- **Minimised licensing costs for end users:** this will broaden the appeal of the tool to potential customers
- **Effective user interface:** this will enable expert users to configure and apply the tool in an efficient manner and will broaden the tool's appeal as part of longer term development, however, the UI design has to be balanced against the need to deliver end-users access to the full functionality of the tool.
- **Minimised overall development costs:** clearly this is in the ETI's interests, but competes with the desire to retain IP to the extent that it requires in house development of functionality that might otherwise be provided by third party solutions
- **Accelerated development timetable:** the ETI's ambition is to have a fully functioning tool available as soon as possible, but this must be balanced against minimisation of development costs and retention of IP

The ETI has counselled a pragmatic approach in balancing these agendas, setting cost effectiveness of the overall tool development as the guiding principle for design decisions, whilst bearing in mind the requirements other potential direct and indirect users.

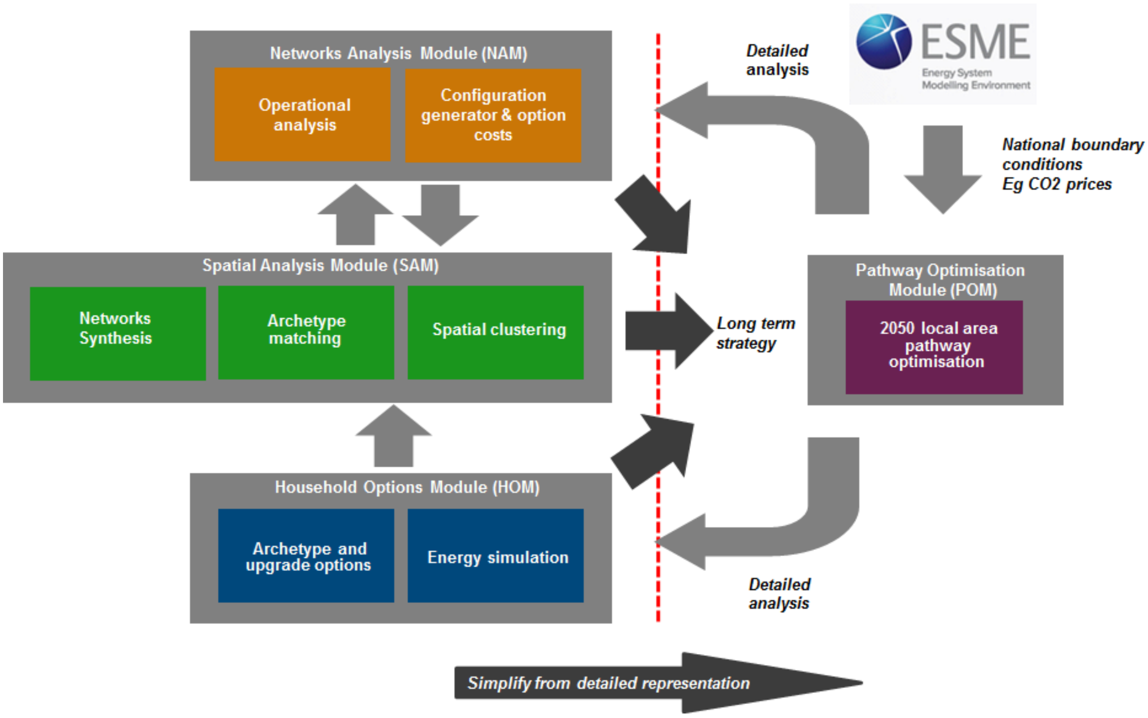
3. Overview of tool

The proposed EnergyPath™ Design Tool can be divided conceptually into four main **modules** as shown in Figure 3-1 below, which include:

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

A high-level overview of each module and their interactions across the tool is provided below, with a more detailed description in sections 5 to 8. The full logic / process flow diagram is shown in the Appendix (section 12).

Figure 3-1 Overview of the EnergyPath™ Design Tool



The primary role of the **HOM** is to generate a database of domestic and non-domestic building archetypes that could be used to represent the existing building stock in an area and the options for evolving this stock over time. The archetypes will be specified by, amongst others, the physical characteristics of the buildings and the heating systems they use (including the resulting profile of final energy demand from the operation of the heating devices), the climate in which the building is located (as identified by the Government Office Region), the occupancy of the building and the costs of converting/upgrading one archetype into another. Given the limited data availability in the non-domestic building sector the process of assessing their energy consumption and upgrade options will be simplified relative to the domestic sector by using performance benchmarks and exogenous profiles rather than dedicated energy simulation. However, for both domestic and non-domestic buildings the framework allows for better primary data to be entered as it becomes available. Consumers will not be modelled as entities within the tool itself. Rather, the flexibility to test the impact of consumer preferences for new technology options and energy using behaviours through exogenous input assumptions or constraints will be included

The primary role of the **SAM** is to create a detailed spatial (GIS-based) representation of the existing local area including key topographical features (roads, rivers, etc), existing energy networks, building locations and other energy system features such as embedded generation. It will assign an archetype (from HOM) to each building. Where the primary data needed to do this is incomplete, it will assign the best match based on available statistical datasets at different levels of spatial aggregation. The SAM also identifies future spatial considerations for the energy system such as the locations where new embedded generation could be sited or allowing users to specify assumptions on the location of new housing developments. The SAM passes spatial information to the NAM to support its analysis of network reinforcement and new build options.

The primary role of the **NAM** is to use data on energy use in buildings, existing networks, and geographic topology to design and cost different configurations of energy networks, for each energy vector (electricity, heat, gas and hydrogen). It outputs spatially varying cost functions for different configurations of each network, which can be used by the POM to compare network build costs and peak load capacity, as part of seeking a cost-optimised pathway.

The SAM and the NAM interact with each other to define a more aggregated spatial representation of the local area, by dividing it into a contiguous set of interconnected *clusters*, chosen to provide a reasonable representation the costs of network options across the different vectors (i.e. by aiming to not bias the data towards one network type or another). This aggregation is necessary to make the optimization problem in the POM computationally tractable, but means that a cluster may, for example, represent a group of streets rather than individual streets. Spatial detail is lost at this stage, but the EnergyPath™ Design Tool is designed to be flexible to allow the definition of the clusters to be changed easily (e.g. adding more clusters to improve the granularity).

Within the tool, the **POM** is the engine that constructs a specific pathway. It receives data from the NAM, SAM and HOM on the representation of the local area energy system and potential future options for evolving this. The POM then takes a more aggregated representation of these options and calculates the choices which produce the lowest cost combination to create a pathway that satisfies a range of constraints and design standards (such as satisfying all householders energy service demands under a range of edge cases).

The national pathway produced by ESME is used to inform some of the **boundary conditions** within the POM, such as carbon or fossil fuel prices, and the availability, price and carbon intensity of transmission-connected electricity. The POM's outputs enable the user to investigate the features of the pathway and key areas of uncertainty.

In addition, the more aggregated POM outputs can also be fed back to the NAM and HOM to validate the feasibility of a resulting pathway, and refine the associated costs.

*It is important to **separate** the way the tool is conceptually constructed, relying on techniques such as mathematical optimisation and Monte Carlo analysis, from the process of developing* *insights to inform more subjective strategic decision making. The following sections describe the separate conceptual elements of the tool in more detail, whereas the wider process of applying the tool is discussed in 4.*

4. Use of the EnergyPath^{TM} Design Tool

4.1. Business process for using tool

4.1.1. Overview

The tool will be sufficiently flexible to accommodate multiple modes of use, which may organically emerge through practical application. For the purpose of illustration however, this section outlines one such process by which an energy system transition plan within an LA area might be developed and implemented. The role of the EnergyPath^{TM} Design Tool within this process is to provide analysis and insight to decision makers to allow them to make strategic choices on the content and execution of the transition plan. The example process is considered at three levels of resolution.

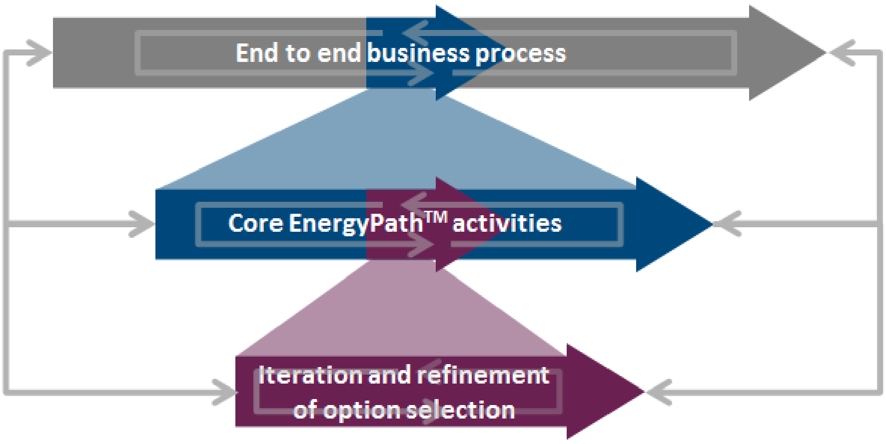
- The **top level** end-to-end process, from the earliest consideration of a transition plan through to execution of projects, and involving the interaction and cooperation of multiple local stakeholders, most particularly the LA and energy distribution network operators.

- An **intermediate level** concerning specifically the core activities surrounding deployment of the EnergyPath^{TM} Design Tool, occurring within the LA and with activities likely led by a team tasked with managing the development of the transition plan.

- The **process of analysis, iteration, and refinement** through which specialist individuals tasked with operating the tool would step in order to produce an initial transition plan (including near term projects) for validation and consultation at escalating levels of authority. This would be the most intense phase of tool operation, though its ongoing use would in practice be likely to continue in a manner intermingled with the broader processes. These tasks are unlikely to occur in a completely linear and isolated fashion. The business model or models pursued by the ETI for commercialisation of the EnergyPath^{TM} Design Tool will determine how exactly this activity is undertaken from a contractual perspective.

These levels are illustrated in the diagram below.

Figure 4-1 Overview of activities in the planning and implementation of a local energy transition pathway



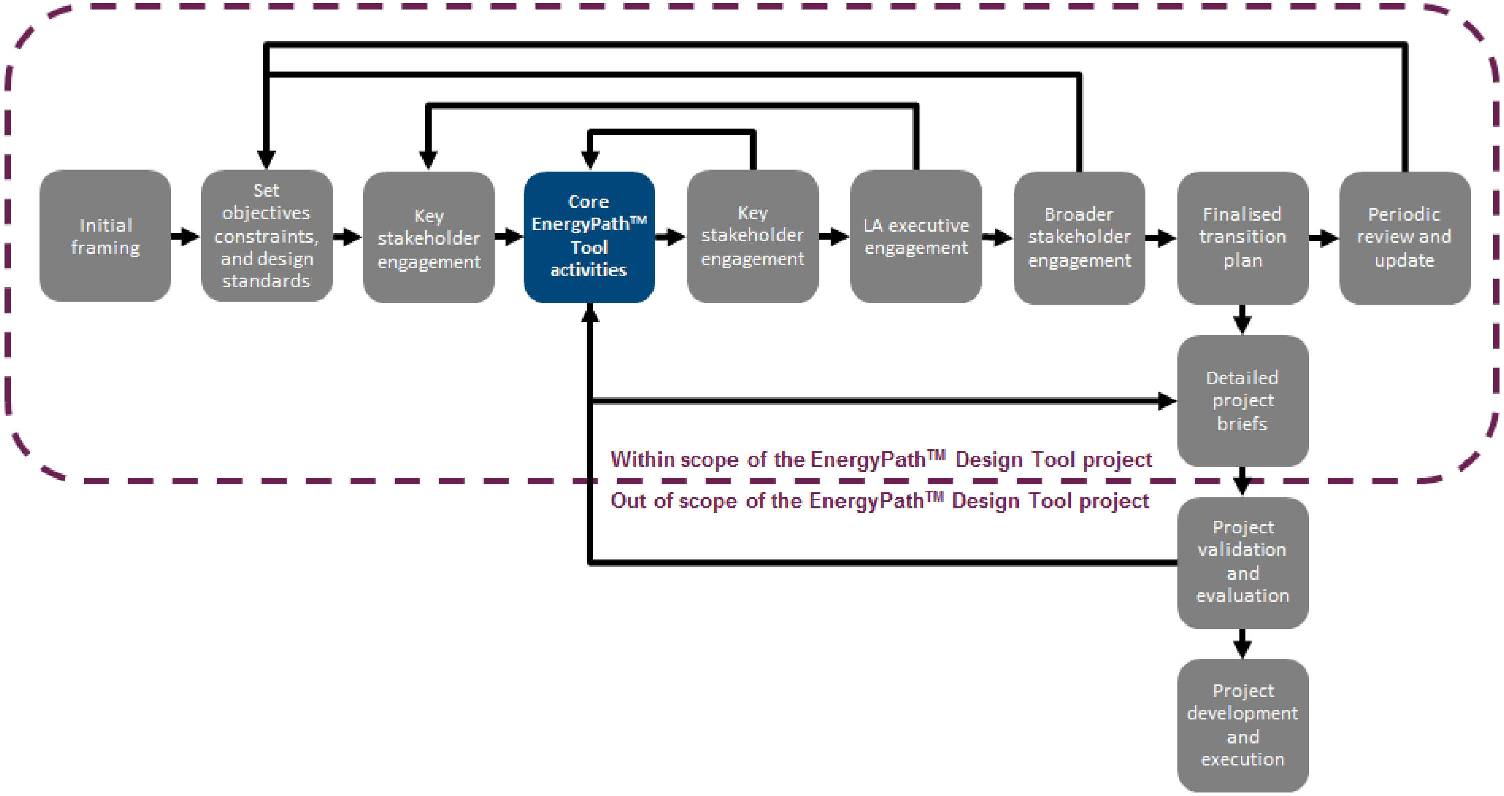
In this section we will refer to the transition plan, describing the plan for the low carbon transition of a local energy system, as being the final product of the end to end business process. This plan would comprise a number of components.

- A defined programme of projects and measures selected for implementation in the near term
- An indication of the main features and alternatives of pathways for progressing transition in the period beyond the implementation of the near term project programme, where uncertainty is too great to allow confident selection and definition of low regret projects; and
- A strategy for monitoring, maintaining and improving the transition plan over time, as time elapses and new data becomes available or is acquired, enabling selection and definition of further projects to facilitate the transition.

4.1.2. End to end process

The full process through which an LA would step, from setting the objectives and targets for a local energy transition plan, through to implementing or initiating the projects which will start to deliver it, is set out in the diagram below.

Figure 4-2 End to end business process for design and implementation of a local energy transition pathway



- **Initial framing:** this step describes activities to provide policy and decision makers with an understanding at a high level of the possible pathways that exist for an area, what particular technologies or options are likely to be viable, and at what range of costs. This will inform the objective and target setting process, and could potentially be informed by a version of the EnergyPath™ Design Tool based on generic local areas (e.g. typical rural, suburban, urban).

- **Set objectives, targets and design standards:** an objective for reduction in carbon emissions may be determined based on an optimal allocation of national level targets between Local Authorities, or via imposition of a carbon price. This might itself be informed through use of the EnergyPath™ Design Tool by a central agency. At the local level, the tool will determine cost effective pathways in this context, and support development of the LA's transition plan. Some LAs may wish to set further targets for the transition plan reflective of their own priorities, most probably concerning social factors. In particular, metrics describing fuel poverty or the level and stability of energy bills might be targeted. Additionally, design standards which viable pathways must accommodate will need to be set, as discussed in section 8.5.3. This in particular will regard inputs to the tool which cannot be forecast or modelled in a reliable manner, but have a direct and significant impact on local energy systems. Some standards may be externally specified by industry norms or regulations, such as the 1-in-20 winter standard already discussed. Other design standards to be incorporated might need to be determined internally by the LA, such as assumptions on demographic change, and how it affects energy usage profiles and building occupancy patterns.

- **Initial key stakeholder engagement:** it will be necessary to engage at an early stage with key stakeholders in local energy provision, which will play a pivotal role in implementing the final transition plan. Most obviously, this will include the owners and operators of the distribution networks for electricity, gas and if applicable, heat. It may also entail large housing providers or the owners or occupiers of significant campuses or installations in the local area. The engagement would involve securing their cooperation in the process, and agreement on the high-level objectives and constraints set.

- **Core EnergyPath™ Design Tool activities:** this step entails preparing the tool, using it to identify viable pathways, discerning a set of projects for implementation in the near term consistent with those long term pathways, and securing support at a certain level of authority within the LA. This is described in greater detail in section 4.1.3.

- **Further key stakeholder engagement:** given the essential role of certain stakeholders in implementing any project or long term transition plan, as discussed, seeking their input on the feasibility and merit of a transition plan and specific projects would be an essential first step after securing internal LA non-executive approval. In practice continuous interaction with such parties is likely to be desirable. Feedback received may necessitate revision and iteration of the tool to accommodate new information or constraints.

- **LA executive engagement:** once a transition plan, including the specific near term projects, has been determined to the satisfaction of the key stakeholders, it is anticipated that executive level agreement within the LA would be sought. This would allow the main decision and policy makers within the LA to satisfy themselves that the plan is consistent with the high level objectives and targets established. Objections at this stage may necessitate a greater degree of revision to the configuration of the tool, such that it would be necessary to re-engage with the key stakeholders to confirm the feasibility of proposed changes.

- **Broader stakeholder engagement:** executive approval of a transition plan is assumed to be the precursor to broader stakeholder engagement. This would involve consultation with the full range of local stakeholders (as identified in section 1.2.2) involved with or impacted by implementation, including local enterprises, community groups and individuals. Communication would encompass the specific projects for implementation in the nearer term, and the longer term element of the transition plan. Evidence of a substantial divergence of public or civic opinion from the parameters assumed in the initial establishment of the tool could entail recasting of the objectives or targets determining the pathway optimisation calculation in the tool, and a significant iteration.

- **Finalised transition plan:** approval at each level of consultation and engagement would result in the finalised transition plan. This will include the set of projects for near term implementation, and the longer term plan for monitoring and managing progress and planning further interventions to deliver the long term transition to decarbonised local energy provision. Some projects may be implemented directly by the LA and some by key stakeholders with which the LA has an established relationship (such as network operators). A large proportion will depend on decisions of private and independent individuals or organisations, for which the transition plan would lay out a strategy of education, engagement and incentivisation.

- **Periodic review and update:** once the finalised transition plan is obtained and implementation of near term projects has begun, long term management of the transition plan will begin. This will incorporate a range of activities, monitoring availability or planning acquisition of new relevant data regarding, for example, local infrastructure, changing local demography, the evolution of technology costs and fuel prices, developments in national level policy and so on. There will also likely be a role for regular review of progress against the plan and redetermination of viable pathways in the light of updated information and completion of projects. This will enable identification and specification of subsequent projects for implementation.

The finalised transition plan will identify a set of discrete projects which are to be implemented in the near term.

- **Detailed project briefs:** the tool will be used at its highest level of temporal and spatial resolution to determine detailed technical specifications and outputs for financial projections. These will provide the basis of project briefs which will launch the process of implementation. In the case of network projects, this will include detailed outputs from the NAM module.

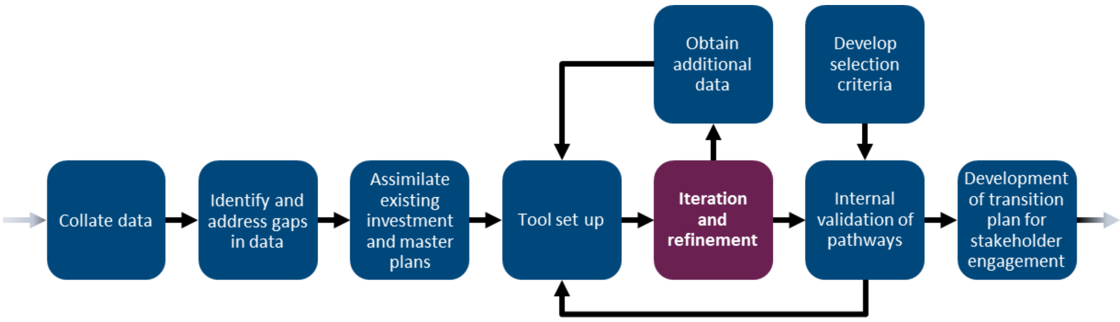
- **Project validation and evaluation:** with a detailed specification of the project in place, an LA would likely perform an initial validation exercise, before committing large investment to implementation. This would include commissioning qualified technical expertise to validate cost assumptions and technical specifications. Financial information provided by the project brief would provide inputs to a financial and business modelling package, being prepared by the ETI as part of a separate work stream. This would facilitate determination of the economic feasibility of the project and inform choices on the sources of capital and commercial strategy employed.

- **Project development and execution:** having scoped a project in detail, and investigated the potential financial and commercial models, the development and execution of the projects would commence. It is recognised that in a large number, or majority of cases, this will not fall within the direct remit of the LA itself, but in that of other stakeholders in the local energy system, as identified in section 1.2.2. This phase will then necessarily involve deep cooperation with these parties. For projects requiring physical intervention, this phase will commence with the commission of Front End Engineering Design (FEED) studies to provide the basis for final design, procurement and construction. For many projects, final implementation might be at the level of private individuals or businesses. In this case LA intervention might be through establishing or reshaping policy incentives or providing information to guide private investment for example.

4.1.3. Core EnergyPath™ Design Tool activities

The diagram below illustrates the steps that would occur within the LA to prepare the tool for use, and prepare its outputs for external consultation. The precise mode of operation and allocation of responsibilities will depend on the commercial model chosen by the ETI.

Figure 4-3 Overview of core EnergyPath™ Design Tool activities



- **Collate data:** it will be necessary to collect and prepare relevant data from various sources to allow the tool to undertake analysis of the local area. In part this will be taken from standard national datasets, and in part from databases on local infrastructure held by LAs or other bodies (for example, details of listed buildings). Close cooperation from network operators or owners of other key assets will be necessary to construct an accurate picture of the local infrastructure.

- **Identify and address gaps in data:** there may be gaps in the available data that would impair the ability of the tool to identify viable and practicable pathways. Additional data may also be required to enable the tool to consider targets set by the LA regarding social or other local environmental factors such as air quality. To address these needs, the LA may choose to commission data specifically, to be gathered through local surveys or questionnaires.

- **Assimilate existing investment and master plans:** LAs may already have in place projects or schemes to develop or influence energy provision in a local area. For example certain areas may be earmarked for development of energy facilities. Development master plans with a broader scope regarding housing and economic development are also likely to be of relevance for example in identifying areas of future development, or existing areas scheduled for redevelopment or repurposing. Distribution network owners may also have associated plans for upgrade of their infrastructure.

- **Tool set up:** having assembled the relevant input data, it will be necessary to compile and format this for entry into the tool, together with representations of the objectives and constraints that have been set to bound the optimisation calculation.

- **Iteration and refinement:** this is the process of identifying, analysing and filtering options within the EnergyPath™ Design Tool, as performed by the user, so as to arrive at a set of projects for implementation in the near term, and indication of the high level features of pathways that deliver long term transition consistent with these projects. This activity is discussed in more detail in section 4.1.4.

- **Obtain additional data:** in the process of operating the EnergyPath™ Design Tool, areas may be identified where gathering additional data would be valuable in refining and reducing uncertainty in the tool's outputs, and aiding the decision making process it supports. This process is described in further detail in section 4.1.4. Where the need and justification for additional data is identified, there would be a process to plan for and undertake its acquisition, and its preparation in a form suitable for the tool's use. It would then be integrated into the tool's configuration.

- **Develop selection criteria:** externally to the operation of the tool, it would be good practice to identify the metrics by which pathways will be judged. This could encompass a broad range of conditions, for example uncertainty on the cost or level of CO_2 savings of the pathway, or the degree of sensitivity to national electricity and fuel prices of the annual cost of energy provision in LA owned housing stock, or the proportion of investment targeted in lower income wards. These may be employed through direct comparison of different pathways, or if quantified limits to the metrics are determined, tests may be formulated that would determine the acceptability or otherwise of a pathway.

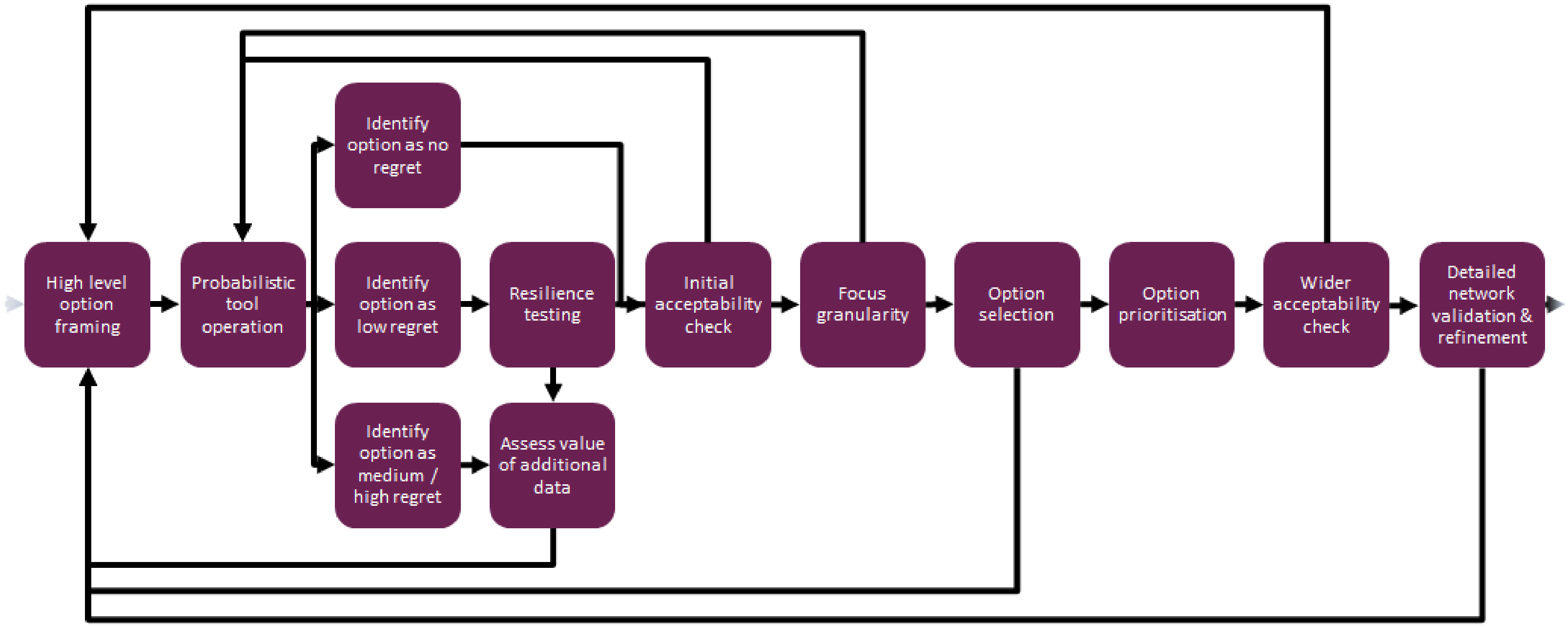
- **Internal validation of pathways:** based on the predetermined selection criteria, approval at an intermediate level of authority within the LA for initial projects and consistent long term pathways would be sought. To the extent that these are not found to meet requirements, it would be necessary to iterate once more, recalibrating the optimisation constraints as necessary.

- **Development of transition plan for stakeholder engagement:** at this stage, a set of initial projects, and the higher level options for progressing transition thereafter, would have been approved. These would require development and documentation in to a more formal transition plan ready for use in consultation with key external stakeholders or presentation to the LA executive. This would contain plans for the ongoing maintenance, management and improvement of the transition plan, and a strategy for the identification and definition of projects for implementation beyond the initial set.

4.1.4. Iteration and refinement of option selection

Due to the computational effort that would be required, it is unlikely to be feasible in one step to obtain a full set of results at the greatest level of spatial and temporal detail and exploring the full range of uncertainty with regards the options. Instead, it is envisaged that temporal resolution, spatial resolution and option uncertainty will be traded off to a certain extent in each operation of the tool, to enable iterative refinement of the solution. The diagram below describes the process the team or individual assigned with operating the tool would follow to identify options, and then specify projects, which will progress the local area transition in a manner consistent with the priorities and constraints set up front. The selected options should also be resilient, such that they would continue to facilitate pathways consistent with the overarching objectives under a range of outcomes regarding uncertainties in the projects themselves (for example, the cost of preparing trenches for the installation of heat network piping), and external factors (for example, electricity prices).

Figure 4-4 Overview of the EnergyPath^{TM} Design Tool modelling steps as part of the refinement of options for a viable local energy system pathway



- **High level option framing:** after configuration of the tool, the first step would likely be operation in deterministic mode, under a number of scenarios. The scenarios would be constructed so as to explore a range of internally consistent outcomes, regarding such factors as national electricity, gas and carbon prices, and potentially others such as technology capital costs. Where the tool selects options in a number of scenarios, they are likely to be strong candidates to be carried forward for further investigation. Determination of how exactly to define an option will require judgement, given the scope for different modes of transition in different locations of an LA area, for variation over time, and for interdependence between options. By considering only a relatively small number of scenarios which still consider a wide range of outcomes however, this problem should remain tractable.

- **Probabilistic tool operation:** having identified options meriting further investigation, the next step will be to consider these more specifically under probabilistic operation of the tool, allowing fuller assessment of the uncertainty around the optimal outcomes of the deterministic runs, and the factors driving selection of the option.

- **Identify option as no regret:** some options may represent sound decisions, in that they reduce CO_2 emissions and / or achieve other objectives in a cost effective manner, under all reasonable circumstances. In this case the option would be identified as no regret, and will be selected for future implementation. Examples may include low cost high effectiveness insulation measures.

- **Identify option as low regret:** some options may be optimal under a wide range of circumstances, that is to say they are selected in a large proportion of the simulations generated by the EnergyPath^{TM} Design Tool in probabilistic operation. These options would be identified as low regret, and would represent strong candidate projects for implementation.

    - **Resilience testing:** the tool user would undertake analysis of the low regret options to determine their resilience. Resilience can be characterised as the "*propensity for an option to remain cost-effective when the tool input assumptions and constraints, in combination or alone, are set so as to produce circumstances that would tend to diminish the cost-effectiveness of the option being considered, or enhance the effectiveness of alternative options*".

    - This may be expressed as a tipping point, especially where an option is sensitive to an input parameter subject to considerable uncertainty. The tipping points are identified through posing questions in the form, for example, of "*what value must occupants of a particular building type attach to visual amenity in order that an air source heat pump is no longer the optimal source of heat provision*?" Having identified this value it would be compared to expectations for the value of visual amenity to inform opinion of the option's resilience. If an option is determined to be resilient it would be selected for more detailed definition and future implementation. If its resilience cannot be determined from the data available, it may be decided that gathering of further specific information may assist decision making (discussed in more detail below).

- **Identify medium/high regret options:** some options may appear to be optimal in a smaller range of circumstances. That is to say that in a significant proportion of the simulations generated by the EnergyPath™ Design Tool in probabilistic operation, the option is not selected. Such options would probably be rejected for implementation in the near term, unless they are shown to be highly sensitive to costs which are known only to a low degree of certainty.

  - **Assess the value of additional information:** in the case that a medium/high regret option is sensitive to costs which are known only to a low degree of certainty, it may be decided that further specific data should be gathered to enable a more confident assessment. This data would be included in the tool's configuration and the process iterated. A similar course may be taken if the resilience of a low regret option cannot be adequately determined without further data.

- **Initial acceptability check:** an initial sense check at this stage would filter out options that, although identified as cost optimal by the tool, may in fact be considered untenable for other reasons, potentially revealing that important practical constraints have not been entered in the model's initial configuration, such as the acceptability of.

- **Focus granularity:** with promising options selected, it will be desirable to investigate them at a more granular level in the tool; that is to resolve their geographical and temporal properties to a greater detail by iterating the tool whilst focusing on particular periods of time or locations in order to enable more precise definition.

- **Option selection:** having been investigated in detail, those options which are considered to be no regret, or optimal under a large range of possible outcomes, have been defined as projects and are selected for implementation. They are fixed in subsequent iterations of the tool, which may alter the propensity of the tool to select other options, allowing them to be selected or rejected, and so on until all resilient low regret options have been identified.

- **Option prioritisation:** the iterative process described above may imply a certain ordering or grouping of projects, where their 'optimality' is co-dependent on the implementation of others. Having defined and selected projects to be carried forward, this step would consider how they should be assembled as a package. For example, some projects may only be implemented once others have been. Some projects may have greater uncertainty associated with them, or tend to lock an area in to a certain pathway. It may be desirable to defer implementation of such projects in order to mitigate against the risk that as circumstances evolve they transpire to be high cost. Careful consideration would also be given to the degree of direct control the LA has on the projects, that is whether they are undertaken directly by the LA, or for example are dependent on decisions of private householders.

- **Wider acceptability check:** at this stage, more detailed checking will be necessary to ensure that, whilst the projects progressed to internal validation are optimal from a cost perspective, they are compatible with wider LA policy or objectives.

- **Detailed network validation and refinement:** a further check will be to ensure that the projects selected are viable given the planned state of local energy networks. In addition, the costs of the potential network options for a given part of the local area could be further refined before incorporation within detailed project briefs. This would occur using the NAM's more detailed analytical capability, as described in section 8.8. *Whilst this step would further enhance the robustness of the projects, it is not intended to be a substitute for the detailed planning and analytical work that individual DNOs would undertake as part of their direct planning and development activities.*

At this stage it is anticipated that a prioritised set of projects, and the high level features of a number of alternative pathways that include the identified near term projects and facilitate the long term transition, would be set and ready for internal validation.

4.2. Key outputs and reporting

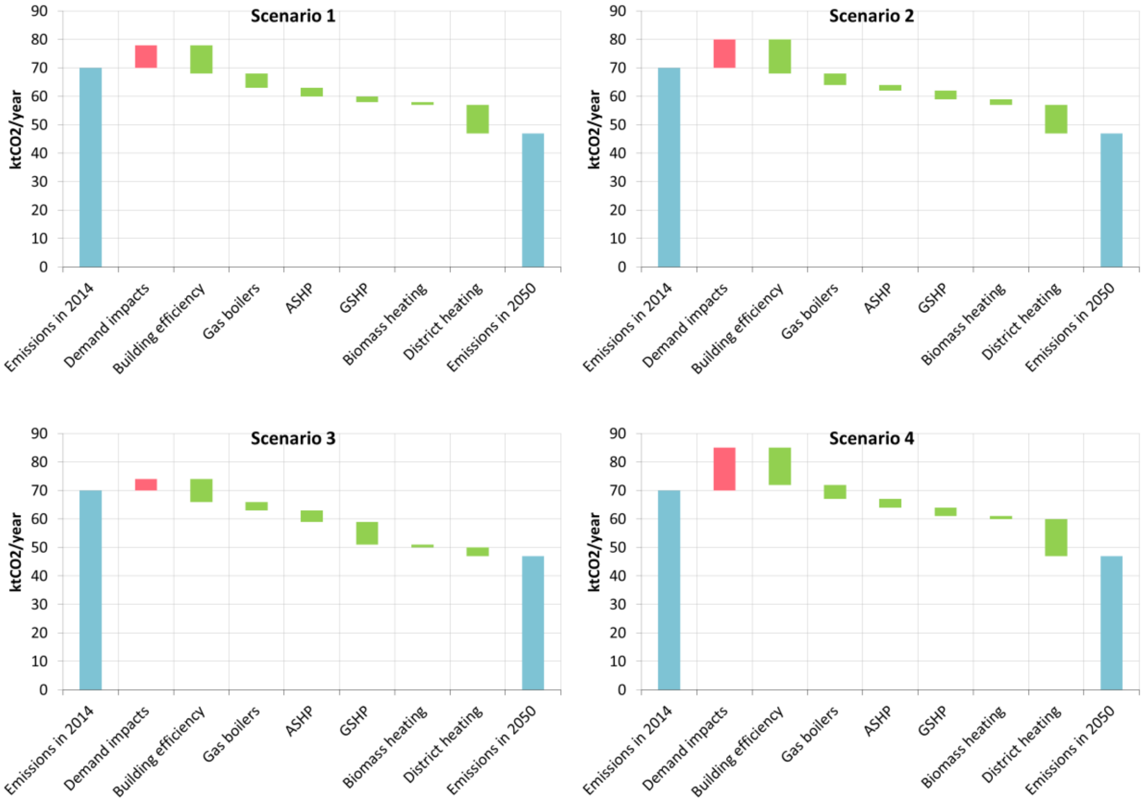
As discussed in section 4.1.1, the tool will be sufficiently flexible to accommodate different methods of use. This section however illustrates some of the outputs and reporting modes that the tool could support by stepping through a process similar to that envisaged in section 4.1.4, and imagining the questions that might be asked and analysis that might be undertaken at each stage. This is necessarily a simplified example, and in practice an LA area may require a much richer data set, making option selection and project definition more complicated.

High level option framing

The first step envisages deterministic operation of the model under a number of scenarios in order to acquire a high level understanding of the options that could comprise a transition pathway, and the areas and circumstances in which there may be competing pathways. The key at this stage would be the ability to summarise data at a level of detail that provides for ready interpretation and communication of the main distinguishing features of viable pathways. One way in which this may be provided is through a waterfall plot illustrating the contribution of different measures to overall emissions reduction over the pathway. If a particular measure makes a large contribution in each scenario, it may be assumed that it is reasonably resilient to different outcomes and is worth investigating.

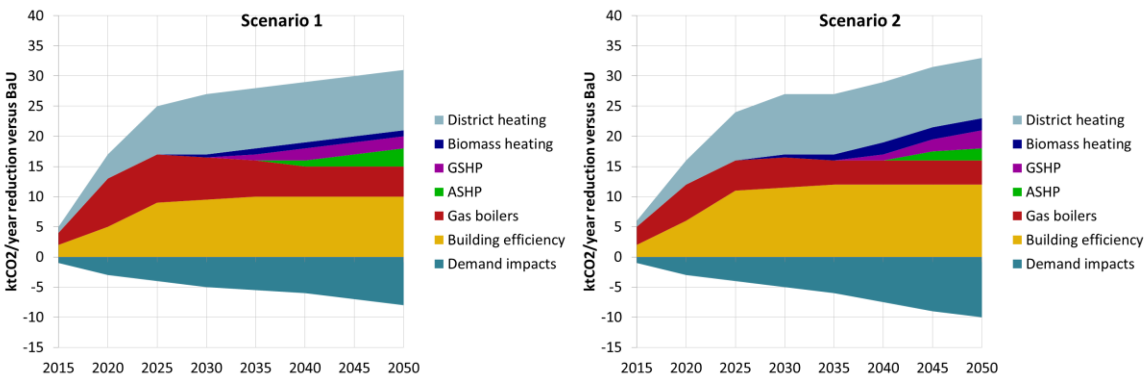
In the example below, the user would observe from the waterfall plot that the building efficiency measure makes a strong contribution to the transition pathway in all of the scenarios considered, and therefore merits further investigation. District heating networks are also selected for large scale deployment in most of the scenarios, but in some cases appear to be displaced to some extent by GSHPs. This option also merits further investigation to understand better the circumstances under which it is preferred, and the circumstances under which GSHPs are the lower cost option.

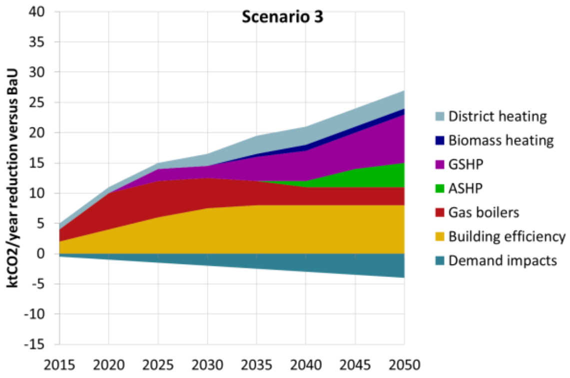
Figure 4-5 Example output: contribution to emissions reduction, 2014 to 2050, by measure

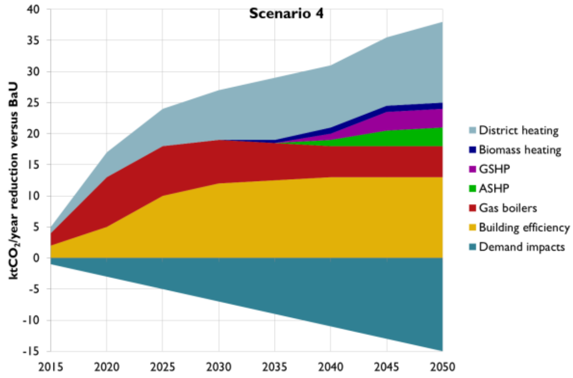


This analysis could be focused on different zones, or aggregations of zones, within an LA area to identify how the pathway may vary geographically. The development of these results over time may also be examined, as in Figure 4-6. These show that deployment of building efficiency measures is made early in all scenarios, reinforcing this option as one of interest. More efficient gas boilers appear to be a useful early option, although longer term their role is reduced. In most scenarios, district heating tends to be deployed at a steady rate over the first decade of the modelled period, whereas GSHPs tend not deployed at scale until the late 2030s.

Figure 4-6 Example output: deployment by measure over time



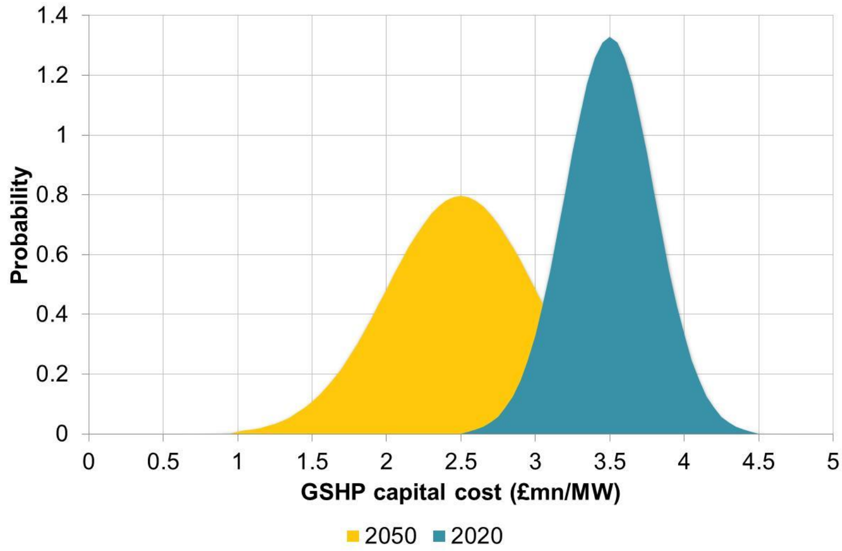




Probabilistic tool operation

Having established that district heating networks and building efficiency appear to be strong options, and that GSHP is favoured under certain circumstances, a more detailed investigation of the uncertainty surrounding these options is appropriate. This is done by introducing a distribution to the inputs and running the tool in probabilistic mode, where the distribution is randomly sampled from over multiple simulations. This will allow the user to build a fuller picture of what factors the options are sensitive to, and how resilient they are to different future outcomes.

Figure 4-7 Example output: CHP capital cost distribution, 2020 and 2050



Option appraisal

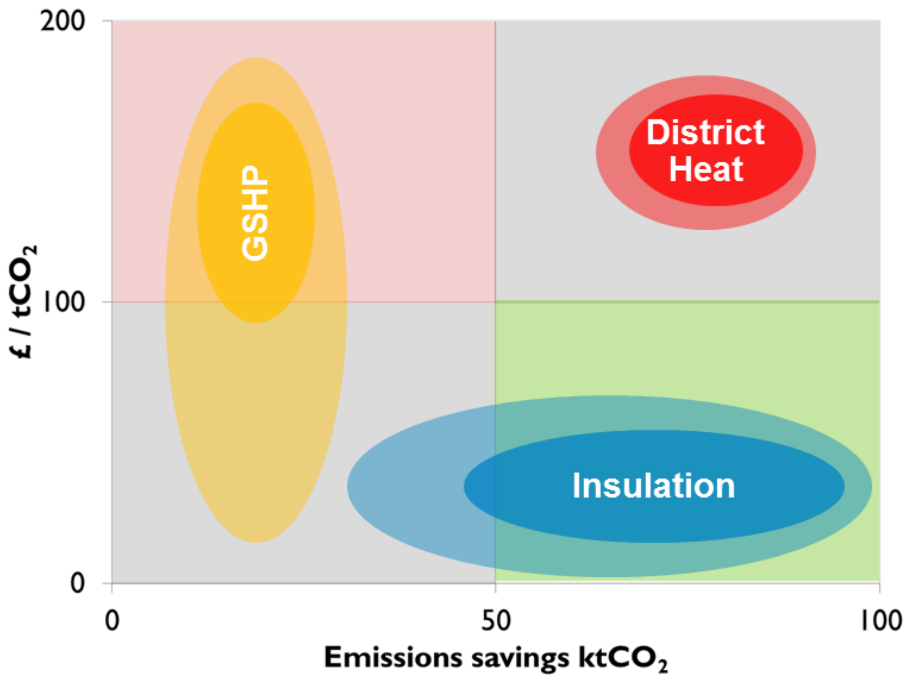
Probabilistic operation brings a richer set of data to the user, allowing him or her to undertake a better appraisal of the options by understanding the impact of uncertainty on the cost-optimal deployment of these solutions, whilst also meeting the requisite constraints and design standards.

The process in section 4.1.1 envisages categorising options as no regret, low regret, or medium/high regret, initially as part of a manual process of using the tool, though it may be possible to automate elements of this in the future. Options which are chosen by the tool in all simulations will be cost-effective under all reasonable circumstances and provide a no regret route to decarbonisation. That is to say that once implemented, no 'reasonable' future eventuality is likely where an alternative option would have delivered equivalent emissions reduction for lower cost. Options frequently chosen by the tool will be cost-effective under a wide range of future eventualities, and would rarely be regretted.

Probabilistic operation allows options to be identified in this way, starting the process of identifying a package of projects representing a course of action which is robust to changing circumstances. An output which could assist this process is a scatter plot of the absolute reduction in emissions achieved by each measure versus BaU, against the per-unit cost of that reduction, over all simulations. Examining where these points accumulate allows categorisation of measures in to those supporting large scale emissions reduction at low costs (likely to be low regret), those supporting limited emissions reduction at high cost (likely to be high regret), and those in between these poles.

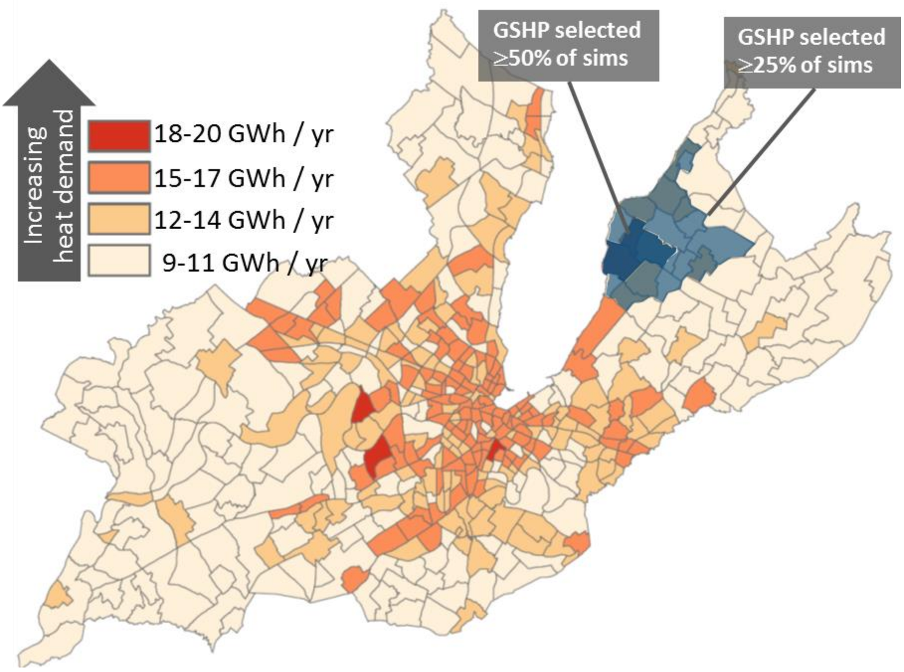
In the example, higher colour density indicates a greater concentration of simulation results in this space. Insulation appears to be no or low regret. It is consistently selected by the tool to provide a large contribution to emissions reduction, and does so at a comparatively low cost. Biomass fired district heating is also consistently chosen by the tool to provide a large contribution to emissions reduction, however it does so at a higher cost relative to insulation. It may be low regret, and as such will require more detailed investigation. GSHPs are selected by the tool to provide a relatively low contribution to emissions reductions in most simulations, but the cost of doing so is subject to significant uncertainty. It may be a higher regret option, but could merit further investigation into costs.

Figure 4-8 Example output: cost effectiveness versus scale for various measures, 2050



Further geographical analysis reveals that in simulations where GSHP deployment is selected, it is limited to a specific area within the LA, which is likely to have distinctive properties, for example, sufficient building outside space, low electricity network reinforcement costs, or the absence of potential anchor loads for heat networks. Further investigation of this option would be concentrated in this area.

Figure 4-9 Example output[9]: geographical distribution of GSHP deployment



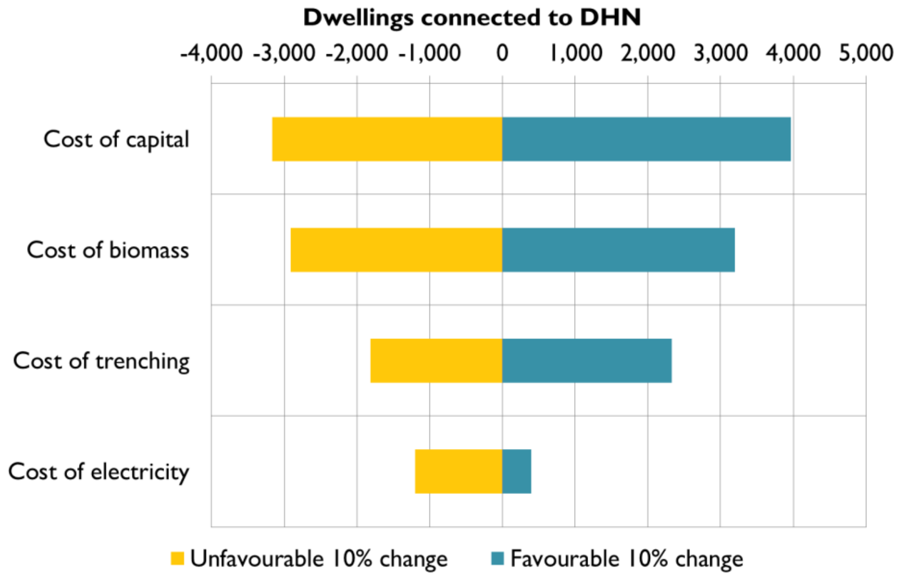
Resilience testing

Where an option is identified as low regret, the user may wish to test and identify more explicitly the circumstances under which it may cease to be the optimal choice, in order to inform decision makers on whether the risk of such a circumstance occurring is one they are prepared to accept. This could be accomplished in part through sensitivity testing, where one variable is flexed whilst others are held constant, in order to determine the prime drivers on the selection or otherwise of particular measures. It may also involve closer examination of results distributions to examine where correlations between factors or patterns exist.

Here, this approach identifies that capital costs is the chief driver of the level of deployment of district heat networks. It could be extended to identify the cost of capital at which a particular heat network ceases to be cost effective. LA decision makers may decide that they are confident that the cost of capital can be held below this level through access to government backed finance or guarantee schemes, and hence they are comfortable with accepting it.

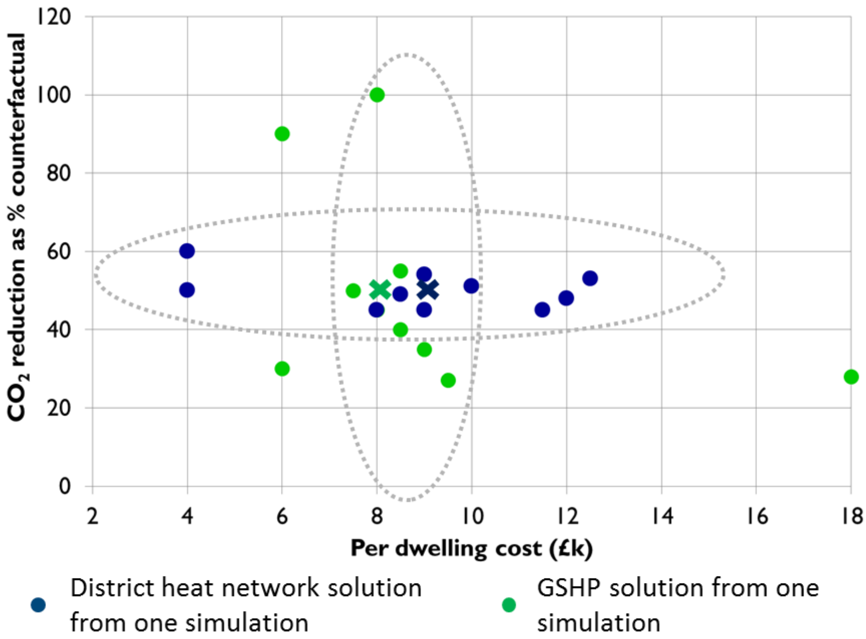
[9] Map adapted from Girardian et al (2010) A GIS based system for the evaluation of integrated energy conversion systems in urban areas, Energy 35 (2010) 830-840

Figure 4-10 Example output: change in the total number of households receiving a DHN connection



More detailed examination of the distribution of results may also reveal further insights. The scatter chart below represents a subsection of the data from Figure 4-8, and considers the area of the LA where there appears to be close competition for provision of heat between a district heat network based solution and a GSHP based solution. It can be seen that the GSHP solution offers a slightly lower cost per dwelling on average, but provides less certainty on emissions reductions and exposes householders to more risk of very high price outcomes. Further investigation may reveal that the most significant drivers are electricity prices and installation costs, and consumer behavior. The former two factors are beyond the control of the LA. The latter is also beyond the control of the LA and is subject to ongoing uncertainty that will continue through the life of the asset. It may therefore be decided that GSHPs represent a higher regret option.

Figure 4-11 Example output: per dwelling cost versus CO_2 savings

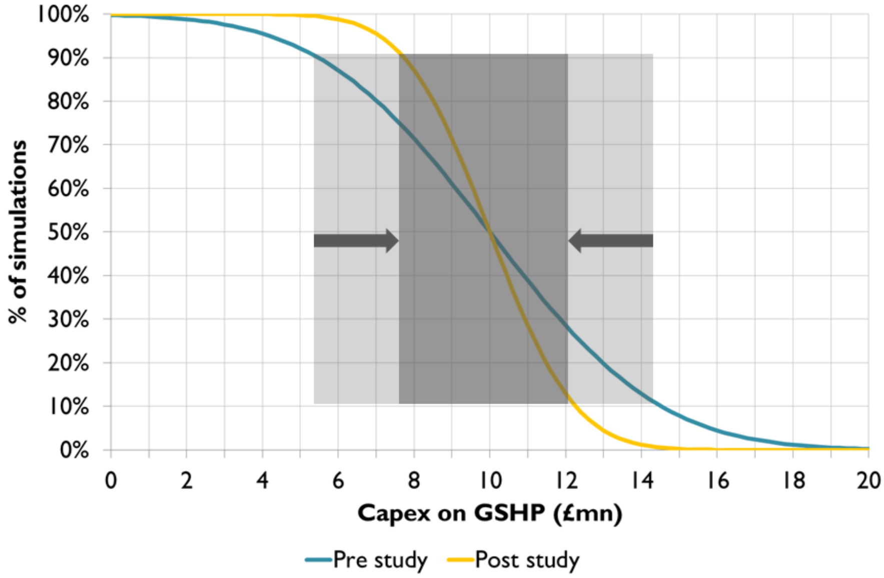


Assess the value of additional information

It may be that the data available to a user is limited initially in some aspects, particularly where it may be dependent on local conditions or where there has been little imperative to gather it prior to consideration of a transition plan. The tool could be used to identify where more data would be valuable in identifying low regret options or rejecting higher regret options, by illustrating how results change when the uncertainty on a particular assumption is reduced or it is assumed to have a slightly higher or lower value.

In the example, it has been determined that whilst GSHPs may provide a lower cost solution in a particular area, but it is subject to risks that the LA may be uncomfortable holding. Before it is rejected the user may wish to confirm that this is a reliable result. Sensitivity analysis showed electricity costs, behavioural factors and installation costs to be the main drivers. Uncertainty around the former two cannot be reduced, but could be for the latter through local ground surveys, for example. If it is assumed that this would reduce the uncertainty on installation costs by 30%, the tool could quantify the extent to which uncertainty on capital costs in this zone is reduced, focusing for example on LA owned properties. This is illustrated in the figure below.

Figure 4-12 Example output: capex on GSHP in LA owned properties



The result might also be calculated when reduced uncertainty and a slightly higher or lower mean value is assumed. It might be found that a slight higher mean installation cost has very little effect on the deployment of GSHPs, but that a slightly lower mean has a positive impact, as illustrated in the figure below. Under these circumstances it might be decided to commission a study in the area identified in Figure 4-9 to acquire a better data set on which to make the decision.

Figure 4-13 Illustrative impact of study on input parameter uncertainty

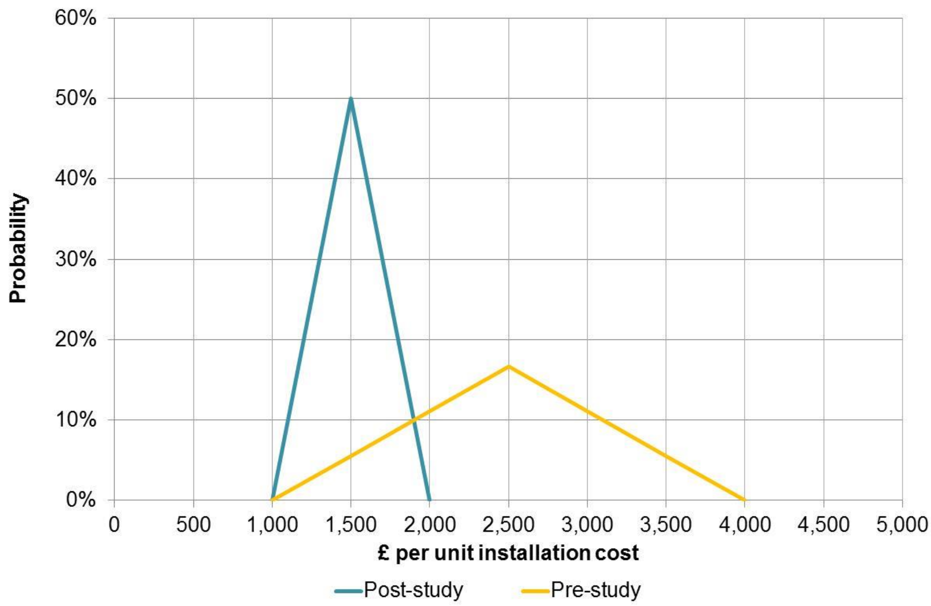
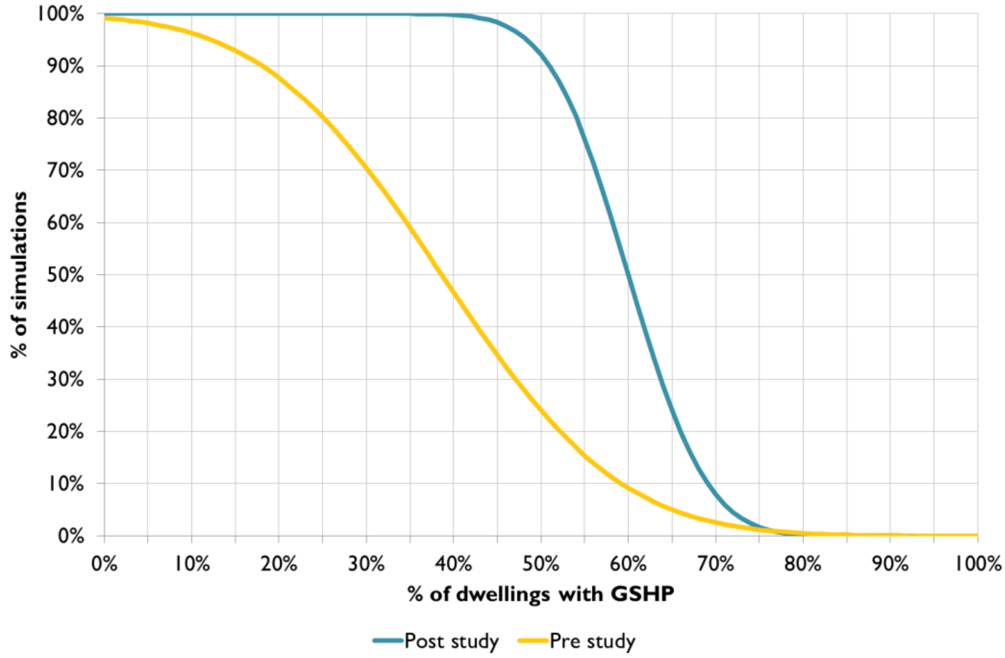


Figure 4-14 Example output: deployment of GSHP pre- and post-study



Initial acceptability check

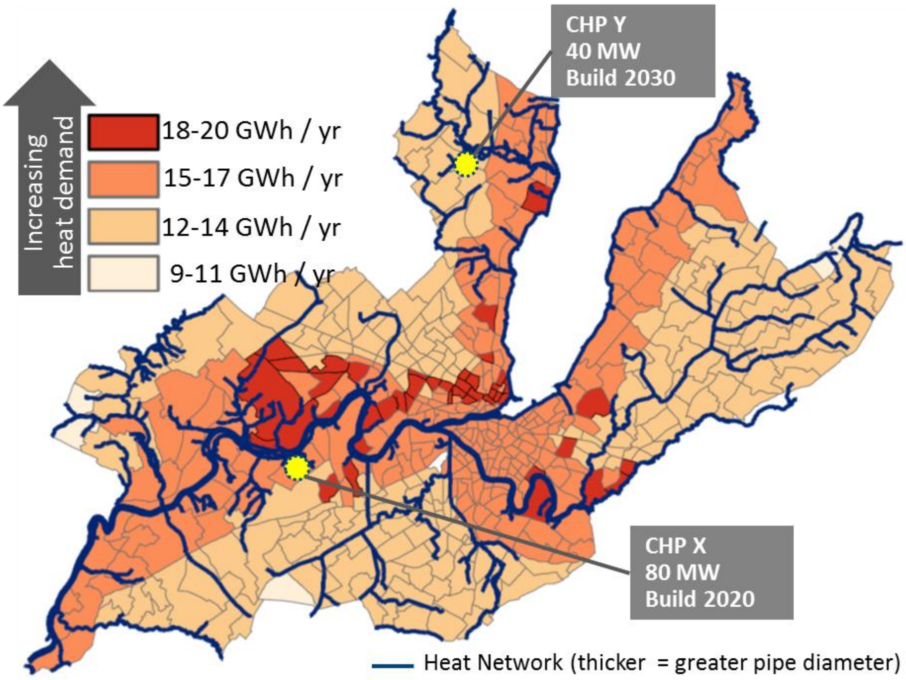
At this point, options that have been identified as warranting further investigation would be subject to an initial sense check to confirm that they are viable, that they obey known constraints, and are aligned to high level policy objectives.

Focus granularity and option selection

As low regret options are identified at a high level and proven to be resilient, an iterative process of "zooming in", in spatial and temporal terms, allows the more exact determination of boundaries and timings of the projects over the pathway to 2050 that would realise these options. These options would then be fixed in a further iteration of the model, which may in turn narrow the breadth of outcomes regarding other options, until the full extent of low regret options that can be identified with a given state of knowledge has been determined.

In the example, as the specifications of an insulation programme are progressively fixed, the details of the district heating network is also resolved in more detail. Whilst the extent of insulation was more uncertain, a CHP plant at location Y was selected by the tool for deployment in 2025. As this uncertainty is reduced and the insulation programme fixed, the optimal deployment date for this CHP recedes to 2030.

Figure 4-15 Example output[9]: heat demand density and CHP plant details



Option prioritisation

Once the projects are selected and defined, projects are prioritised and programmed to form part of the broader transition plan. The projects that are lowest regret, or those that are prerequisite enablers to other projects (such as insulation or grid upgrades), are likely to be selected for implementation first. Later deployment will be desirable for options that tend to lock in a pathway, or exclude options subsequently.

LAs may also wish to implement first those projects over which they have direct control or at least strong influence, such as those within LA owned buildings, to secure early progress. Other projects requiring implementation by private and independent decision makers are likely to involve a more complex strategy of long term engagement, influence and incentivisation, and be subject to significant uncertainty regarding timing.

Projects where the CO_2 emissions reduction effectiveness is subject to a significant degree of ongoing uncertainty, driven for example by consumer behavior, may need to be managed carefully over the duration of the transition plan to avoid over reliance on their effectiveness to deliver overall targets.

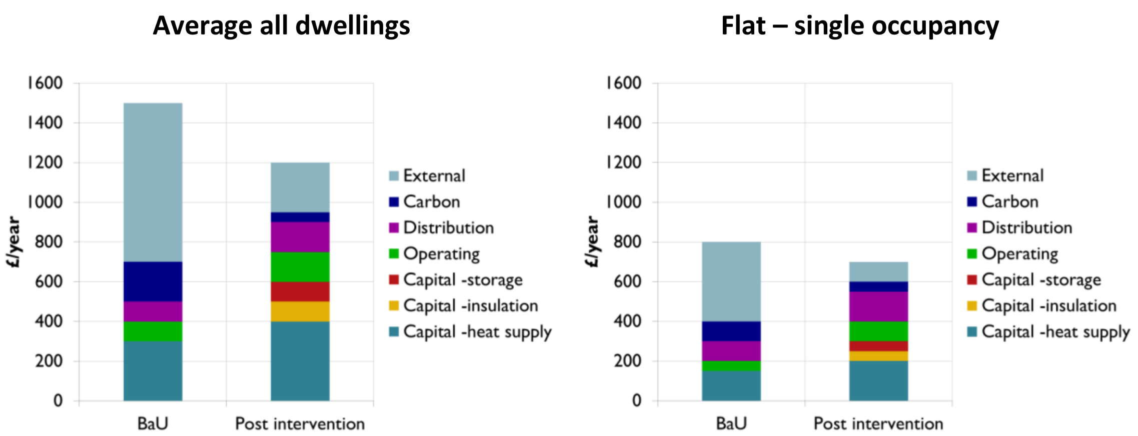
Wider acceptability check

At this stage a programme of projects, and high level options for subsequent pathways, has been identified. The EnergyPath™ Design Tool is focused on the building, network and other technology choices necessary for the development of a least-cost local energy system transition plan, subject to the impact of carbon pricing or carbon constraints. Social factors may be included to the extent that they can be parameterised as costs or constraints, but this may not always be feasible. Therefore, at this stage it is appropriate to check performance of the pathways against wider targets. In particular, this is likely to include the cost of energy provision.

The majority of the tool's outputs can be presented in a GIS format, and potentially combined with other existing GIS datasets. For example the cost of household energy could be combined with a socioeconomic GIS dataset of income levels to understand the potential implications for fuel poverty. The tool could also enable estimates of other factors, such as noise, job creation or air pollution, by relating back to proxy drivers in the pathways. High level summary of the most important parameters may also be useful to assist communication.

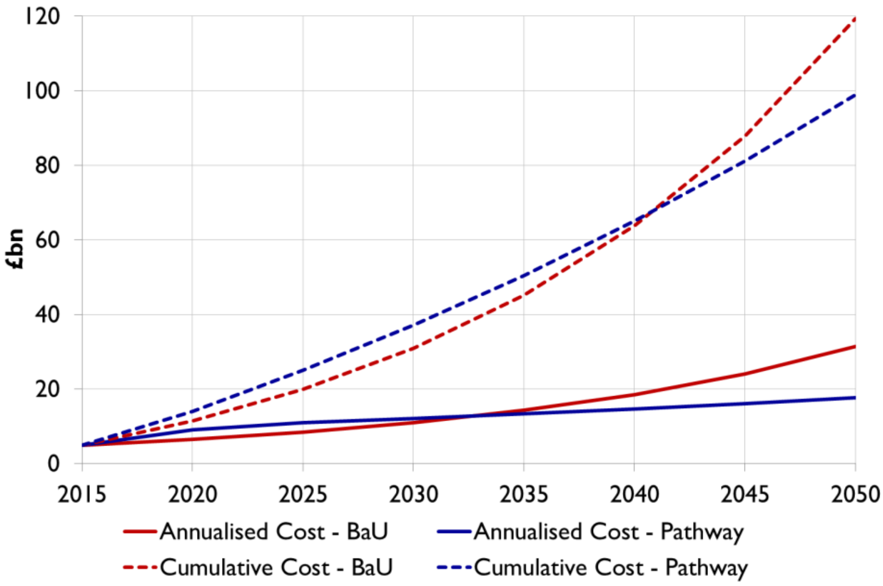
In the example, a particular zone within the LA with a high level of social deprivation has been selected. The annual cost of energy provision for all dwellings and single occupancy flats has been plotted in 2025, under BaU circumstances and after implementation of the initial programme of projects. Costs are reduced in both cases.

Figure 4-16 Example output: annual cost of energy provision, by dwelling type



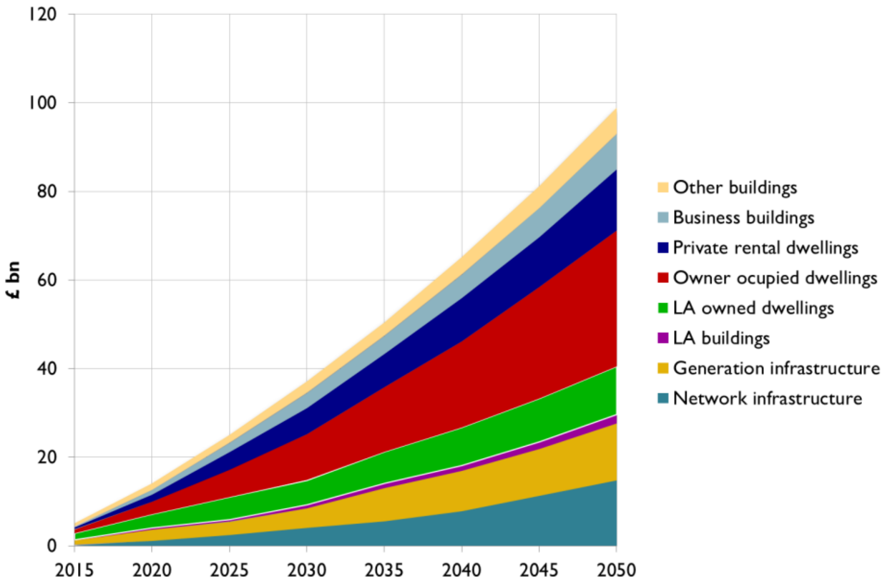
Summaries of headline figures have also been prepared for internal communication purposes, such as that below, illustrating cumulative and annualised cost under BaU and one of the high potential pathways.

Figure 4-17 Example output: total cost of energy provision, annualised and cumulative



Analysis has also been undertaken of the locations of the investment to be undertaken in a high potential pathway, as an indicator of the proportion of the pathway decisions lying in the jurisdiction of different types of actor. This will be an important consideration in developing the final transition plan. Figure 4-18 illustrates that a relatively small part of the pathway would be directly enacted by the LA, limited to that occurring within LA owned housing stock and municipal buildings and facilities. A somewhat larger portion would be concentrated within the remit of a small number of network operators and large scale investors, with whom the LA may have an established relationship and a degree of influence. A large portion, approximately 60%, would however be disaggregated across the domain of numerous private individuals, businesses and other organisations free to make independent decisions. The effectiveness in reducing CO_2 emissions of investments in these locations once made is also likely to be dependent to a large degree on uncertain consumer behavior. The LA may well identify this as a point of risk in a transition pathway based on such a pathway, but consider it unavoidable given the nature of the local area. It may therefore embark on a process of engagement locally, and potentially with national government to ensure it is equipped with the necessary powers to incentivise decision making in the required manner.

Figure 4-18 Example output: investment by actor type

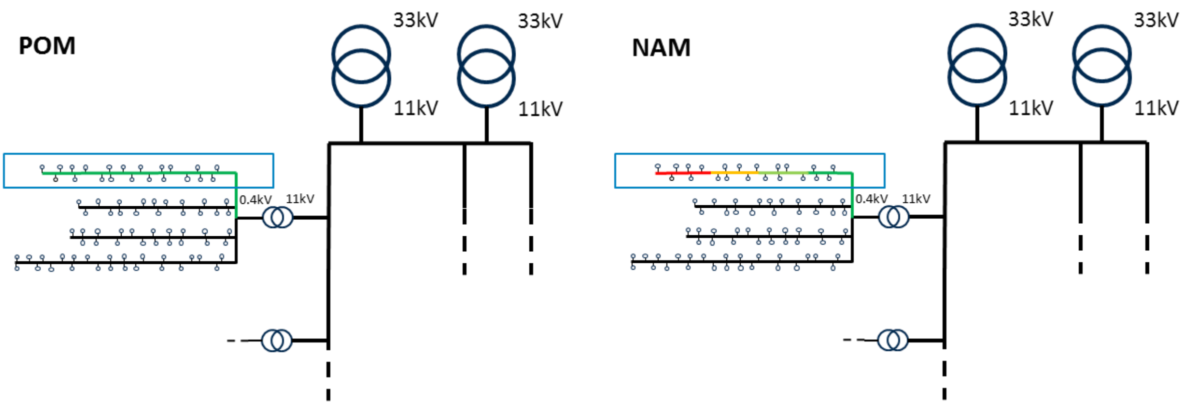


Detailed load flow

The final confirmatory step for validation of a pathway is assurance that the planned capacity of energy networks is sufficient to accommodate the anticipated loads. This will be done in the NAM, which includes a load flow modelling capability. It is also envisaged that the network options considered given the wider evolution of load in the area will be refined using the more disaggregated analysis possible in the NAM.

In this example, when a zone of the LA area where stress is greatest is subjected to load flow modelling in the NAM, it emerges that voltage drops below the design standard threshold in winter peak conditions by 2025. The configuration in the transition plan will be updated so that when reinforcement occurs in 2020, a larger diameter line in this part of the network is installed, in order to reduce impedance. The increased cost is found not to be sufficient to alter the wider choices the tool makes in this area.

Figure 4-19 Example output: network schematics



Project details

At the end of the consultation phase, when projects are approved for implementation, the tool will be used to generate technical specifications for projects and data that will enable financial modelling. This will include figures for such factors as the quantum and timing of capex and fixed opex, generation or energy savings (as appropriate) and their value based on a shadow price and output calculated for each time slice and period.

The example below sets out summary financial details for the initial CHP plant linked to the district heat network (plant X in Figure 4-15) and identified for commissioning in 2020.

Table 4-1 Example output: CHP plant X summary financials

	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
Capacity operational (MW_e)				80	80	80	80	80	80	80	80	80
Capex (£ mn)	5.0	35.0	25.0	15.0								15.0
Generation electrical (GW_e)				341	340	339	362	370	359	358	383	391
Generation thermal (GW_{th})				316	322	312	318	325	301	322	345	351
Electricity price (£/MWh)				41.00	42.00	43.00	41.00	41.00	43.00	44.00	42.00	42.00
Heat price (£/MWh)				19.00	19.00	20.00	20.00	20.00	22.00	21.00	20.00	20.00
Electricity revenues (£ mn)				14.0	14.3	14.6	14.9	15.2	15.5	15.8	16.1	16.4
Heat revenues (£ mn)				6.0	6.1	6.2	6.4	6.5	6.6	6.8	6.9	7.0
Opex (£mn)				12.0	12.2	12.5	12.7	13.0	13.2	13.5	13.8	14.1

Figure 4-20 provides illustrative financials for this plant over its assumed 25 year life, including periodic refurbishment capex, and the cumulative discounted free cash flow to the firm. At a 10% discount rate (pre-tax real), indicative of private sector return requirements for such a project, it is apparent that the NPV for the project is below zero. This may indicate that the project may need to be undertaken within the public sector, or more likely indicate the level of subsidy required for the project to be financed using private capital.

Figure 4-20 Example output: CHP plant financial summary

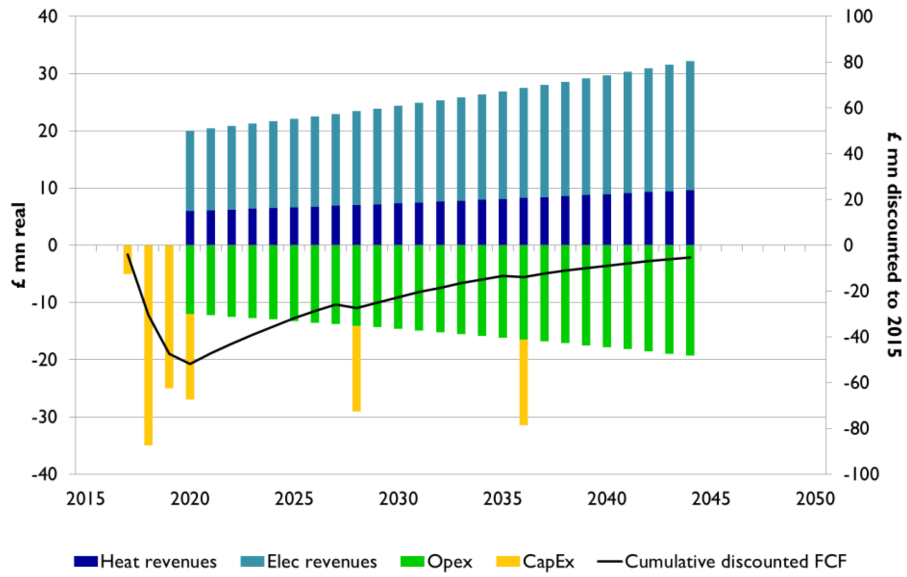
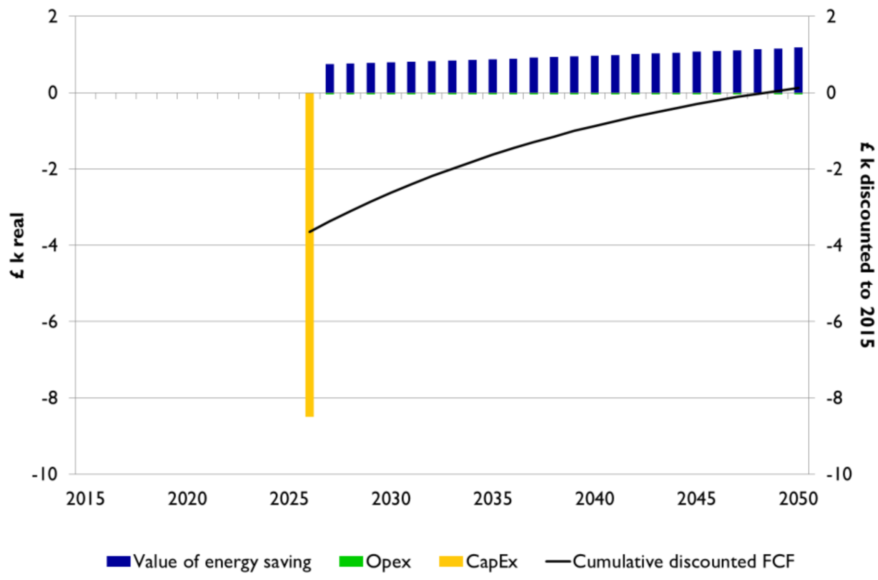


Figure 4-21 illustrates summary financial data over an assumed 25 year life time for Solid Wall Insulation, which is selected for installation later in the 2020s in a relatively remote area of the LA dominated by detached and privately owned properties. An 8% discount rate has been used for consumers in this case. At this rate, the measure provides a small positive NPV for the householder. However the extent to which this calculation holds true is likely to be highly dependent on the behavior, lifestyle and circumstances of that household. Such factors can only be approximated by the tool based on larger populations, and are subject to great uncertainty and variation over time.

Furthermore personal discount rates are also highly uncertain, and so the extent to which individual households make an investment decision in the case of a given financial projection is also subject to uncertainty. Therefore LAs are likely to approach such “projects” with caution, seeking carefully to understand the interaction of consumer behavior and incentive schemes (such as grants or Feed in Tariffs for heat under the Renewable Heat Incentive, RHI) with take up of particular measures. To some extent this can be captured in the configuration of the EnergyPath™ Design Tool, for example by representing availability of the RHI for a technology as a reduced cost. Detailed representation of consumer investment decision making is not however component to the tool’s cost optimisation approach, and some “post processing” of results may be necessary for an LA to determine a level of private take up of a given measure on which it is content to rely for its transition plan.

Figure 4-21 Example output: SWI summary financials



Accommodating changes arising from consultation or validation exercises

In the process of consultation, it is probable that new constraints or objectives may be identified which will affect the optimisation calculation, and lead to changes in the pathways determined. Some examples, and how these might be accommodated and impact on the tool's outputs are briefly discussed here.

- Work by the LA's environmental protection unit has identified a protected species on a piece of land selected for development of a CHP plant in a draft transition plan, and has begun a process to protect it from development.
  - In the configuration of the tool, the zone containing the proposed plant is now identified as being unavailable, constraining it from being selected for development. When the tool is re-optimised with this constraint, it is found that district heating remains the favoured solution, but now two smaller CHP plant are selected for development in alternative locations. There is a relatively small increase in costs associated with reduced efficiency of the overall pathway.
- Internal consultation within the LA reveals that the updated economic development master plan proposes a new office park on the outskirts of a settlement.
  - Accommodating demand from the new office park will require reinforcement of the electricity network at the 33 kV level from 2018, when the development is assumed to be completed. This is included in the known network developments, and the tool is re-optimised on this basis. The upgrade provides headroom for a group of dwellings which had previously remained on gas based heating until 2045 to implement electric heating solutions from 2025, enabling decommissioning of the gas network in this area to be brought forward.

- Consultation with the local water utility reveals that they are planning investment in an anaerobic digestion facility at a sewage treatment plant.
  - When this new heat source is included in the known development of local infrastructure, district heating becomes the optimal solution for heating provision in nearby housing, and investment in electricity network reinforcement is postponed.
- The LA has decided that it can no longer afford to subsidise operation of a swimming pool, which will be forced to close.
  - When this anchor load is removed from the tool, a district heat network in this area is no longer the lowest cost solution. The tool tends to select a more complex range of measures in the area, including selected investment in more marginal forms of insulation (such as SWI), earlier electricity grid reinforcement and transition to electric heating, and some use of biomass boilers. The tool user undertakes more detailed analysis of these options to confirm their resilience and to resolve project details to a higher level.

4.3. Practical illustration of key modes of use

This section is designed to illustrate how the user of the tool would practically address particular issues that may arise, or pertinent questions that may be posed by decision makers, as part of the process of using the tool to assist in design a local energy transition plan. Such issues include:

- How engagement with the community and technical engineering expertise may impact and be integrated within the business and transition plan design process
- How the tool can inform the degree of headroom included in designs, so as to build infrastructure which is resilient to different outcomes
- How the existing plans of LA or network owners can be accounted for
- How the tool can assist consideration of energy system choices over which an LA can exercise a degree of influence or control versus those it has more limited control over (e.g. consumer uptake behaviour in owner occupied or rented properties)
- How the tool can assist consideration of investment decisions made in centralised monopoly service providers (such as DNOs) versus those by individual private households exposed to free market competition; and
- How the tool can assist consideration of the impact of consumer behaviour on developing a robust transition plan.

As a hypothetical scenario, it is imagined that during the phase of assimilating existing investment and master plans in the example in section 4.2, it was determined that the LA has already developed plans to build a limited district heat network in a certain area, based around anchor loads provided by LA owned facilities.

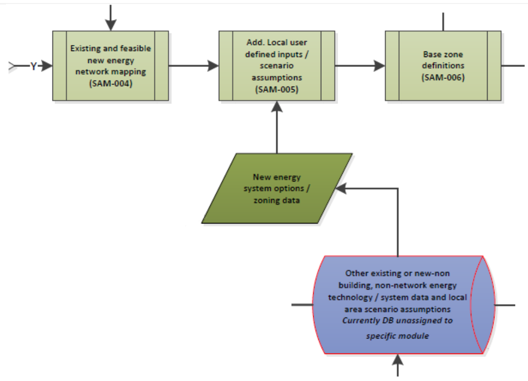
This option would be fixed upfront by entering data in 3 areas of the tool[10]:

- Its locational details would be entered via the SAM – e.g. the zoning of the heat sources, pipe network and connection points to heat loads, etc
- The integrated database interface to define the associated characteristics of the heat network – e.g. the size, operating lifetime, costs (investment and operating), etc
- User-defined constraints with the POM to force the deployment of this network at the right time within the pathway analysis (i.e. so it is user-controlled rather than an open economic decision). This would cover both the network itself and the upgrade of the LA owned buildings providing the anchor load

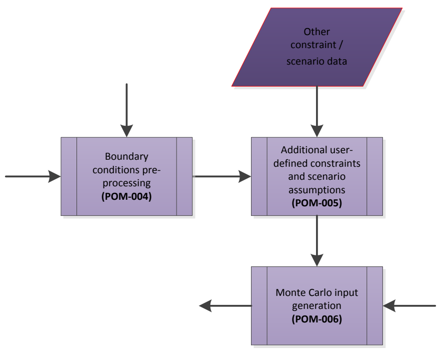
Figure 4-22 Business process and tool flow excerpts: including existing development plans

Business process excerpt

SAM excerpt



POM excerpt

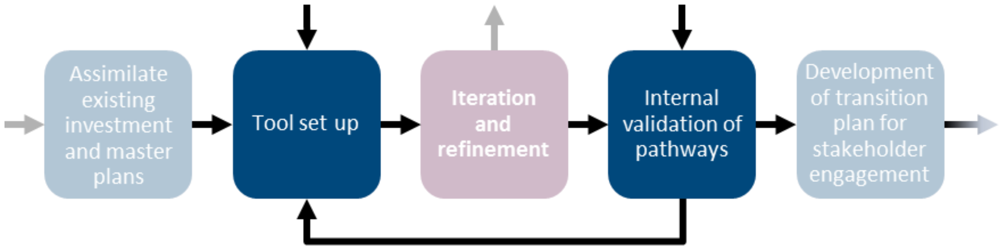


When the model is developed and the POM run, the results indicate that the cost optimal solution is for this network to be expanded to include an adjacent, mainly privately owned residential, district. When this plan is presented for internal approval, the steering committee express concern over what level of take up can be expected in this district, given householders are free to choose whatever heating solution they prefer. They ask the team responsible for use

[10] As part of longer-term development a streamlined interface would be created to ensure that all data inputs and configuration settings across the entire tool can be managed as efficiently as possible.

of the EnergyPath^{TM} Design Tool to further investigate the sensitivity of this option to different uptake outcomes.

Figure 4-23 Business process excerpt: iteration after internal validation

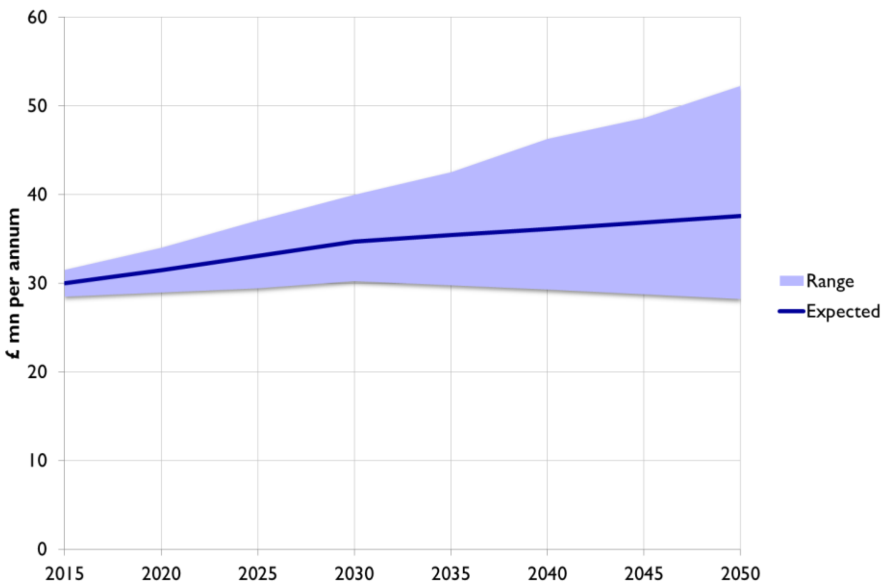


The EnergyPath^{TM} Design Tool team investigates by forcing in the POM different levels of uptake of heat network connections in the district. They identify that below 40% uptake, all other things being equal, the POM stops selecting this option in favour of further electricity grid reinforcement.

The team also investigates under what conditions this uptake rate may transpire, assuming that if DHN connection presents a positive investment case at a private consumer level discount rate, the barrier to uptake is an uncosted “hassle factor”. In the POM, this is explored by assigning different levels of cost to consumer hassle, converted from an equivalent cost per hour the consumer spends in handling the administration of a heat network connection or enduring degraded domestic conditions during installation. It is found that at a cost of hassle of £140/hr, uptake in the district falls below the critical 40% level.

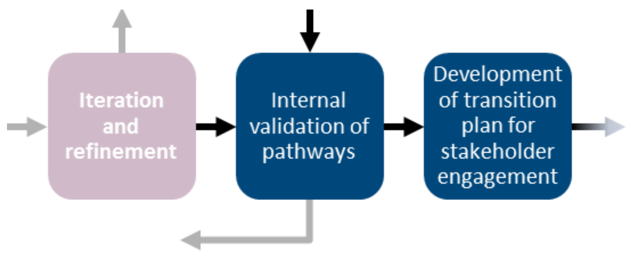
The team identifies this factor as a possible subject for future investigation. By examining the total system annual costs the team is also able to identify the value of increasing uptake of heat network connections. At the minimum uptake of 40%, total costs are at parity with the competing electric solution, at 60% take up annual costs are reduced by £15 mn and at 80% take up annual costs are reduced by £40 mn.

Figure 4-24 Example output: annual system costs with one standard deviation range around expected cost of consumer hassle



The team presents their findings for internal approval. The steering committee are now satisfied that there is sufficient understanding of how consumer take up and other factors such as energy prices impact this particular project. Whilst energy prices are beyond the LA's control, it is considered that DHN connection uptake is to some extent controllable. The LA has powers to vary council taxes within a certain range, and can support financing of home energy measures as a Green Deal provider. As part of the development of the transition plan, a policy unit within the LA is instructed to consider how these levers can be combined to incentivize, in an equitable manner, uptake of DHN connections and insulation in the area.

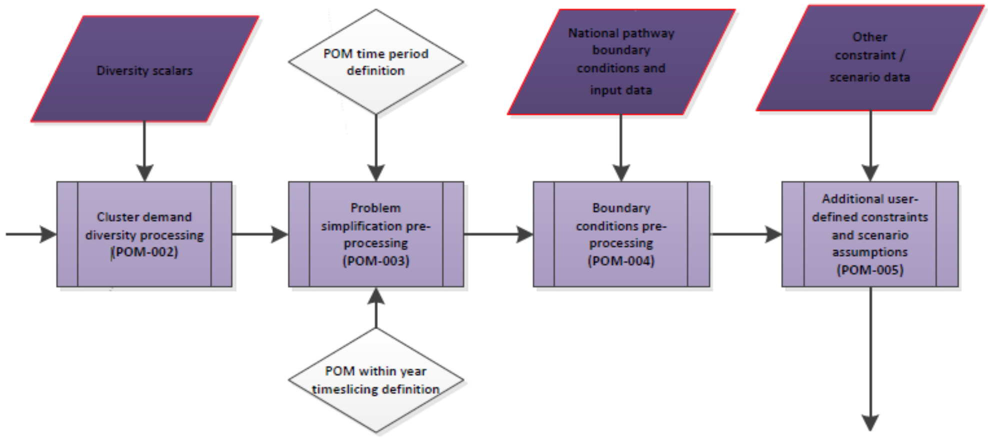
Figure 4-25 Business process excerpt: transition plan development



The DHN project is approved for inclusion in the plan for further consultation with key stakeholders including the regional electricity DNO, who use the analysis to consider their plans for grid reinforcement in the area. By studying the detailed outputs of the NAM when DHN uptake is fixed at different levels in the POM, they gain insight into what level of resilience against outturn DHN uptakes would be provided by different levels of electricity network reinforcement – i.e. if fewer customers connect than expected and peak electricity demand is higher than anticipated.

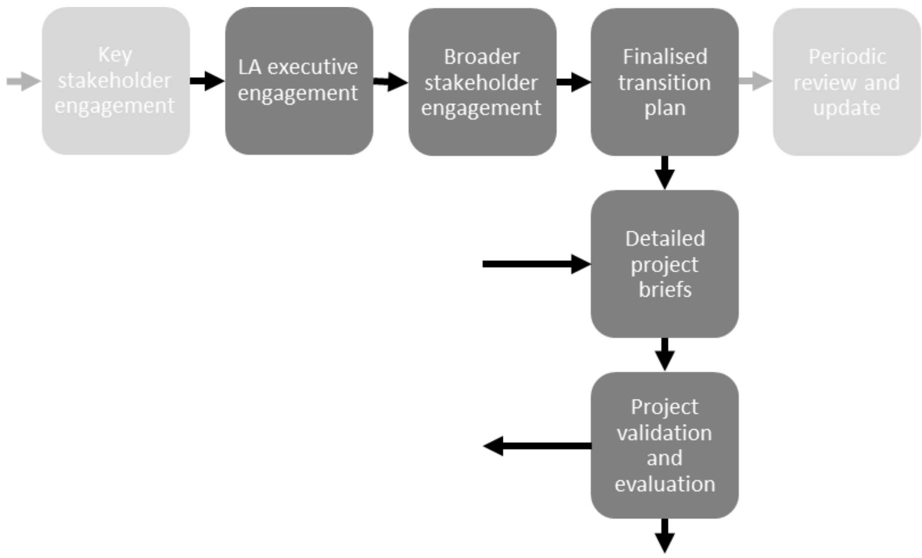
To further inform this analysis, the DNO also considers the impact on reinforcement requirements in the area of the network when different diversity scalars are assumed when calculating peak demand. In dialogue with the LA, the DNO determines a sensible level of headroom to include in its future reinforcement and asset replacement plan in the area, which is *now consistent* with the boundaries and residual uncertainty associated with the DHN plan.

Figure 4-26 Model flow excerpt: setting diversity scalars and other constraints in the POM



The projects are now reviewed and approved by the LA executive and put to the public at large for consultation, and eventually included in the final transition plan. The DHN project (comprising the district heat network and to service it two energy centres, one based on biomass fired CHP and one on GSHP), now begins the process of implementation. The first step is to acquire expert review and validation of the project design and costs, before committing more substantial funds to a full FEED.

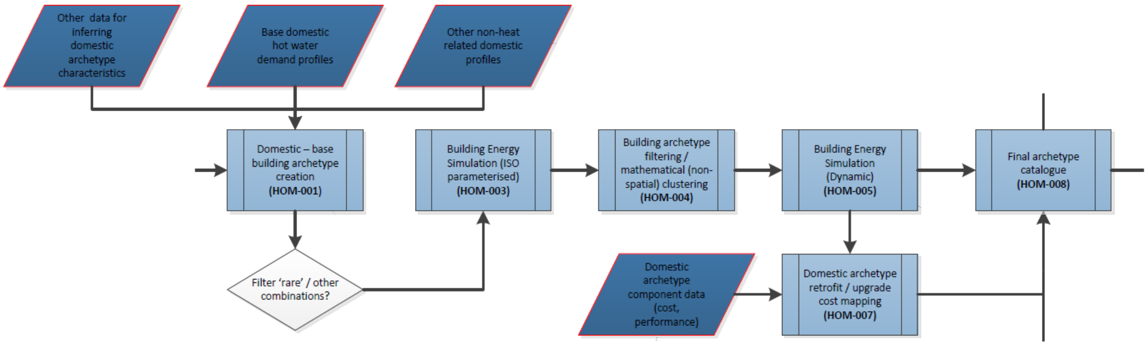
Figure 4-27 Business process excerpt: consultation to project validation



The LA commissions an engineering and planning consultancy to review the expected demand for the system, the technical specifications, and the cost estimates. The review undertaken determines that a GSHP based energy centre is not feasible at the site selected as the density of settlement thereabouts is too great to enable a sufficient length of ground collector to be laid. Additionally, in surveying the housing stock in the area to assess demand, a building type has been identified using materials and construction techniques peculiar to the area. The properties of this stock are such that energy efficiency can be improved at a lower cost than other housing of a similar age and size. As a result, the EnergyPath^{TM} Design Tool team makes a number of amendments to the tool set up.

- In the HOM, details of the new domestic building archetype are entered, in the existing state and in various states of upgrade.
- In the SAM, locations of this building type are identified where known. The area initially identified for GSHP deployment is now flagged as being unsuitable for this technology, and hence this information will automatically flow through and the option will be excluded in subsequent analysis in the POM.

Figure 4-28 Model flow excerpt: establishing new building archetypes in the HOM



The POM is rerun with these new conditions. The new solution identifies a greater degree of insulation in the newly identified building archetype as being cost optimal. This in turn enables the heat demand of the network to be satisfied by a single enlarged biomass fired CHP plant. Plans for the larger CHP plant are validated by the LA's advisors, and expected costs found to be within the range projected by the tool. The plan then moves into the development phase with commissioning of a full FEED study.

4.4. Interfaces

The focus of effort for Stage 2 of the programme should be to produce a functional tool, suitable for an expert level user. As such, development of a fully integrated Graphical User Interface (GUI) is not within the scope of this functional specification, however, this will where possible make use of existing 3rd Party interfaces (such as those in GIS packages). And create 'non-polished' interfaces necessary to streamline the operation of the tool and view inputs and outputs.

In the longer term, further development of a GUI would make the use of the tool by expert users more 'efficient'. The development of the tool during Stage 2 therefore should follow best practice and do nothing that would prejudice against the later integration of a more 'polished' GUI.

The UI design also has to be balanced against the need to deliver end-users access to the full functionality of the tool. This could be undertaken via a set of more simplified primary control screens with limited functionality and 'hidden' administration screens providing access to the full set of functionality, or alternatively semi-automated 'wizards' to guide a non-expert user through the more complex process steps and input assumptions.

4.5. Other end user requirements

4.5.1. Scenario and sensitivity management

In applying the tool to a local area, a range of processes will need to be followed, as discussed in section 4, including data input and infrastructure analysis. The user will typically configure a range of scenarios and sensitivities, and be running the model in multiple modes (including deterministic, probabilistic for analysing uncertainty, and constrained runs for resilience testing). This will be managed through a control interface, providing the capability to manage the process flow, set up specific runs and sequences, and manage results sets.

The model will ensure that configuration control files are generated as part of the tool use, to ensure an audit trail and reproducibility. The resulting data, from deterministic use, bespoke scenario testing and probabilistic operation will be archived and managed so as to make it readily accessible to multiple users. The tool will have the ability to automate the execution of a standard package of uncertainty analysis.

Data storage

The EnergyPath™ Design Tool will have the ability to store *all* data generated in simulations, labelling and archiving it in a fashion that allows easy access and management. In practice, the extent to which this is desired may be limited by the availability of data storage and the performance of host hardware. Accordingly, the tool will facilitate straightforward adjustment of settings regarding data storage, allowing operation in a mode suitable for the task being undertaken and the performance of the machine used.

For example, the full energy supply/demand balancing results for every within day time period for every simulations will constitute a significant quantity of data. Whilst this data might be used as part of post-processing to create final summary results, it may be unnecessary to store the full set of disaggregated results as these could always be reproduced at a later date if required.

As discussed in the (D2) Design Architecture deliverable, it is envisaged that the tool is linked to the ETI's wider PLM (Product Lifecycle Management) system for the SSHP, which coordinate the overarching data management and audit requirements.

4.5.2. Audit

The successful application of the EnergyPath™ Design Tool will depend on a very large amount of underlying data. This will come from a wide range of sources, including nationwide data sets held by both private and government sponsored entities, locally focused datasets and plans, asset databases held by network companies, and assessments of technology components. These datasets will over time evolve and be updated, or replaced or supplemented with new sources.

A data catalogue will be produced to identify and map these different data elements, and be structured so as to ensure ease of ongoing maintenance. When the tool is operated, the specific versions of each dataset used will be logged, to ensure reproducibility and provide an ability to track data provenance and produce audit trails.

The specification of the tool with regards these requirements will be informed by work on data management and auditing being undertaken as part of the ETI's wider SSH project, and hence the final EnergyPath^{TM} Design Tool, will need to be compatible with this wider audit framework.

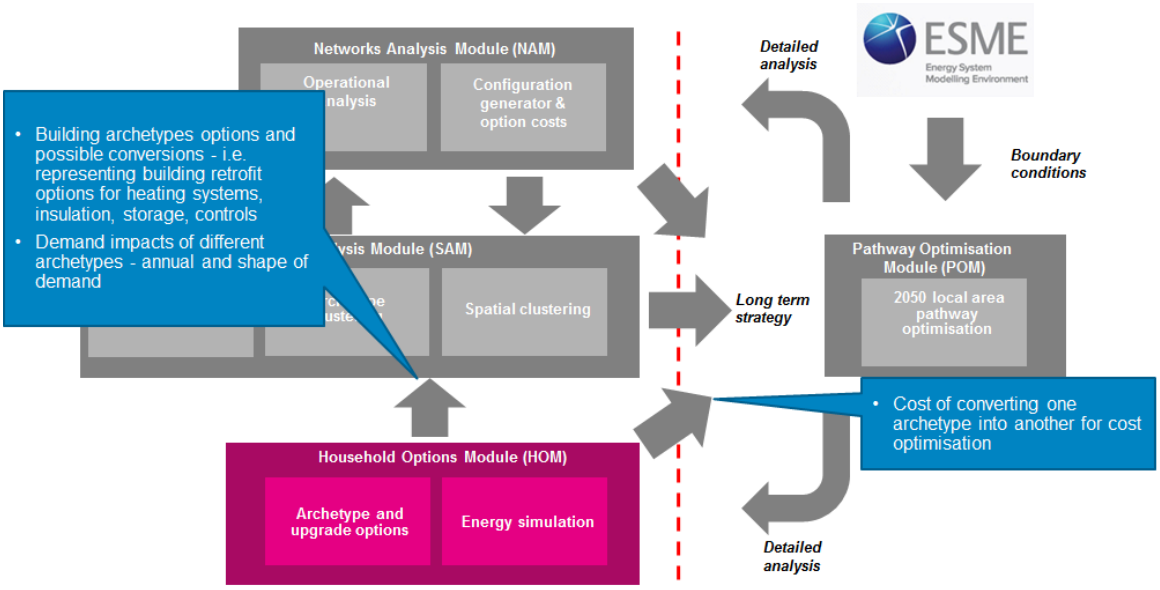
5. Household Options Module (HOM)

5.1. Overview

The Household Options Module (HOM) is one of the four key components of the EnergyPath™ Design Tool. The HOM provides the building archetypes (***both domestic and non-domestic***) required for other software components to characterise the existing building stock within Local Authority areas under assessment, and the options for upgrading this stock in future. 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.

The HOM also captures important parameters regarding the behavioral interactions of building occupants with their energy using equipment and systems, which is passed to the Pathway Optimiser Module (POM) to enable the uncertainty surrounding different technology choices to be estimated.

Figure 5-1 HOM context



The characteristics of buildings are represented in the HOM using a so-called "archetype" approach. This is analogous 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 is a deliberate data architecture design decision to help simplify the final pathway optimization undertaken within the POM (see section 8). Estimates of energy demand from domestic building archetypes are derived from bottom-up engineering simulations.

While both domestic and non-domestic buildings are represented by the HOM, the balance of complexity within the model is tilted towards the domestic sector. At a high level, archetype approaches are taken for both sectors, although different approaches to classifying the archetypes and estimating energy demand are used. This is due to real-world constraints regarding available data and knowledge of how to characterise the UK non-domestic stock and its energy use, which are described in more detail in section 5.5. In short, the diversity in built forms, usage patterns, climatic control and equipment provision between non-domestic buildings of even the nominally same type (such as offices) is so large that creating bottom-up engineering simulations of a "typical" building of any nominal class has only limited real-world applicability. This necessitates a more statistically informed approach, as described in Section 5.3.

A challenge for the HOM is collapsing the vast matrix of 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 LA. This is achieved through the use of mathematical problem decomposition (clustering[11]) techniques to reduce the problem dimensions down to those building performance metrics which most strongly affect the desired outputs. In other words, where the characteristics of a number of archetypes are very similar, it is possible to aggregate them into one archetype with limited loss of information. To support this decomposition process the HOM uses a fast parameterised 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[12].

Once the 'short list' of the most material domestic building archetype has been created, bottom-up dynamic thermal modelling of space heating and hot water is then used to generate possible demand profiles for a number of 'characteristics days[13]' at 30-minute granularity (electricity demand profiles for EVs, lighting and other appliances would be added from exogenous assumptions). This would be undertaken for a range of different heating technology options for the archetypes including gas boilers, heat pumps, hybrids, etc and consider the optionality introduced by features such as storage.

The number and definition of characteristic days that can be considered is flexible and could 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 particular standard, for example, to meet a 1-in-20 or a 1-in-50 cold winter day.

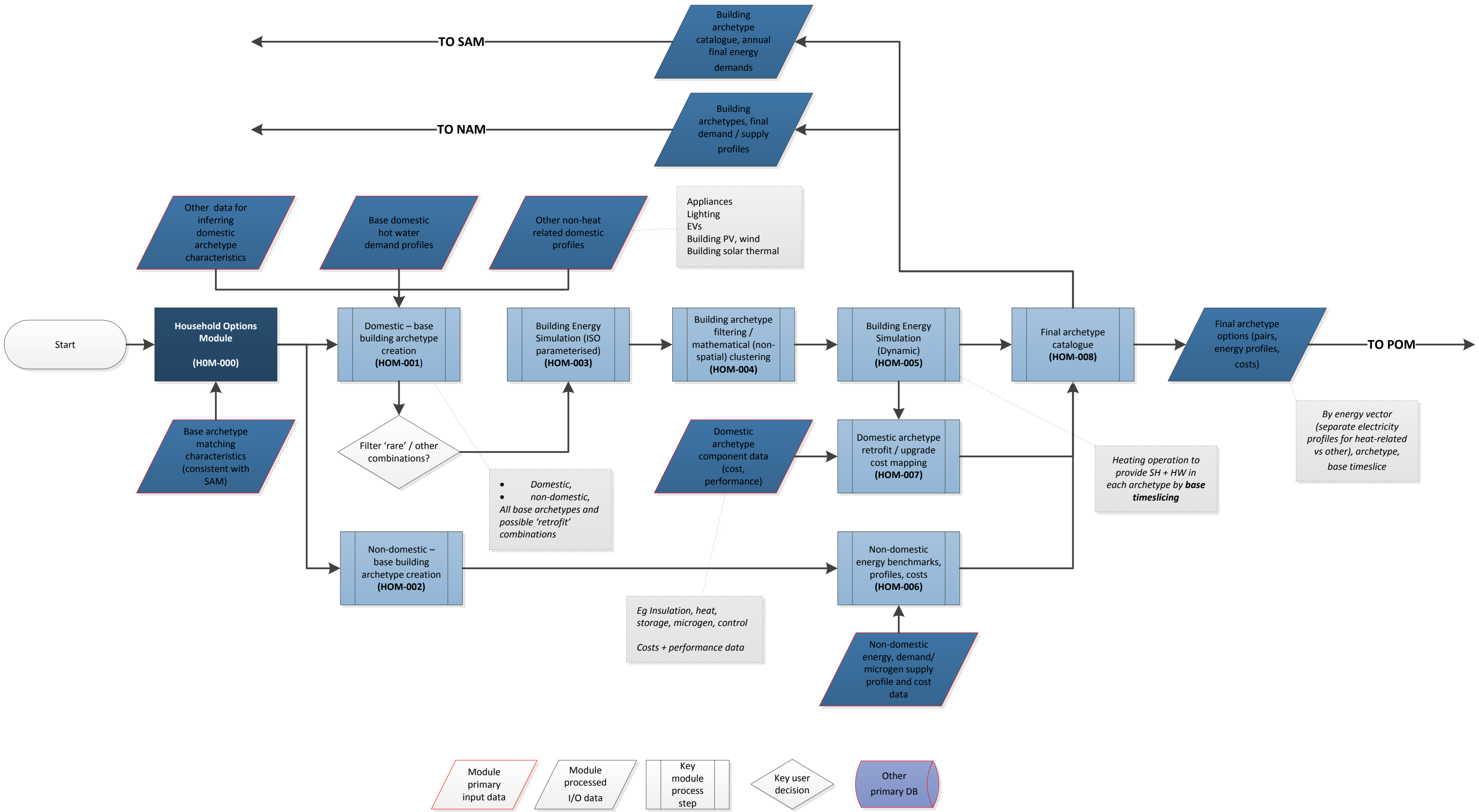
5.1.1. Module diagram

The diagram below shows the key logic / process steps for the HOM, the equivalent for the full tool covering all modules is shown in section 12.

---

[11] 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 EnergyPath™.
[12] For these calculations, EnergyPath™ employs an ISO 13790 model based on the UK Standard Assessment Procedure (SAP) methodology
[13] As part of this assessment it will be necessary to simulate a period of time around the day of interest to account for e.g. the slower variation in thermal mass

Figure 5-2 Key HOM logic / process steps



5.2. Key outputs

Key outputs from the Household Options Module include:

- The final catalogue set of archetypes used to represent a local area and the valid pairs to convert one archetype into another, for example to represent a valid improvement in efficiency or a different heating system.

- Parameterised final energy demand profiles for different building archetypes (which vary be user-defined characteristic days and can vary over timeperiods[14]) for different energy vectors on individual characteristic days, which are provided to the NAM to help establish the peak demand requirements for zonal load flow modelling and against which the POM can perform demand/supply balancing.

- The base timeperiod annual energy consumption of individual building archetypes for different energy vectors, which are provided to the SAM to support the validation of the archetype matching process

- The combined costs of retrofitting building fabric, alternative heating technologies, storage and controls which are passed to the POM for the purposes of numerical cost optimisation.

[14] For example, to reflect rising mean internal temperatures

Table 5-1 Output data summary

ID	Data field	Destination	Purpose	Data granularity	Uncertainty parameters
1	Final archetype catalogues and valid conversion pairs	SAM / POM	SAM – support building archetype matching process POM - cost optimisation	Potentially around <10^3 final archetypes referenced against full set of valid archetype characteristic combinations Matrix of valid combinations to convert one archetype into another	N/a
2	Final energy demand profiles by archetype	POM / NAM	POM - supply / demand balancing NAM – to help understand maximum possible peak demand to bound load flow analysis	• Electricity (heat-related), electricity (EV), electricity (other), gas, network hot water, and other final energy vectors (biomass, H2) • By each archetype (some archetypes have >1 profile to reflect flexibility of operation of some devices/storage), • By characteristic days • By half-hour within day granularity • By timeperiod (e.g. if want to reflect changing mean internal temperatures over time)	Uncertainty in profile shapes for electricity (heat), gas, network hot water mostly influenced by: • External temperature on characteristic days • all related factors listed for "ID 2: Energy consumption by archetype" Uncertainty in profile shapes for electricity (other) mostly influenced by • Occupancy [by default occupancy will be inferred from floor size, but the user will be able to override this as part of the simulation process to test different sensitivities] • Power rating and number of electrical appliances • Occupant operation of electrical appliances
3	Annual final energy consumption by archetype	SAM	SAM – support validation of building archetype matching process	• Electricity (heat), electricity (other), gas, network hot water, and other final energy vectors (biomass, H2) • by each archetype, • for the base timeperiod	Uncertainty in consumption for electricity (heat), gas, and network hot water is mostly influenced by: • Building thermal conductivity • Building inner thermal mass • Occupancy • Occupant heating hours • Occupant heating temperature Uncertainty in consumption for electricity (other) influenced by the same factors as listed for "ID 1: Final energy demand profiles by archetype"
4	Archetype retrofit / Conversion Costs	POM	POM cost optimisation	• Costs of converting one archetype into another considering separate retrofit options for building fabric and heating systems for each archetype, • Costs will vary by vintage of retrofit to	Uncertainty in retrofit costs influenced by: • Complexity in required alterations to building fabric (labour/materials) • Complexity in required changes to heating system (labour/materials) • Uncertainty inherent in matching starting condition of a dwelling to its

				account for changing costs (e.g. of heating systems in different timeperiods)	archetype representation

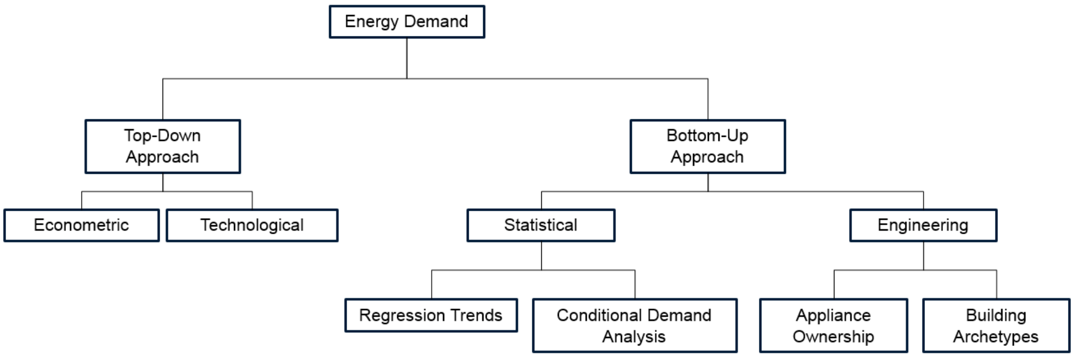
5.3. Key functional requirements

5.3.1. Characterising building archetypes

The EnergyPath™ Design Tool is a strategic design tool for investigating technology transitions at Local Authority scale. A key functional requirement for the HOM is therefore to represent energy demand from buildings to a sufficient level of detail to derive insights on possible technological change.

Approaches to modelling residential sector energy demand can be characterised broadly as either top-down or bottom-up methods. Tools such as the EnergyPath™ Design Tool seek to represent the outcomes of endogenous technological change at a detailed level and so a bottom-up approach is required to allow for explicit representation of the housing stock.

Figure 5-3 Broad Approaches for Characterising Residential Energy Demand (Based on Swan & Ugursal, 2009[15])



All complex models require some abstraction from reality in order to ensure that the problems considered remain computationally tractable. The approach taken for the Household Options Module (HOM) is to create a large dataset containing pre-defined housing archetypes and information on their estimated energy demands, against which data on real-world addresses can be matched. The housing stock of the United Kingdom is diverse and spans a wide range of construction methods, architectural styles, form factors, and vintages amongst its 26 million unique homes.

A challenge for defining archetypes in the HOM is to reduce the number of archetypes so that only the most material characteristics for simulation of the required outputs are being modelled. The approach taken is described in detail in Section 5.5 below.

5.3.2. Estimates of demand and within day profiles

A key functional requirement for the HOM is to estimate the final energy demand profiles of individual building archetypes for a set of characteristic days, such as average seasonal days across the year or may extreme cases such as a 1-in-20 cold winter day, which reflect a design

---

[15] Swan, L. G., & Ugursal, V.I. (200(. Modeling of end-use energy consumption in the residential sector: A review of modelling techniques. Renewable and Sustainable Energy Reviews, 13(8), 1819-1835

standard that the viable pathway must be able to accommodate. The profiles are used by the NAM to help understand the maximum likely demand seen on a section of a network and within the POM as part of the overall supply / demand balancing process.

Bottom-up approaches to modelling residential energy demand can be broadly characterised as either engineering or statistical methods:

- Statistical approaches broadly use empirical data on real-world energy use, such as meter readings, to derive observed relationships between phenomena like internal temperatures and the use of energy in homes over time.
- Engineering approaches seek to represent physical processes that occur in buildings such as fuel combustion from heating appliances and thermodynamic interactions between the heating system and the structure.

The science of modern building energy modelling is dependent on a mixture of both statistical and engineering-type approaches due to challenges associated with capturing real-world data on how people inhabit and use energy in their homes. The approach taken for the HOM therefore combines statistical and engineering approaches for estimating different end-user energy demands:

- Space heating demand is estimated using an engineering approach that uses a detailed technical simulation of thermodynamic processes in different building archetypes.
- Non-heating electricity and hot water demand are derived from parametric methods that use statistical estimates of energy using behaviour linked to occupancy. More detail is described in Section 5.5.3 below.

Temporal granularity is a key consideration in dynamic simulation models, and covers two key dimensions.

- ***The number and type of characteristic days.*** This is intended to be a flexible element of the HOM, subject to computational constraints within the dynamic simulation process discussed in 5.5.3). As a minimum however, the peak design conditions for the end-user building systems need to be captured. This varies for different sectors and for different end-use demands. For domestic heating plant in the UK, for example, the peak design condition is likely to be a winter day, while for commercial buildings with electrical air conditioning systems, the peak could fall on a summer day. This suggests that at least 2 characteristic days are required for dynamic simulation of all building archetypes. End-users may want to evaluate design conditions that are more onerous than typical cold and hot periods, such as heat waves or exceptionally cold 1 in 20 year winters.
    - A distinction is often made between profiled energy demand on weekday and weekend periods so design conditions may also need to take into account the day of the week as well as the season.
    - Finally, dynamic simulation of characteristic conditions typically requires an extended period of between 4-8 days in advance of the target day to be represented in order for transient effects such as thermal mass response and storage to be captured adequately.

- ***The temporal granularity within day.*** These can vary from simple on- or off-peak blocks across the day down to very fine resolution (e.g. <1 minute) intervals. Very high time resolution steps are more appropriate for detailed engineering system design tools rather than software used for evaluation of strategic technology deployment potential, like the EnergyPath™ Design Tool. Greater time resolution also has implications for computational performance i.e. program run times. A key early activity within Stage 2 will be to test the levels of performance required for the dynamic simulation elements of the HOM

The combination of characteristic day and within day period is referred to within the EnergyPath™ Design Tool parlance as a *timeslice.* The level of granularity in the timeslicing ultimately depends on the resolution needed to understand potential trade-offs in pathway design within the model.

For example to understand how desirable a hybrid boiler is, it is important to consider the combination of peak shaving (e.g. to avoid network reinforcement) within day versus overall emissions and direct running costs, which depend on operation across the year. As heat demand swings substantially across the year (along with other factors such as the cost of electricity) sufficient temporal detail (e.g. seasonal / month) is necessary for the model to understand the trade-offs in running the gas versus heat pump component. Without this detail the user must instead make an exogenous assumption about the balance of operation of these components across the year. Similarly, within day there must be sufficient granularity to understand where the electricity peak is likely to arise given overlapping demand profiles and the ability to shift load across the system, either from the operation of hybrid boiler or from other options such as within building or network connected storage.

As a result we are initially proposing to model the following typical timeslices, subject to performance limitations on the HOM dynamic simulation model

- Four typical seasons by working and non-working days (in addition to this the user can define a number of additional non-typical characteristic days to represent design standards)
- 30-minute time resolution within day

It is unlikely to be tractable to represent this level of detail within the POM optimisation problem, alongside significant spatial and pathway timeperiod detail. However, having the HOM produce the final energy demand profiles at this level of resolution gives the user (via the POM - see section 8.6.3) flexibility over how the final timeslices are defined for pathway analysis. The user can easily collapse or exclude certain timeslices and understand to what extent the solution changes – e.g. if you retain 4 seasons and 1 peak day versus collapsing this to 1 peak day and 1 typical day to represent the entire year, or aggregating the 30 minute periods into 3 blocks within-day.

Separately, having the fine resolution within day is also likely to be beneficial for the detailed analysis step taking the high-level POM network solutions back into the NAM (see section 8.8) for further refinement. It is important to note that in the initial analysis of network options (which are then used in the POM) the NAM does not need the full load profiles explicitly, just an understanding from the profiles of the likely highest load on different parts of the network to inform its test configuration process. However, when refining the final network designs it will be important to granular time step testing and using 30 minute contiguous profiles across a number of characteristic days is comparable with other detailed network tools such as WPD's Falcon Sim model.

It is important to note that the HOM produces a 'standalone catalogue' of building archetypes and their associated final energy demand profiles in *isolation* for each building. Once buildings have been allocated spatially and different pathway upgrade options (e.g. in terms of electrifying heating) it is then important to understand demand diversity impacts between buildings in a localised area, to avoid overestimating peak demand levels and by extension network/generation reinforcement requirements. *This cluster-level diversity is discussed further in section 8.6.2.*

5.3.3. Costs of retrofitting / converting building archetypes

Domestic

A key output from the HOM is the cost of different retrofit options for individual building archetypes, as making changes to the existing building stock is one of the many options available to the POM. The POM needs this information in order to make solution choices based around the costs of the simulated housing options.

Retrofit options include making improvements to the thermal efficiency of the buildings by improving the building fabric with measures such as insulation, and deploying alternative heating technologies such as different types of electric heat pump in lieu of conventional gas boilers[16]. Not all building retrofit options will be applicable to all building archetypes. For example, buildings which have solid walls by definition cannot be selected for cavity wall insulation, and ground source heat pumps are not likely to be realistic options for individual above-ground flats.

While changes to the heating system are effectively binary choices (the system is changed or it is not), in reality, even for a single housing archetype, there are many degrees of fabric retrofit which are possible. For example, loft insulation could be added as a stand-alone measure, or in combination with wall-insulation. This is an important optimisation problem to consider, as the costs of implementing measures in concert may be lower than making multiple discrete interventions.

The approach taken in the EnergyPath™ Design Tool is to create a flexible framework to define archetype retrofit combinations – i.e. one archetype is converted into another at a given cost with resulting changes to its final energy demand profiles (e.g. due to different heating systems or improved insulation). What the conversion actually represents could be a minor incremental improvement or a whole house retrofit. The number of retrofit permutations that can be accommodated will be determined as part of performance testing requirements in Stage 2.

Non-domestic

Modelling the retrofitting of non-domestic buildings offers additional challenges over doing the same for the domestic stock. Non-domestic buildings are significantly more diverse than

[16] Where a particular vector is no longer available e.g. the gas grid is decommissioned the cost of converting appliances (e.g. gas to electric cooking) would also need to be accounted for.

dwellings in terms of their energy use, even within buildings of nominally the same type. For example, buildings identified or classified as offices range from low-rise naturally ventilated structures that have often been converted from former dwellings, industrial or agricultural buildings, to purpose built commercial tower units. The information technology, lighting, heating, ventilation and cooling (HVAC) strategies and systems employed in office buildings can vary widely with age, occupancy and tenancy, sometimes even on different floors or wings of the same structure. The result is that even when the form and function of two non-domestic buildings are similar, the energy demand and the available pathways for energy efficient retrofit are diverse and resistant to characterisation or archetyping.

As described in Section 4.6.1, HOM estimates energy demand for non-domestic buildings using a statistical inference approach as opposed to bottom-up engineering-led methods. At the time of writing, there are no large scale studies on the performance and costs of non-domestic retrofitting of UK buildings that can draw broad, statistically robust generalisations about costs and or resultant performance improvements. However, the best short term (<10 year) prospects for better understanding of these issues is likely to come from the outputs of future field trial studies and other exogenous data sources rather than innovations in bottom-up modelling.

It is therefore proposed that non-domestic archetype retrofit in HOM, takes a statistical approach. The model will be constructed with the functionality to accept future information on energy retrofit performance and costs, with the intention that these are populated in future by external data. While no hard data would be available for populating these model elements initially, the relevant input fields could still be used as part of scenario analysis and sensitivity testing within the EnergyPath™ Design Tool.

Possible approaches for achieving this level of functionality that will be investigated in Stage 2 include:

- The ability to define discrete non-domestic archetypes in the HOM with different levels of performance, following a similar framework to that used for domestic buildings i.e. retrofitting causes a building to move from one archetype class to another, with an associated cost
- The use of a scalar function or functions for energy demand reduction linked to representative cost curves in the POM, applied to aggregate demand from the HOM

5.3.4. Capturing ‘householder behaviour’

User behaviour is a strong determinant of energy use in and ultimately accounts for much of the observed variation in levels of energy demand and patterns of energy use between otherwise similar buildings. A key functional requirement of the HOM is to capture an estimate of this variation from individual building archetypes. Consumer preferences also directly relate to the acceptability of different technologies and implicitly affect their installation costs.

The approach taken for representing householder behaviour in the HOM is:

- To capture uncertainty from energy-using behavior in buildings by estimating the upper and lower boundaries of demand from a combination of endogenous probabilistic simulation and exogenous ranges determined from literature

- To handle consumer preferences as indirect exogenous inputs, as noted in section 2.1

5.3.4.1. Energy-using Behaviour

Building occupants operate the energy using appliances in their homes to meet their daily needs and comfort requirements, which vary between individuals and households. Variation in how people operate their power and heating systems is driven by factors such as individual preferences and social and cultural norms (washing clothes every day, expectation of thermal comfort levels etc). While it is difficult for any single building to estimate the exact pattern of energy use that will arise on a given day, it is possible to estimate a measure of the spread around the typical or average day. Options for doing so are summarised below, *the approach to uncertainty across the tool is discussed more in section 9.*

End-Use Energy Demand	Characteristic	Method for Estimating Uncertainty	Details
Space Heating	Consumption	In principle could generate distributions endogenously using HOM simulation models varying internal temperature desired, occupant hours of heating, etc	Monte Carlo analysis of space heating consumption with dwellings at different internal temperatures with different heating system types
	Peak Demand		
Hot Water	Consumption	Apply exogenous distribution of demand from external studies.	Sources include EST/Defra Measurement of Domestic Hot Water Consumption in Dwellings, 2008, or DECC, Defra, EST Household Electricity Use Survey (HEUS) 2011
	Peak Demand		There is only limited potential to estimate the uncertainty from the hot water peak demand curves from freely available data, but HEUS does have examples of variation between working days and holiday days that could be used as a proxy range until future studies
Appliance Electricity Use	Consumption	Apply exogenous distributions of demand from detailed external studies.	Examples include the Household Electricity Use Survey (HEUS) and the Energy Follow Up Survey (EFUS) 2011
	Peak Demand		Statistical analysis of variation in electrical load profile data for non-electrically heated dwellings using external sources (see data acquisition deliverables in section 4.4.1)

5.3.4.2. Consumer Purchase Behaviour

Different households in otherwise identical buildings with the same heating system and electrical appliances may exhibit different patterns and levels of energy use, as described above. Another important factor to consider in demand modelling is that different households demonstrate variation in terms of their appliance ownership and how this changes over time. Consumer preferences may result in different households choosing to buy greater or fewer electrical appliances, to invest in microgeneration systems, or to choose to drive an electric vehicle (EV). This means that over time, two initially identical buildings could eventually arrive at a future time period where they possess different levels of insulation and completely different heating systems.

Carrying out a detailed modelling exercise to estimate the pattern of future investments in energy using technologies over time between different building and household combinations is complex. This requires discrete choice econometric methods or an explicit agent-based approach to understanding the influence of technology availability and cost, market structure, supply chains, and end-user demand on what may be bought and when. This level of complexity is best approached by using tools that have been specifically designed to investigate these issues, and it is not proposed to build-in the full level of functionality required to generate the required insights within a strategic technology pathway tool like the EnergyPath™ Design Tool.

This does not mean consumer preferences cannot be captured in the EnergyPath™ Design Tool, but it does mean that the onus is on the end user to understand and represent the effects of consumer choices by varying the inputs to the model. The effects of different consumer preferences can be compared and contrasted through establishing different sets of scenario inputs, informed by data and insights derived from other tools and studies. For example:

- A scenario where consumers exhibit a strong preference for a certain technology can be examined by constraining the model to force a minimum level of uptake.
- Another example is to change the modelled costs of different technologies to reflect their implicit consumer value, such as representing the "hassle" costs associated with the installation of a technology that is hard to integrate and involves disruption to daily life i.e. representing the negative value associated with unpopular technologies.

5.4. Key inputs

The key inputs for the HOM are summarised below. All inputs are ultimately required for creating the building archetypes, performing energy demand simulations for domestic and non-domestic buildings, and determining retrofit costs. Data inputs to the HOM can be separated into:

- Information that is required for matching building archetypes with address-level data from the SAM. These are the *class or type identifiers* that can be used for characterising buildings when detailed local survey information is not available to provide a *complete* set of data for an individual building.
- Information that cannot be matched by the SAM, but which is nevertheless required for building archetype energy simulation. This information often needs to be inferred statistically from best available third party datasets. For example, building construction details (wall and window types) are rarely available from primary spatial data, and are typically inferred based on characteristics such as the building age.
- Information that is purely required for the purposes of building energy calculations. In a few cases this information is exogenously user-defined, which implies that the end user must make specific assumptions to frame their investigation from a series of choices that are defined in POM, for example the desired internal temperature.

It should be noted that there is no location specific data in the HOM at this stage. However, the characteristics used to create the archetypes and undertake the energy simulations must be both consistent with data used in other modules, in particular the SAM, (e.g. the age categories of buildings) and cover the range of possible conditions that could need to be modelled in different locations (e.g. typical external temperature profiles in different UK regions).

The tables below describe what is technically required within the tool and how it would be used. The relative importance of the primary data inputs and other factors such as costs and licensing restrictions, particularly where they must be purchased from an external provider, are discussed in deliverable (D4) Data Acquisition Plan.

Table 5-2 Input data – archetype identification characteristics for SAM matching – information from primary data

ID	Data field(s)	Input type	Purpose	Granularity	Source
1	Property Type	Primary	Key for determining morphology of building, in particular external facing wall area and volume, which are material for energy use	Per domestic archetype ~5 classes. Examples: Flat, terraced, bungalow, semi-detached, detached	HOM definition based on Experian ConsumerView classes
2	Floor Area Banding	Primary		Per domestic archetype ~5 classes. Examples: <50m², 50-69m², 70-89m², 90-109m², 110m²>	HOM size bands based primarily on English Housing Survey data, but can be adjusted as appropriate to match across datasets
3	Property Age	Primary	Age on its own does not give energy use, but important for matching many other morphological characteristics	Per domestic archetype ~7 classes Examples: pre-1870, 1871 - 1919, 1920 - 1945, 1946 – 1954, 1955 - 1979, post-1980, post-1990	HOM definition based on Experian ConsumerView classes
4	Government Office Region	Primary	Location determines factors such as external temperatures in summer/winter, and solar insolation, so this is material for space heating energy use and possible microgeneration from photovoltaics	Per domestic archetype ~11 classes. North East, North West, Yorkshire and The Humber, East Midlands, West Midlands, East of England, London, South East, South West; also Wales, Scotland	HOM definition based on Office of National Statistics and Ordnance Survey MasterMap
5	Tenure	Primary	Tenure can be a useful predictor of retrofit levels and appliance ownership	Per domestic archetype~3 classes Social, Owner Occupied, Private Rented	HOM definition based on Experian ConsumerView classes
6	Non-Domestic Floor Area Band	Primary	Determines size of building	Per non-domestic archetype, classes TBD Examples: <99m2, 100-199m2, 200-299m2, 300-399m2, 400-499m2, 500-599m2, 600-699m2, 700-799m2, 800-899m2, 900,999m2, 1000m2>	HOM definition based on Valuation Office Agency (VOA) classes
7	Non-Domestic Activity Class	Primary	Determines energy use intensity and pattern of use	Per non-domestic archetype ~300 classes	HOM definition based on Valuation Office Agency (VOA) classes

Table 5-3 Input data – archetype identification characteristics for SAM matching - information that can be inferred from aggregated sources (postcode, LSOA, MSOA, LA, GOR, national scale etc.)

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Rural/Urban Classification	Primary	Determines wind exposure, so material for space heating energy use	Per domestic archetype ~3 classes. Examples: Urban, Suburban, Rural	Source-dependent	Source-dependent	Imputed from third party sources possibly Experian data, or calculated in SAM based on address point density	Source-dependent
2	No. of Storeys	Primary	Key for determining morphology of building, in particular external facing wall area and volume, which are material for energy use	Per domestic archetype ~3 classes. Single Storey, Two Storey, Two Storey	EHS reporting provide no data on the accuracy of this metric or how it was collected, analysis of raw data needed to quantify uncertainty	England	Imputed based on English Housing Survey data (2010 Homes Report Annex Table 1.8, 1.9, 2011 Homes Report Annex Table 2.9) matching number of storeys above ground and the presence or not of a basement to "ID 11: Tenure" *NB: Ordnance Survey experimental height dataset or commissioned LiDAR surveys may offer an additional source. Similarly GeoInformation group data has primary data on building height and number of inferred stories, but this is not comprehensive across the UK.*	Modelled data – address level information imputed from regional statistics *But as noted in the previous column, as primary data becomes more widespread this characteristic could be matched directly as per the other items in the table above.*
3	Wall Type	Primary	Key for determining thermal efficiency and retrofit options, material for energy use	Per domestic archetype ~4 classes. Examples: Cavity Filled, Cavity Unfilled, Solid Wall Uninsulated, Solid Wall Insulated Externally, Solid Wall Insulated Internally	"	"	Imputed based on English Housing Survey data (2011 Homes Report Annex Table 1.16) matching wall construction to "ID 1: Property Type", "ID 3: Property Age" and "ID 11: Tenure"	Modelled data – archetype characteristic is imputed from a GOR level survey

4	Window Type/Glazing Type	Primary	Key for determining thermal efficiency and retrofit options, material for energy use	Per domestic archetype ~8 classes. Examples: Mixed Types, Single Glazed Wood Casement, Single Glazed Wood Sash, Single Glazed UPVC, Single Glazed Metal, Double Glazed Wood, Double Glazed UPVC, Double Glazed Metal	"	"	Imputed based on English Housing Survey data (2011 Homes Report Annex Table 1.20) matching window type to "ID 3: Property Age"	"
5	Loft Insulation Thickness Band	Primary		Per domestic archetype ~5 classes. Examples: 0-50mm, 51-100mm, 101-150mm, 151-200mm, 200>	"	"	Imputed based on English Housing Survey data (2011 Homes Report Annex Table 4.6) matching insulation thickness to "ID 11: Tenure"	"
6	Existing Primary Heating System Type	Primary	Key for determining space heating and hot water use, demand profiling	Per domestic archetype ~20 classes. Examples: Gas Boiler with Storage, Gas Combination Boiler, Electric Resistive Heating, Oil Heating, Solid Fuel + various new heating system types	"	"	Imputed based on English Housing Survey data (2011 Homes Report Annex Tables 4.2 and 4.8, matching fuel types to "ID 11: Tenure" and/or 2010 Homes Report Annex Table 2.2 matching Age/Type to off-peak electricity supply)	"
7	Secondary Heating System Type	Primary	Key for determining space heating and hot water use, demand profiling	Per domestic archetype Examples: Electric Resistive Heating, Solid Fuel	"	"	Imputed based on English Housing Survey data (2011 Homes Report Annex Table 4.19) matching secondary heating to "ID 11: Tenure"	"
8	Heating System Controls	Primary	Key for determining space heating and hot water use, demand profiling	Per domestic archetype Examples: Time Clock Control, Cylinder Thermostat	"	"	Imputed based on English Housing Survey data (2011 Homes Report, Annex Table 4.12) matching whether or not a house would benefit from heating system controls to "ID 1: Property Type"	"
9	Hot Water Storage	Primary	Key for determining space	Per domestic archetype	"	"	Imputed based on English	"

	Type		heating and hot water use, demand profiling	Examples: Tank Size / Type combinations			Housing Survey data (2011 Homes Report, Annex Table 4.12) matching whether or not a house would benefit from hot water insulation or not to "ID 1: Property Type"	
10	Hot Water Tank Insulation Type	Primary	Key for determining space heating and hot water use, demand profiling	Per domestic archetype Examples: Foam, mineral wool jacket	"	"	"	"
11	Hot Water Tank Insulation Thickness Band	Primary	Key for determining space heating and hot water use, demand profiling	Per domestic archetype Examples: None, 0-160mm, 160mm>	"	"	"	"
12	Number of Baths/Showers	Primary	Material for hot water consumption	Per domestic archetype Examples: 1, 2=>	"	"	Imputed based on English Housing Survey data (2011 Homes Report, Annex Table 2.7) matching whether or not a house would benefit from hot water insulation or not to "ID 11: Tenure"	"
13	Type of microgeneration installation	Primary	Affects final energy vectors	Per domestic archetype Examples: None, Solar PV, Solar Thermal, Micro-Wind	"	"	Can possibly be obtained from Germserv Microgeneration Certification Scheme Database	Address level information may not be released and microgeneration may need to be imputed from regional or large area scale data, at best
14	Non-Domestic Building Height Band	Primary	Non-domestic energy simulation	Per non-domestic archetype Examples: 1 storey, 2 storey, 3 storey, 4 storey, 5 storey, 5 storeys>	Source-dependent	Per non-domestic archetype	Ordnance Survey experimental height dataset or commissioned LiDAR surveys are the major potential sources of information. Similarly GeoInformation group data has primary data on building height and number of inferred stories, but this is not comprehensive across the UK.	Source-specific

Table 5-4 Input data – variables for building energy and retrofit cost simulation that cannot generally be matched to spatial datasets

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Component cost and performance of archetype upgrades – building fabric	Primary	Cost overall conversion	Based on ETI optimising thermal efficiency study – considers • Insulation options Decline in costs over time	High, medium and low costs	UK	ETI - Optimising Thermal Efficiency of Existing Housing (OTEoEH) Study – component costs	Triangular distribution of uncertainty only. No documentation on methodology for arriving at cost estimates.
2	Component cost and performance of archetype upgrades – heating systems	Primary	Cost overall conversion	Based on ETI Smart Systems and Heat Work Area 1 data – considers • Heating systems • Storage • Controls Decline in costs over time	Likely that performance and cost data will fall within upper and lower bounds	UK	ETI – Smart Systems and Heat Work Area 1 ETI ESME Model Macro-DE project Various published DECC, CCC, other sources	SSH WA1 does contain performance data for different technologies, but whether these are specified as ranges or single numbers, which technologies are covered, etc. needs to be investigated
3	Component cost and performance of archetype upgrades - microgeneration	Primary	Affects final energy vectors	Per archetype. Examples: Solar PV, Solar Thermal, Micro-Wind	Source-dependent	N/A	Performance and cost data to be collated from published estimates	Source-dependent
4	Target Mean Internal Temperature	Primary	Domestic energy simulation	Single Point	User-defined input	Per Domestic Archetype	Targets from CIBSE guidance	Uncertain central estimate, allow user to vary for sensitivity/resilience testing
5	Occupancy levels	Secondary	Domestic energy simulation	Number of occupants and pattern of occupancy	Inferred from floor area data	Per Domestic Archetype	Inferred	Uncertain central estimate, allow user to vary for sensitivity/resilience testing
6	Electrical Appliance Consumption Band	Primary	Key for determining electrical energy demand	Per domestic archetype. Examples: <3,000 kWh, 3,000-7,500 kWh, 7,500 kWh>	Elexon dataset does not provide information on uncertainty, may be possible to infer	Per Domestic Archetype	Elexon stratification or Intertek Household Electricity Survey for DECC	Elexon data doesn't include information on validity or validation. Applicability of Home Electricity Survey data beyond original 250 dwelling sample is unknown

					statistically from Home Electricity Survey			
7	Ventilation System	Primary	Most UK properties are naturally ventilated, possible to infer others from heating system/type/age/size – generally newer properties	Per domestic archetype Examples: Naturally ventilated, mechanically ventilated	User-defined input	N/A	Assumption that existing properties are naturally ventilated	No known high level sources at national/regional level record mechanically ventilated properties
8	Shower Type	Primary	Material for energy demand vector	Per domestic archetype Examples: gas, electric	User-defined input	Per Domestic Archetype	Imputed based on primary heating system – can expect non-gas heated properties will not have gas heated showers, for example	No known high level sources at national/regional level record whether showers are electric or not in UK properties
9	External Temperature on Characteristic Days	Primary	Domestic energy simulation	Hourly or better for up to 5 characteristic periods of at least 5 days each (to allow for thermal mass effects)	Uncertainty inherent in future projections of weather conditions	Per Domestic Archetype	Synthetic design weather data such as the CIBSE Design Weather Years or output from the UKCP09 Weather Generator	See notes on uncertainty
10	Domestic Electrical Appliance (Non-Heating) Demand Profile - covers - EVs - Appliances and lighting	Primary	Domestic energy simulation	Half hourly per archetype for up to 5 characteristic periods of 1 day each	Uncertainty inherent in source information will filter into the HOM	Per Domestic Archetype	Elexon Load Profiles or other sources identified as part of data acquisition deliverables. Project partners WPD have data on diurnal electrical demand profiles DECC Home Electricity use Survey Data ETI partners E.ON and EDF Energy also possess data on diurnal electrical demand ETI data on EV charging profiles	Elexon documentation is vague regarding accuracy of profiles at the household level – likely to be broadly valid in aggregate but not for individual households Need to isolate non-heating demand from overall electrical profile in some cases (e.g. Elexon class 1 profile are very similar to gas-heating properties from DECC HEuS survery)

Baringa

11	Domestic Hot Water Demand Profile	Primary	Domestic energy simulation	Hourly per archetype for up to 5 characteristic periods of 1 day each	See limitations	Per Domestic Archetype	Energy Saving Trust Domestic Hot Water Study or Household Electricity Use Survey	The applicability of the EST profile beyond the 120 dwelling study sample is not known
12	Non-Domestic Electrical Energy Consumption	Primary	Non-domestic energy simulation	Non-Domestic Archetype specific	Uncertainty of overall energy use can be quantified statistically from original data frequency distributions, but not necessarily by end-use	Four Towns Study Area	Sheffield Hallam University Four Towns Longitudinal Study, CIBSE TM46	Applicability of data beyond original 700-building sample is not known
13	Non-Domestic Heat Energy Consumption	Primary	Non-domestic energy simulation	Non-Domestic Archetype specific	"	"	"	"
14	Non-Domestic Electrical Appliance (Non-Heating) Demand Profile	Primary	Non-domestic energy simulation	Half hourly per archetype for up to 5 characteristic periods of 1 day each	Uncertainty inherent in source information will filter into the HOM	Per Domestic Archetype	Elexon Load Profiles or other sources identified as part of data acquisition deliverables. Project partners WPD have data on diurnal electrical demand profiles ETI partners E.ON and EDF Energy may also possess data on diurnal electrical demand	Elexon documentation is vague regarding accuracy of profiles – likely to be broadly valid in aggregate WPD / E.ON / EDF Energy to advise
15	Non-Domestic Space Heating and Hot Water Demand Profile	Primary	Non-domestic energy simulation	Hourly per archetype for up to 5 characteristic periods of 1 day each	Dependent on source	TBD	Project partners Ramboll to advise on diurnal load shapes for non-domestic heat demand ETI partners E.ON and EDF Energy may possess data on diurnal gas demand	Applicability of data may be restricted to certain sub-domains depending on its source.

5.5. Key HOM logic / process steps

5.5.1. Base building archetype creation [HOM-001/2]

5.5.1.1. Domestic [HOM-001]

As indicated in Section 4.4, a wide range of physical housing characteristics are important for determining estimates of energy use in dwellings. The first step in the process for creating domestic building archetypes is to determine the various combinations of characteristics that will be modelled:

- Building characteristics that are available in SAM datasets and cannot generally be inferred (see Table 4-2) provide the high level matrix of possible archetype permutations
- For each of these permutations, other SAM matching characteristics[17] can be inferred from aggregate datasets, such as national housing surveys (see Table 4-3). Statistical analysis of this data can be used to determine whether there are any combinations of defining characteristics that do not exist in the housing stock or which are statistically not significant at a user-defined level (for example <1% of the stock). For example, modern new-build flats in urban areas which have oil fired heating are technically possible to construct, but are arguably not useful to simulate in the EnergyPath™ Design Tool, as there may be few or no real-world examples, nor is it likely to be a potentially attractive technology for retrofit in a smart heat system. Cavity wall construction was introduced to the UK in the 19th century, so buildings constructed prior to this period should only have solid walls, etc. Combinations of retrofit options that appear improbable can also be filtered at this stage, such as:
    - Heating systems that require external fuel storage (such as LPG or oil heating) in individual apartments
    - Heating systems with thermal storage tanks in small properties (e.g. <50m^2) that have already been designed with or converted to use combination boilers
    - Heating systems that require ground-floor access on properties that may not have ground-floor access, such as ground-source heat pumps in individual apartments
    - Roof-mounted micro-generation systems on properties that may not have roof access, such as individual apartments
- In turn, a large number of inputs dwelling archetypes need to be inferred for energy simulation purposes. In the absence of any better information for an energy simulation input, the HOM will adhere to SAP and ISO13790 defaults which follow the reference

[17] For example, there are no existing national datasets that show precise “Wall Type” at an individual address level, but data at a national level does enable wall type to be inferred by property age, type and tenure. This is a limitation of current data availability. If additional address level data on building characteristics becomes available there is no barrier in principle to including this when generating the high level matrix of characteristics.

standard. Statistical analysis of the geometric characteristics of buildings from housing survey data and architectural literature will also be used to establish key inputs such as the number of rooms and their spatial inter-relationships – past work has revealed that these tend to be strongly unimodal i.e. there is a typical layout that can be archetyped. Identical assumptions used in the parameterised simulation will also be reflected in any dynamic simulations.

- The process of inferring large numbers of energy simulation inputs from available data will introduce uncertainties into the EnergyPath^{TM} Design Tool when the SAM matches individual addresses to the best available archetypes. It may also be possible in many cases to quantify this uncertainty. For example, no data may be available that indicates how many electric showers a household has, so this is likely to be a user-defined input backed by assumptions. The extremes of the distribution are found by modelling a binary choice – either all the hot water for bathing is provided by an electrical vector or it is not. The uncertainty introduced into the resulting consumption can then be estimated by taking the end use energy consumption for domestic hot water from the showering/bathing sub-model and applying it to either electricity or gas/DH hot water demand, thus giving the two extremes.

5.5.1.2. Non-domestic [HOM-002]

Physical characteristics that are used in characterising non-domestic buildings for the purposes of energy use simulation are described in section 5.4. Non-domestic buildings are significantly more difficult to characterise than the domestic stock because of the heterogeneity of built forms and activities, and also because the majority of spatial data sources record the addresses of premises rather than physical buildings. It is often the case in the UK that individual premises span multiple buildings or a single building may contain multiple premises. Ideally, non-domestic energy consumption for use in the HOM should be directly measured from primary data where this is possible, i.e. direct meter readings, outputs from building energy management systems, or actual end-user billing information.

Where direct measurement is impossible, an archetypal representation of the non-domestic stock will be necessary to estimate energy use from the non-domestic stock. Two possible methods are identified:

- **"Self-Contained Unit (SCU)" Method:** floorspace should be matched to physical buildings in the SAM to improve the positional accuracy of energy demand. Techniques for achieving this are currently under development at a number of UK academic institutions and require the use of building height data, which is typically obtained from LiDAR surveys.

- **Aggregate Floorspace Method:** In other cases, only an aggregate approximation of the floorspace within an individual SAM cluster may be possible.

For both the physical and aggregate floorspace model, the non-domestic stock will be assigned class and activity types based on Valuation Office Agency Special Category (SCat) coding, which is then used for estimating energy demand, as described in section 5.5.4.

5.5.1.3. Other non-heat related electricity

As indicated by the data inputs described in Section 5.4, the HOM addresses non-heat related electricity demand / supply largely through the application of exogenously determined inputs for consumption and profiling. These will be separated into:

- EV charging profiles, so that the assumptions around charging profile and implications for spatial cluster diversity in the POM can be more easily processed (see section 8.6.2), drawing on ETI data initially.
- All other non-heat related electricity (lighting and appliances)
- In addition, the HOM will contain profile, cost and other relevant data for building-scale micro-generation technologies such as solar PV. These do not lead to additional archetype classifications, but a flags are applied to say whether the building is appropriate or not for the technology. The POM can then decide whether it is cost-effective to deploy these technologies (limited by data from the SAM on e.g. estimated roof area for each building in the area).

Lighting energy consumption for dwellings is calculated as a function of floor area within ISO13790 but actual appliance use is derived from statistical analysis of observed data. Sources for estimating dwelling appliance energy consumption include the Defra, DECC and EST funded Household Electricity Survey (HES) and the DECC Energy Follow Up Survey (EFUS). Non-domestic buildings will have their electrical energy consumption estimated on a kWh/m2 basis as described in Section 5.5.4.

Profiles for lighting and appliances can be obtained from a number of sources, such as Elexon profiles for non-metered customers, which are used in the UK Balancing and Settlement Code[18]. It is also anticipated that ETI member organisations may be in a position to provide the ETI with diurnal profile data suitable for use in HOM.

5.5.2. Building archetype dimensional reduction for simulation [HOM-003/4]

A cursory examination of the number of identifying characteristics and the example states for each characteristic (see tables 4-2 and 4-3) reveals a large combinatorial problem posed by the number of possible permutations of domestic building archetypes.

However, many of these possible archetypes might have similar levels of energy performance. Mathematical 'cluster' analysis is used to group building types with similar key performance characteristics together, thus reducing the number of simulations that needs to be performed to a computationally tractable number *(note that mathematical clustering here is not related to the broader concept of spatial clustering used throughout the rest of the EnergyPath™ Design Tool).*

[18] Elexon data segments out homes that are primarily electrically heated from those that are not. The shape of the Elexon electricity demand profile from the gas-heated properties (Class 1) looks very similar to the shape of the electricity demand for gas-heated properties and those with secondary electric heating the Household Electricity Use Survey. This means the shape of an Elexon load profile for a gas heated house is a good proxy for lighting/appliance load.

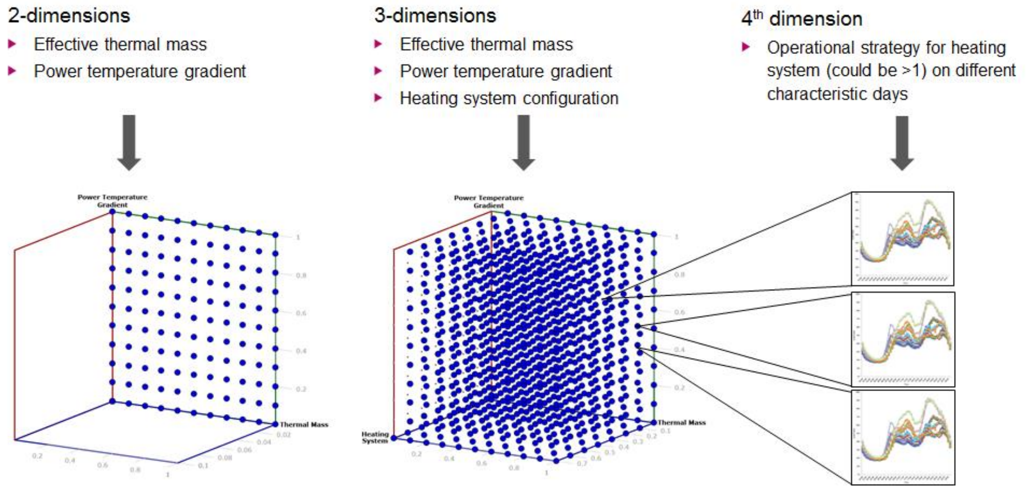
‘Cluster’ analysis takes the full spectrum of possible user-defined building permutations and represents them by only the most material characteristics that affect their required outputs and the trade-off in choices seen by the POM. The most material outputs which affect the energy consumption of the building and the diurnal load profile of demand are:

- The thermal efficiency of the building fabric and its overall exposed surface area, which can be expressed as a Heat Loss Rate in W/K.
- The effective (short-term) Thermal Mass of the building which can be expressed in J/K
- The type of heating system installed in the building
- The operational strategy employed by the heating system (use of pre-heating, storage, hysteresis tolerated, etc.)

These four outputs map to a 4-dimensional problem space which will be populated with a user-defined number of archetypes under the following process:

- A housing stock database is created to represent the characteristics of UK housing at the national level.
- The Heat Loss Rate (HLR) and Thermal Mass (TM) for all of the buildings in the housing stock database are simulated using a parameterised ISO 13790 compliant method based on the UK’s Standard Assessment Procedure (SAP) developed for compliance with the European Energy Performance of Buildings Directive (EPRD)
- Algorithmic decomposition is used to cluster simulated buildings into groups based on relative density and a user-specified number of target clusters. It is not possible in Stage 1 to pre-judge which algorithms will be appropriate for meaningful decomposition of the problem but example approaches include:
  - Variations of Lloyd’s algorithm, also known as k-Means clustering
  - Density-based algorithms like DBSCAN/OPTICS
- Each building archetype, expressed as a point within the HLR/TM problem space, can have multiple heating systems associated with it, giving a 3-dimensional grid. In turn, each of the heating systems can have multiple operational strategies (see section 4.5.3.1). This gives the 4-dimensional problem space.
- The retrofit of buildings is represented by moving from one point on the 3-dimensionsal grid (HLR/TM/heating system combination) to another point, with parameterised costs calculated for the conversion steps taken.

Figure 5-4 Overview of dimensional problem space



These clustering techniques have already been applied across a number of studies, but typically considering fewer dimensions. The additional complexity in this case arises from the number of dimensions considered, and as a result the associated increase in data to be managed.

A number of flexibility options are proposed for the data architecture of the model.

- Not only can the main clustering metrics and their range bucketing be changed, but the algorithmic decomposition algorithm used is also flexible.
- The building stock database used can be enriched with additional survey information, such as locally gathered datasets. The influence of these changes can then be automatically propagated through to the model.

A key task in Stage 2 of the project will be to review the clustering options presented by the housing stock database, which will initially be based on the English Housing Survey (EHS), and identify the meaningful partition ranges along each axis. These remain to be determined but the illustrative table below offers a simple illustration of how the clustering approach can reduce the number of simulations that are required.

Table 5-5 Comparison of basic versus mathematically clustered characteristic states

Basic Dwelling Characteristics	Example Number of States	Mathematically Clustered Characteristics	Number of Example States
Property Type	5	Heat Loss Rate	10
Floor Area Banding	5	Effective (short-term) Thermal Mass	10
Property Age	7	Heating System Type	10
Government Office Region	9	Heating System Operational Strategy	3
Rural/Urban Classification	3		
Tenure	3		

Number of Storeys	3		
Wall Type	3		
Glazing/Window Type	7		
Loft Insulation Thickness Band	5		
Primary Heating System Type	10		
Secondary Heating System Type	2		
Heating System Controls	2		
Hot Water Storage Type	2		
Hot Water Tank Insulation Type	2		
Hot Water Tank Insulation Thickness Band	3		
Ventilation System	2		
Electrical Appliance Consumption Band	3		
No of Baths/Showers	2		
Shower Type	2		
Total Permutations*	51,438,240,000	**Total Permutations**	3,000

*Not all combinations are feasible (for example 2-storey and 3-storey bungalows by definition, do not exist), so this direct combinatorial comparison exaggerates the magnitude of example combinations, but is a useful illustration of the scale of the problem.

Static domestic building energy simulation to support process [HOM-003]

To support the dimensional reduction process a very fast static building energy model is needed. It is not possible to use the same dynamic energy simulation model (see section 5.5.4) which is used to generate the within day profiles as this is not sufficiently fast to assess the large number of permutations.

The HOM will use a parameterised ISO13790 model used as part of the above clustering process (as described in the (D2) Design Architecture Deliverable this will re-use and integrate the ETI's existing SAP model). This determines annual energy consumption for each archetype, but not peak demand conditions.

5.5.3. Domestic building dynamic energy simulation [HOM-005]

A dynamic building simulation tool is proposed for determining the possible operational profiles of individual heating systems and the consequent final energy demand for electricity, gas, or heat provided from a district heating network. The key outputs to the model beyond the HOM will be the sample curves for each characteristic day (typical and edge cases), and three summary statistics: how much energy can be time-shifted during the day, by how many hours, and what the impact on peak power is.

The structure of the EnergyPath^{TM} Design tool data architecture, which is designed to solve a large and complex optimisation problem, necessitates a trade-off between detail and complexity when modelling the "smart" aspects of the system, such as in-building energy storage and the choice of different operational profiles for load shifting at a network scale. These design choices are particularly relevant for electrical heating options such as heat pumps, as future grid-reinforcement costs are likely to be a function of the extent to which intense peaks can be mitigated or avoided through these types of option.

The ideal method of determining design choices of this nature would be to employ an optimised dispatch algorithm for all of the individual systems in every building on the network simultaneously, along with information about key constraints such as electricity pricing, and the electrical engineering constraints (voltage drop, thermal limits etc.) on the network. However, the architecture for the EnergyPath™ Design tool compartmentalizes all of these components separately for reasons of computational tractability.

Building operation is determined essentially off-model in HOM, network engineering constraints are modelled in the NAM, while the impacts of diversity and overall supply / demand balancing are determined in POM. This means that when carrying out dynamic simulation in HOM to determine the operating profile for say, a heat pump, the model has no ex-ante scenario information about what may be happening at a grid level to try and mitigate peaks, such as the use of price signals or active load control.

The approach taken in the HOM for providing optionality in the dispatch of heating devices will be to provide more than one profile for some archetypes, giving the POM a range of design choices to select from. For example, an archetype with integrated storage may have a number of profiles representing different degrees of possible load shifting from the storage (while still meeting the desired internal comfort requirements). The POM would then assess, which of this profiles is most cost-effective from the overall system perspective. This implies some degree of automated control at the system level. If instead it is assumed that consumers override the controls or do not install these in the first place, fixed profiles with no optionality could be used, and alternatives applied as part of sensitivity testing.

A number of key factors are material for the operational strategy employed by dwelling space heating systems and the resulting load profile shapes of the final demand vectors (gas, electricity, biomass etc.). The number of load profiles to be developed for each archetype on each characteristic day may be computationally constrained and needs to be balanced against the requirement to include other problem space dimensions as described above in section 4.5.2. Operational strategies are determined by considering a range of factors as outlined in the table below.

Table 5-6 Drivers of operational strategies

Key Optionality Drivers	Process Approach
External temperature	External temperatures can be determined for each characteristic day being modelled in the dynamic simulation. As part of determining the characteristic day information it will be necessary to model a longer period (e.g. a week before) to understand the implications for thermal mass.
Internal temperatures	Target thermal comfort temperatures during occupied periods will be specified for individual archetypes, with the distribution of uncertainty around the central estimate determined as described in section 4.3.4.1
Heating system configuration	A limited number of configurations are possible for space heating and hot water delivery. These include: • Primary heating only • Primary heating with secondary hot water heating

	• Primary heating with secondary hot water and space heating
Size of primary heating system	For each archetype and each heating system type the system size can be matched to heat demand following relevant guidance from engineering best practice e.g. CIBSE design guides. In the majority of cases this will result in a limited range of system sizes that are capable of meeting the demand requirements for each archetype.
Size of secondary heating system	Secondary heating systems are typically sized to charge hot water storage when the primary system cannot reheat the cylinder quickly enough to meet demand requirements. Reheat times can be estimated for different archetypes based on the size of the property, the likely number of hot water using systems, and the size of the thermal store.
	Secondary heating systems are also used for "boosting" the space heating output when the primary system cannot meet demand. Hybrid-type systems where one of the heating devices has a low delivery temperature may be designed intentionally in this fashion.
Relative size of primary/secondary heating system for hybrids	There are a limited number of hybrid systems on the market that have standardized sizes for their respective components, which will be used as a guide when matching the heating system to individual building archetypes.
Suitability of building structure for pre-heating	Pre-heating does not need to be considered for thermally lightweight buildings because there are no benefits to employing this strategy.
Size of storage tank	Domestic storage tanks come in a limited number of standard sizes. The size of the tank can be determined as a function of the primary and secondary heating systems and the size of the property, which are both effectively constraints on tank size.

These factors will need to be explored for each archetype under an automated process. While an almost infinite number of variations are technically possible, it is likely that only a few strategies are anticipated to be of material interest and can be determined through the identification of important parameters (i.e. those that the system is most sensitive to) as well as through logical deduction.

Other modules in the EnergyPath™ Design Tool will principally be using the final load profiles for the purposes of demand-supply matching in different time slices. The most detailed time-of-use pricing based on supply and transmission constraints is determined by ESME, and is currently set to 5 unequal periods. Dynamic micro-simulation will provide half-hourly diurnal profiles, as well as peak power, but the number of simulations to be run will be in part constrained by the number of time-of-use pricing periods. Cluster analysis of the outputs of multiple micro-simulation runs will be used to parameterize the key outputs that are used by the rest of the model: the maximum reductions in peak power; and the amount of energy that can be shifted within and between time-of-use periods.

Developing an in-house building dynamic simulation tool for ETI is likely to exceed the time and budget constraints the project. A number of possible 3rd party tools have been evaluated and described in the (D2) Design Architecture deliverable.

The computational load represented by the need to perform dynamic building energy simulations for a large number of archetypes remains to be determined in Stage 2. A number of strategies exist for overcoming this challenge:

- Simulating a lower number of archetypes and populate a relatively sparse problem-space grid through interpolation
- Parallelisation of calculation procedures using virtual machines hosted with cloud computing platforms such as Amazon EC2 or Microsoft Azure

5.5.4. Non-domestic building energy calculations [HOM-006]

Dynamic energy simulation is not proposed for non-domestic buildings. The heterogeneity of built forms and activities in the UK non-domestic stock makes estimation of diurnal load profiles and their validation against real-world data using bottom-up engineering methods an intractable problem within the scope of the EnergyPath^{TM} Design Tool project.

The UK does have a certified ISO13790 model for non-domestic buildings, which is known as SBEM (Simplified Building Energy Model). Despite its name however, SBEM is an intensely data heavy model that requires data that are not available from national datasets and which are for all intents and purposes impossible to infer or cross reference against at an address level with any degree of precision from third party sources.

As a result, a statistical approach (rather than a bottom-up engineering led approach) is proposed for estimating energy non-domestic demand in the HOM. The HOM will estimate non-domestic energy demand by matching individual non-domestic archetypes against high level annual kWh/m^2 benchmarks, derived from studies such as the Sheffield Hallam Four Towns Longitudinal study, CIBSE TM46, CIBSE Guide F, and their future equivalents. This follows best available practice in building services engineering for estimating demand at a concept design stage. The per m^2 data is be combined with the primary floor area data and non-domestic building categorisation data to create a unique estimate for each building in a local area as described in section 6.5.1.

The accuracy achieved by benchmarking energy consumption on an area basis is dependent on the quality of the input data, which are derived from empirical observations. Existing datasets are widely acknowledged to have shortcomings relating to the age of the collected data and the number of observations. This is an area for ongoing research and the HOM data architecture will be flexible enough to accommodate new information as it becomes available.

Statistical data used for energy profiling of non-domestic buildings is held by a number of bodies such as utility companies and specialist consulting firms, a number of which are closely affiliated with the ETI and partner organisations participating in the development of the EnergyPath^{TM} Design Tool.

The above approach does not preclude the use of more realistic demand profile data as this becomes available from survey data within a given LA area, particularly for non-domestic public properties.

5.5.5. Building archetype retrofit / upgrade costs [HOM-007]

The costs of building archetype retrofit will be determined under the following process:

- Retrofit transition mapping will take place after the initial filtering of non-viable archetypes (e.g. no 2-storey bungalows) but prior to mathematical problem dimension reduction to the key material characteristics as described in section 5.5.2.

- Retrofit transition mapping is envisaged as an automated process which will be used to determine which starting archetypes can be changed into which other archetypes in the database. This is a one-to-many mapping problem. Key principles will include:

- Separate mapping of interventions in different retrofit elements, such as wall insulation, loft insulation, window type, heating system type.
- Building fabric elements are presumed to have their thermal performance improved rather than degraded e.g. single glazing would optionally be replaced with double glazing but not the reverse.

- Retrofit costs from the ETI Optimising Thermal Efficiency of Existing Housing (OTEoEH) project will be used to determine the range of costs for each intervention (the OTEoEH project database gives upper and lower boundaries for individual component elements), and separates material and labour costs on a per unit/area basis).
  - The data itself does not provide any indication of economies of scale, but it would be possible to assign scalars to reflect e.g. potential reductions in labour costs from large-scale roll-out and use the proposed functionality within the POM to reflect this within the pathway trade-offs (see section 8.7)

In the event that the mapping problem reaches any computational limits then prioritisation of retrofit options can be informed by discussions with individual project Local Authority partners.

5.5.6. Final archetype catalogue [HOM-008]

The final output from the HOM is an archetype catalogue dataset, with each line representing a different option for the POM to choose from. The final archetype catalogue will be comprised of:

- A large number of physical combinations of different building characteristics (archetypes), with statistically insignificant (where the significance level is user-defined) and impossible combinations of characteristics removed
- For each line, final energy demand by vector (gas, electricity, network-supplied hot water) and load profile shapes for a number of characteristic days and operating profiles. These will have been determined for each archetype through dynamic simulation following a mathematical dimensional reduction process. This means that in some cases, buildings with different characteristics may show identical estimated profiles, as the differences introduced by their physical characteristics do not result in mathematically significant differences in load shape.
- Within-day data on non-heat related electricity loads, such as EV charging profiles or PV supply profiles, will be stored in HOM.
- A list of technically viable archetype transitions (which archetypes can be transformed into which other archetypes) alongside associated costs.

5.6. Output validation checks

A number of output validation checks are proposed to be carried out on the HOM outputs as part of the application of the tool itself:

- Comparing ISO13790 outputs against the same properties modelled in other ISO13790-compliant packages, such as the Macro Distributed Energy (Macro-DE) housing sub-model

- Comparing dynamic outputs against observed load profiles where available.

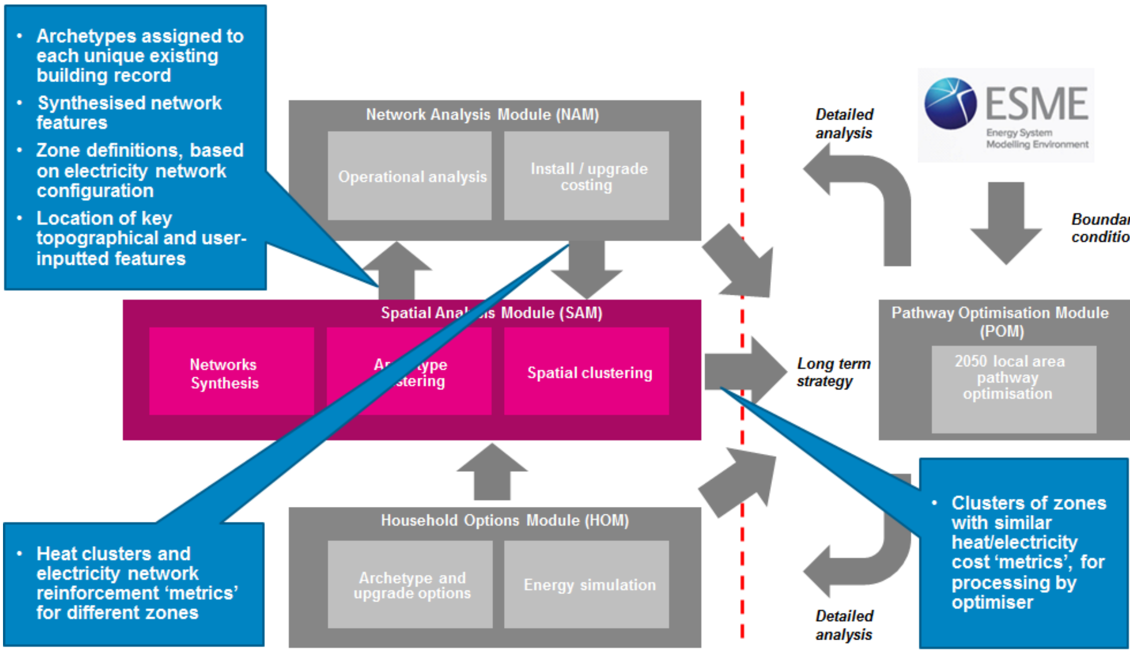
6. Spatial Analysis Module (SAM)

6.1. Overview

The role of the Spatial Analysis Module (SAM) is to create a detailed representation of the local area that contains sufficient spatial data to enable the Networks Analysis Module (NAM) to assess the costs and feasibility of potential network reinforcements and new build, and the Pathway Optimisation Module (POM) to make well-informed choices on the optimum energy strategy for the area and elements within it (e.g. buildings, networks etc.). As part of its internal processes the SAM needs to be able to simplify the spatial representation to make the optimisation process tractable.

In order to provide the Networks Analysis Module (NAM) with the inputs it requires, the SAM is designed to develop geographic information system (GIS)-based layers containing detailed spatial and topographical information about the local area, synthesise a representation of the electricity distribution network serving the area (and possible routes for new network topology) and assign building archetypes (defined in the Household Options Module (HOM)) to each individual building address in the area.

Figure 6-1 SAM context



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. The outputs of the NAM are then combined with the detailed spatial information and used to aggregate basic 'zones' (e.g. individual streets) into homogenous 'clusters' where the constituent zones are spatially contiguous and have similar network upgrade costs.

The spatial information associated with each of these clusters and the number of each building archetype contained within each of them are then passed onto the Pathway Optimisation Module (POM) for whole system energy cost optimisation.

In addition to existing spatial data, the SAM also provides the interface for the user to enter future spatial data (i.e. that will occur as part of a scenario) or future options that could be utilised or impact on the pathway design. This could include, but are not restricted to geothermal or biomass availability, conservation areas restricting archetype upgrades, possible sites for large heat source sites or embedded generation, sites for potential new build housing, etc. This information is stored in the GIS layers and passed to POM for the pathway analysis.

The main steps undertaken in the SAM are summarised below and described in more detail in subsequent sections.

Assign building archetypes to addresses

The SAM uses address level data on the location and type of buildings and combines it with other data sources containing detailed building attribute characteristics to develop a detailed stock database that defines the key attributes of every residential and non-domestic building in the area. This information is used to match individual building addresses to their nearest representative archetypes defined in the HOM.

Synthesise a representation of the electricity distribution network

Where real data on the spatial layout of the LV distribution network, i.e. the network of LV feeders, is not available from local Distribution Network Operators (DNOs), the SAM will synthesise a representation of the network based on available data. GIS layers containing individual building locations and road network data are combined with data on the location of distribution substations (which it is assumed the DNO will provide) and used by the SAM to synthesise a representation of the existing low voltage electricity distribution network, using standard GIS software network analysis packages. The detailed topographical information of the network, including lengths of individual low voltage feeders and the connectivity of buildings to different distribution substations, is sent to the NAM for heat and electricity network upgrade cost calculations.

The process of generating the above network topology also creates information on the possible routes for new networks (e.g. district heat networks).

Generate clusters

The NAM uses heat network costs to cluster, i.e. spatially group, individual ‘zones’ (defined by the highest geographical granularity for network cost calculations e.g. individual feeders). This creates several clusters in the LA, where each cluster represents either an area suitable for a district heat network or a potential ‘representative component’ of a district heating network (the POM can decide if and how to connect these clusters together as part of an overall heat network). The list of remaining un-clustered zones and the electricity network cost data is passed by the NAM to the SAM. The SAM uses standard GIS spatial clustering tools to combine zones which are spatially contiguous and have similar network upgrade costs into aggregate clusters. These clusters and the spatial information of all the zones within them are then passed to the POM.

Additional local information

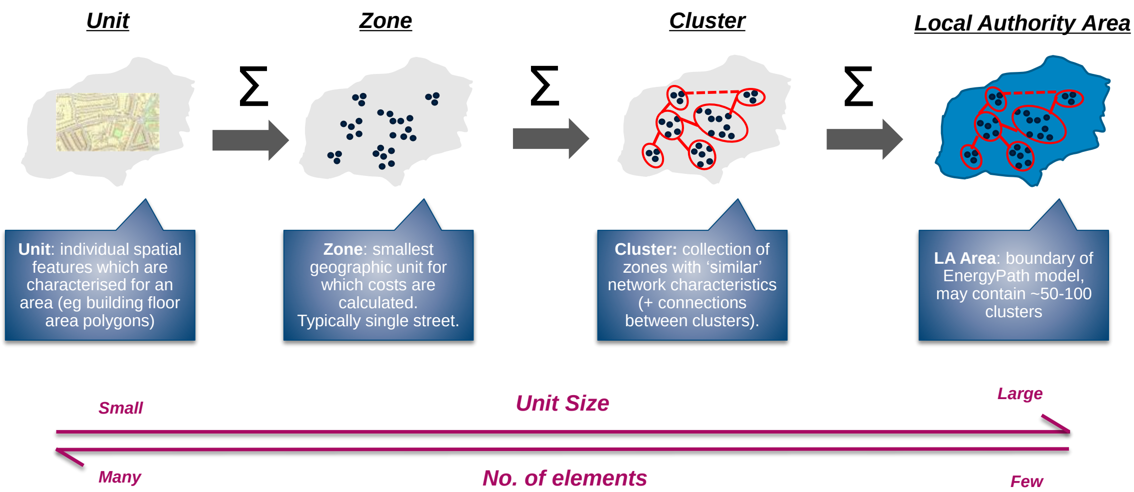
The SAM also saves additional information about other local features of the energy system as defined by the user using external data sources (e.g. planning permission for new builds and embedded generation in a LA). This could represent existing features, those which are already planned for a given future timeperiod, or a set of constraints on possible options (e.g. a new CHP can only be located in zone X or Y, but Z due to planning restrictions). This additional spatial data is sent to POM for inclusion in the pathway analysis.

6.1.1. Key spatial elements

The key aim of the SAM is to collect and synthesise detailed data on buildings, topology and resources at a zone level. The starting point is a 'unit' or an individually identifiable energy system feature within a local area (e.g. each individual building floor area polygon will be mapped using OS data). A 'zone' represents the highest geographical granularity at which detailed network cost calculations are performed in the NAM, typically a street which contains a number of building units.

The higher the geographic granularity, the more accurate will be the network upgrade cost calculations. However, increasing granularity results in larger numbers of costed zones in a LA area and in increasing computational processing demand by the POM as it optimises for every zone. Therefore these zones are aggregated into clusters and represented by aggregate network upgrade costs, in order to ensure computationally feasible optimisation within the POM.

Figure 6-2 Definition of zone and cluster within a local authority area



6.1.2. Zone definition

Zones represent the highest geographical granularity at which building, network, resource and topographical information is collected to be passed to the NAM. The NAM uses data on the network connectivity and zone characteristics (e.g. list of unique LV feeders, buildings connected to each and their distances to each feeder/substation) to develop network options.

In order to facilitate the accurate definition of network upgrade costs in the NAM, zones are based on the smallest component of the network, which is aimed to be an individual street level

LV feeder as indicated in the diagram below. However, it is important for the SAM to reflect how these zones are connected together at high levels of aggregation such that the overall network topology features are respected within any spatial analysis (see section 7.3.5 for further discussion):

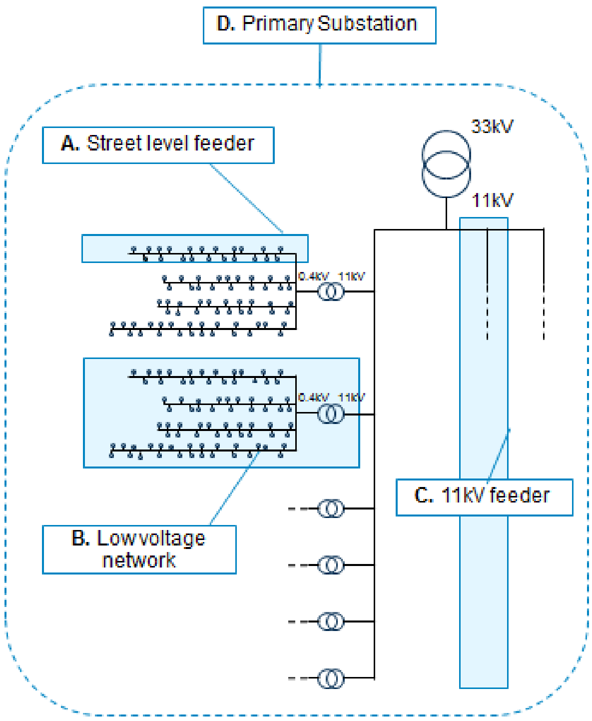
A. Street level feeder that connect buildings on roads to the nearest distribution substation;

B. Low voltage network (Distribution substations that have several LV feeders)

C. 11kV feeder (HV feeders that connect several distribution substations to the nearest primary substation);

D. Primary substation that has several HV feeders.

Figure 6-3 Identified zone building blocks levels reflecting electricity network connectivity



6.1.3. Cluster definition

Clusters represent grouped zones that are spatially contiguous, have homogenous network upgrade costs and are connected to the same electricity network component i.e. LV feeders connected to the same distribution substation, distribution substations (and all of their LV feeders) connected to the same HV feeder or HV feeders (and all of their distribution substations) connected to the same primary substation. The purpose of clustering is to create fewer spatial zones that have homogenous network upgrade costs, in order to minimize the computational burden on the POM.

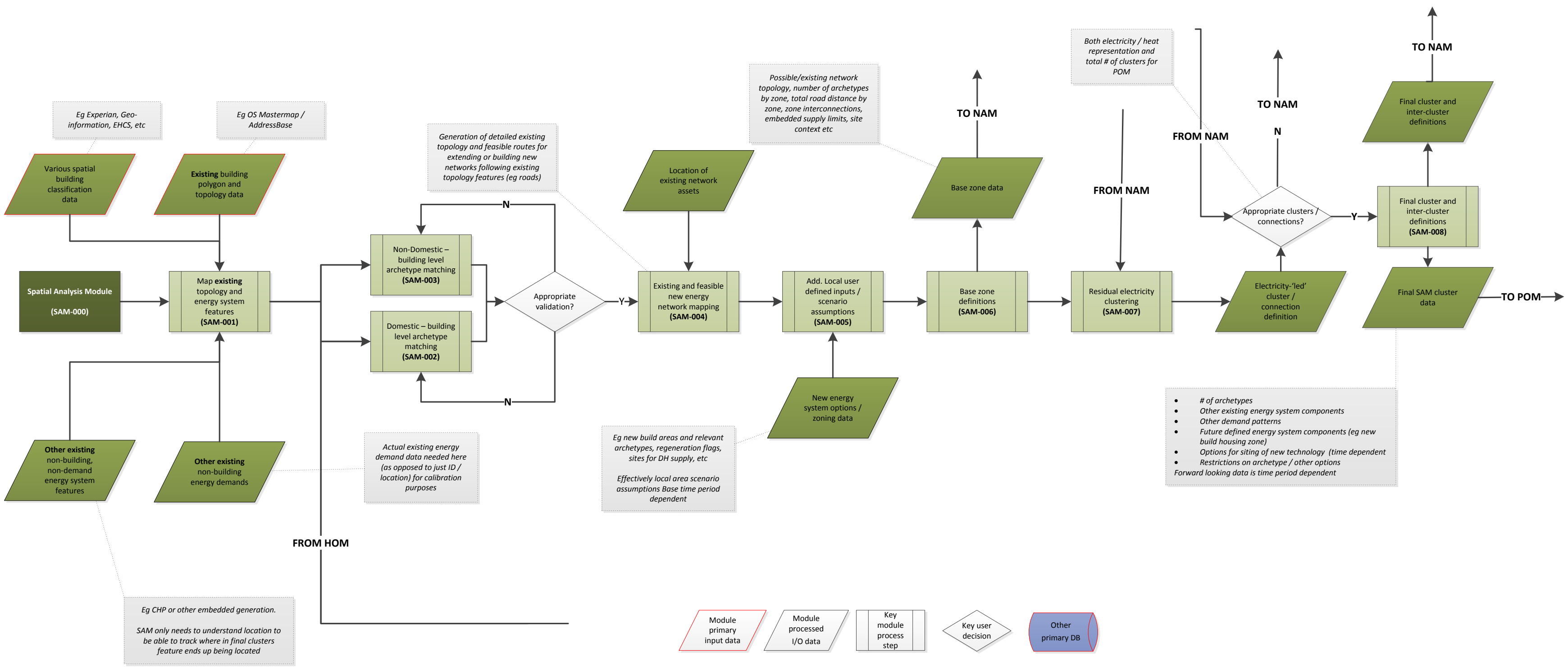
The clustering of zones is carried out in two steps.

- Initially, the NAM clusters zones based on the heat network costs to create clusters that minimise the heat delivery cost. This creates several clusters in the LA, where each cluster represents either an area suitable for a district heat network or a potential ‘representative component’ of a district heating network that could be connected with other clusters. See section 7.5.5.
- The remaining unused zones are then clustered by the SAM using standard spatial clustering algorithms in GIS software to create additional clusters based on homogenous electricity network upgrade cost.

6.1.4. Module diagram

The diagram below shows the key logic / process steps for the SAM, the equivalent for the full tool covering all modules is shown in section 12.

Figure 6-4 Key SAM logic / process steps



6.2. Key outputs

The main output of the SAM is a list of clusters with homogenous network upgrade cost functions. Simple locational information associated with each of these clusters, the number of each building archetype contained within each cluster and any technology suitability flags associated with the local topography or user-inputted data, are also passed onto the Pathway Optimisation Module (POM) for whole system energy cost optimisation.

In order to enable the NAM to carry out its functions, the SAM also produces a series of intermediate outputs, as listed below.

6.2.1. Building archetype database

The SAM outputs a record associated with each individual address (both domestic and non-domestic) within the LA area, with a building archetype assigned to each address, referencing the building archetypes defined in the HOM.

6.2.2. Synthesised energy network

Electricity network

SAM outputs a representation of the local electricity distribution network to the NAM. Where real-life data on the layout of the electricity network is not available from local DNOs, synthetic data on the network layout is generated within the SAM. Specific outputs generated for each building within the LA area include:

- ID of nearest LV feeder
- ID of nearest distribution substation
- Distance from nearest distribution substation
- ID of HV feeder
- ID of connected primary substation

NAM uses this information to generate electricity network upgrade costs curves.

Gas network

SAM provides information on the potential extension of existing gas networks. This involves calculating the distances of off gas buildings to the nearest on gas sites along the road network. These lengths are then used within NAM to cost the gas network extension.

Heat network

SAM provides information on the length of roads within each zone, the archetypes associated with each road and the distance of each building to its adjacent road. This data is used in the NAM to help cost the installation of new heat network.

6.2.3. User-inputted spatial topology, resource and future planning

SAM creates a GIS layer with topology data for the LA. It also acts as the interface to store data on local resource availability (e.g. geothermal, biomass, embedded generation), future new build developments and additional local characteristics (e.g. conservation zones) based on external data provided by the user (e.g. loading GIS or excel based files from new development plans, proposed power generation, etc.). This data is passed to both NAM / POM, either in the form of flags associated with each cluster (e.g. altered technology suitability), or as individual elements for analysis.

6.2.4. Zone definitions

This is based on the smallest component of the electricity network to be costed in the NAM, as well as the outputs of the local electricity distribution network synthesis, the SAM outputs zone definitions assigning a zone ID to each component of the synthesised electricity distribution network.

The SAM would also provide further zone context information which could be used for the network costing within the NAM, such as the site context (e.g. see section 7.3.4) via a measure of building density per hectare.

Table 6-1 Output data summary

ID	Data field	Destination	Purpose	Data granularity	Uncertainty parameters
1	Final building archetypes	NAM/ POM	NAM peak demand estimation for load flow modelling POM building archetype conversion choices and supply / demand balancing	Locational information, archetype, distance to nearest feeder and substation (including IDs for both) and zone ID for each building within the LA area.	-
2	Synthesised electricity network	NAM	Provides data required by the NAM to generate electricity and heat network upgrade cost functions	Locational information, as well as Individual LV and HV feeder ID and length, and IDs for each additional network component, including LV and HV substations (and which feeders they are connected to).	-
3	Gas network extension	NAM	Provides data required by the NAM to generate cost for connecting off gas buildings to the gas grid	Distance of individual buildings to the gas grid along the road network	-
4	Heat network	NAM	Provides data required by the NAM to generate cost for installation of heat networks	Locational information and length of roads in individual zones, the archetypes associated with each road and the distance of individual buildings to the adjacent road	-
5	User-inputted data on building stock constraints	POM	Modifying the technical potential for building upgrades based on constraints such as conservation areas, listed heritage sites etc.	Locational information provided for individual buildings or for selected geographical areas using GIS layers	-
6	User-inputted data on resource availability	POM	Determining the suitability of technology implementation based on local resources e.g. biomass, geothermal sites etc.	Locational information about the geographic extent of local resources, provided using GIS layers or excel based databases	-
7	User-inputted data on future development plans	POM	Providing the spatial and temporal information for future infrastructure plans e.g. new build, network extension, embedded generation etc	Locational information provided using GIS layers for the LA and temporal information provided as a separate input	-
8	User-inputted data on existing embedded generation	POM	Determining the net loads on the network by POM	Locational information on the type and size of generation using GIS layers or Excel databases for the LA	-
9	Zone definitions	NAM	Provides data required by NAM to perform initial heat network based clustering	The zone definition is provided in the form of a zone ID assigned to each electricity distribution network component. Secondary outputs will include simple locational information, the total road lengths within each	-

				zone and the archetypes associated with each road.	
10	Cluster definitions	NAM/POM	Needed by NAM to derive the aggregated electricity and heat network cost metrics. POM will use the clusters to perform system level optimisation	Simple cluster locational information, building archetype numbers within each cluster and any technology suitability flags associated with either local topography or user-inputted data.	-
11	Zone site context	NAM	Used to adjust the base network cost options e.g. if they are urban, suburban, rural, London. Based on a measure of building density per hectare	Zone level information	-

6.3. Key functional requirements

The key functional requirement of the SAM is to process a series of complex local building, electricity network and geographical datasets, and to combine these into a limited number (e.g. less than 100) of simplified 'clusters', containing all the information that is required for the POM to carry out its optimisation calculations. To do this the SAM is closely linked to the NAM with two-way input/output data flows, given the importance of spatial topology on the understanding of network reinforcement and new build options

In order to do this, it must be able to perform a number of processing steps, as defined below.

6.3.1. Understand the location of existing building archetypes

In order for the POM to make choices as to the balance between network upgrades or new build options and individual building upgrades, the SAM must output simplified data on the type and location of buildings within the LA area.

In order to do this and due to the lack of detailed building-level data for every building within a LA area, the SAM must combine detailed GIS-based local area representations, with other local or higher-level datasets on the attributes of the building stock in order to assign a distribution of standard building archetypes (defined within the HOM) to the address level data. The output is a database of standard building archetypes, for which the POM can reference the HOM in order to understand relevant energy demand patterns.

The process used to develop the database of standard building archetypes must be flexible enough to accommodate existing or future local known building information.

6.3.2. Synthesise electricity network features where data is limited

In order for the NAM to evaluate network upgrade functions within the LA area, the SAM must first synthesise a representation of the existing local electricity network – to include an understanding of which buildings are connected to which low voltage (LV) and high voltage (HV) substations and the length of feeders between them.

The analysis must use available data for inputs, including basic information provided by local DNOs (e.g. substation location) and known local topographical features (e.g. building locations, road locations and lengths, etc.).

Again a flexible approach must be adopted to allow for the potential availability of detailed local electricity distribution network data, provided by local DNOs.

6.3.3. Allow the user to input a wide range of spatially-related data

The SAM must allow users to input location-specific data, which will affect the choices made in the POM. This may include:

- Data that could affect technology suitability for a given area, e.g. local biomass/geothermal resource, conservation areas, etc.
- Data that may prevent the implementation of certain network/building upgrade decisions, e.g. topographical barriers including rivers, etc.

- Data that could inform a decision on the development of future network extensions, e.g. areas ear-marked for future development, or known planned developments.

6.3.4. Zone definitions

In order for the NAM to perform a heat network cluster analysis and to define electricity network reinforcement cost functions based on contiguous sections of the electricity distribution network, the SAM must first define a series of zones, which are used to break the LA area into manageable areas. Zone definitions must be passed to the NAM to enable the network costing analysis.

6.3.5. Clustering zones

In order to constrain the computational burden for the optimisation in the POM, the SAM must combine any zones which have not been clustered in the NAM, into fewer, larger clusters of zones with homogenous electricity network upgrade costs and electricity network connectivity (this does not imply that the entire cluster is upgraded at the same time). These clusters and the key information describing them (e.g. building archetypes, simple locational information, etc.) must then be passed to the POM.

6.4. Key inputs

Primary inputs

SAM uses several key external datasets to develop an enriched database of existing and future spatial elements in the LA. These datasets are listed below:

1. Ordinance survey (OS) datasets of the local authority including:
    - Mastermap layer containing detailed topographical data
    - AddressBase Premium with location of individual buildings/dwellings and their associated Unique Property Reference Number (UPRN)
    - Integrated Transport Network (ITN) layer with detailed information about the existing road network
2. GeoInformation Group address level database for existing domestic buildings
3. Experian database at address/street/postcode/postal sector level for domestic buildings
4. English Housing Survey (EHS), Living in Wales (LiW) and Scottish Housing Condition Survey (SHCS) dataset for residential sector at government office region (GOR) level
5. Valuation Office Agency (VOA) database for non-domestic buildings at address/postcode level
6. DNO data on the local electricity distribution network, including:
    - Location for distribution and primary substations

- Number of LV and HV feeders per distribution and primary substations

7. Local user defined inputs, to include, but not restricted to:

    - Building stock or other technology constraints used to alter the technical potential for these options. This can be based on inputs related to individual or groups of buildings, or geographic constraints loaded as GIS layers, such as conservation areas, areas not suitable for ground source heat pump deployment, etc.

    - Resource availability used to determine the suitability of technology implementation based on local resources e.g. biomass, geothermal sites etc., loaded as GIS layers representing each local resource.

    - Existing embedded generation used to determine the net loads on the network by POM, loaded as a database of known locations.

    - Future development plans that provide the spatial and temporal information for future infrastructure plans e.g. new build sites, network extension plans, future embedded generation, etc. These can be loaded as a database of known future developments and locations, or as GIS layers representing new build sites, with associated characteristics.

Secondary inputs

SAM relies on a number of databases that are generated internally within the EnergyPath™ Design Tool, in different sub-modules. These are:

1. List of archetypes and their attribute characteristics as defined in the HOM

2. The electricity network upgrade costs for individual zones as calculated in the NAM

3. The list of zones not used in the initial clusters based on the heating network in the NAM.

The tables below describe what is technically required within the tool and how it would be used. The relative importance of the primary data inputs and other factors such as costs and licensing restrictions, particularly where they must be purchased from an external provider, are discussed in deliverable (D4) Data Acquisition Plan.

Table 6-2 Input data

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Building footprint area, address, UPRN, building use,[19]	Primary	Matching the building to external databases with building attribute characteristics	Defined for residence/hereditament at address level	-	All addresses in the UK	GIS layers of OS MasterMap, AddressBase Premium and potentially Points of Interest	-
2	Domestic Age, Non-domestic type, UPRN/address,	Primary	Identifying the basic building attributes	Defined for all residential dwellings at address level Similarly for non-domestic buildings	Not available for all residential dwellings or non-domestic buildings	All residential addresses in the UK	GeoInformation Group	Only provides around 70% coverage in UK, mainly concentrated in urban areas, more limited coverage for non-domestic buildings
3	Age, tenure, type, location, GOR, fuel, address	Primary	Identifying the basic building attributes	Defined for all residential dwellings at address/post code/postal sector level	May not be available at address level and therefore a distribution at aggregated level will be applied	All residential addresses in the UK	Experian ConsumerView data	Accuracy at address level may be limited, this would require applying aggregated post code/postal sector level data
4	Age, tenure, type, location, fuel, glazing, loft type/insulation, wall type/insulation, heating technology, number of storeys	Primary	Identifying the detailed building attributes	Defined for sample residential dwellings at GOR level	Data available for sample buildings (around 16k for England)	All GOR in the England, Wales and Scotland	English Housing Survey , Living in Wales , Scottish Housing Condition Survey)	Building characteristic distribution in the sample data at GOR level is applied at address level
5	Wall and loft insulation	Primary	Identifying the technical potential for retrofit measures	Defined as averaged insulation levels per street	Based on insulation installation data till 2012	UK	Energy Savings Trust Home Analytics	
6	Type of non-domestic	Primary	Classification of non-	Defined for all	-	Whole of UK	Valuation Office	-

[19] Ordnance Survey experimental height dataset may become available in future to help determine non-domestic building height

	building and floor area by each storey		domestic buildings for matching with archetypes in HOM	hereditaments at address/postcode level			Agency	
7	Topographical features e.g. rivers, greenspaces,	Primary	Identifying local features that act as constraints or impact cost of network upgrade	Defined at individual zone level	-	Whole of UK	GIS layer of OS Mastermap	-
8	Other local area conditions	Primary	Other local area conditions provided by bespoke GIS datasets (e.g. ground conditions which may impact on network cost scalars applied in NAM)	GIS dataset referenced at zone level for NAM	-	LA	LA	-
9	Existing heat source, embedded generation, etc.	Primary	Allow POM to optimise heat network	Defined for all known sites in the LA	-	LA	User	Relevant GIS layer or excel based files may not be available
10	Planned new-build developments, new heat sources, etc.	Primary	Allow POM to consider new developments in its optimisation of available options	Defined for all planned sites in the LA	-	LA	User	Relevant GIS layer or excel based files may not be available
11	Listed/heritage buildings in conservation area, geothermal, biomass resources	Primary	Identifying geographical constraints for application of retrofit/technology measures to building stock	Defined for all known constraint areas within the LA	-	LA	User	Relevant GIS layer or excel based files may not be available
12	Road node coordinates, road length	Primary	Synthesising the LV electricity network	Defined for all roads adjacent to existing buildings	-	All roads adjacent to buildings across the UK	GIS layer of OS Integrated Transport Network	
13	Distribution substation coordinate/address	Primary	Synthesising the LV electricity network	All substations defined in the LA under consideration	-	Defined by the user for the area under study	User/DNO	Relevant GIS layer or excel based files may not be available
14	Primary substation coordinate/address	Primary	Synthesising the HV electricity network	All substations defined in the LA under consideration	-	Defined by the user for the area under study	User/DNO/OS points of interest	Relevant GIS layer or excel based files may not be available

								GIS layer of OS Points of interest does not provide full coverage of UK
15	Zones already used in the heat network based clusters	Secondary	Clustering of unused zones	All zones in the LA	-	LA	NAM	-
16	Electricity and heat network cost metric	Secondary	Clustering of unused zones	All zones in the LA	-	LA	NAM	-
17	Building archetype definition	Secondary	Matching UPRN to an existing archetype from HOM	All feasible combinations of attribute characteristics	-	Residential and non-domestic sector	HOM	-
18	Gas distribution network data	Primary	Topology used to inform gas network extension, upgrade, conversion costs	GIS layer	-	GB	National Grid operated DNOs from MAP viewer software, other DNOs provide directly	-

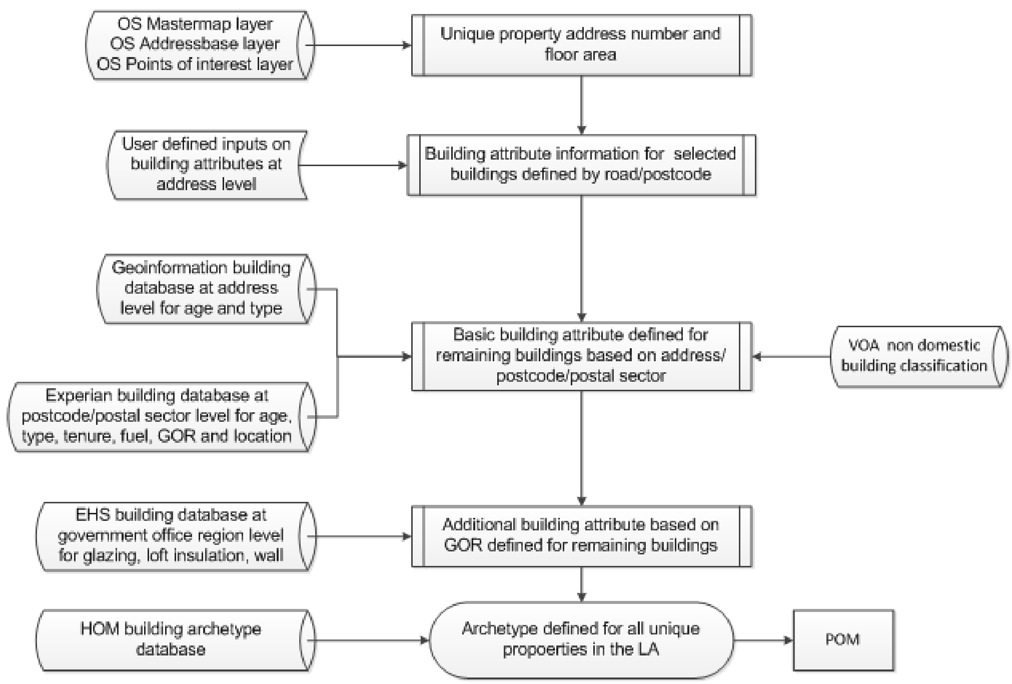
6.5. Key SAM logic / process steps

6.5.1. Map existing topology and energy system features [SAM-001]

The first step is to map and configure all of the available spatial data for the local area as outlined in section 6.4 including GIS layers for building polygons, road networks, etc.

6.5.2. Building-level attribute definition and archetype matching [SAM-002/3]

Figure 6-5 Building archetype analysis process steps



6.5.2.1. Residential sector [SAM-002]

SAM will use OS AddressBase Premium and MasterMap layers to populate a database of existing building stock with UPRN, address, floor area and building type. This provides detailed spatial information for all the residential buildings in a LA. At this stage the user is able to define attribute characteristics (include building age, type, tenure, location, fuel, glazing type, loft insulation, wall construction, number of storeys, heating technology, etc.) at building level, where known, for residential buildings in a LA. This can be done by loading an external dataset in the required format or by manually entering the attribute values for every known building.

The following sections assume that certain national level datasets are available, however, the underlying process of matching is generic such that data at different levels of spatial aggregation, covering different building archetype characteristics could be included within the process.

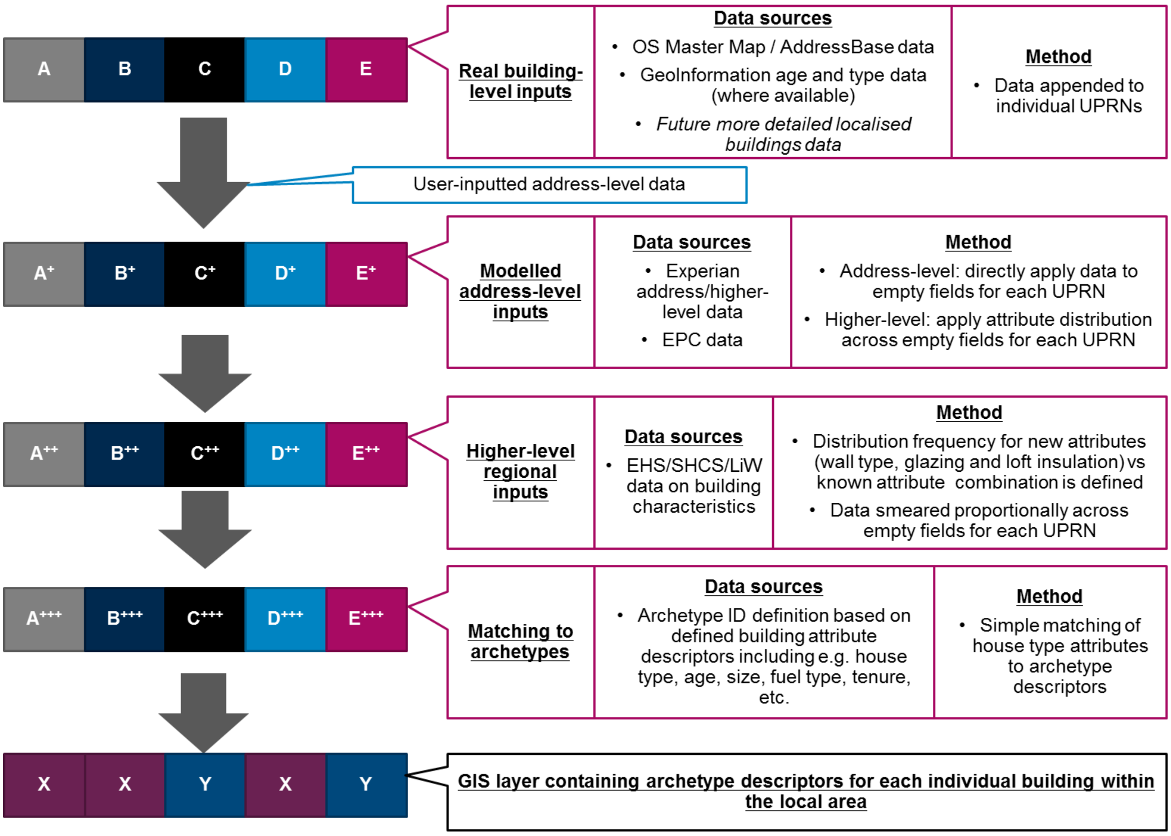
Next, GeoInformation is used to populate age and type attributes for residential buildings where available and where these have not already been defined by the user. This is done by matching the UPRN between the OS and GeoInformation datasets. For buildings whose age and type attribute is not defined by the user and is also not covered in the GeoInformation dataset, Experian modelled data is used to populate the age and type.

With the age and type now defined for all the buildings in the LA, Experian data (which may be at a dwelling/street/postcode level) is used to assign the tenure, location, GOR and fuel for all the dwellings not already defined by the user. If higher level (i.e. not address-level) data is used here, the distribution of these additional attributes is applied proportionally across the buildings in the area in question. *Any user-inputted data that has already been provided at an address level has precedence over these aggregated higher level datasets*. The distribution of additional attributes of tenure, location and fuel are adjusted based on the user inputs so that the final attribute of the buildings i.e. the combination of age, type, tenure, location and fuel matches the e.g. Experian dataset at the more aggregated level.

Lastly, additional data on glazing level, loft insulation, wall construction, number of storeys and heating technology is populated for all the dwellings, by applying GOR-level distributions from the EHS, LiW and SHCS datasets, for all dwellings where these attributes have not been defined by the user. This distribution of wall type, glazing and loft insulation are applied based on the defined attributes of age, type, tenure, location and fuel.

Through the above steps, the detailed building attributes for all residential dwellings are defined at address level. These attributes are based on datasets available at address level as well as higher geographic granularity (e.g. postcode/postal sector/ GOR). The distributions of the attribute characteristics still conform to the aggregated distribution, based on the already populated attributes at lower geographic granularity e.g. GOR level distribution of wall type, glazing and loft insulation combinations is applied based on the known attributes of age, type, tenure, location and fuel type. Since the chosen attributes and the characteristic values are matched with those used in HOM, the resulting archetypal classification of residential building is always available in HOM database.

Figure 6-6 **Schematic of the step-wise approach to defining attributes for each individual UPRN and assigning residential building archetypes**

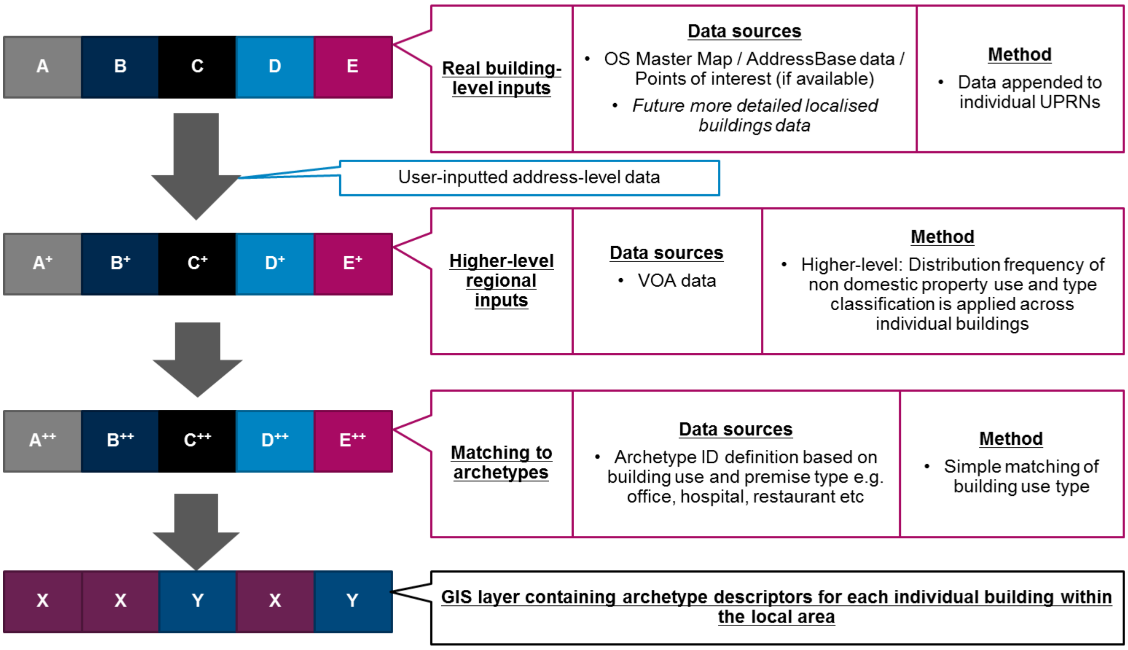


6.5.2.2. Non domestic sector [SAM-002]

The SAM will use OS AddressBase Premium and Mastermap layers to populate an existing building stock database with UPRN, address, floor area, building type and use and use (high-level non-domestic classification, or more detailed data where available).

At this stage the user is allowed to define non-domestic use type at building level, where known, for non-domestic buildings in a LA. This can be done by loading an external dataset in the required format or by manually entering the use type for every known building. The VOA dataset is then used at address level to define the building use classification of all buildings (e.g. office, hotel, school etc) not already defined by the user (this may be augmented by using OS Points of Interest data to provide more detailed usage categories). Finally, these classifications are then matched to the archetypes defined in HOM for non-domestic sector based on building use type.

Figure 6-7 Schematic of the step-wise approach to defining attributes for each individual UPRN and assigning non domestic building archetypes



6.5.2.3. *Validation*

SAM can perform a validation check to ensure the archetypal classification process results in an accurate representation of energy use in the LA based on historical metered energy demands at various levels of geographic granularity.

The implied annual energy (heating and electricity) demand for the LA will be calculated based on the archetypal classification of the buildings and the energy demand defined in the HOM. This could then be compared with the annual energy consumption (and peak demand on characteristic days, if available) obtained from the actual metered data to inform the user of the difference between implied and actual metered data.

DECC currently release this data at LSOA level for domestic consumption and MSOA level for non-domestic, but discussions with DECC during Stage 1 have indicated that it may be possible to obtain this at postcode level, subject to commercial confidentiality in the non-domestic sector and resource to support the processing of the data.

If this difference is significant, this would signal to the user that additional, more localised data would benefit the analysis, in order to more accurately reflect the condition of the local building stock, e.g. LA-specific data on the distribution of glazing, wall type and loft insulation in the local area, to better reflect the LA rather than averaged data at GOR level from EHS/LiW/SHCS. This should result in a better representation of LA building stock, matching more closely with the metered consumption data.

If additional data is not available and the difference between the available metered data and modelled annual energy consumption is higher than the maximum user defined limit, the model could apply a calibration factor to energy consumption in each area for which actual aggregated consumption data is known e.g. by suggesting a reduction/increase in the heating and electricity demand as stated in HOM for the relevant archetypes or by changing the frequency of the archetypes.

6.5.3. Network topology synthesis [SAM-004]

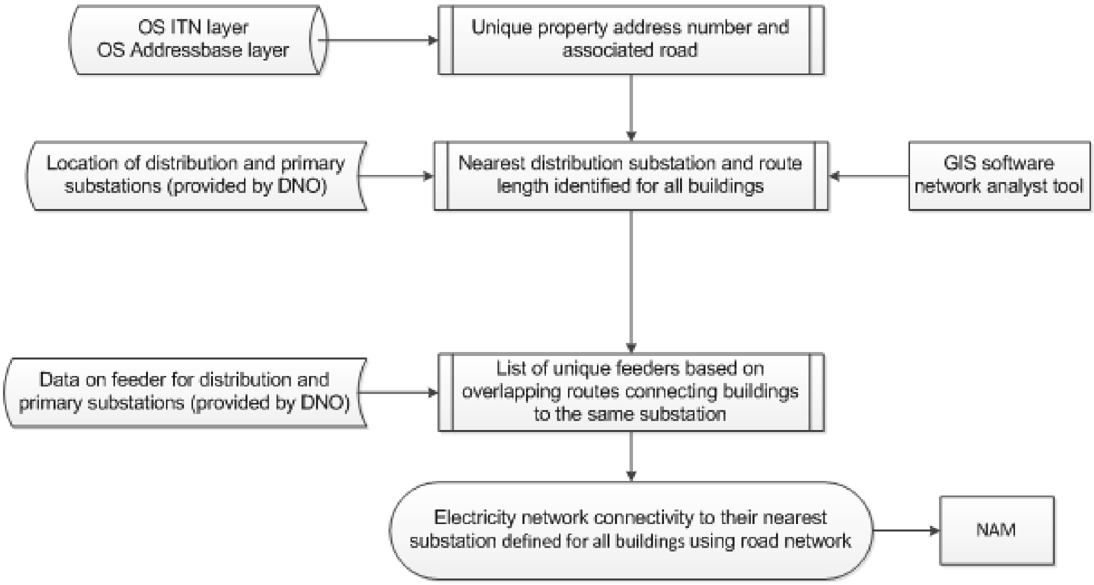
SAM synthesises an electricity network for the LA, if DNO data on electricity network is not available, by using localised building, road and DNO data. Firstly, OS AddressBase Premium and ITN GIS layers are used in GIS software to populate building nodes and the road network. GIS functionality is used to assign each building to its adjacent road using spatial analysis. At this stage the user is required to load DNO data on the location of existing distribution and primary substations as a GIS layer or excel based file.

If the DNO data on location of substations is not available, OS Points of Interest GIS layer will be used to populate the locations if available, however this does not provide complete coverage of substations in the whole of UK. If substation location is still not available, the number of substations will be based on the total connections to buildings in the LA (using typical connections per substation figure) and they will be assumed to be located equidistant in the LA under consideration.

In addition to the location of the substations, the data set from DNO may also contain information on the number of LV and HV feeders per each distribution and primary substation respectively. The user can also define the locations for future distribution and primary substations that will form the network for new build developments or are based on existing network development plans of DNO.

The GIS network analyst tool is next used to identify the closest distribution substation to each individual building, based on the route along the roads taken to connect to the nearest distribution substation.

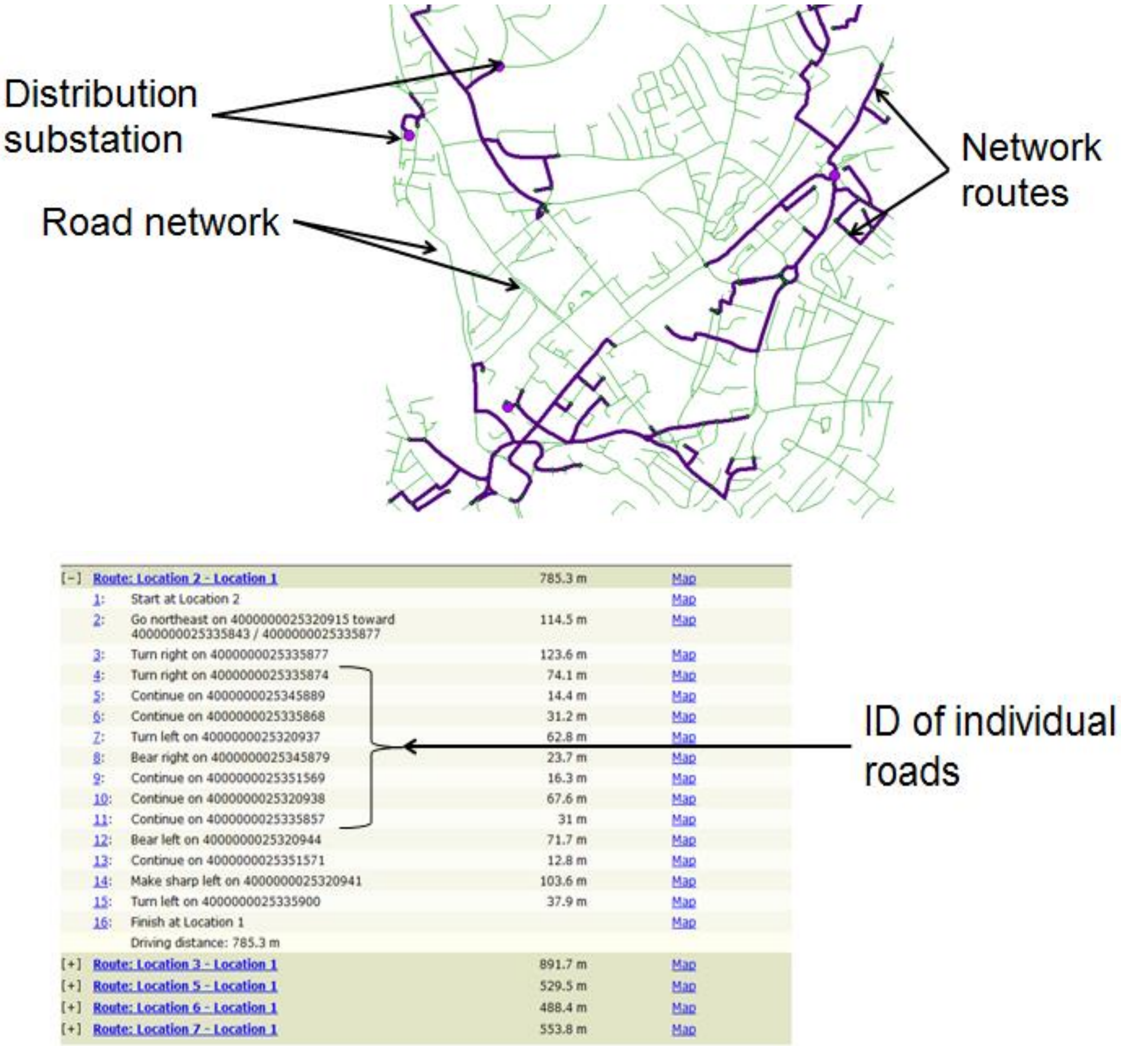
Figure 6-8 Electricity network synthesis process steps



This information, along with DNO data on the total LV feeders if available, is processed to determine the unique LV feeders for each distribution substation, the associated residential and non-domestic buildings and their distances to the nearest distribution substation. This process is then repeated to connect the distribution substations to their nearest primary substation via the road network routes and the route information, along with DNO data on the total HV feeders if available, is processed to determine the unique HV feeders for each primary substation, the associated distribution substations and their distances to the nearest primary substation. In addition data on the roads and the length is also calculated for heat network calculations within the NAM.

Where existing network topology data is already available from the DNO e.g. detailed GIS layers or simple distribution and feeder connectivity data for individual buildings, the dataset would be used for the known buildings and network synthesis only applied for areas not already covered in the DNO data.

Figure 6-9 Example of network synthesis



Thus a detailed database is created which is passed to the NAM for electricity network cost calculation. This contains the following information each building in the LA:

1. Building archetype ID
2. Adjacent road ID
3. Distance from road
4. Connected LV feeder ID
5. Connected distribution substation ID
6. Distance to distribution substation
7. Connected HV feeder ID
8. Connected primary substation ID

9. Distance of distribution substation to primary substation

In addition, the SAM also generates other spatial information necessary for the NAM to cost other network reinforcement of new build

- Information on the potential extension of existing gas networks. This involves calculating the distances of off gas buildings to the nearest on gas sites along the road network. These lengths are then used within NAM to cost the gas network extension.
- The length of roads within each zone, the archetypes associated with each road and the distance of each building to its adjacent road. This data is used in the NAM to help cost the installation of new heat network.

6.5.4. Additional local user-inputted data [SAM-005]

The SAM allows the user to input location-specific data, which will affect the choices made in the POM. This is done through a simple user interface, allowing the user to upload GIS layers or add other information directly to the existing GIS datasets. Data inputted may include the following information:

ID	Data	Purpose
1	Building or technology option constraints	*Determining the technical potential for building upgrade based on constraints like conservation areas, listed heritage sites etc, or restricting possible sites for new embedded generation*
2	Resource availability	*Determining the suitability of technology implementation based on local resources e.g. biomass, geothermal sites etc*
3	Future development plans	*Providing the spatial and temporal information for definitive future infrastructure plans e.g. new build, network extension, embedded generation etc*
4	Existing embedded generation	*Determining the net loads on the network by NAM / POM*

6.5.5. Zone definitions [SAM-006]

The SAM defines zones, which are used to break the LA area into manageable areas for the NAM. Zones are the smallest building blocks and are defined using electricity network characteristics. SAM assigns zone IDs to each component of the electricity network, as described above. We have identified 4 possible building blocks, increasing in size, however, the aim it to define the zones at the most granular street level feeders.

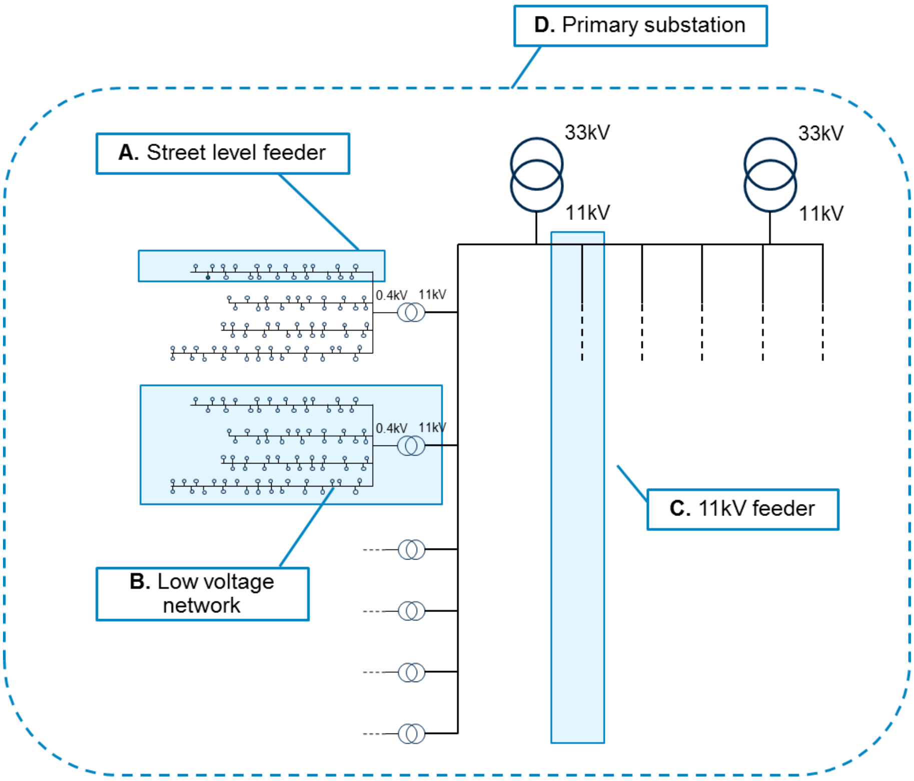
A. Street level feeder

B. Distribution substation

C. 11kV feeder

D. Primary substation

Figure 10   Identified building blocks



As part of the zone definition process the SAM will need to assign energy system features to the appropriate zone, for example, each unique building on a street level feeder would be assigned to that zone.

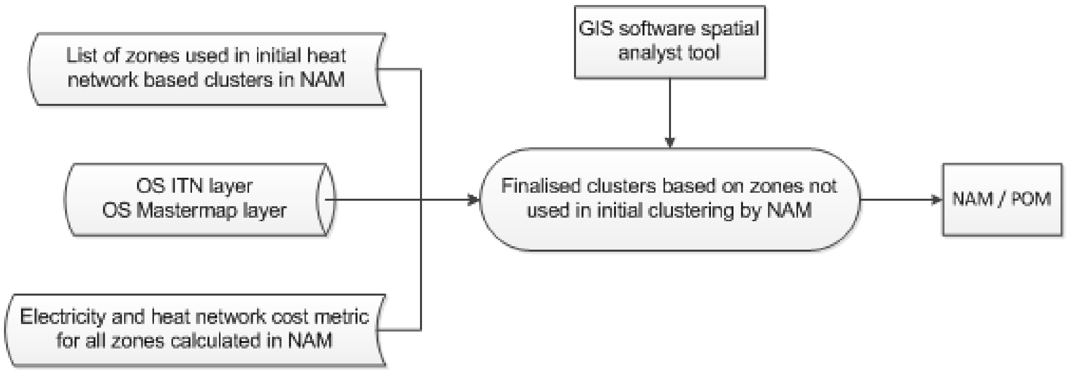
For some energy system features such as an existing (or potential site) for a large energy centre or non-domestic building (e.g. hospital), it may be more appropriate to assign them to their own zone, with an inter-zone connection to the relevant part of the electricity network. Flags would be assigned to specific larger-scale features that should automatically be assigned their own zone.

Finally, there will be a series of other energy system features such as a potential biomass resource which are identified as their own polygon within the spatial map and effectively become their own zone as part of the definition process.

6.5.6. Clustering [SAM-007/8]

Clustering is performed in the EnergyPath™ Design Tool to produce spatially contiguous regions with homogenous network upgrade costs.

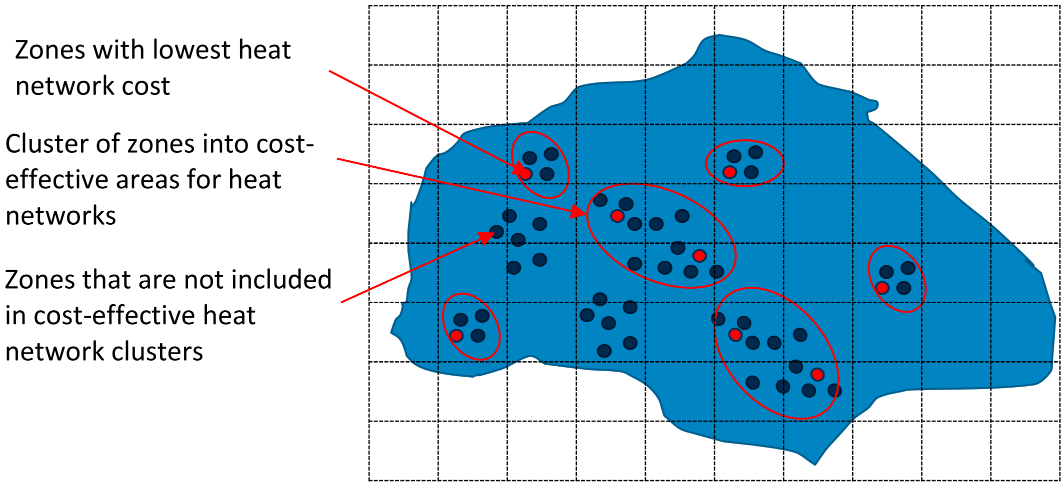
Figure 6-11 Clustering of zones process steps

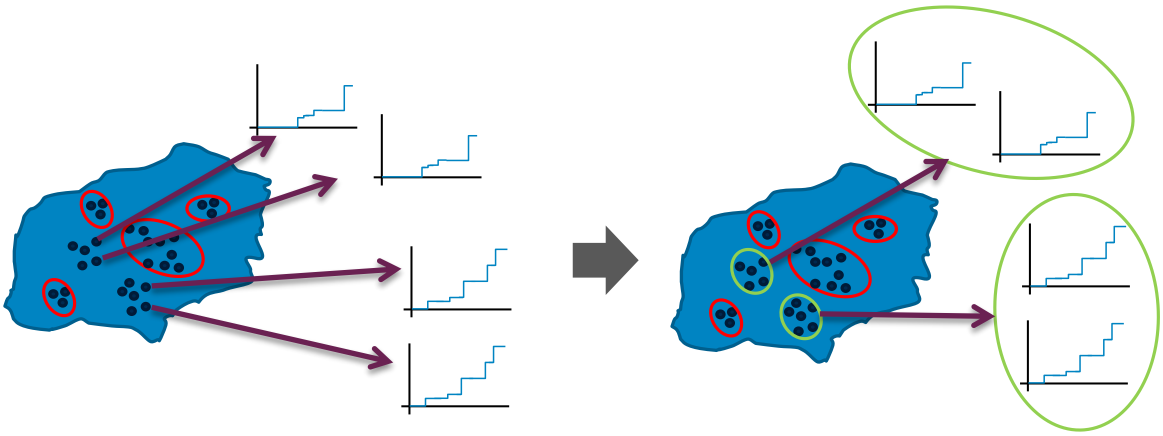


An initial 'heat-led' clustering process is performed in the NAM. This aims to balance the representation of both heat and electricity network costs within the cluster boundary, but is led via the exploration of possible heat network configurations, given the greater degrees of freedom for a new build network. This process is described further in section 7.3.5.

The zones that are not included within initial heat network cluster boundaries are then passed by the NAM to the SAM. These zones, along with their network cost metrics (based on the cost curves from the NAM), are then clustered by the SAM using GIS spatial clustering algorithms to generate clusters of spatially contiguous zones that have homogenous electricity network upgrade costs (see section 7.5.6 for further discussion of cost metrics).

Figure 6-12 Overview of zones and clusters within local area





These additional clusters and the list of zones within each cluster are then passed back to the NAM to generate aggregated network cost curves for each cluster. The clusters and their aggregate cost curves are then used by POM for the system level energy cost optimisation.

6.6. Proof of concept activity

Three key proof of concept activities have been performed to test the feasibility of proposed methodologies for the SAM. These are summarized briefly below, please see the Appendix section 13.1 for further details.

6.6.1. Household archetype definition and matching

A semi-automated Excel-based tool has been developed to create a detailed building stock attribute database. This tool utilises sample data for Exeter available from the OS website to populate a building database with UPRN and addresses. The tool contains the functionality for the user to define detailed attribute characteristics at address level for all known buildings. The tool then uses dummy data from GeoInformation and Experian to define the basic attributes of age, tenure, type, location, GOR and fuel. EHS data is then used to smear known glazing, wall type and loft insulation data across the buildings. Using the detailed attribute definition, an archetype is assigned to each UPRN. Illustrative output of this prototyping model is shown below:

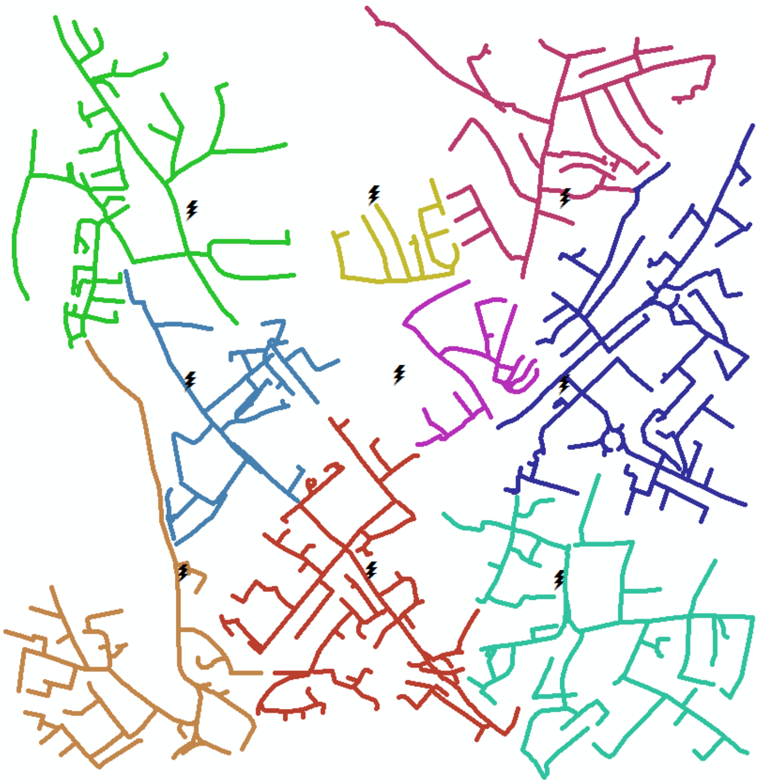
Figure 6-13 Archetype definition based on the attribute characteristics

				Building attributes							
buildingNumber	throughfareName	postcode	Final archetype ID	Fuel	Location	Size	Tenure	Age	Wall	Glazing	Loft insulation
117	OKEHAMPTON ROAD	EX4 1ER	16	Gas	Rural	Flat	Social	Recent	CWU	double	0-50 mm
119	OKEHAMPTON ROAD	EX4 1ER	21	Gas	Rural	Flat	Social	Recent	SW	single	0-50 mm
1	NAPIER TERRACE	EX4 3EZ	21	Gas	Rural	Flat	Social	Recent	SW	single	0-50 mm
1	ELDERTREE GARDENS	EX4 4DE	22	Gas	Rural	Flat	Social	Recent	SW	single	51-100 mm
37	LONGBROOK STREET	EX4 6AW	22	Gas	Rural	Flat	Social	Recent	SW	single	51-100 mm
0	BLACKBOY ROAD	EX4 6ST	26	Gas	Rural	Flat	Social	Recent	SW	double	0-50 mm
37	PRIORY ROAD	EX4 7AP	28	Gas	Rural	Flat	Social	Recent	SW	double	101-150 mm
39	PRIORY ROAD	EX4 7AP	29	Gas	Rural	Flat	Social	Recent	SW	double	151-200 mm
1	WELLSWOOD GARDENS	EX4 1RH	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
10	WELLSWOOD GARDENS	EX4 1RH	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
30	LOWER NORTH STREET	EX4 3EU	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
0	NEW BRIDGE STREET	EX4 3JW	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
11	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm
17	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm
18	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm

6.6.2. Network analysis

GIS software functionality has been explored to test the possibility of automated synthesis of electricity network. This involves using the OS AddressBase premium, ITN and MasterMap layers in GIS and using the network analyst tool functionality to calculate the routes linking each building to its nearest distribution substation and each distribution substation to the nearest primary substation along the existing road network. The detailed routing information is then processed to derive a list of feeders, their connectivity to buildings and the distances of each building to the nearest distribution substation. An illustrative example of a synthesized network developed using the prototype model is shown below.

Figure 6-14 Synthesised electricity network developed using OS sample data for Exeter



6.6.3. Cluster analysis

GIS software functionality has been explored to test the possibility of automated clustering of zones into aggregated clusters. This involves using the OS AddressBase premium, ITN and MasterMap layers to populate the detailed spatial information for a LA. In addition illustrative data on electricity network cost, to be provided by NAM, is used. The GIS clustering algorithm is used to generate clusters that contain spatially contiguous zones with homogenous network cost. *The ArcGIS clustering algorithms only allows spatial clustering based on a single metric*.

This single numeric value has to define:

1. The network upgrade cost as calculated in the NAM

2. The network connectivity constraints to ensure that zones are not clustered together in ways that inappropriate breach the network topology

This is achieved by processing the cost metric into a single meta-variable based on the network connectivity (e.g. the cost metric is adjusted to reflect the connection of zones to distribution substation, HV feeder and Primary substation) and use in initial clusters (all zones used in initial clusters get the same cost metric of their relevant cluster ID).

The result of the connectivity constraints is that they take precedence over the cost metric in determining final clusters, thereby reducing the number of possible clustering outcomes. Thus the initial clustering is determined by the network connectivity while the cost metric(s) from the NAM is/are a secondary clustering parameter. Thus all the distribution stations on a HV feeder need to be clustered before two HV feeders can be clustered together. The final clustering results in individual zones being grouped together in spatially contiguous clusters that have similar network upgrade cost metrics and obey network connectivity clustering rules.

6.7. Output validation checks

The final outputs on the building archetype database from SAM will be validated against the input datasets and any other external datasets available to the user. This involves comparing the detailed building stock archetype database against datasets used as inputs at various level of geographic granularity (e.g. Experian, EHS, LiW, SHCS etc). This would require comparing aggregated stock attribute distribution with the input dataset to verify:

1. The stock levels of glazing, loft and wall type with the GOR level data

2. The stock levels of age, tenure, type, location and fuel with the Experian data at post code or postal sector level

In addition, there will be functionality to compare the distribution of the individual attribute characteristics or their combination within an archetype at a user defined geographic granularity with additional datasets or survey results carried out at LA or sub LA level.

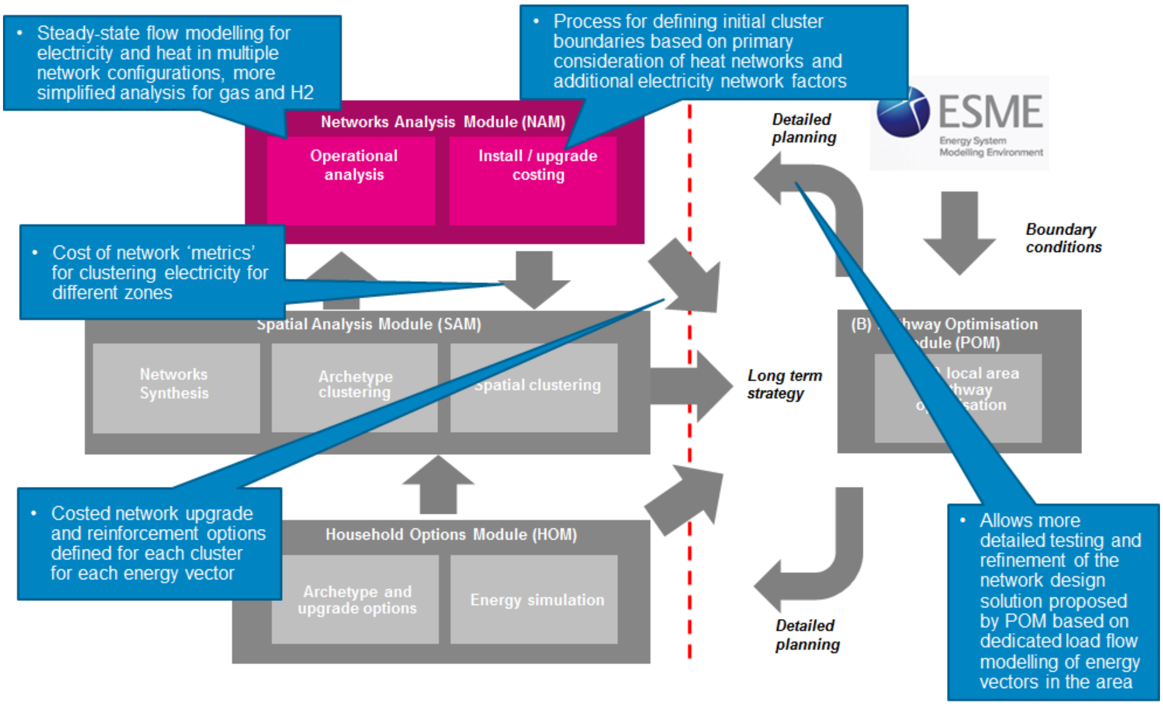
7. Network Analysis Module (NAM)

7.1. Overview

The Network Analysis Module (NAM) provides the detailed analysis for quantifying potential network upgrade and new build options (their characteristics and costs) associated with an increase in energy peak demand requirements across the energy vectors under consideration: electricity, heat, gas and hydrogen. These options are then passed to the POM as part of the pathway analysis.

The key interactions between NAM and the other modules are shown in Figure 7-1 below.

Figure 7-1 NAM context



As shown in the Figure above, NAM contains two key components, each with a distinct set of requirements:

- The 'operational analysis and heat-led clustering' tool, which is mainly responsible for performing load flow studies to simulate the operation of energy networks (existing and potential) in the area under investigation, as well as for performing the 'heat-led' clustering process to define initial cluster boundaries. The NAM interacts closely with the SAM and together they help to define the final area cluster boundaries, trying to ensure the cost of network options are as representative as possible across all vectors within a cluster, and that the total number of clusters is tractable within the POM optimisation.

- The 'test options generator and costing tool', which is responsible for generating costed options for all zones, clusters and inter-cluster connections within the area under investigation. These cost functions are based on network upgrade or new-build cost functions depending on the reinforcement package to be considered;

In summary, the main requirements from NAM are the following:

- Materiality of load flow modelling: NAM should perform load flow modelling studies across the different energy vectors to ensure accurate allowable capacity limits are used when calculating reinforcement cost functions. All network constraints which can materially impact load flows (and hence costs) across each energy vector must be taken into account.

- Ensuring that existing and potential energy networks are accurately represented: NAM should ensure that all existing and potential energy networks are represented to a sufficiently high accuracy in terms of their physical and geographical characteristics, together with an accurate spatial and temporal representation of potential demand and generation options.

- Producing cost functions for a range of network reinforcement options: NAM should output cost functions for different network options for each energy vector. Cost functions must compare network reinforcement costs and maximum allowable capacity (load/generation) which will then be passed to POM for system optimisation. Moreover, NAM should also ensure that the number of network options tested by the scenario costing tool, as well as the number of network reinforcement options sent to POM, must be kept tractable computationally.

- Materiality of network component costs: Similarly to the point above, NAM should ensure that the cost functions produced for a given network reinforcement option accurately reflect the costs of component upgrades or additions, accounting for local factors (geographical, economic etc.) which may affect these

- Producing 'heat-led' clusters with specific boundary constraints: A key responsibility of NAM is to cluster together zones primarily using heat network characteristics ('heat-led' clustering), however with boundary constraints set by the electricity network and combined with other potential threshold metrics such as size of cluster and electricity network costs:

    - It is important to note that the overarching process used to define clusters is trying to ensure that the network costs within them are as well represented as possible for **all** energy vectors and do not unfairly bias the solution in POM in a particular cluster

- The 'heat-led' nomenclature simply refers to the technical modelling process by which initial clusters are defined and does not imply a preference for a district heating solution in a particular cluster[20]

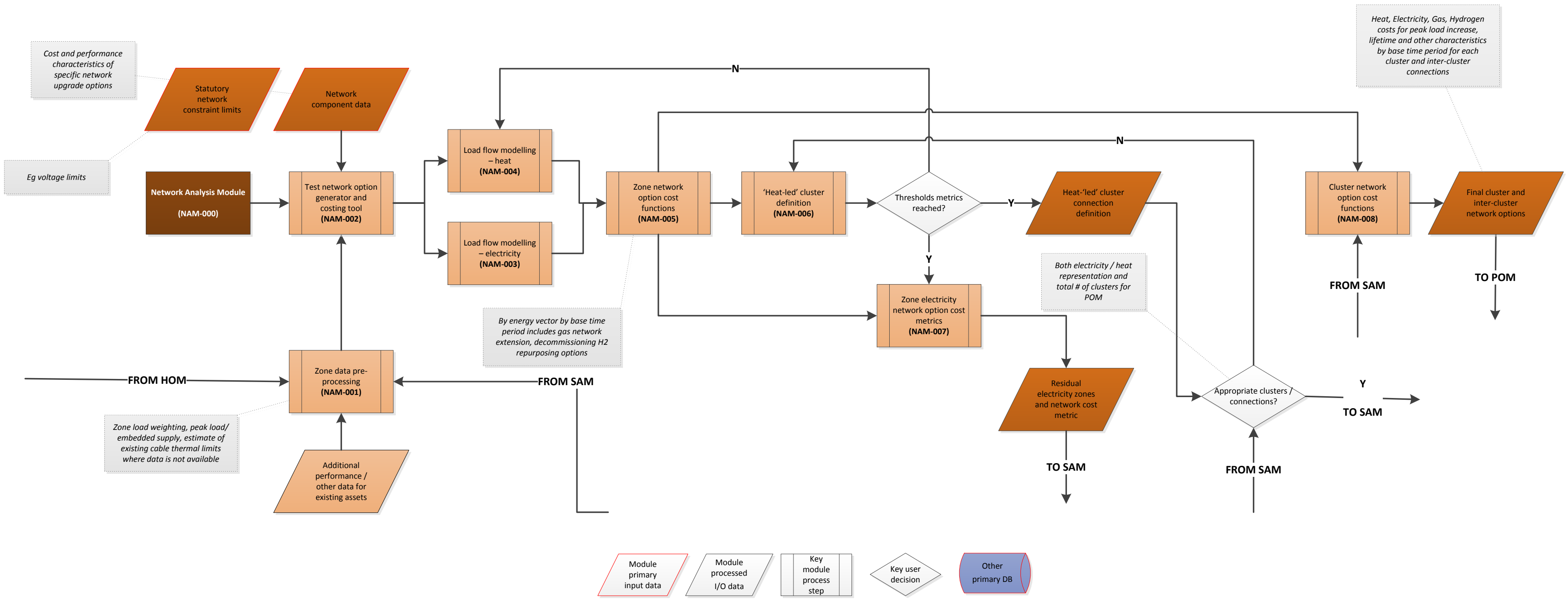
- Detailed testing and refinement of proposed POM design solution: Finally, NAM should allow more detailed testing of the feasibility and refinement of the higher-level network design solution proposed by POM.

7.1.1. Module diagram

The diagram below shows the key logic / process steps for the NAM, the equivalent for the full tool covering all modules is shown in section 12.

[20] This is ultimately a cost-based decision for the POM, or alternatively, the user can force-in particular network designs and test their resilience under a range of sensitivities

Figure 7-2 Key NAM logic / process steps



7.2. Key outputs

This section describes the key outputs from NAM, where they are used in the overall solution and for what purpose.

The key outputs from NAM can broadly be summarised in two categories:

1. NAM outputs related to cluster definitions: The NAM is responsible for defining the final cluster and inter-cluster connections that are used by POM for the optimisation of the overall energy system. Some zones are clustered using heat network characteristics ('heat-led' clustering[21] which takes place in NAM) while others are clustered using electricity network characteristics ('electricity-led' clustering which takes place in SAM). Once the clustering process has been completed, the final cluster and inter-cluster definitions are then sent to POM for the final optimisation process. The final inter-cluster definitions include both existing as well as potential new inter-cluster connections. The following outputs are included in this category:

    a. 'Heat-led' cluster definitions;

    b. Zone cost metrics for the 'electricity-led' clustering process (i.e. simple metrics derived from zone cost functions that are used to cluster together zones with similar electricity network cost functions);

    c. Final cluster and inter-cluster definitions.

2. NAM outputs related to cost functions: The NAM is also responsible for outputting cost functions at the intra-cluster and inter-cluster level for each energy vector. These cost functions compare reinforcement/new build costs and maximum allowable capacity for each reinforcement option and are used by POM for optimisation of the energy system. NAM is also responsible for producing cost functions at the zone level (which are then used for determining intra-cluster cost functions) however zone cost functions are not sent to POM as they are not at a high enough level of aggregation. The following outputs are therefore included in this category:

    a. Intra-cluster cost functions for all energy vectors;

    b. Inter-cluster cost functions for all energy vectors.

NAM outputs are summarised in Table 7-1 below.

[21] Heat-led clustering takes place based on the concentric-circles heat network calculations in NAM, however with boundary constraints set by the electricity network.

Table 7-1 Output data summary

ID	Data field	Destination	Purpose	Data granularity	Uncertainty parameters
1	'Heat-led' cluster definitions	SAM	After the 'heat-led' clustering has taken place in NAM, the NAM is then responsible for sending these cluster definitions to SAM. NAM is also responsible for identifying the zones which have not been clustered under the 'heat-led' approach. These zones are also passed to the SAM for the final 'electricity-led' clustering process.	The 'heat-led' cluster definition includes a list with the following data: - A list with all zones (including buildings and relevant network components within each zone) that have been identified as being part of a 'heat-led' cluster. This includes: ▪ Zone ID; ▪ Cluster ID. - A list with all other zones (including buildings and relevant network components within each zone) that have not been identified as being part of a 'heat-led' cluster – these zones are then clustered in SAM using the 'electricity-led' clustering process; ▪ Building ID (and archetype); ▪ Network component ID; ▪ Zone ID; The cluster definitions are timeperiod independent as they do not change over time	
2	Zone cost metrics for the 'electricity-led' clustering process	SAM	The provision of zone cost metrics to SAM occurs for 'electricity-led' clustering to take place. This is to ensure that zones with similar electricity network cost functions are grouped together.	Zone cost metrics are simple numerical metrics that describe how expensive electricity network reinforcement for a particular zone would be, based on the derived zone cost functions. Zones with similar cost metrics are then grouped into clusters during the 'electricity-led' clustering process (Section 7.5.6).	Min, mode, max for key elements of network cost as per section 7.3.4
3	Intra-cluster cost functions (electricity)	POM	Final intra-cluster cost functions (for electricity) to be sent to POM for the optimisation process	These represent options for allowable load/generation (kW) with associated costs (£) for different electricity network reinforcement options within the cluster – inter-zone meshing options also considered here at the MV level (but not at the LV level). See section 7.5.7. Costs are comprised by: - Annuitised capital expenditure (fixed) for intra-cluster electricity network reinforcement or expansion - Fixed operational and maintenance costs (fixed) for intra-cluster electricity network reinforcement or expansion - Variable operations and maintenance costs (dependent on utilisation and hence final values are calculated in POM) for intra-cluster electricity network	As above

				reinforcement or expansion These cost functions are provided for each timeperiod vintage – i.e. as costs may change in future build years.	
4	Inter-cluster cost functions (electricity)	POM	Final inter-cluster cost functions (for electricity) to be sent to POM for the optimisation process	As above for electricity network meshing options	
5	Intra-cluster cost functions (heat)	POM	Final intra-cluster cost functions (for heat) to be sent to POM for the optimisation process	As above for heat networks [based on pipe size (mm)/temperature (°C)] & cost (£) and allow POM to make optimal decisions with regards to the sizing of the heat distribution network.	
6	Inter-cluster cost functions (heat)	POM	Final inter-cluster cost functions (for heat) to be sent to POM for the optimisation process	As above for heat networks [based on pipe size (mm)/temperature (°C)] & cost (£) and allow POM to make optimal decisions with regards to the sizing of the heat distribution network.	
7	Intra-cluster cost functions (gas)	POM	Final intra-cluster cost functions (for gas) to be sent to POM for the optimisation process	These represent pairs of maximum allowable load (kW) [based on number of customers not already connected to the gas network, and connection size] & cost (£) and allow POM to make an "all-or-nothing" decision on extending the gas network to cover off-gas grid buildings, to decommission it for operational costs savings, or to do nothing. Costs are comprised by: - Annuitised capital expenditure (fixed) for extending the gas network to cover off-gas grid buildings - Fixed operational and maintenance costs (fixed) for extending the gas network to cover off-gas grid buildings - Variable operations and maintenance costs (dependent on utilisation and hence final values are calculated in POM) for extending the gas network to cover off-gas grid buildings - Decommissioning costs for decommissioning the existing gas distribution network - Fixed operational and maintenance cost savings (fixed) for decommissioning the existing gas distribution network - Variable operations and maintenance cost savings (dependent on utilisation and hence final values are calculated in POM) for decommissioning the existing gas distribution network These cost functions are provided by time period.	
8	Inter-cluster cost functions (gas)	POM	Final inter-cluster cost functions (for gas) to be sent to POM for the optimisation process	These represent pairs of maximum allowable load (kW) [based on number of customers not connected to the gas network, and connection size] & cost (£) and allow POM to make optimal decisions with regards to inter-cluster gas network meshing options.	

				Costs are comprised by: - Annuitised capital expenditure (fixed) for inter-cluster gas network meshing - Fixed operational and maintenance costs (fixed) for inter-cluster gas network meshing - Variable operations and maintenance costs (dependent on utilisation and hence final values are calculated in POM) for inter-cluster gas network meshing These cost functions are provided by time period.	
9	Intra-cluster cost functions (hydrogen)	POM	Final intra-cluster cost functions (for hydrogen) to be sent to POM for the optimisation process	These represent pairs of maximum allowable load (kW) [based on number of customers in the network, and connection size] & cost (£) and allow POM to make an "all-or-nothing" decision on repurposing an existing gas distribution network to allow it to be able to transport hydrogen, or to do nothing. Costs are comprised by: - Annuitised capital expenditure (fixed) for repurposing an existing gas distribution network - Fixed operational and maintenance costs (fixed) for repurposing an existing gas distribution network - Variable operations and maintenance costs (dependent on utilisation and hence final values are calculated in POM) for repurposing a gas distribution network These cost functions are provided by time period.	
10	Inter-cluster cost functions (hydrogen)	POM	Final inter-cluster cost functions (for hydrogen) to be sent to POM for the optimisation process	These represent pairs of maximum allowable load (kW) [based on number of customers in the network, and connection size] & cost (£) and allow POM to make optimal decisions with regards to inter-cluster hydrogen network meshing options. Costs are comprised by: - Annuitised capital expenditure (fixed) for inter-cluster hydrogen network meshing - Fixed operational and maintenance costs (fixed) for inter-cluster hydrogen network meshing - Variable operations and maintenance costs (dependent on utilisation and hence final values are calculated in POM) for inter-cluster hydrogen network meshing These cost functions are provided by time period.	

7.3. Key functional requirements

The following sections outline the key functional requirements of the NAM in more detail, *the logic / process required steps to undertake these are outlined in section 7.5.*

7.3.1. Capture most materiality issues within load flow modelling

NAM should be able to perform load flow modelling studies across the different energy vectors to ensure accurate allowable capacity limits are used when calculating cost functions. This is because detailed load flow modelling cannot take place in POM as it needs to be done at a lower level of granularity. These capacity limits are determined by the local supply and demand conditions in the energy network under consideration, as well as by the identified key network constraints. For the majority of networks capacity limits are expected to be set by peak demand conditions, however it is also possible that network reinforcement or new build could be triggered by net exports when embedded supply outstrips demand (for example for a distribution network with very high penetration of small scale PV).

Table 7-2 Key constraints for electricity and heat networks

Network type	Component	Key constraints
Electricity	Network lines (LV, MV)	Voltage, Thermal
Electricity	Transformers (primary, secondary)	Thermal
Electricity	Meshed feeders	Voltage, Thermal
Heat	Network pipes (HT, MT, LT)	Energy delivered, operating temperature, losses
Heat	Pumps	Energy delivered, operating temperature, losses

Electricity network modelling

Electricity network modelling focuses on electrical distribution networks up to (and including) 33kV as network reinforcement at higher voltage levels is likely to be impacted by additional factors that are outside the scope of this project. These factors include demand from large industrial customers, developments of large power stations etc. The baseline evolution of the 33kV network will need to be captured by exogenous scenario assumptions. However, as mentioned in 8.5.2 it will still be necessary within the POM to account for the ability to expand the network if driven by local area conditions.

The following constraints that can impact network reinforcement costs are considered, those which are not considered and the rationale for this are described further below:

- Voltage limits at the 33kV, 11kV and 230/400V level (as well as any other intermediate levels such as possibly 6.6kV for some UK distribution networks);
- Thermal limits of all three-phase and single-phase overhead lines and underground cables in the system;
- Thermal limits of all primary and secondary transformers (likely to be 33/11kV and 11/0.4kV) in the system.

These network constraints can be alleviated using a range of possible network reinforcement options. The following key network reinforcement options are considered by NAM:

- Replace, incremental upgrade or add a new distribution network line (underground cable or overhead line);
- Replace, incremental upgrade or add a new distribution transformer;
- Replacement, upgrade or addition or new meshed feeders[22].
- Distributed generation (including incurring costs to curtail those generators as a means of avoiding network reinforcement);

A range of other options, primarily 'smart' options which affect the level and shape of supply or demand on the network are captured as part of the POM pathway analysis and are assessed in parallel with the above options. They include:

- Demand side response;
- Energy storage management;
- Other smart grid technologies - substation automation and dynamic thermal rating systems[23].

Table 7-3 Network reinforcement options – electricity networks[24]

Option – Electricity networks	Voltage limits	Cable thermal limits	Transformer thermal limits
Replace, incremental upgrade or add a new network line	✓	✓	
Replace, incremental upgrade or add a new transformer			✓
Meshed networks	✓	✓	✓
Distributed generation	✓	✓	✓
Demand side response [in POM]	✓	✓	✓
Energy storage management [in POM]	✓	✓	✓
Other smart grid technologies [in POM]	✓	✓	✓

The focus of the modelling described here is on steady-state conditions, with fault analysis not included as part of the detailed load flow simulations. NAM should, however, attempt to relate

[22] Meshing options are only considered at the 33kV and 11kV level. A mesh can be created by closing one or more open points in the distribution network to resolve the network issue under consideration.
[23] In principle any option can be defined as long as it is possible to create the relevant cost function for the POM.
[24] Note: the load flow model will only consider generic load and generation and the range of load considered will cover existing to extreme conditions

normal running criteria to n-1 contingency conditions. This will be undertaken by reducing the headroom provided by the relevant options to reflect n-1 conditions to ensure the required build meets this standard.

Due to the focus on strategic local area design, only network constraints that can materially impact load flows (and hence costs) across each energy vector will be considered. The focus of the modelling described here is on steady-state conditions, with fault analysis not included as part of the detailed load flow simulations as mentioned above. Similarly, power quality, stability issues and harmonic analysis are also not considered here. The rationale for excluding these factors is outlined in Table 7-4.

Table 7-4 Materiality of electricity constraints that are not modelled in detail

Parameter	Definition	Key factors	Why is it excluded?
Steady-state voltage unbalance	In a three-phase system, voltage unbalance takes place when the magnitudes of phase or line voltages are different, when the phase angles differ from balanced conditions, or both.	– Phase distribution of power demand – Phase distribution of power generation – Power factors – Distribution line impedances	DNOs currently pay little attention to voltage unbalance (governed by Engineering Recommendation P29). For the purposes of the EnergyPath™ Design Tool, it would be very difficult to know which phase existing customers are connected to, as well as to make meaningful assumptions with regards to which phase new customers may connect to.
Fault analysis	In a power system a fault is any abnormal electric current. Design of systems to detect and interrupt faults is the main objective of power system protection and can be an important design consideration.	– Fault currents – Type of fault – Location of fault	In general, fault analysis is more applicable for MV distribution networks (where traditionally the majority of faults have occurred) rather than LV distribution networks. Detailed fault analysis would be required in order to be able to predict whether fault levels in a distribution network have exceeded allowable limits. This would significantly add to the overall modelling complexity, and would also have cost implications in terms of the load flow software required.
Harmonic analysis	Power system harmonics are created by non-linear devices connected to the power system. High levels of power system harmonics can create voltage distortion and power quality problems.	– Non-linear power demand – Non-linear power generation – Distribution line impedances	Detailed harmonic modelling would be required in order to be able to predict whether harmonic distortion levels in a distribution network have exceeded their allowable levels (as governed by Engineering Recommendation G5/4-1). This would significantly add to the overall modelling complexity, and would also have cost implications in terms of the load flow software required. It should also be noted that if the harmonic producing loads are small in relation to total

			load, then harmonics will not present the most limiting network constraint.
Power system stability	The ability of a power system to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact.	– Magnitude of disturbance – Rotor angle stability – Frequency stability – Voltage stability	As with harmonic analysis, detailed stability modelling would be required in order to model the stability of a dynamic distribution network. We feel such analysis is outside the scope of the EnergyPath™ Design Tool and would have significant implications in terms of complexity and cost.

District heat network modelling

District heat network modelling focuses on modelling the heat production facilities, pumps and pipes, storage facilities and connected buildings within the area under investigation. The focus is on thermal modelling rather than detailed hydraulic modelling, with turbulent fluctuations and leakage not considered in the modelling approach. Similarly, it is also assumed that fluid characteristics like density and heat capacity are constant. The rationale for excluding these factors is outlined in Table 7-5.

Table 7-5 Materiality of heat network constraints that are not modelled in detail

Parameter	Definition	Key factors	Why is it excluded?
Turbulent fluctuations	In fluid dynamics, turbulence refers to a flow regime characterised by chaotic property changes. The operating conditions of a district heat network are not perfectly stable and involve thermal and hydraulic transient regimes, in particular, temperature waves combined with temperature fluctuations	– Thermal transient conditions – Hydraulic transient conditions – Sudden temperature changes	The dynamic behaviour of district heat networks under varying operating conditions can be important in determining the safety and cost efficiency of the overall solution. Literature review suggests that simplified approaches provide a good approximation of overall operating conditions, particularly for representing the time delay in a system. Inaccuracies can occur, however, when trying to predict the temperature value at a specific time during the emergence of the temperature changes. We feel such analysis, however, is outside the scope of the EnergyPath™ tool and would have significant implications in terms of complexity and cost.
Energy theft & leakage	Theft (for gas and electricity mainly) refers to energy losses that cannot be accounted for. Leakage refers to fugitive emissions from leaks in old pipes.	– Energy theft – Leakage	These typically form a very small part of overall energy delivered. Whilst they are very hard to model, some approximations may be used to take them into account.
Pressure drop caused by valves and fittings	A parameter affecting pressure drop in piping systems is pressure loss in the fittings and valves of the system. Typically, the calculated head loss caused by the valves and fittings within a pipe segment is expressed as an additional length of pipe that is added to the actual length of pipe when calculating pressure drop.	– Resistance coefficients and flow coefficients of valves and fittings	For piping systems within production facilities, the pressure drop through fittings and valves can be greater than that through the straight run of pipe itself. In long pipeline systems, however, which are the focus of the EnergyPath™ Design Tool, the pressure drop through fittings and valves can typically be ignored with minimal loss in accuracy.

The selection of pipes and pumps in a district heat network has important cost implications, but also important operational implications in terms of mass flow rates, velocity, operating temperature and heat losses, flow pressure drop and pump electrical energy consumption which need to be taken into account by the load flow model.

The aim of NAM is to determine, in terms of engineering feasibility, the minimum size of pipes and pumps that can be selected subject to a set of network constraints (e.g. typical maximum engineering velocities) and then pass the economic trade-offs to the POM, where the final size selection takes place. The following components are considered directly in the NAM as part of the network feasibility testing:

- Low Temperature (LT) pipes (temperature range 70°C to 85°C), Medium Temperature (MT) pipes (pressurised, temperature range 80°C to 115°C), High Temperature (HT) pipes (pressurised, temperature range 120°C to 200°C)
- District heating pumps.

The wider set of components, which make up the overall district heat network are considered in the POM (alongside the above network options produced by the NAM)

- Heat production plant (CHP or heat-only)[25] and overall temperature of the network
- Heat accumulators and storage facilities
- Wider topology of the network (radial, meshed, ringed) by connecting clusters within the local area together

The final network design is selected by the POM in accordance with the fundamental economic principles of heat network design shown the figure below, as some of the key factors that must be considered (e.g. the effective cost of the electricity to run the pumps) are only available as part of the pathway analysis undertaken by the POM.

It is important to note that temperature difference (ΔT) is assumed to be constant for all of the network infrastructure within a cluster as a whole, but the overarching temperature level can be selected by the POM given the heat supply options. As a result, this implies an interdependent set of network choices that needs to be taken into account when designing the overall process.

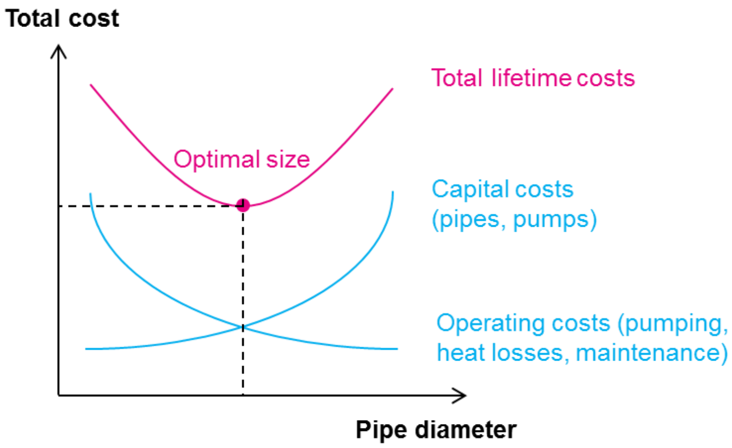
To a large degree, ΔT will depend on mass flow rates and on the velocity of the district heat network, which in turn will depend on the size of the pipes/pumps selected. Mass flow rates and velocity will in turn determine the flow pressure drop in the network, which is typically measured for the most remotely connected customer. The most common term used is “flow pressure drop per unit length” or “target pressure loss” (TPL) which refers to the combined pressure loss of the supply and return piping. Typically pumps and pipes are also selected based on maximum TPL for the most remote consumer.

As Figure 7-3 shows, smaller pipes have lower capital costs but higher operating costs (including losses). Maximising ΔT at design conditions can reduce total investment – this is because

[25] For the purposes of the NAM modelling a free, unlimited heat source is considered for the purpose of testing network feasibility

reducing flow temperatures can significantly reduce heat losses in the figure. Similarly, however, it is important to ensure that savings in heat losses are not offset by increased pumping costs.

Figure 7-3 Economic design principle for district heat networks



In terms of overall network topologies, two-line radial systems are predominantly used for small and medium-sized district heat networks. Meshed or ringed networks allow for incorporation of several heating stations and offer reliability benefits however due to longer and larger pipelines the investment cost is also typically greater. The choice between different network topologies (radial, ring or meshed) also takes place in POM based on the economic trade-offs described above.

In summary, the following key heat network options are modelled as outlined in the table below.

Table 7-6 Network reinforcement options – district heat networks

Option – Heat networks	Energy delivered	Operating temperature	Losses
Replace, incremental upgrade or add a new network pipe	✓	✓	✓
Replace, incremental upgrade or add a new pump	✓	✓	✓
Heat production plant and temperature of the overall network [in POM]	✓	✓	✓
Heat storage [in POM]	✓	✓	✓
District heat network topology (radial, ring, meshed) [in POM]	✓	✓	✓

Gas and hydrogen network modelling

NAM does not undertake detailed flow modelling for gas and hydrogen network to determine their suitability for a given cluster. Instead, the proposed approach is to represent these options, along with their associated costs and benefits, in POM by considering:

- Extension of the gas distribution network to cover off-gas grid buildings;
- Decommissioning some parts of the existing gas distribution network;

- Repurposing an existing gas distribution network to allow it to be able to transport hydrogen;

Dedicated new build hydrogen networks are excluded for Stage 2, with gas network repurposing considered as the only potential option, given the expected costs of new network, but these could be included in future.

Given the binary nature of decisions on decommissioning[26] an existing gas distribution network or repurposing it for transporting hydrogen, these are evaluated for a given cluster as an "all-or-nothing" decision and is not considered at the zone level. It would also be possible to reflect other combinations, such as the impact on decommissioning costs if a heat network is being installed at the same time.

7.3.2. Ensuring that energy networks are accurately represented

NAM should ensure that all existing and potential energy networks are represented to a sufficiently high accuracy in terms of their technical, physical and spatial characteristics (using the detailed topographical and siting information from the SAM), together with an accurate spatial and temporal representation of *the potential* demand and generation options in the area under investigation.

All network data are provided to NAM by SAM as explained in Section 6.2. In practice this involves the provision of the following key data types:

1. Distribution of archetypes across individual buildings in local area;

2. Synthesised physical layout of local electricity distribution networks (and data to assess future potential network layouts), based on available data and GIS algorithms;

3. Existing and planned discrete sites for e.g. CHP deployment, new developments;

This information is used for two key purposes:

- Ensuring that all spatial information about the area under investigation is taken into account – this is important so that all parameters that can materially affect network component costs are included (Section 7.3.4);

- Ensuring that all physical information about the area under investigation is taken into account – this is important so that suitable boundaries (for example building blocks, rivers, cliffs, railways etc.) are used during the clustering process but also in order to capture load flows from production to consumption as accurately as possible;

It is important to note that the level of profile of embedded supply and demand (and associated peak demand or maximum export conditions) are ultimately determined in the POM, and are contrasted against the range of network reinforcement new / build options from the NAM.

---

[26] Decommissioning a gas network will result in capital costs but also in operating cost savings and these economic trade-offs will be simulated in POM.

Hence the NAM needs to model the full spectrum of load / supply conditions across the different parts of the network, such that a set of network options are available for any conditions that could be observed in the POM. The NAM can use data from the SAM/HOM to help inform the boundary of these conditions – e.g. it will know the number of each building archetype by zone, and the maximum heat load without any diversity effect to estimate the highest peak load that could reasonably be seen (e.g. if all buildings used electric resistive heating).

In practice, the network modelling will require engagement with the local relevant parties according to the specific energy vector modelled (e.g. DNOs for electricity networks, heat network operators for existing district heat networks etc.) to ensure that the existing networks modelled in NAM are calibrated to a sufficient level of real world detail[27].

As gas and hydrogen networks are not modelled in detail, the key information required by NAM is whether the area under investigation is currently being supplied by a gas network, and how close to one is it if not. Depending on potential energy requirements, the tool also considers the optimal size of pipes required to extend the gas network to cover off-gas grid buildings, or to repurpose existing gas networks to be able to transport hydrogen.

7.3.3. Producing accurate cost functions for a range of network options

One of the key requirements of NAM is to be able to output cost options for different network reinforcement and new build options and for each energy vector (electricity, heat, gas, hydrogen). Cost functions must compare network reinforcement costs and maximum allowable capacity (load/generation) for system optimisation in POM e.g. when trading off reduced emissions through increased use of heat pumps versus the cost of reinforcing electricity networks to accommodate increased electrical load.

For existing network expansion the cost functions need to characterise the expansion costs plus any material operational costs and losses. For new build networks the full costs of installation are included, along with the operational costs and losses and the costs (where relevant) for later expansion or increased geographical penetration.

The values for network options that will be calculated by NAM are[28]:

- Capital costs;
- Fixed operational and maintenance (FOM) costs;
- Variable operations and maintenance (VOM) costs;
- Operational lifetime and age of the asset (if existing).

---

[27] If data is not available for a particular network component, generic values will be used based on expected network design.
[28] The component costs of the option are intended to reflect the economic resource costs of the options (see section 8.5.1) and are not on the same basis as e.g. the price control costs that are submitted to Ofgem, which include factors such as pension costs (which may vary by DNO). However, it would be possible to post-process the network component costs from the EnergyPath™ Design Tool to account for these additional factors.

- Construction/deployment period where this is >1 year.

FOM can be added to a capital expenditure and considered as one metric. On the other hand, however, VOM is dependent on utilisation (rather than peak capacity) and as such its final value will be determined by the POM depending on system dispatch optimisation. Finally the age and technical lifetime of reinforcement options is required in order to accurately represent the trade-offs between load-driven and age-driven reinforcement.

Three important issues need to be taken into account when producing cost functions:

1. Network costs must accurately represent the cost of covering each MW of local load (or local generation for situations where network reinforcement could be triggered by net exports to the network);

2. Where network reinforcement options have dependencies or are mutually exclusive this must be taken into account;

3. The number of network options tested by the NAM, as well as the number of network reinforcement options sent to POM, must be kept relatively low to ensure a tractable optimisation. As such, due to the complexity of the optimisation problem, only network constraints that can materially impact costs are considered.

7.3.4. Capturing material uncertainties in network component costs

The final cost functions produced by the NAM (see section 7.5.7) are constructed from individual component costs (e.g. the cost of a wire or the labour cost of installation). These costs may be uncertain due to a number of factors, which have been identified, largely based on the ETI 2050 Energy Infrastructure Outlook project, and can be summarised as follows:

- Site context: the site context is important in determining land or access rights costs, transportation costs, as well as other costs such as costs related to street works, planning and consents costs etc; and from the above project is driven primarily by whether the area is classed as urban, suburban, rural or London[29].

- Material costs: these refer to the costs of purchasing the network component under consideration and can depend on commodity prices, global/national/regional supply and demand conditions, foreign exchange fluctuations, learning curves for future costs etc.

- Labour costs: Labour costs depend on regional labour costs, the skills availability in a particular region, as well as on the complexity of the installation under consideration.

- Plant costs: The availability of suitable plant to support the installation of the network would also have an impact on overall costs.

- Installation costs: The scale of installation could also have an impact on costs and this will depend on a complex inter-relationship between the sizing and capacity of the

[29] Rural (typically defined as locations with <30 dwellings per hectare); Suburban (typically defined as locations with 30-60 dwellings per hectare); Urban (typically defined as locations with >60 dwellings per hectare); London.

overall installation and the various system costs. In general, large installations could bring system costs down through economies of scale and through avoiding duplication of some costs and labour.

- Ground conditions – excavation difficulty: degree of difficulty expected to be encountered in the excavation of trenches and holes during construction.
- Ground conditions – ground contamination: excavated material which requires specialist handling and disposal as a result of chemical contamination contained within the soil.
- Ground conditions – ground water conditions: water requiring intermittent or continuous pumping during construction operations to keep excavated areas safe and dry.

These factors are summarised in Figure 7-4 below. For heat distribution networks we have also identified the height difference between the heat source and the supplied customers as an additional factor that could potentially impact total costs. Indicatively, for a hilly area where the difference between the heat source and the last customer is 100m, costs for the heat distribution network would be expected to increase by roughly 3% due to the increased need for suitable heat exchangers. This has a relatively low impact on heat network costs, we have decided not to explicitly model height difference however it could be added in the future as an additional parameter, for example, by incorporating Ordnance Survey contour layers into the representation of the local area.

Figure 7-4 Key identified cost scalars

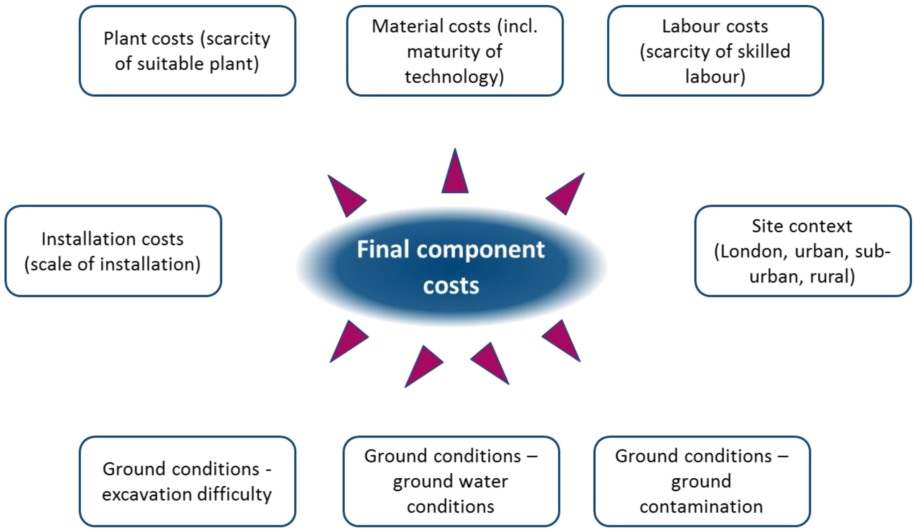


Table 7-7 summarises the approach for how the NAM takes the identified component cost modifiers into account:

- Material costs, labour costs, plant costs are all grouped together and captured as a set of discrete high/medium/low values – i.e. creating multiple sets of network options with different costs. These could be explored in the POM via sensitivity testing or used to set the triangular distributional parameters for Monte Carlo simulation of the network costs.

The approach to uncertainty across the EnergyPath^{TM} Design Tool is discussed further in section 9.

- Installation costs reflect possible economies of scale from large versus small-scale installation. This can be reflected in the decision making trade-offs in the POM via the use of MIP (Mixed Integer Programming) see section 8.5.2 for further details.

- Site context is modelled as a separate cost scalar, with information from the SAM (e.g. building density per hectare in each zone) used to flag the appropriate cost scalar for the NAM.

- Ground conditions (excavation difficulty, ground contamination and ground water conditions) can affect some network components more than others (for example underground cables more than overhead lines) and for those that they do, costs could vary very significantly. This information would be input ideally as a GIS layer flag (or direct cost scalar) for different zones by the Local Authority. The information would then be used to adjust the network component costs.

Table 7-7 Treatment of identified cost scalars

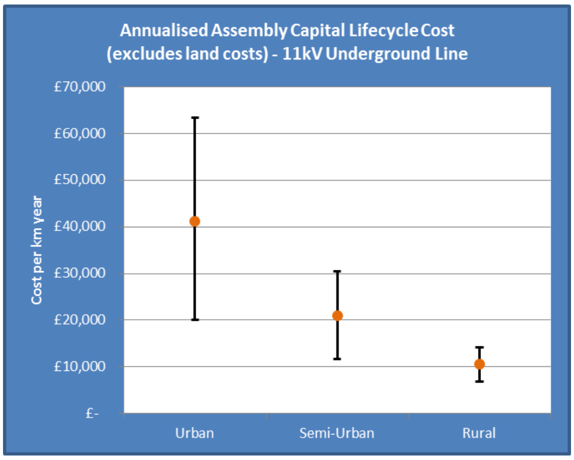
Parameters	Allowable Values	Treatment	Data exchange
Site context	Rural Sub-urban Urban London	Capture range of uncertainty by using different cost scalars for London, urban areas, sub-urban areas and rural areas.	SAM to send information to NAM regarding the site context of the area
Material costs	Low Central High	Capture range of uncertainty by modelling a range of possible outcomes (representing different market variances) with regards to material costs .	Discrete high/medium/low values for each cost function sent to POM
Labour costs	No skilled labour scarcity Central Skilled labour scarcity	Capture range of uncertainty by including a range of possible outcomes (representing different market variances) – these will be captured under the “material costs” category.	As per material costs, combine high across material, labour, plant costs
Plant costs	No scarcity of suitable plant Central Scarcity of suitable plant	As with “Labour costs”.	As per material costs, combine high across material, labour, plant costs
Installation costs	Large scale of installation Central Small scale of installation	Scale of installation aims to capture economies of scale. This is can be captured in the POM via the use of MIP (see section 8.5.2	Cost component for installation separated from other network cost functions and captured within the POM to reflect economies of scale
Excavation difficulty	Soft ground – no rock or hard material Intermittent rock / hard material Prolific rock / hard material	This will not be explicitly modelled at this stage as this information would not be available from a GIS package – the range of uncertainty will be captured under “material costs”. It will be possible, however, to better reflect these costs as appropriate once this information is available.	Flagged by the Local Authority who would then adjust costs appropriately

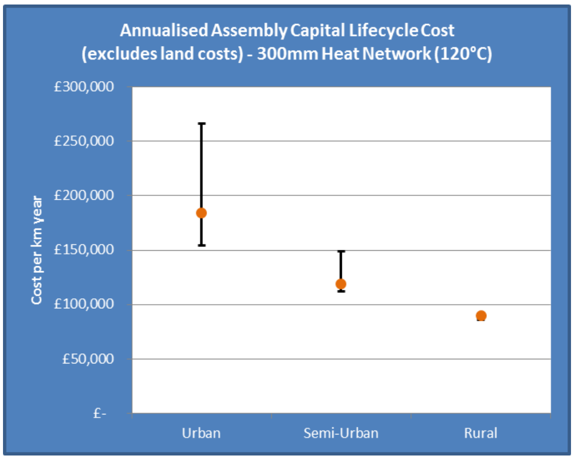
Ground contamination	Clean and inert ground Mildly contaminated ground Heavily contaminated ground	As with "Excavation difficulty".	As with "excavation difficulty"
Ground water conditions	Little or no ground water Intermittent dewatering required Continuous dewatering required	As with "Excavation difficulty".	As with "excavation difficulty"

For illustrative purposes, Figure 7-5 below show the annualised cost of an 11kV underground cable (rated at 6 MVA) and a 300mm underground heat network (with an operating temperature of 120°C). The annualised cost is calculated as the initial cost plus repetitive refurbishment costs at replacement cycles, plus abandonment cost, and assuming a 40 year asset lifetime.

It can be seen that a significant range of uncertainty may exist for both components considered here, particularly if they are installed in urban areas. The challenge will be to include this uncertainty in the modelling approach in order to allow for different market variances to be represented (see section 9). However, this should be narrowed down to the extent possible using the most accurate technical, spatial and physical information for the area under investigation.

Figure 7-5 Annualised cost of an 11kV underground cable (rated at 6 MVA) and a 300mm underground heat network (with an operating temperature of 120°C)





7.3.5. Producing 'heat-led' clusters with specific boundary constraints

Aggregating spatial zones into a smaller number of clusters is required in order to reduce the number of options available to the optimiser, thus ensuring that the optimisation problem is tractable given finite computing power. The final clusters used in POM should contain a reasonable representation of new and existing network options for both electricity and district heat networks within the same geographic boundary (and should also include a set of simpler options for gas grid extension, decommissioning or repurposing to hydrogen) so as not to bias the network and other trade-offs assessed in the POM.

At this stage, we believe it is important that the chosen approach will provide sufficient flexibility to ensure that the trade-offs and the economies of scale associated with developing and operating heat networks are fully taken into account, but which neither significantly underestimate or overestimate the electricity network costs in the cluster. This 'heat-led' approach is undertaken given that there are more relatively more 'degrees of freedom' associated with defining a new build heat network compared to reinforcing an existing electricity network.

It is also important to note that the aim of the 'heat-led' cluster definition is not to define a sufficiently large enough cluster that represents an entire network, but a potential component cluster of this. By seeing all trade-offs between all clusters simultaneously based on the data from the NAM, the POM can then make the decision to develop heat networks across multiple clusters, connect these clusters together, as well as to an appropriate heat source; all as part of creating an overall economically viable heat network.

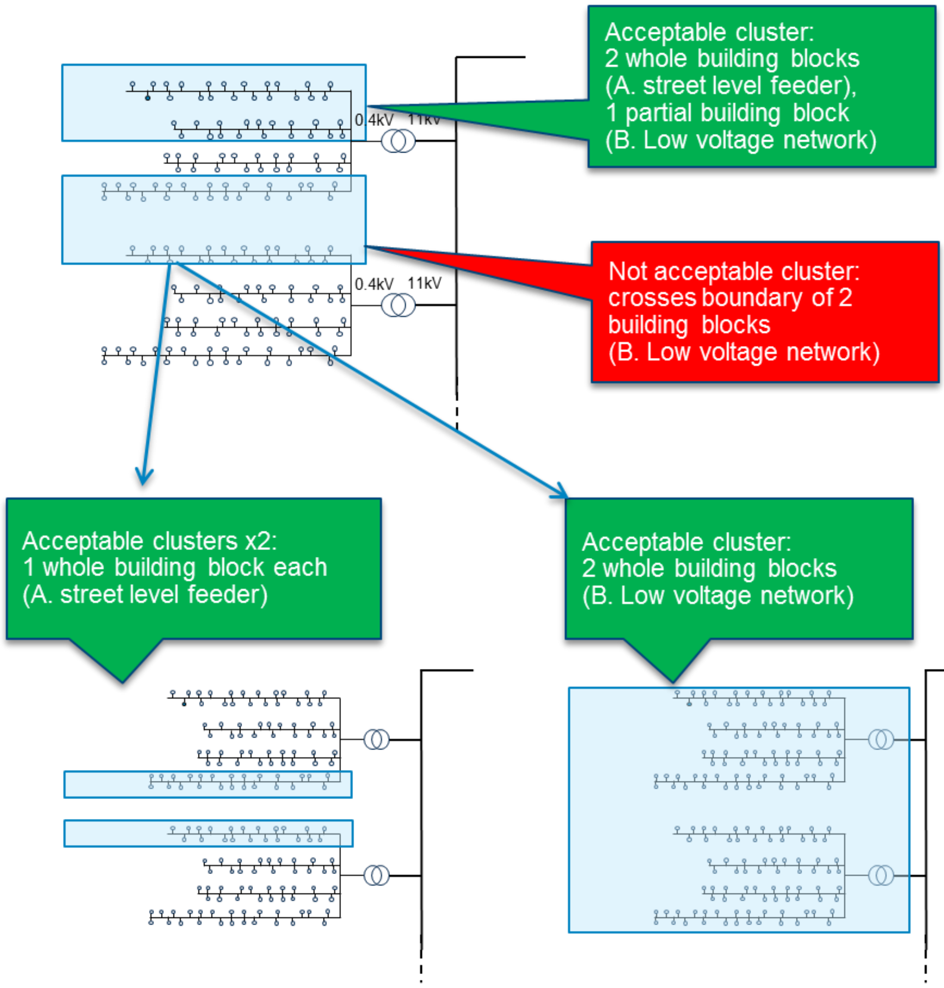
To balance the trade-off between the representation of heat and electricity network costs, we would expect to use a combination factors to define the final boundaries of the different clusters as part of the automated process. However, we would aim to develop the flexibility for the user to easily apply these factors in different combinations and with different values – i.e. to configure the automated boundary finding process. The process is described in more detail in section 7.5.5, but the factors would consider:

- The implied cost for developing a heat network within the cluster
- The maximum size of the cluster (e.g. number of buildings)
- An equivalent cost for upgrading the electricity network within the cluster which can could be contrasted with the heat network value in parallel to help ensure both costs are representative
- Physical boundary constraints which respect existing electricity network topology, described in further detail

As described in Section 6.1 zones are the smallest building blocks for clusters, and are defined using electricity network characteristics: these could be the LV distribution feeder supplying a street, the LV distribution network (which would include a number of LV distribution feeders) etc.   Clusters of zones are gathered together in the first instance primarily using heat network characteristics, however with boundary constraints set by the electricity network. This ensures that the derived clusters do not cross electricity network boundaries.

A high-level schematic of the approach is shown in Figure 7-6 below showing a variety of acceptable and unacceptable clusters based on the methodology described above.

Figure 7-6 Cluster boundary constraints using building blocks



Some zones may have limited heat network potential, and so will not be clustered using heat characteristics. All zones remaining un-clustered following the 'heat-led' approach are clustered in SAM using electricity network cost characteristics, and will observe the same boundary constraints.

7.3.6. Allowing more detailed analysis of solutions proposed by POM

Detailed load flow modelling cannot take place in POM as it needs to be done at a lower level of granularity. Once the network reinforcement solutions for each cluster have been selected by the POM, it is important to be able to perform a more detailed testing of these solutions to ensure their feasibility and potentially refine the solution to feed into more detailed project briefs for the local area. This is described further in section 8.8.

In practice, this more detailed analysis would ensure that the proposed solutions for electricity and heat networks do not violate any of the identified network constraints (e.g. voltage limits) and that the cost functions that represent these solutions accurately reflect the costs that are likely to be incurred. For gas and hydrogen networks this validation is not necessary as the physical constraints underpinning the design of these networks are not modelled in detail in NAM.

7.4. Key inputs

This section describes the key inputs to NAM, where they are used in the overall solution and for what purpose. The key inputs to NAM can broadly be summarised in four categories:

1. Primary inputs: These are primary inputs to the NAM and do not require interactions with other modules. The following inputs are included in this category:

    a. Generic network component database, i.e. a database containing all relevant information (costs and operational parameters) for the network model components modelled in NAM;

    b. Any applicable cost scalars or flags, i.e. a list with any applicable scalars that will be used to inflate or deflate the baseline component costs depending on the factors identified in Section 7.3.4;

    c. Statutory network constraint limits, i.e. a list with any limits to be applied on the identified network constraints after which network reinforcement or upgrade is considered to be required, such a drop below the minimum acceptable voltage level;

    d. Static spatial data from SAM or GIS control, such as user-inputted GIS reflecting the site context (e.g. building density per hectare) used to apply a number of the relevant costs scalars in b. above.

2. Inputs from SAM: These are inputs from SAM related to geographical information about the area under investigation. The following inputs are included in this category:

    a. Overarching area information, i.e. spatial and physical information about the area to be used to determine appropriate cost functions and clusters;

    b. Intra-zone information, i.e. geographical and technical information about existing and potential network components in the zone, as well as the number of each building archetype within the zone;

    c. Inter-zone information, i.e. geographical and technical information about existing and potential inter-zone connections for all energy vectors. This could reflect the ability to mesh the two zones together if there is no physical reason they could not be connected;

    d. Final cluster and inter- cluster definitions, i.e. a final definition of all cluster and inter-cluster connections (i.e. existing or feasible clusters which can be connected together) in order for NAM to derive cost functions for the final 'electricity-led' clusters.

3. Inputs from HOM: These are inputs from HOM related to demand profiles. The following inputs are included in this category:

    a. Archetype demand profiles to help estimate maximum load per zone;

b. Archetype generation profiles to help estimate maximum likely embedded generation per zone.

4. Inputs from POM: These are inputs from POM related to more detailed testing and refinement of the network reinforcement solutions provided by the module (Section 7.3.6). The following inputs are included in this category:

   a. Network design solutions proposed by POM.

   b. Evolution of load supply / demand profiles within a cluster over the pathway (and within year/day)

The tables below describe what is technically required within the tool and how it would be used. The relative importance of the primary data inputs and other factors such as costs and licensing restrictions, particularly where they must be purchased from an external provider, are discussed in deliverable (D4) Data Acquisition Plan.

Table 7-8 Input data – primary inputs

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Generic network component database	Primary	A database containing all relevant information (costs and operational parameters) for the network model components modelled in NAM.	Electricity networks: Costs and operational parameters for distribution lines, transformers, generation units, energy storage units and smart technologies Heat networks: Costs and operational parameters for pipes, pumps, heat production plant and heat accumulators and storage facilities Gas and hydrogen networks: Costs and operational parameters for gas and hydrogen pipes. The generic network component database will be based on the ETI 2050 Cost Database.	See below	Include all modelled network components, across all energy vectors and across the UK	External (mainly based on ETI 2050 Cost Database)	-
2	Any applicable cost scalars	Primary	A list with any applicable scalars that will be used to inflate or deflate the baseline component costs depending on the factors identified in Section 6.3.4.	Cost scalars for the key factors ("component cost rate modifiers") that can materially affect network reinforcement costs. These include: - Site context - Material costs, labour costs, plant costs, installation costs - Ground conditions (excavation difficulty, ground contamination, ground water conditions) The cost scalars will also be based on the ETI 2050 Cost Database.	Cost scalars to account for the uncertainty (e.g. high / low) in the generic central component costs	Include all modelled network components, across all energy vectors and across the UK	External (mainly based on ETI 2050 Cost Database)	-
3	Statutory network constraint limits	Primary	A list with any limits to be applied on the identified network constraints after which network reinforcement is considered to be required.	Statutory limits for the modelled network constraints, including: - Statutory voltage limits at the LV level (electricity) - Statutory voltage limits at the MV level (electricity) - Operating temperatures (heat) - Target pressure loss (heat) - Operating pressure (gas, hydrogen)	-	Include all relevant network constraints, across all energy vectors and across the UK	External (based on Engineering Recommendations and other technical sources)	-

4	Technical characteristics of existing network components	Primary	SAM contains spatial information on where components are located, but NAM needs information on current headroom in current and future years to understand spectrum of requirements for network reinforcement options	Technical information about all existing network components across all energy vectors (cable/pipe sizes, transformer ratings, remaining lifetime etc.)	-	All existing network components	DNO, LA	May not be available for all parameters (e.g. thermal limits on all cables) and will have to be estimated

Table 7-9 Input data (from SAM)

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Overarching area information	Secondary	Providing spatial and physical information about the area under investigation in order to be used to determine appropriate cost functions and clusters.	<u>Spatial information:</u> providing information about the area under investigation around all parameters affecting component costs & cost of network reinforcement. This includes: - Site context - (if available) – Any specific information about the area that could impact material costs, labour costs, plant costs, installation costs - (if available) - Ground conditions (excavation difficulty, ground contamination, ground water conditions) <u>Physical information:</u> providing physical information around suitable boundaries for clusters according to the four electricity building blocks identified (street level feeder, low voltage network, 11kV feeder, primary substation) as well as other more generic physical boundaries (rivers, cliffs, railways etc.) Spatial granularity by zone	Ground conditions likely to be highly uncertain until a direct survey is undertaken	Include all modelled network components and buildings within the area under investigation	SAM Site context estimated from building density in SAM Other parameters entered as separate GIS layer	-
2	Intra-zone information	Secondary	Geographical and technical/physical information	<u>Geographical information:</u> providing geographical information about the zone, which includes:	-	Include all modelled	SAM	-

			about existing and potential network components and buildings in the zone;	- <u>Electricity</u> - Total street distance & distance from building to distribution substation (for all buildings); - <u>Heat</u> - Total street distance & distance from the local heat source or heat distribution system (the heat load-weighted point calculated in NAM using this information must lie on a road); - <u>Gas/Hydrogen</u> - Total street distance & distance from building to gas distribution system (for all buildings); - <u>All energy vectors</u> - Potential locations for new-build network components (for ex. potential location for a new transformer or a new heat source etc.). - <u>Building archetypes</u> – number of each archetype within the zone - <u>Embedded generation</u> – applicability of different types and maximum potential size of embedded generation <u>Technical/physical information:</u> providing technical/physical information about the zone, which includes: - Information around whether the buildings in the zone are currently connected to the gas distribution system.		network components and buildings within the area under investigation		
3	Inter-zone information	Secondary	Geographical and technical/physical information about existing and potential inter-zone connections for all energy vectors	<u>Geographical information:</u> providing geographical information in terms of distance between existing inter-zone connections and potential new inter-zone connection options – e.g. between two meshing points following road network (including connections to zones located in different clusters) for all energy vectors <u>Technical/physical information:</u> providing technical/physical information with regards to the ratings of these connections and any other relevant technical characteristics	-	Include all modelled network components and buildings within the area under investigation	SAM	-
4	Final cluster and inter-cluster		A final definition of all cluster and inter-cluster connections in order for NAM to derive	The final cluster definition includes a list with the following data: - A list with all clusters (and the archetypes across	-	Include all clusters and inter-cluster	SAM	-

	definitions		cost functions for the final 'electricity-led' clusters.	buildings in each cluster) as identified following the 'heat-led' and the 'electricity-led' clustering process. ▪ Building ID (and archetype); ▪ Network component ID; ▪ Zone ID; ▪ Cluster ID. - A list with all existing and potential new inter-cluster connections; ▪ Cluster ID; ▪ Connection ID and type of connection.		connections within the area under investigation		

Table 7-10 Input data (from HOM)

ID	Data field(s)	Input type	Purpose	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Local demand profiles	Secondary	Estimate likely maximum load conditions possible within zone to help bound load flow modelling	Building archetype half hourly-demand profiles for different characteristic days	Range of discrete demand profiles representing e.g. high/medium/low drivers of demand	All buildings within the area under investigation	HOM	-
2	Local generation profiles	Secondary	Estimate likely maximum embedded supply conditions possible within zone to help bound load flow modelling	Half hourly-supply profiles for different characteristic days	Range of discrete profiles representing e.g. high/medium/low output	All generation units within the area under investigation	HOM	-

Table 7-11 Input data (from POM)

ID	Data	Input	Purpose	Granularity	Uncertainty	Coverage	Source	Limitations

	field(s)	type			parameters			
1	Network design solutions	Secondary	Receive solutions proposed by POM and perform more detailed load flow testing to ensure accuracy and feasibility and refine solutions.	Receive proposed network reinforcement solutions for the modelled network components within a given area. Solutions for each timeperiod and each cluster	Multiple design solutions from Monte Carlo simulation within POM	All clusters within the area under investigation	POM	Data at cluster level
2	Supply / demand profiles	Secondary	As above	Supply/demand profiles Solutions for each timeperiod, timeslice and each cluster	Multiple design solutions from Monte Carlo simulation within POM	All clusters within the area under investigation	POM	Data at cluster level

7.5. Key NAM logic / process Steps

The next sections describe the key steps within NAM that fulfill the modelling requirements described in Section 7.3.

7.5.1. Zone data pre-processing [NAM-001]

Relevant spatial and physical information is first received from the SAM (see section 7.4) with regards to the area under investigation, at a zonal level of detail, along with information from the HOM on the demand and generation profiles associated with archetypes in the area. Zones are the smallest building blocks for clusters, and are defined using electricity network characteristics.

We have identified 4 levels of the zone building blocks, increasing in size, which must be respected as part of any final cluster definition (see section 7.3.5, Figure 7-6):

A. Street level feeder that connect buildings on roads to the nearest distribution substation;

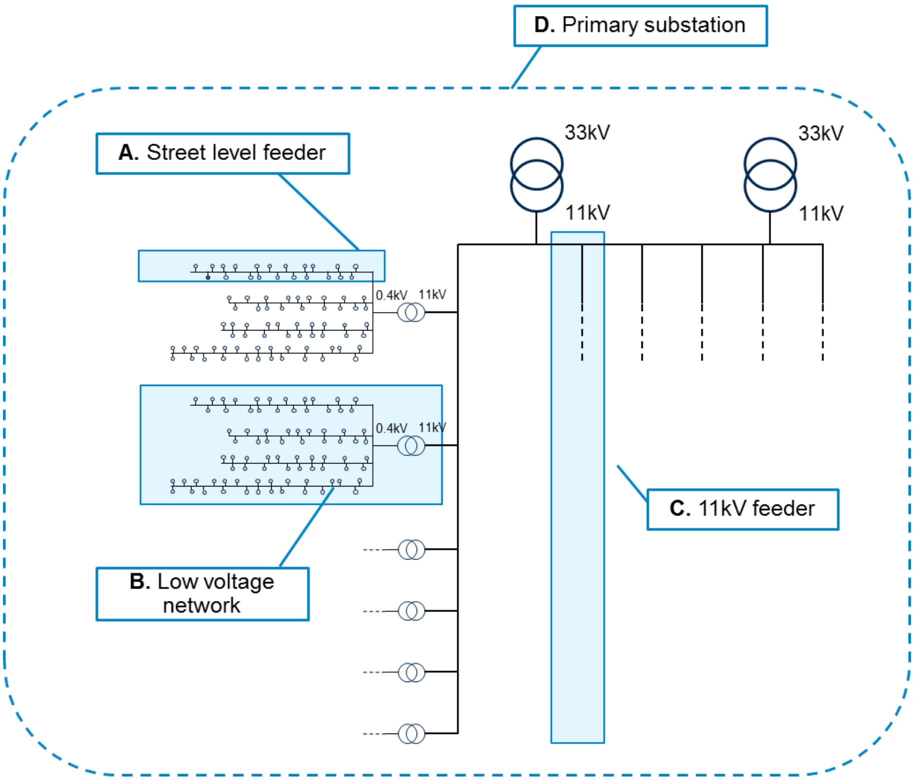
B. Low voltage network (Distribution substations that have several LV feeders)

C. 11kV feeder (HV feeders that connect several distribution substations to the nearest primary substation);

D. Primary substation that has several HV feeders.

These blocks are shown schematically in Figure 7-7 below.

Figure 7-7 Identified zone building blocks levels reflecting electricity network connectivity

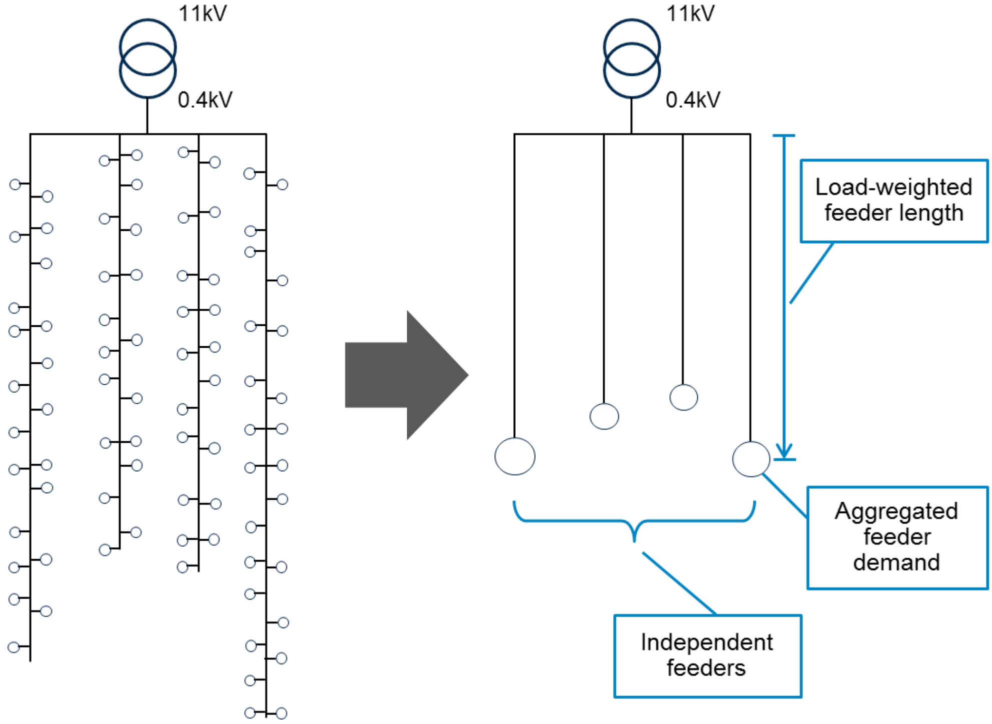


The aim is to maintain the base zone definition for the load flow modelling at the lowest level of spatial granularity (i.e. at street level) within the NAM in order to maximise accuracy of the derived cost functions. However, given the number of load flow studies that need to be undertaken to test the full range of network reinforcement and new build options, it is likely to be necessary to simplify the network representation slightly, as described below.

Electricity network modelling

For electricity this is undertaken first by collapsing the representation of individual buildings (units) on a street-level LV feeder (rated at 0.4kV) to a single load weighted distance point as shown in Figure 7-8. Where the LV feeder is split via a breeches joint or link box, the simplification would be undertaken separately for the split and main feeder branch.

Figure 7-8 Simplified street level feeders



In this simplified representation load flow calculations can be performed on each feeder independently to assess peak load / cost characteristics for reinforcement options for each feeder cable[30]. Care must be taken:

- When selecting the impedances to be used for the load-weighted feeder cables, care must be given to ensure that the voltage drop at the load-weighted point closely resembles the voltage drop that would be expected at the most remote network point of the actual network;
- To ensure that any potential thermal bottlenecks (for example parts of the network where smaller-sized cables are used or if a section of cable has a lower than normal rating because it is within a duct) are also identified and taken into account in the modelling solution – e.g. by modelling it as a different cable type.

As part of the pre-processing, the maximum possible load/supply conditions per zone is estimated (e.g. by using the maximum non-diversity adjusted building archetype demand profiles on the most extreme characteristic day) to help bound the number of load flow studies to be considered.

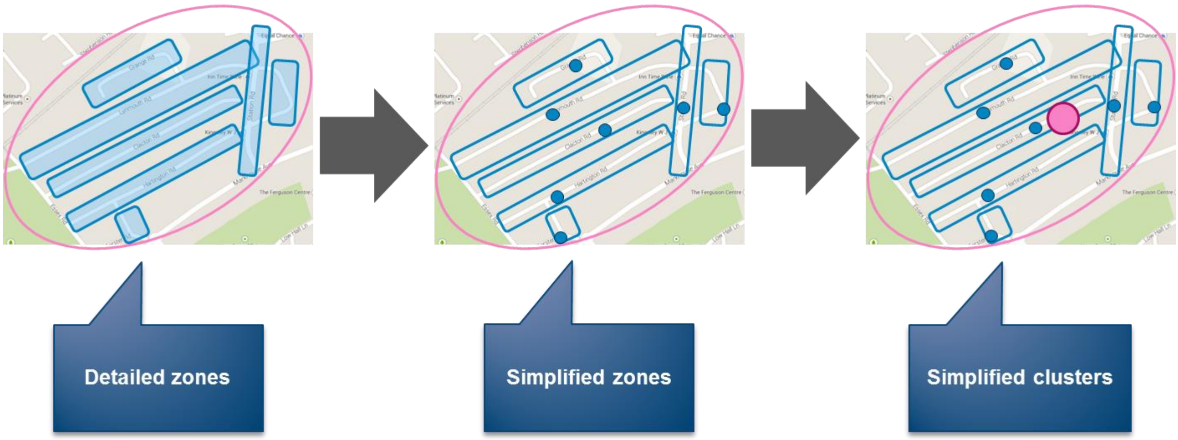
District heat network modelling

[30] As part of early development this approach will be validated against the result using the full load flow modelling. In addition, if it is computationally feasible to remain at the individual building unit representation this would be maintained.

In a similar manner to electricity, a number of pre-processing steps are undertaken to help generate a representation of a potential heat network configuration. The first step is to calculate the maximum heat load-weighted centres for each, with the constraint being that the load-weighted points must be located on a road, which represents the route of the assumed future heat network in that zone.

These are then used to calculate the maximum load-weighted centres for a number of zones within a cluster, again with the constraint being that those points must be located on a road within the cluster (Figure 7-9). The broader process of defining the cluster boundaries is described in section 7.5.5, but the key point here is that the cluster load-weighted centre is a function of both the number and shape of the zones within the cluster itself, and will need to be re-calculated if either of these change; the zone load-weighted centre is, however, static.

Figure 7-9 Simplified heat network – total load and load-weighted centers

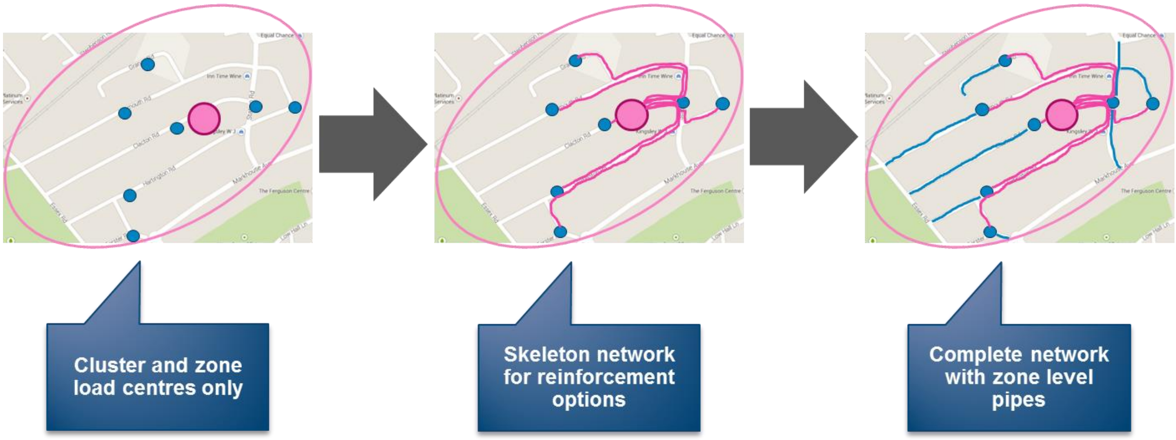


It is assumed that the heat distribution system (which could be an inter-cluster source) or the heat source directly supplying the cluster connect at the cluster load-weighted centre point and then supply all the connected zones within that cluster.

A skeleton heat network topology is then formed by developing pipe networks from the cluster load-weighted centre point to individual zone load-weighted centre points using minimum road distance. The network reinforcement/expansion options use this skeleton heat topology, but the POM then evaluates a range of different pipe/pump sizing options to determine optimal sizing for the cluster[31]. Finally, it is also assumed that the size of the intra-zone network does not change. In practice, this means that the size of the pipes supplying the customers connected along a zone is the same as shown in Figure 7-10.

[31] It is assumed that full details of the original network outline would be stored, so that this can be refined further as part of the detailed network analysis of higher-level solutions from the POM.

Figure 7-10 Simplified heat network – heat network skeleton



7.5.2. Test options generator and costing tool [NAM-002]

Electricity and heat

For a given spatial area (e.g. zones at the different building block levels for electricity) this step generates a series of network reinforcement/new build options which must be tested within the load flow modelling tool. The components are drawn from a master component database list which is valid for each relevant section of the network (e.g. cable upgrades for feeders or transformer upgrades for substations).

Via an automated process, all valid combinations of reinforcement options are added to the network configuration and a series of iterative load tests[32] are undertaken on that configuration until network constraints on the relevant part of the network are breached (e.g. thermal limits are breached or voltage drops below a given threshold).

Where these limits are breached reflects the maximum allowable load or embedded supply export, enabled by these network options. The options are then mapped to the underlying costs and characteristics within the component database to understand the associated:

- Capital and installation costs (these may be adjusted to reflect the locational and other uncertainties outlined in section 7.3.4 – as result the final options may have a discrete set of high/medium/low costs which can be explored as part of the overall uncertainty analysis – this is discussed further in section 9)
- Fixed operational and maintenance (FOM)
- Variable operations and maintenance (VOM)
- Operational lifetime of the asset.

An example of this process is outlined in the Appendix section 13.2.2

[32] From a maximum net export position, where relevant, through to maximum load.

Testing all possible combinations of network components in all feasible configurations to their failure point is clearly a computationally intensive process, however, this process can be simplified in a number of ways:

- As outlined in section 7.3.1 dynamic / active network options (such as storage) are handled in the POM, hence the number of options to be examined in the NAM is reduced.

- More careful selection of the number of load tests via use of search algorithms (e.g. half interval/binary search) or by inferring potential bounds from the underlying component characteristics.

- By separating out testing of network options, which are effectively independent of others (e.g. radial lines and transformers) from others which are interdependent (e.g. meshing) and hence must be tested in conjunction with other network options.

The process of creating the final cost functions/options for use in the POM from this basic process is complex and is discussed further in section 7.5.4.

For the vast majority of distribution networks it is anticipated that network reinforcement will be triggered during periods of 'maximum demand – minimum embedded generation'. This is because the most extreme system conditions in terms of network flows are likely to be experienced during Winter peak demand, although for some areas (particularly for areas with high penetration of air conditioning units in the South) this could also be during Summer peak demand.

Conversely, for zones with significant potential for distributed generation it is possible that network reinforcement (particularly to prevent high voltages) could be triggered during periods of 'minimum demand – maximum embedded generation' if the total installed capacity of these units is greater than peak demand.

For example this could be the case during a particularly sunny day in an area with significant penetration of small-scale PVs. These zones should be identified and the network reinforcement options produced for them should cover both 'maximum demand – minimum embedded generation' as well as 'minimum demand – maximum embedded generation' system conditions as shown.

For non-building scale embedded generation it will be important to use information from the SAM on the potential location and maximum installable capacity (e.g. of small scale wind) to test network reinforcement configurations, which are likely to be required in significant net supply cases. Where it is clear that an electricity network reinforcement option is only being driven be larger-scale embedded supply, this could ultimately be linked to peak supply from certain technologies within a cluster in the POM, as opposed to the peak demand net build-scale generation, which would be used to drive the majority of network option decisions.

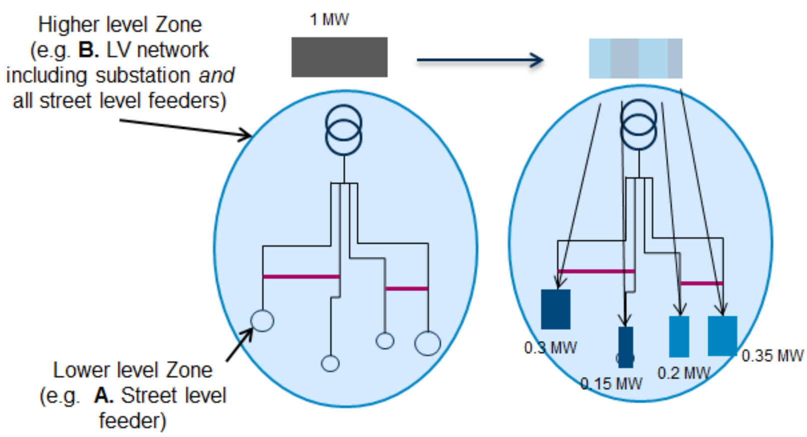
Aggregating electricity load moving from lower to higher level zone building blocks

When performing the load flow studies, it is necessary to distribute the incremental load (for example due to heating) amongst the individual zones connected at a lower level building block to the next level up (e.g. a street level feeder connected to a low voltage distribution substation). For electricity networks this is required in order to accurately test voltage and thermal constraints on the connected lines.

This is necessary because when the final cluster boundaries are determined (which reflect the different zone building block levels within them) the POM has no understanding of where load is developing spatially within the cluster, as this information has been lost to simplify the problem. It only understands the aggregate peak load (net of embedded generation) due to choices it is making within the cluster and must trade this off against the cost of network reinforcement to support this peak (see Figure 7-11).

Hence the way the underlying network cost functions are constructed for a cluster (based on the zones within it) must superimpose a similar approximation to ensure that the way the POM ultimately interprets the cost functions is consistent.

Figure 7-11 Example of aggregating load from lower to higher level zone building blocks



There are a number of ways to approximate this load distribution ex-ante, for example:

- Lower level Zones with lowest cost first – this would be comparable to a best-case scenario where incremental load would accrue at the parts of the networks with the lowest reinforcement costs (for example due to DNO incentives);
- Lower level Zones with highest cost first – this would be comparable to a worst-case scenario where incremental load would accrue at the parts of the networks with the highest reinforcement costs;
- Equally distributed amongst lower level Zones – this would assume that the same level of incremental load would be added to each zone within the cluster;
- Based on zone's potential peak load share of cluster – this would assume that the level of incremental load added to each zone would be based on its potential peak load share (which in turn is likely to depend on the number of customers connected in that zone);
- Distributed based on a lower level Zone's current load share – this would assume that the level of incremental load added to each zone would be based on its current load share (which in turn is likely to depend on the number of customers connected in that zone);

The same basic principles for aggregating load apply as you move further up the zone building block levels – i.e. from **B.** Low Voltage Network to **C.** 11kV feeders to **D.** Primary substations

Our current view is that the last option (based on current load share) is likely to be the least distorting default in terms of the final POM solution. Ideally a metric which reflected the likely propensity to electrify by zone would be used (i.e. trying to pre-judge the outcome of the pathway analysis in the POM). We will investigate further the extent to which it is possible to estimate such a metric in Stage 2, however, it is complicated by the myriad of factors which are ultimately traded-off to determine the degree of electrification.

Gas and hydrogen network modelling

Unlike electricity and heat, it is not proposed to undertake steady-state network flow modelling in Stage 2 for gas and hydrogen (although this could be added in future drawing on the same basic process above). Instead the NAM produces "all-or-nothing" cost option for the following options for each zone

- Extending the gas network to cover off-gas grid buildings. This is undertaken by using data from the SAM to understand the number of off-gas grid buildings in a zone and the distance to the nearest section of the existing gas grid following the road network. The cost per metre is multiplied by that of the relevant pipe size (assumed to be the same as the connection point) adjusting for any double-counting of distance where multiple off-gas grid buildings would be connected via the same route.

- Decommissioning the existing gas network for operational costs savings. As above, data on the existing network (e.g. length) within the zone would be combined with per unit estimates of decommissioning costs to create a cost for the zone as a whole

- Re-purposing the existing gas network to allow it to be able to transport hydrogen. As per decommissioning data on the existing network within the zone would be combined with per unit costs for re-purposing to hydrogen (the costs for changing appliances is estimated within the HOM and based to the POM as part of the building archetype data and potential hydrogen supply sources are evaluated in the POM)

7.5.3. Load flow modelling [NAM-003/4]

The steady-state load flow modelling helps to understand how network constraints (such as thermal limits or voltage as outlined in section 7.3.1) are affected under different configurations of network reinforcement / new build options under different load conditions. For the majority of networks capacity limits are expected to be set by peak demand conditions, however it is also possible that network reinforcement could be triggered by net exports when supply outstrips demand.

For electricity networks:

- For simple radial networks (i.e. for networks where meshing options are not considered) load flow calculations can take place in the network reinforcement 'wrapper' rather than in a dedicated network load flow modelling software. This allows the overall process to be considerably faster.

- For systems which exhibit meshing (i.e. at intra-cluster and inter-cluster level) dedicated network load flow modelling will be used to ensure that load flows are modelled as accurately as possible.

For heat networks:

- Load flow calculations (following the concentric circles methodology) will take place using a dedicated network load flow modelling software in order to ensure that all physical parameters underpinning heat networks are accurately modelled when testing different network designs.

Examples of the application of the load flow modelling for electricity and heat are outlined in the Appendix section 13.2.

For gas and hydrogen networks, as mentioned in the previous section, load flow modelling for these energy vectors is not performed in NAM, with all calculations performed in the network reinforcement 'wrapper'.

7.5.4. Zone level electricity network option cost functions [NAM-005]

Once the electricity network reinforcement / new build options have been tested via the load flow modelling they need to be converted into cost functions for the relevant zones. These need to consider the relevant components at each *level* of the zone building blocks A to D in section 7.5.1 (this is important because all final cluster boundaries definitions have to respect this underlying topology definition).

These zonal functions are aggregated later to a cluster level function and passed to the POM for use in the pathway analysis (as described in section 7.5.7). *For heat this process is different* as there are intermediate steps to first define sensible cluster boundaries and then evaluate the heat network cost options within the cluster as a whole (see section 7.3.5) rather than individually for a zone and then aggregate up.

We have identified five possible methods for collating the electricity network options at a given peak load of increasing accuracy and complexity. These options are:

1. Incremental cost function (lowest cost reinforcement option to cover each MW of capacity);

2. Discrete additive options (both for £ cost and for MW capacity headroom);

3. Discrete additive options with careful selection of cost order of packages – i.e. where we need to represent dependency (X needs to be done before Y) - X must be cheaper than Y to operate properly within the POM's least cost optimisation;

4. Separate 'sets' of cost options e.g. cable options vs transformer options being driven off same peak MW – this would allow us to resolve the £/MW headroom interdependency issue for everything but meshing;

    b. Potentially this could be overcome by separately testing two configurations of the network: (i) with meshing; and (ii) without meshing.

5. Separate discrete options with interdependencies (e.g. X -> Y -> Z) as best way to capture interaction on £ and MW headroom across different packages of options particularly meshing.

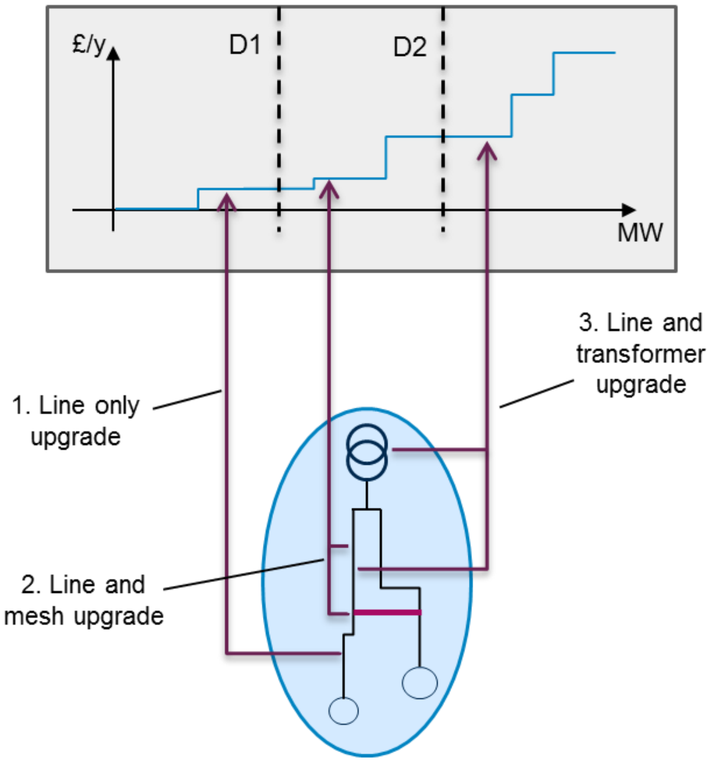
Due to the limitations of the simpler options we propose to focus on Options 4b and Option 5 in Stage 2. The main difference between these two is whether inter-dependent options (such as meshing) are captured endogenously within the POM optimisation (Option 5) or need to be tested via sensitivity (Option 4b). The former requires the use of integer programming conditional constraints, which although feasible (as demonstrated in the POM proof of concept work – see section 8.7) has a sizeable performance impact.

Method 1: Incremental cost function

The first method is to consider a simple cost function where all the available network options are considered to be incremental. The only strength of this option is reduced complexity, however its fundamental weakness is that it does not account for dependencies between different network reinforcement options due to its incremental representation.

For example, having selected reinforcement option #2 in an early timeperiod (i.e. to reinforce a network line and also create a mesh) to meet capacity D1, we would not dismantle the mesh to select reinforcement option 3 (line and transformer) to meet capacity D2 in a future period. In this regard reinforcement options 2 and 3 should be mutually exclusive and importantly the full cost of option 3 should be incurred to implement it, however this cost function representation would misleadingly only consider the incremental cost of moving from option 2 to 3.

Figure 7-12 Incremental cost function (lowest cost reinforcement option to cover each MW of capacity)

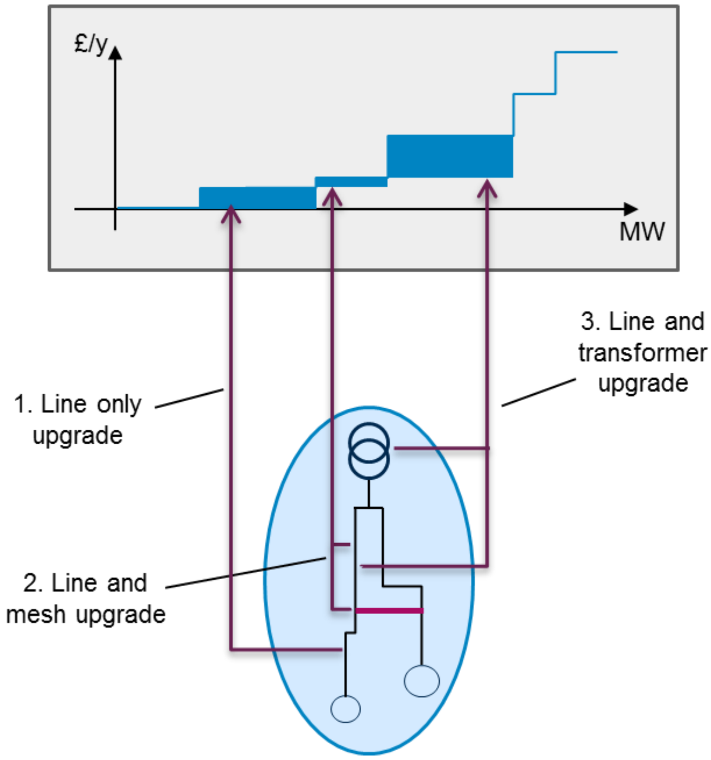


Method 2: Discrete additive options

The second method is to only reflect a set of discrete network reinforcement options in the function, which are genuinely additive in terms of their headroom – i.e. no options can be included which are mutually exclusive. The benefit of this is that the incremental costs are more reflective of the nature of the upgrade costs.

However, the key weaknesses are that this limits the available network options which can be represented and that an incremental upgrade may be lower cost compared to a pre-requisite enabling upgrade. For example in the figure below, on an incremental basis option #2 (i.e. to reinforce a network line and also create a mesh) is cheaper than option #1 (i.e. to reinforce a network line only) and as a result the optimiser will always choose option #2 before option #1, however, in reality option #1 must be undertaken *before* option #2 is available.

Figure 7-13 Discrete additive option (both for £ cost and for MW capacity headroom)

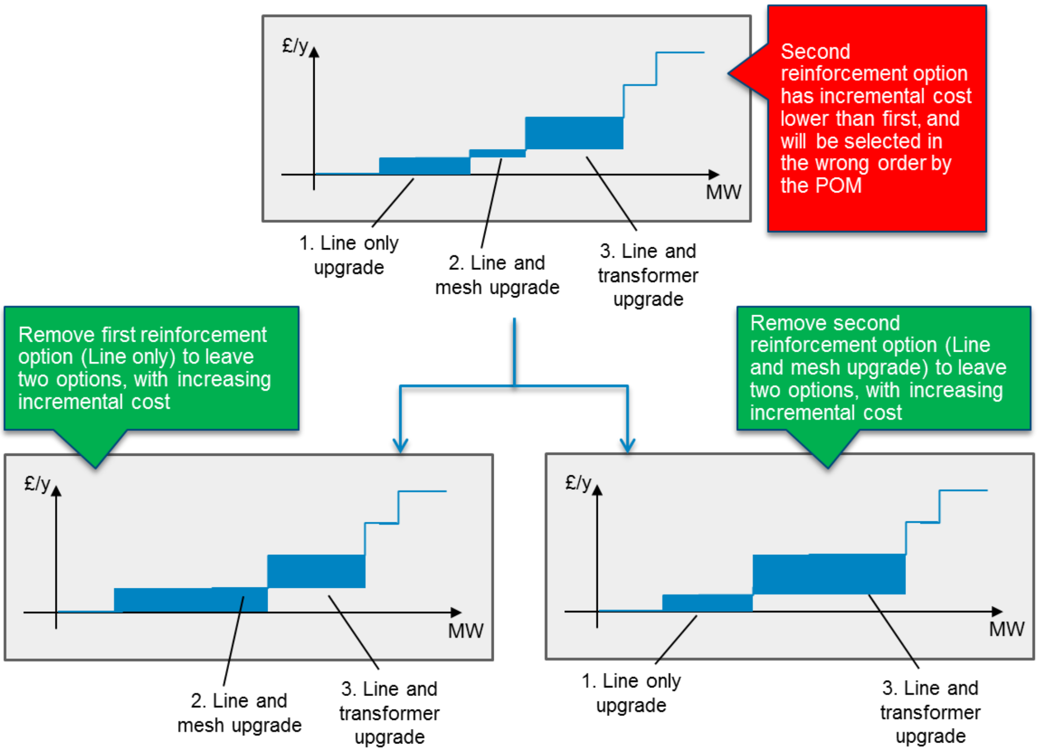


Method 3: Discrete additive options with selection of cost order of packages

This is a refinement of Method 2 by removing or adjusting previous options to ensure that network reinforcement options are monotonically increasing in terms of costs and the appropriate order of network upgrades. As shown in Figure 7-14, there are two main ways to achieve this:

- On the left hand side - aggregate both the enabling option #1 (Line only) and the incremental upgrade option #2 (Line and mesh upgrade) into a single option (thereby losing the differentiation between them)
- On the right hand side – remove option #2 the incremental mesh upgrade

Figure 7-14 **Discrete additive option with selection of cost order of packages**



There are still two key issues with this method:

- The first is that this process could remove a large number of potential network reinforcement options (especially for relatively densely populated areas where radial networks are used but where meshing could represent a viable alternative). This could therefore push up the cost of reinforcing the network from the true least cost solution.
- The second issue is that this option also does not account for dependencies – for example the new combined options #1 plus #2 are still mutually exclusive with #3 due to meshing, however they appear as additive options.

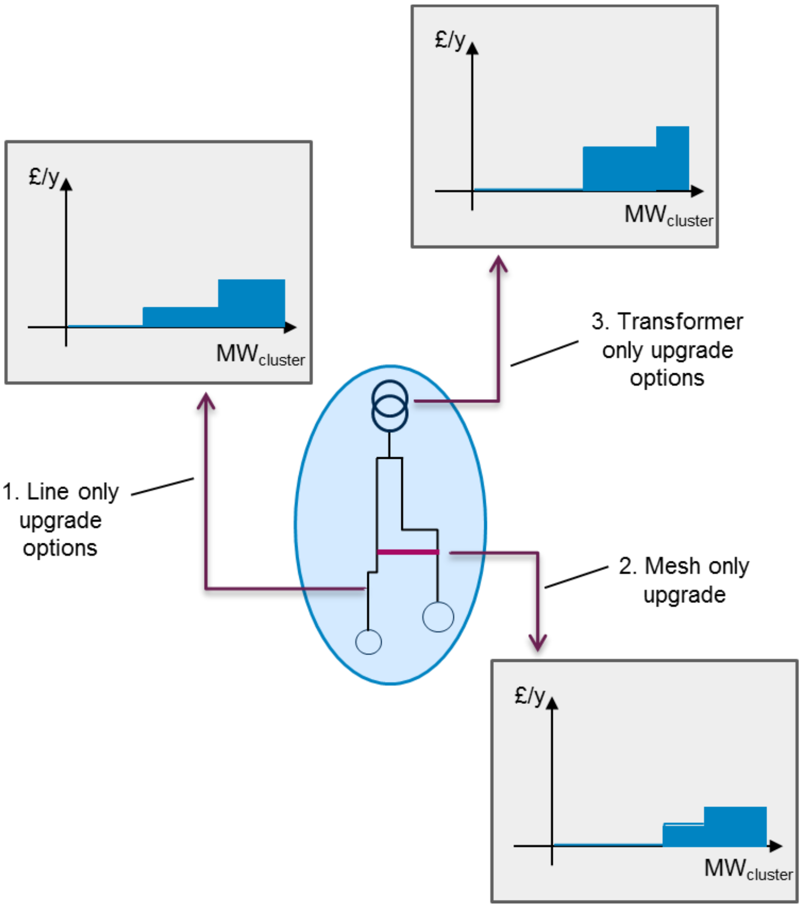
Method 4: Separate 'sets' of cost options for the same peak MW in cluster

This method is an extension of Method 3 based on the notion that network reinforcement functions for some network components can be considered independently – for example for radial lines and transformers. This is because each of these components will 'fail' (i.e. will need to be reinforced) at a load flow independent of the other network components that are connected in series. For each independent component, reinforcement options are assessed, and ranked by cost and allowable capacity.

Hence within the POM, a given cluster peak load is effectively driving reinforcement across a number of independent network reinforcement functions *in parallel.* The way the peak load drives each function will need to be scaled based on the distribution of load as discussed in section 7.5.2.

The key issue with this approach is that where meshing is present the component costs are no longer independent – the peak allowable load is a function of the reinforcement of radial lines and meshing links (e.g. where voltage drop constraint is binding). As a result, this means that meshing as a network reinforcement option needs to be considered ex-ante by the tool.

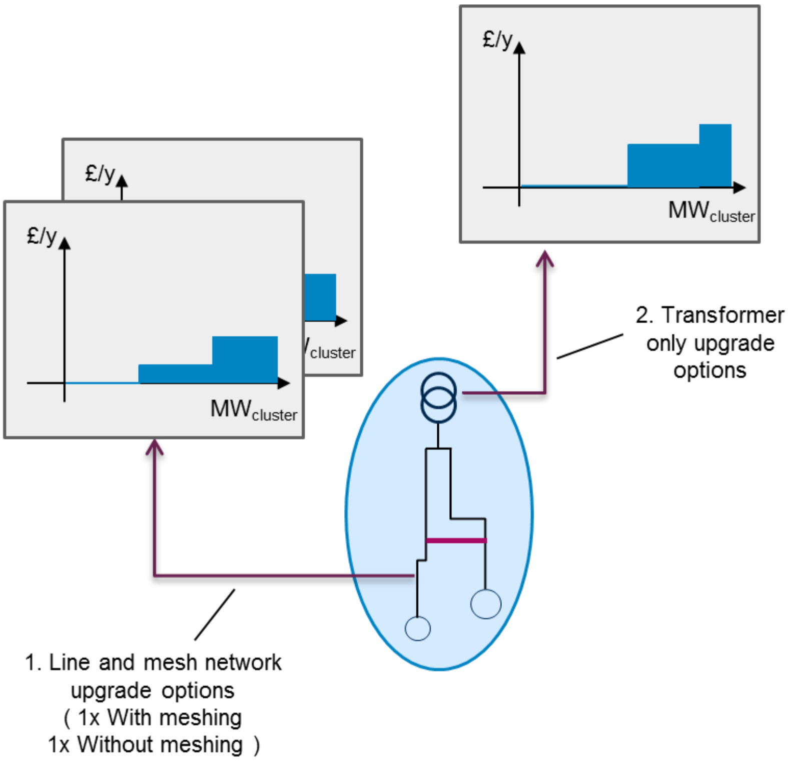
Figure 7-15 Separate 'sets' of cost options for the same peak MW



Method 4b: Separate 'sets' of cost options for the same peak MW – with/without mesh

As an extension of method 4 it would be possible to test the impact of meshing by creating two sets of network cost functions in all relevant cases (i.e. with and without meshing) and run two discrete sensitivities in the POM.

Figure 7-16 Separate 'sets' of cost options for the same peak MW – with/without mesh



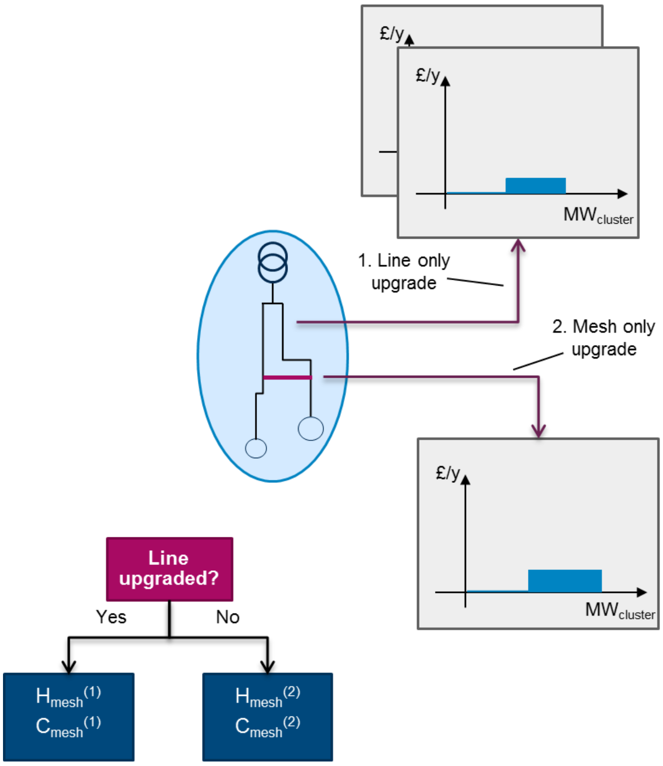
Method 5: Separate discrete options with interdependencies between years

This method is based on the notion that for many network reinforcement combinations the costs and/or increased capacity headroom associated with the upgraded components cannot be summed linearly as they are inter-dependent (e.g. the options are either mutually exclusive or a higher cost reinforcement option is first needed to enable a lower cost truly incremental upgrade). This can be illustrated using the following examples:

- Reinforcing a network line would give an incremental increase in capacity headroom of H_{line} at a cost C_{line}
- Creating a meshed link would give an increase of H_{mesh} at a cost C_{mesh}
- Performing both options, however, would not result in the sum of the above in terms of capacity headroom, i.e.
  - $H_{line + mesh} \neq H_{line} + H_{mesh}$
- In terms of costs, this example would likely lead to similar costs being accrued (i.e. $C_{line + mesh} = C_{line} + C_{mesh}$), however there are other situations where it is possible that costs would not be the same, for example if savings were realised by reinforcing both a transformer as well as a line (i.e. $C_{line + transformer} \neq C_{line} + C_{transformer}$)

The incremental head room H_{mesh} from creating a meshed link is dependent on whether the line has been upgraded or not. A tree of "if" statements (reflected by conditional constraints using Integer Programming) could therefore be used to capture these dependencies within the optimiser (this has been tested within the POM proof of concept work see section 8.7). As a result this is the most accurate approach for reflecting network cost functions, but may require a significant number of conditional values to be created, which may not be computationally tractable.

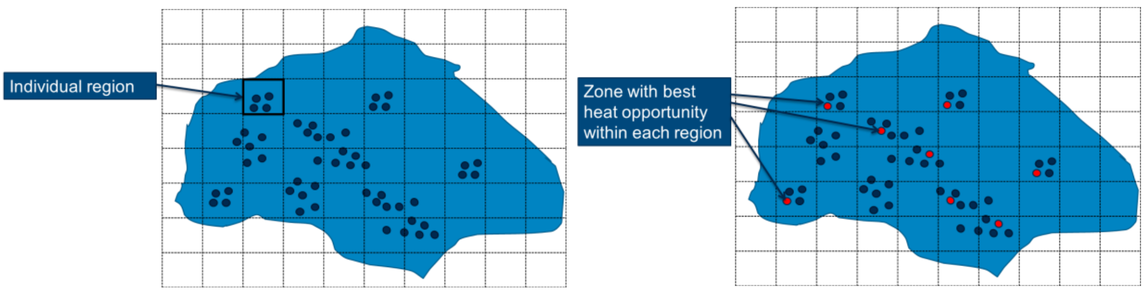
Figure 7-17 Separate discrete options with interdependencies between years



7.5.5. ‘Heat-led’ cluster definition [NAM-006]

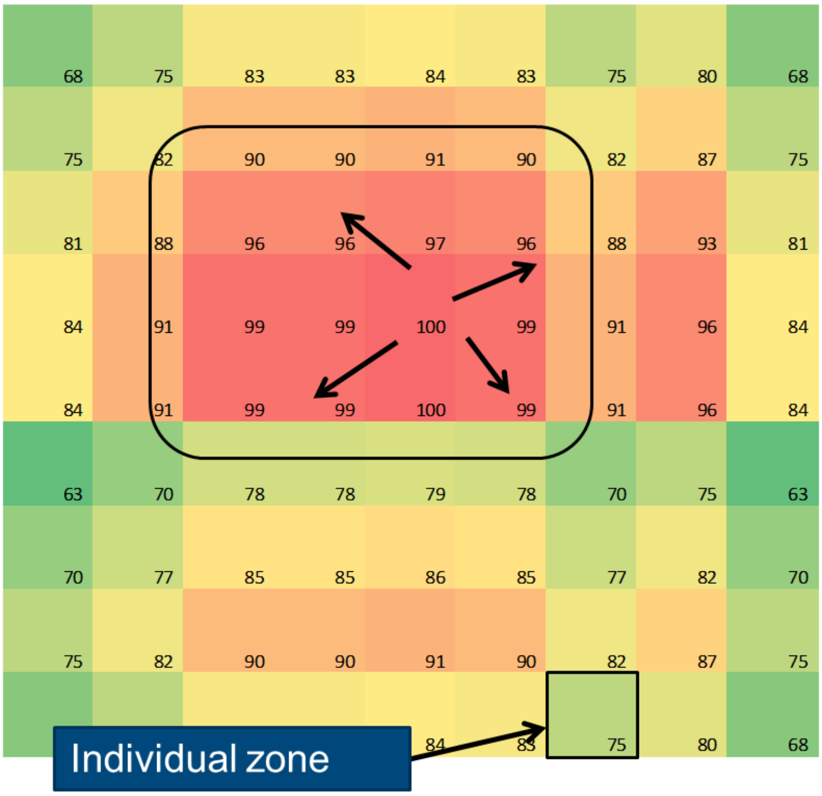
This section explains the ‘heat-led’ clustering process that takes place in the NAM. The first step is to split the area under investigation into ~50-100 grid squares based on the maximum likely clusters that can be computed in the POM. Zones are clustered by expanding outwards from a central zone in each grid square with the highest identified heat network potential, based on the metrics discussed further below.

Figure 7-18 Heat clusters – identifying heat ‘hotspots’



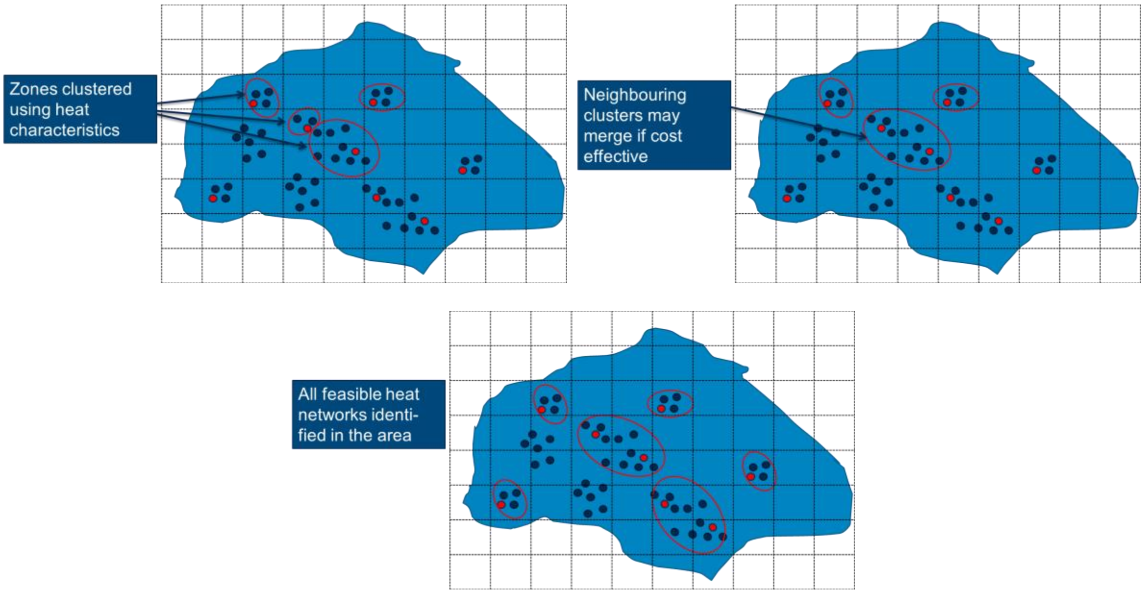
Starting from the central zone, additional adjacent zones are added to the 'test' cluster and the key metrics are re-assessed at each step as the cluster boundary expands (see Figure 7-19). At some threshold value for the metrics (e.g. rising £/MWh in terms of costs of the heat network) the cluster boundary is fixed. This mimics the real-world engineering approach of considering how linear heat density changes as the boundary of a potential heat network is changed.

Figure 7-19 Illustration of concentric circles methodology for the 'heat-led' clustering process



As part of the process of expanding the cluster boundary and re-testing the metrics for all zones within the 'test' cluster it is possible that these zones may merge with nearby test clusters. This process is continued until additional zones result in no improvement in the 'performance' of the identified clusters, and all feasible heat networks zones have been identified in the area (see Figure 7-20).

Figure 7-20 Heat clusters – growing heating networks and determining 'heat-led' clusters



We have identified two possible primary cost metrics for determining how suitable the cluster definition is for a *potential* heat network that would be contained within it as outlined in the table below. Neither of these metrics need to consider the cost of the heat supply source at this stage or the final heat connection point from 'plate heat exchanger to building' as these are a function of decision making in the POM. Hence the LHNC is a function of the main distribution and transmission pipe network only.

- 'Detailed' Levelised Heat Network Costs (LHNC) (£/MWh)
  - This would include an assessment of the necessary distribution pipe network and connection to transmission network / heat source.
  - Detailed steady-state modelling would enable pipes to be individually sized based on expected peak load and the total cost would be spread across the typical annual demand for all connected buildings in the cluster
  - The load weighted centre for the transmission connection would also be re-evaluated as the cluster grows (as outlined in section 7.5.1). Due to the high cost of transmission connection, larger areas benefit from spreading this cost over a larger MWh pool.
- 'Simple' LHNC (£/MWh)
  - Unlike the detailed approach, this simplification would not require detailed load flow modelling for each 'test' cluster definition as instead all zones would be assumed to have the same sized pipe network, sized to meet peak load.
  - In addition, it is assumed that the transmission network connection into a cluster remains the zone with the highest load (i.e. position does not change as

cluster increases), although the size of the transmission connection may still change with increasing load.

Both approaches will be investigated in Stage 2, however, the detailed levelised cost approach would add considerable complexity as detailed flow-modelling would have to be undertaken at each stage of 'testing' the cluster boundary definition.

Combined heat, electricity and other threshold metrics for defining boundaries

Regardless of whether a simple of detailed approach is used to calculate the heat network cost metric, a key conceptual consideration is around the threshold metrics that will be applied during the 'heat-led' clustering process to determine the boundaries of a 'heat-led' cluster.

This needs to consider not only the heat network 'suitability', but other factors such as the representation of the electricity network (so as not to bias the representativeness of the costs of either network in the cluster) and the maximum number of possible clusters. Potential threshold metrics that could be used for these purposes include:

- LHNC threshold for developing and operating the heat network in £/MWh (based on either simple or detailed calculation), which could be applied in a number of ways e.g.
  - Rising £/MWh, or rising for X consecutive expansions[33]
  - A cost gradient boundary, rather than a full inflection points of rising £/MWh costs.
- Total number of buildings in the cluster, or total heat demand (peak and annual) in the cluster to help physically bound the size of the cluster (and by extension limit the number of clusters to be evaluated in the POM);
- The physical boundary constraints of the electricity network, which affect both the way a cluster can expand to incorporate new zones as described in section 7.3.5, but which could also form hard threshold constraints
- An equivalent levelised electricity network cost within the expanding cluster as per that for heat, which could be contrasted in parallel – e.g. where the LHNC cost is still decreasing, but the gradient of the electricity metric is changing rapidly; this would indicate that the representation of the electricity network cost in the cluster was being sacrificed at the expense of the heat representation

At this stage, we believe it is important that the chosen approach will provide sufficient flexibility to ensure that the trade-offs and the economies of scale associated with developing and operating heat networks are fully taken into account, but which neither significantly underestimate or overestimate the electricity network costs in the cluster. To balance the trade-off between the representation of heat and electricity network costs in a semi-automated

[33] E.g. if the NAM only expands outwards considering 1-zone blocks it is likely that some economies of scale of having developed a larger heat network will be missed. Conversely, if NAM expands outwards considering a very large number of zone blocks then we would not be able to appropriately capture the optimal decisions for some of these zones.

fashion, we would expect to use a *combination of these threshold metrics for different clusters*, but would aim to develop the flexibility for the user to easily apply one or more and change the threshold settings.

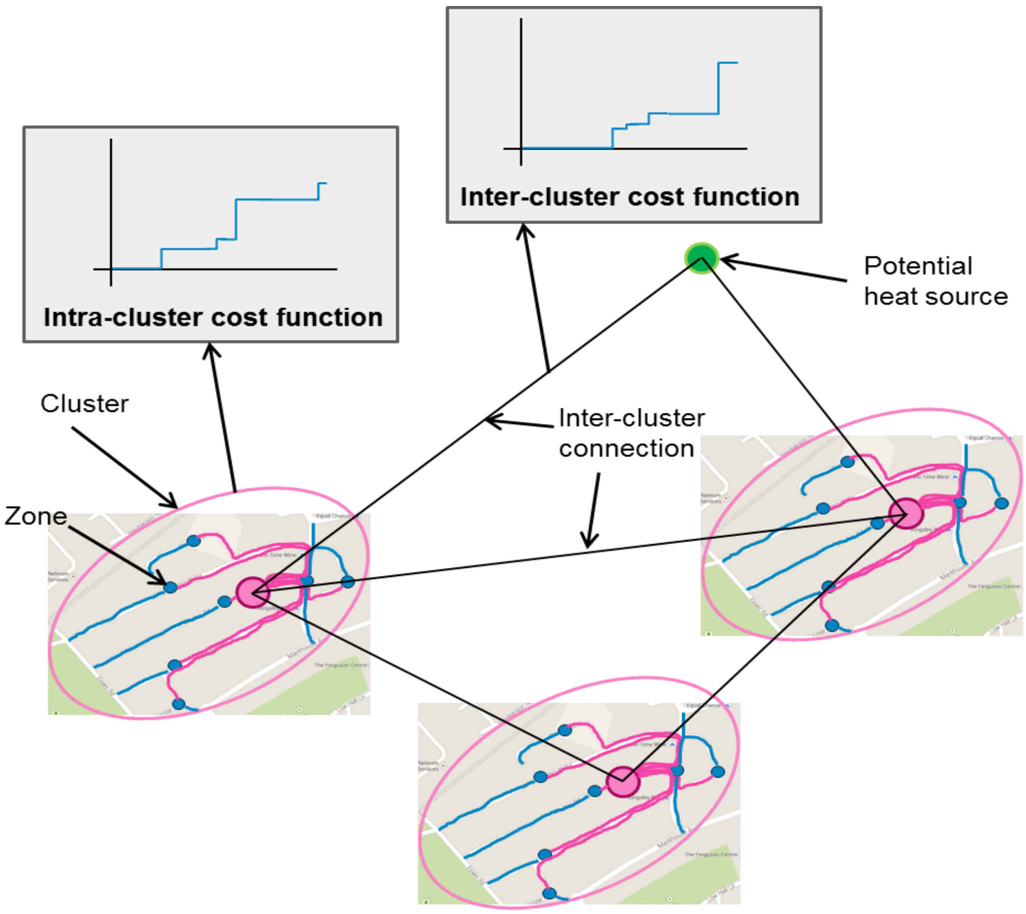
It is also important to note that the aim of the 'heat-led' cluster definition is not to define a sufficiently large enough cluster that represents an *entire network*, but a *potential component cluster of this*. By seeing all trade-offs between all clusters simultaneously, the POM can then make the decision to develop heat networks across multiple clusters, connect these clusters together, as well as to an appropriate heat source; all as part of creating an overall economically viable heat network (e.g. see Figure 7-21, section 7.5.7).

In practice, predicting the optimal set of network choices before modelling these decisions in POM will be challenging and to some extent depend on the topology of each Local Authority area. Expert user judgment will still be needed to define the initial clusters, examine the POM results and refine the clusters definition as part of the overall process of using the EnergyPath™ Design Tool. The above process aims to enable the user to undertake this as efficiently as possible.

Final cluster-level network cost options for heat

It is important to note that even if the 'simple' heat cost metric is used as part of the intermediate steps to define the cluster boundaries, *once the final boundaries have been defined the detailed load-flow approach (described in section 7.5.2) is still undertaken for the cluster as a whole and potential inter-cluster connections*; to assess the network cost options for the POM as accurately as possible. This includes the clusters defined by the heat-led process and the residual clusters defined by the SAM (see next section) as it is important to have costed options for heat networks development in *all* clusters (along with the equivalent for other energy vectors).

Figure 7-21 Developing cluster reinforcement cost functions for heat

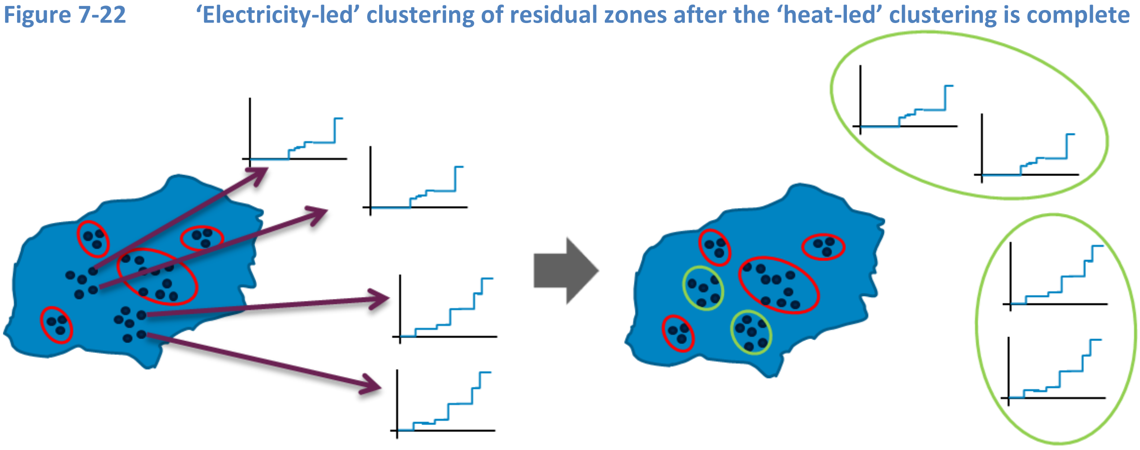


As described in section 7.3.1, the purpose of the detailed analysis is to evaluate not only the minimum pipe size (and appropriate pump size) for a given heat load within a cluster, *but also* outline a set of trade-off options (e.g. pipe size vs pump size vs network temperature) as these are ultimately economic trade-offs to be evaluated within the POM.

7.5.6. Zone network option electricity cost metrics [NAM-007]

As part of the heat-led clustering process described above, some zones may remain unallocated to a cluster. These residual zones are clustered based on their electricity costs using the GIS spatial statistical grouping technique in the SAM outlined in section 6.5.6. To enable this, the electricity cost functions created for each zone must be converted into a single cost metric, as the GIS grouping functionality is limited to this, to describe how expensive the costs of reinforcement are.

Figure 7-22 ‘Electricity-led’ clustering of residual zones after the ‘heat-led’ clustering is complete

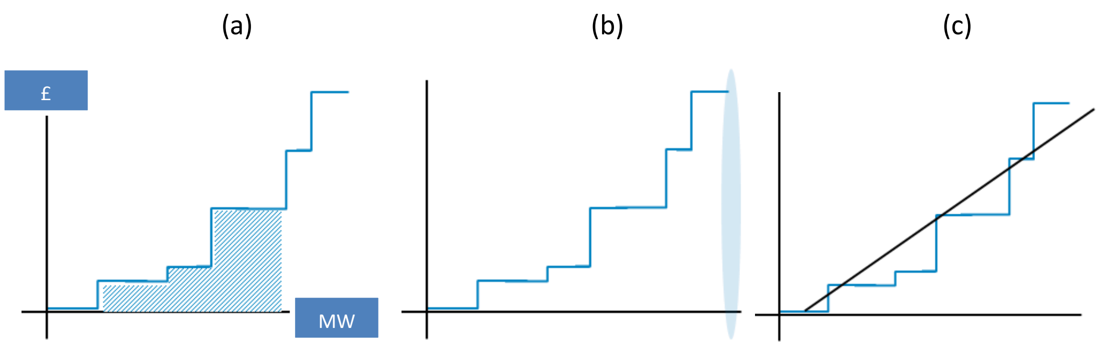


Possible zone network reinforcement cost metrics could include (Figure 7-23):

- The area under the derived cost function, potentially normalised by the MW values (a);
- The electricity network reinforcement cost at potential peak load (b);
- The derived cost function gradient (c);

These metrics (and potentially others) will be explored further in Stage 2 with real world sample data, however, the aim is to embed the most representative ‘shape’ of the cost of reinforcement. This means that a) and c) are likely to be more appropriate than b), and a) may be less sensitive to the value used to fix the biggest reinforcement option than c). As producing the different metrics is likely to be a fairly trivial calculation after the initial cost functions are produced, it would be possible to include multiple metric options and allow the user to decide which to utilise.

Figure 7-23 Possible zone network reinforcement cost metrics



In addition, the definition of a cluster cannot change between time periods within the POM hence if the zone cost functions and subsequent metrics evolve significantly over the pathway to 2050, the electricity-led clustering process must account for this. For example, the final metric used may be the average of the peak reinforcement costs (option b above) across all time periods under consideration.

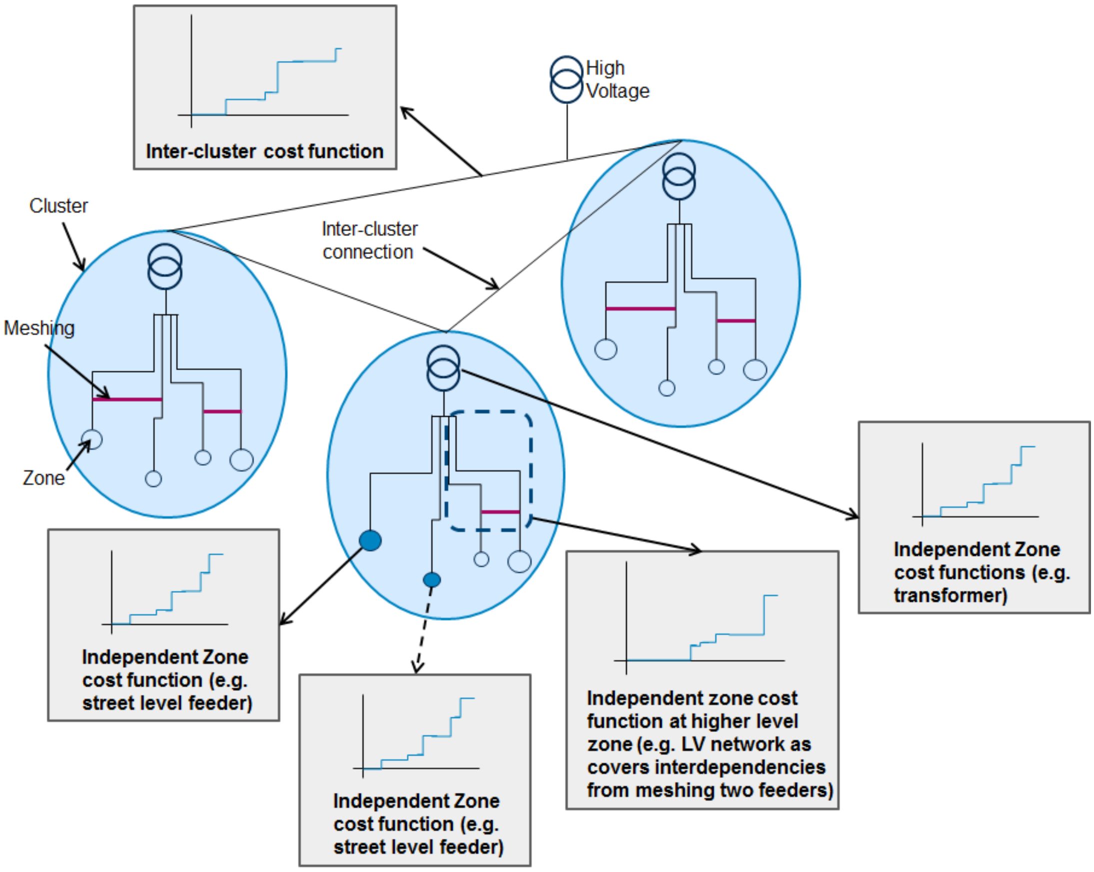
7.5.7. Cluster-level network option cost functions [NAM-008]

As the POM only operates at a cluster level of spatial granularity the final step is to aggregate any remaining zone level functions within the final cluster boundaries. For heat networks this is not necessary as the final cost functions are only assessed at the cluster level as described in section 7.5.5.

For gas and hydrogen networks the NAM produces "all-or-nothing" cluster cost options for (i) extending the gas network to cover off-gas grid buildings; (ii) decommissioning the existing gas network for operational costs savings; (iii) repurposing the existing gas network to allow it to be able to transport hydrogen. Given the lumpy nature of these decisions, cost functions at the cluster simple reflect the additive costs of all zones contained within the cluster, for each of the 3 options.

For electricity, the final cluster cost functions are constructed using the *independent* sets of zone cost functions that have already been developed as well as the cost functions for all inter-cluster connections (existing or potential) as shown in Figure 7-24.

Figure 7-24 Developing network cluster cost functions for electricity



As described in section 7.5.2, when undertaking the original load flow modelling, it is necessary to make a simplifying assumption with respect to how load at different levels of network connectivity (e.g. street level feeder) is distributed when aggregating up to the next level of the network (e.g. LV network feeders and substation). This is because the POM only calculates a single peak load (net embedded generation) value for the cluster as a whole.

For network options represented by cost functions at lower network levels (e.g. LV network feeders) driving reinforcement directly from the cluster-level peak load is not appropriate and would over-estimate the required reinforcement. Hence some of the zone cost functions themselves have to be adjusted before they used at the cluster level so that their interpretation as a function of cluster level load is appropriate.

If cluster load is assumed to be distributed amongst zones using a fixed % for each zone, the zone cost function can be "stretched" to be converted from MW_{zone} to $MW_{cluster}$

- For example if zone load = X% of cluster load, then $MW_{cluster} = \frac{MW_{zone}}{\text{X\%}}$ for each zone.

Importantly the load distribution used *must reflect that used in section 7.5.2 as part of the original load flow modelling.*

7.6. Proof of concept activity

The key tasks within the NAM proof of concept study can be summarised as follows (further information is provided in the Appendix – Section 13.2):

- Test and help to select a suitable commercial network load flow tool that can provide the basis of the operational analysis required in NAM
    - A commercial network flow model considerably simplifies development and allows us to focus on interactions between the load flow model and the scenario costing tool in order to produce cost functions;
    - During the testing phase we have considered a number of criteria including: (i) performance; (ii) ability to handle GIS data sets; (iii) ability to model multiple energy vectors; (iv) ease of inputting and outputting data (particularly in Excel format); (v) support services provided by the software developer; (vi) cost
    - The two most suitable software tools that have been identified and tested are Siemens' PSS SINCAL and BCP Switzerland's NEPLAN.
    - Wider assessment of the two packages is described in deliverable (D2) Design Architecture
- We have carried out a range of simple load flow studies under different networks to understand what the most limiting network constraints are likely to be;
- Based on these load flow studies we have then run a simple internal process whereby a number of mock up test configurations and network reinforcement options were evaluated for a given area;
- Based on the above we have then created a set of representative network cost functions for the area in question in order to describe the obtained costs for the range of potential peak demands;

- Finally, we have also considered how the 'concentric-circles' methodology could work in practice by outlining the process by which 'heat-led' clusters can be created based on their heat network characteristics.

7.7. Output validation checks

As explained in section 7.2, the key outputs from NAM can broadly be summarised in two categories:

1. <u>NAM outputs related to cluster definitions:</u> Some zones are clustered using, primarily, heat network characteristics ('heat-led' clustering) while others are clustered using electricity network characteristics ('electricity-led' clustering).

2. <u>NAM outputs related to cost functions:</u> The NAM is also responsible for outputting cost functions at the intra-cluster and inter-cluster level for each energy vector. The NAM is also responsible for producing cost functions at the zone level (which are then used for determining intra-cluster cost functions) however zone cost functions are not sent to POM as they are not at a high enough level of aggregation.

For both of these categories, output validation checks must be undertaken to determine whether the outputs from the module are fit-for-purpose for the overall modelling solution. This is because it is necessary to determine whether the derived cluster definitions and cost functions appropriately represent all energy vectors considered here, and particularly so for both heat and electricity which are modelled in detail.

In summary, the key checks that must be undertaken include:

- Ensure that all derived clusters respect the boundary constraints set by electricity network characteristics (as explained in Section 7.3.5). Clusters must be a collection of whole building blocks, or part of a single large building block, but cannot be a collection of partial building blocks;

- With regards to the derived cost functions it is also necessary to validate their accuracy in terms of: (i) materiality of load flow modelling; and (ii) materiality and uncertainty of network component costs. In practice, the former means checking that all network constraints that could have a significant impact on costs are taken into account for each test load flow simulation, whilst the latter means checking that any other material physical or geographical factors that may affect component costs are included in the modelling approach, which is likely to be undertaken in conjunction with the Local Authority.

- Finally, NAM is also responsible for performing more detailed testing of the solutions proposed by POM to ensure their feasibility and to potentially refine the network solutions. This is because detailed load flow modelling cannot take place in POM as it needs to be done at a lower level of granularity. For Stage 2 it is proposed that this is a largely manual, expert-user processing step. Over the longer term, it would be possible to automate the way the results from the POM are passed back to the NAM to simplify the validation process. However, it is unlikely to be appropriate to automate all subsequent refinements as this will, in part, be a result of expert user insight as opposed to an automated calibration.

Baringa

8. Pathway Optimiser Module (POM)

8.1. Overview

The EnergyPath™ Design Tool will support strategic local area energy system prioritisation and planning, focused in particular on choices for buildings and networks, alongside local area energy generation and storage, over the pathway from now to 2050.

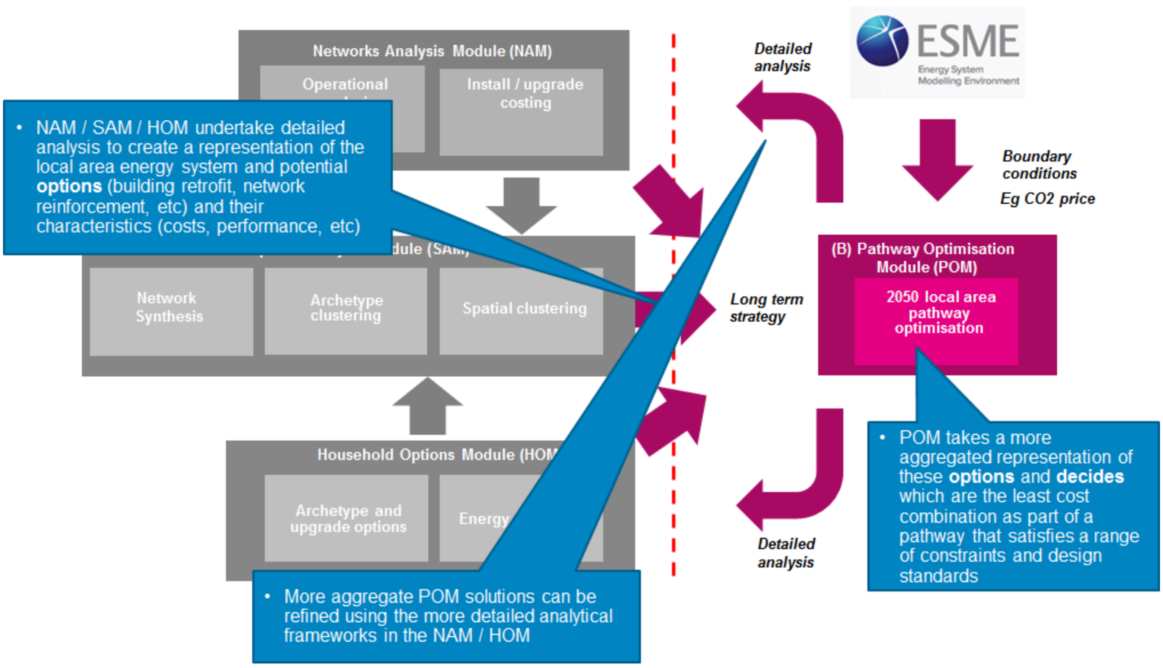
As described in the previous sections the HOM / SAM / NAM modules help to create a detailed spatial picture of the existing local energy system, along with the ***options*** available for evolving the system over time such as building retrofits or installing a new district heat network. The other modules also create the necessary data about the options (cost, performance and various other characteristics) so that the POM can compare and trade-off these options against each other. However, to make the pathway analysis a tractable problem the level of detail needs to be simplified, for example, by reducing the spatial detail from individual units or zones (see Figure 6-2) to a smaller set of clusters.

As shown in Figure 8-1, POM is effectively the automated decision-making engine focused on

- **What *options*** should we deploy, **where** and **when**?
- For some ***options***, **how** should we utilise these once deployed?

The process needs to consider whether the options that are chosen as part of an overall viable energy system pathway for a local area are collectively appropriate (e.g. cost-effective and feasible) and satisfy various other goals (ensure households are provided with sufficient comfort, is consistent with meeting our climate change targets)

Figure 8-1 POM context



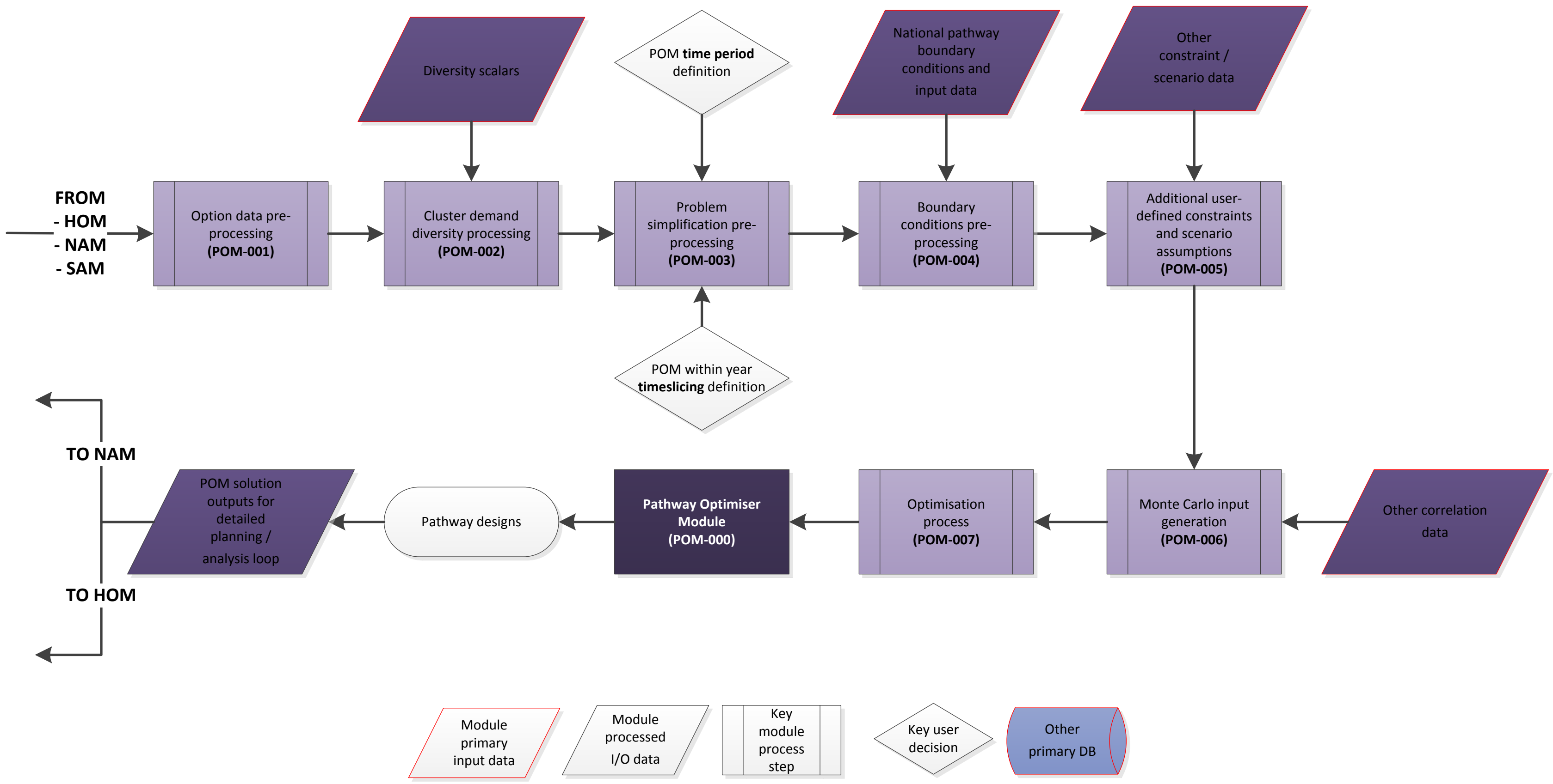
To make the problem of comparing available multiple options more tractable, their representation in the POM must be simplified relative to the other modules, for example, dividing the area into relatively small number of clusters (e.g. 50 clusters versus 1000 zones). Simplification naturally involves a loss of detail and may then affect the options chosen. To counter this two key features are proposed:

- In-built flexibility to vary easily the level of aggregation/simplification and understand how the solution changes as you move along this spectrum
- A detailed analysis loop whereby the more aggregated solution from the POM can be investigated in more detail in the underlying NAM / SAM modules, for example:
    - Exploring the feasibility of the aggregate network reinforcement solution for a cluster from the POM back in the NAM, by using the more detailed network flow modelling and consideration of underlying zones and network topology within the cluster

8.1.1. Module diagram

The diagram below shows the key logic / process steps for the POM, the equivalent for the full tool covering all modules is shown in section 12.

Figure 8-2 Key POM logic / process steps



8.2. Key outputs

The POM provides the main outputs from the EnergyPath^{TM} Design Tool, which are ultimately used to help inform decision making by Local Authorities and other key stakeholders. This is discussed in more detail in section 4, but at a high-level the key outputs from POM must be able to be re-packaged and interpreted such that they help understand:

- What are the high-level features of the overall area viable energy pathways?
  - Deployment of specific network, building, other options (or more broadly ***'types'*** of pathway design such as heat network focus versus electricity reinforcement)
  - At what point should they be built over the pathway to 2050
  - At what point should they be built over the pathway to 2050
  - What are the direct impacts such as emissions or energy 'costs'?
  - When POM outputs are combined with socioeconomic or other data, what are the distributional impacts, e.g. the potential effects on fuel poverty?
- What are the key geographical and underlying features of the pathway, in particular where in the local area are specific options being deployed?
- What are the key areas of uncertainty (this is discussed in more detail in section 9)?
  - How confident are we that something is a 'low regrets' option within a pathway (in terms of the cost of the energy system or other metrics), as well as how resilient or sensitive the option is to changing key input assumptions (such as the fossil fuel prices rising faster than anticipated)
  - How do we understand what the value / impact is of reducing uncertainty by undertaking further primary research (e.g. surveys to better understand the cost of retrofitting local buildings)?
- What does the pathway design mean in terms of specific projects?
  - The outputs need to contain sufficient information on individual options, or groups of options, that can be packaged into potential projects. This should cover data on the type of option, capital and operating costs, size or number, deployment dates, etc. This would be used to target more detailed FEED (Front End Engineering and Design) studies to enable the specific projects to undertaken.
- In addition, a number of POM outputs are required for the detailed analysis loop in the NAM and HOM (see section 8.8 for further) details. These are mentioned explicitly in the output table below, but for example, include the cluster-level network options selected over the pathway and the evolution of peak demand and the associated electricity demand profile within the cluster itself.

The table below outlines the fundamental outputs of the POM, which can be processed and combined in a number of different ways to provide the insights required and generate some of the more specific outputs described in section 4.

Essentially, all information regarding all the options and choices made in the POM (see section 8.5.2) are available in the outputs (e.g. what, where and when a technology option is deployed). These can then be combined with other static input parameters (e.g. costs of deployment or emission factors) to understand the impacts of the selected options (e.g. total costs of deployment of pathway or CO2 emissions).

Further optimisation-specific information can also be generated such as the shadow (marginal) price of key variables such as electricity, and the slack remaining on constraints (see section 8.5.3) to understand which of these are *binding* for a given pathway in the POM.

Table 8-1 Output data summary

ID	Data field	Destination	Purpose	Data granularity	Uncertainty parameters
1	Stock of building archetypes and conversions	• End-user	• Components of pathway design	Stock of archetypes that exist in given timeperiod as well as timeperiod of conversion of archetype pairs (from / to) in each • Cluster	Simulated outputs – 1 for each set of Monte Carlo inputs used
2	Choice of building operating profile	• End-user • HOM	• Operation within pathway design • Detailed analysis (test whether simplified operating profile chosen appropriate given overarching system parameters such as electricity costs)	Where >1 operating profile, data by • Archetype vintage • Cluster • Timeperiod • Daily timeslices	As above
3	Stock of network options and network new build options	• End-user • NAM	• Components of Pathway design • Detailed analysis (test whether network design options are feasible, most appropriate given overall load evolution in cluster)	Stock of network options that exist in each timeperiod as well as timeperiod of build in each • Cluster • By each energy vector (electricity, heat, gas)	As above
4	Energy / supply demand balances	• End-user • NAM	• Components of Pathway design • Detailed analysis (test whether network design options are feasible, most appropriate given overall load evolution in cluster)	Energy inputs and outputs to each technology option (building, embedded generation, etc) by • Cluster • Timeperiod • All timeslice Intercluster network flows by • Timeperiod • All timeslice By extension other necessary results can be calculated • energy consumption, • resource use • CO2 emission results, • Peak demand Note that level of timeslicing may vary by product (e.g.	As above

				full within day for electricity/heat, daily for gas)	
5	Stock and new of other energy system features (CHP, DHN boilers, network storage, etc.)	• End-user	• Components of Pathway design	Stock of technology options that exist as well as year of build each • Cluster • Time period By each energy vector (electricity, heat, gas)	As above
6	Operation of other energy system features	• End-user	• Operation within pathway design	For flexible technologies, the level of utilization by • Technology vintage • Cluster • Timeperiod • Timeslice (level of timeslicing dependent on the products the technology sees as an input/output)	As above
7	System costs	• End-user	• Understanding of pathway impact	Underlying cost components of deployed options (e.g. CAPEX, FOM, VOM, fuel costs) can be calculated for each option by • Cluster • Timeperiod	As above
8	System shadow prices	• End-user • HOM	• Understanding of pathway impact • Detailed analysis (test whether simplified operating profile chosen appropriate given overarching system parameters such as electricity costs)	Data granularity defined by granularity of constraint which generates shadow price – i.e. if imposed by timeperiod, cluster, timeslice – results will be available on this base An example is electricity within the main S / D balancing constraint, which would be available at the most detailed level of temporal and spatial granularity	As above
9	Constraint slack	• End-user	• Understand which constraints are *binding* the pathway design, and if not how close are they to binding	Data granularity defined by granularity of constraint – i.e. if imposed by timeperiod, cluster, timeslice – results will be available on this base	As above

8.3. Key functional requirements

8.3.1. Optimisation

For a given set of input assumptions to the POM (cost of available options or wider exogenous assumptions such as fuel prices) there are likely to be too many trade-offs to explore all possible combinations of options manually, to understand what the 'best' combination is for a pathway.

An analogy is the DECC 2050 calculator pathway calculator[34], which contains approximately 40+ levers and 4 basic settings for each lever. Rationalising this limited set manually is time consuming enough without adding further complexity to the choices such as when or where over the pathway an option is undertaken, which is of critical important for the EnergyPath™ Design Tool. To add further complexity, the significant uncertainty over multiple input parameters over the pathway to 2050 means that resolving the 'best' combination of options needs to be repeated under multiple sets of input assumptions.

Optimisation is considered to be the most appropriate technique to help resolve the huge number of possible trade-offs. It is effectively just a 'calculator' that can resolve all possible combinations of choices simultaneously, whilst maximising or minimising some objective, such as profit or cost. Constraint can be added to restrict inappropriate combination of choices to reflect real world issues or design standards that must be achieved, e.g.:

- "No building of technology X before year Y in location Z"
- "CO2 emissions must remain below limit of X"

Provided the underlying optimisation model is a good 'representation' of the problem this then allows the user to focus on how the optimiser design output varies under different sets of inputs assumptions. For example, in the context of the EnergyPath™ Design Tool a good representation of a local energy system is one which ensures that supply and demand are balanced across the system appropriately (meeting peak demand requirements) and, as an direct result, ensure that householders' energy service demands and comfort requirements are met.

8.3.2. Efficient uncertainty analysis

'What-if' analysis will be a fundamental part of the process of using the POM, for example:

- Undertaking single deterministic runs changing one or more input assumptions and exploring the impact on the pathway design
- Extending this simple what-if analysis to understand the resilience of a pathway design (by fixing key elements of the solution such as the deployment of a heat network) and understanding how the outputs of the POM (e.g. the cost of the energy system) change in response to different input assumptions (higher gas prices)

---

[34] http://2050-calculator-tool.decc.gov.uk/pathways/11111111211111111111111111414111111111111111111111111/primary_energy_chart

However, there a large number of possible input parameters that could be changed and a wide range of potential uncertainty on these parameters over a 30+ year pathway, in particular the cost of technology or fuel prices.   The POM therefore also needs to allow the user to explore a wider set of uncertain input parameters as efficiently as possible.

It is proposed to enable this via Monte Carlo simulation of a number of key input parameters. This would create a number of *sets* of input parameters each of which would be optimised deterministically. The outputs would comprise a *set* of results from which distributions, rather than discrete numbers could be established.

Finally, we are not proposing to undertake a formal *stochastic optimisation* within the POM as opposed to the *probabilistic simulation* process we have described above. This is for two main reasons:

- The rationale for applying a formal stochastic optimisation is to understand the optimal 'hedging decision' in the near term given future uncertainty, and recourse strategies should the future turn out to be different to that expected[35] - i.e. the best decision making given what we *are expecting* to happen. By contrast the uncertainty which needs to be explored in the EnergyPath™ Design Tool is subtly different and is primarily about understanding the overall pathway design that a local area should be *aiming for*, over a long-term pathway, considering the inherent uncertainty over this extended time horizon in the round

- The problem size increases exponentially, quickly becoming intractable and limiting the number of variables that can be considered stochastically, whereas a probabilistic simulation scales linearly with the number of discrete optimisations to be undertaken (and which can potentially be solved in parallel)

8.3.3. Accurate energy system representation

The more accurate the representation of energy system choices (and by extension trade-offs) in the POM the better the optimiser solution, in particular the accuracy with which the costs and the fundamentals of energy system operation are reflected.

A more accurate representation is generally achieved via combination of more granularity and adaptations to the structural representation of the optimization problem, such that it mimics reality as closely as possible:

- **Granularity** – increased granularity can be achieved by having a more detailed spatial representation (more clusters each of which covers a smaller area), more timeperiods (i.e. annual rather than 5- or 10-yearly steps on the pathway), or more timeslices within year (more characteristic days and periods within day)
  - Finer clusters offer more accurate information on network reinforcement/new build costs

[35] It is often used for decision making around energy storage – i.e. do I inject or withdraw now given my expectations about future energy prices

- More timeperiods offers a more accurate representation of how costs will evolve over time and the impact on the timing of investment (as opposed to averaging these costs over a number of years)
- More timeslices within year offers a more accurate representation of energy balancing and by extension the costs of operating the system. In particular there is a trade-off between POM choosing how best to operate a flexible device given the wider system conditions and pre-judging its operation. This is important to understand the suitability of more flexible technologies such as hybrid boilers, storage, micro-CHP and the impact on the demand profile, net peak demand[36] and emissions

- **MIP (Mixed Integer Program)** representation for some of the optimisation variables which can provide a better representation of the cost of options than an **LP** (Linear Program) representation, for example:
  - *Lumpy* investment of some options (e.g. new large scale CHP) as under an LP formulation the POM could build any idealised fraction of a new plant it wanted, which would likely underestimate the costs
  - Discrete (and potentially interdependent) choices such as being able to fully repurpose the gas grid in a cluster to use H2, which in reality is likely to be 'all-or-nothing' decision. The alternative under an LP representation is an idealised partial repurposing of the network in a cluster, which is not feasible from an engineering standpoint
  - Economies of scale, which it is not possible to represent within a standard least cost LP formulation. This is likely to be important at a local level to understand better the trade-off between piecemeal investment spread over a longer term horizon (as there is time preference to delay costs where possible) versus widespread early expansion with higher upfront costs, but which is potentially cheaper when considering the entire pathway

The basic representation of an energy system structure within the POM is outlined in Figure 8-3 and is essentially the same as that within the ESME model or MARKAL/TIMES energy system models. Starting from the left hand side are a number of individual energy resources (e.g. biomass, gas, wind, etc). These can be used directly in an end-use technology (e.g. gas in a building archetype) to produce the required energy services (e.g. space heat, hot water, etc) or converted via intermediate routes into a number of energy carriers (e.g. gas in a power plant to produce electricity) which are then used in the final end-use technology. A number of intermediate conversion steps can exist (e.g. the earlier electricity step could be used to electrolyse hydrogen which is then used in a boiler or fuel cell). As part of the transport of resources and energy carriers across the system network technologies may also exist, such as an electricity distribution network.

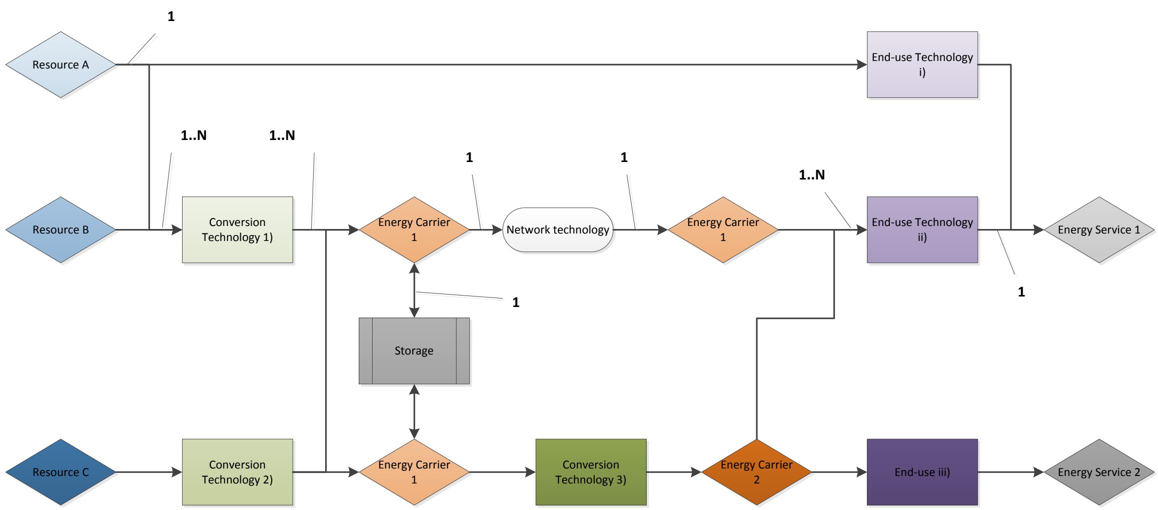
[36] I.e. peak demand in the cluster net of embedded generation, and the resulting implications for network reinforcement and new build

In theory the conversion technologies can reflect a many-to-many relationship with inputs and outputs. I.e. they can take in one or more inputs (resources or energy carriers) and produce one or more energy carrier outputs (e.g. in a co-fired CHP plant). Similarly, end-use technologies may be able to use more than one input (e.g. in a hybrid gas boiler). In addition, storage technologies exist which can temporarily store and release an energy carrier or resource (e.g. electricity or gas).

The POM decides what resources to use, what technologies to build (conversion, storage, network, end-use), how these are operated (and indirectly what energy carriers to produce), such that *all* energy service demands specified within the scenario are satisfied. The key constraint that must be satisfied across the entire energy system structure as part of this is that supply must also equal demand for all other energy carriers and resources.

For example, if 100 units of energy carrier 2 are required by end-use iii (to produce sufficient quantities of energy service 2) these 100 units must be produced by conversion technology 3. By extension sufficient inputs of energy carrier 1 must be available to feed into conversion technology 3, with the exact quantity dependent on conversion technology 3's efficiency. The process extends back through the chain such that, a sufficient quantity of energy carrier 1 must have been produced from elsewhere and so on, back to the original resources, which must also be available in sufficient quantity.

Figure 8-3 Example of basic energy system model structure



For further details of a standard energy system modelling structure please refer to the ESME Functional Specification document.

8.3.4. Other

Two other key related requirements for the POM module are that it is

- Easily scalable with respect to
  - Adding more 'options', which can be considered as part of the trade-offs in the creation of a pathway (e.g. building archetypes, network options, fuel types, etc)

- Granularity in terms of the numbers of spatial clusters, timeperiods and timeslices. This has significant implications for POM performance, but flexibility to simplify (or increase) the complexity of the problem will be built directly into the EnergyPath™ Design Tool and is discussed further in section 8.6.3. This flexibility is important because it is difficult to say ex-ante where additional detail matters most and may vary by local area given differences in underlying data.

- Reasonable performance limits
  - POM is by definition trying to resolve a highly complex optimisation problem, but to allow 'reasonable' user interaction this leads to overnight solving (~15 hours) as a typical working threshold[37]. For probabilistic simulations there is also the ability to distribute the optimisation problems across more than one 'machine or cloud instance' in parallel, but this incurs additional cost
  - The scalability described above along with other routes to simplify the optimisation problem, allows performance to be tailored and the POM used in different ways. For example, a 'simpler' representation of the local area energy system can created and more simulations explored to understand the implications on the pathway design due to a wider range of uncertainty, versus a very detailed representation for 1 deterministic scenario.
  - MIP optimization in particular can lead to a significant performance decrease and we propose to retain a parallel MIP/LP representation for key variables as a further simplification option. But, as a result we are not considering the use of **NL-MIP** (Non-Linear Mixed Integer Programs) even though this may in some circumstances allow us to better represent reality[38].

8.4. Key inputs

The SAM / NAM / HOM modules provide the majority, but not all of the input information necessary for the POM, which can be categorised into 3 main areas:

- Direct inputs from a single module, which provides all necessary data
  - This may include the parameters necessary to undertake Monte Carlo simulation of the inputs
- Partial inputs from other modules, which need to be combined to provide a full set of necessary data
  - The SAM primarily provides spatial IDs/Flags for the components of the energy system (existing or potential options) that are contained within a cluster such as the number a given archetype. These are combined with other datasets (e.g.

---

[37] This is broadly comparable to an ESME pathway simulation run with ~100 simulations

[38] For example representing a non-linear cost function in its original form versus a piecewise-linear approximation of this function.

from the HOM) to create all the information that is needed by the POM (e.g. the costs and demand profiles associated with the base archetype and possible conversion options)

- Primary data inputs including
  - Those necessary for pre-processing of POM options (such as diversity factors or other technology information on existing non-building, non-network energy system features)
  - National pathway boundary conditions from ESME
  - Others user-defined scenario inputs which are not already generated as part of other modules such as technology / option build rate constraint values or Monte Carlo correlation factors

The tables below describe what is technically required within the tool and how it would be used. The relative importance of the primary data inputs and other factors such as costs and licensing restrictions, particularly where they must be purchased from an external provider, are discussed in deliverable (D4) Data Acquisition Plan.

Table 8-2 Input data

ID	Data field(s)	Input type	Purpose / sub-fields	Granularity	Uncertainty parameters	Coverage	Source	Limitations
1	Cluster-specific ID / flag information	Secondary	Provides the POM with the data on number of clusters existing / possible interconnections between clusters and ID/ flag information on • # of each existing building archetype in cluster X • Spatial restrictions on archetype retrofit / other options as inputted by the user within the SAM • Other existing non-building / non-network energy system component (IDs only) • Future energy system components that ***will*** be constructed (e.g. new build housing zone and base archetype) • Future energy system options for siting of new technology that ***could*** be constructed (e.g. new CHP, energy storage, etc) or with possible local resources, which could be exploited (e.g. biomass, waste, geothermal) I.e. the SAM provides information on where energy system features are sited by cluster (existing, future deployment or potential future option) which can be combined with other data to create the full set of characteristics	Data by cluster, by timeperiod	N/A	By cluster and inter-cluster	SAM	Cluster definition naturally loses spatial information within it
2	Database of option characteristics for non-building, non-network options (*both* existing and new)	Primary	Combined with the ID flag provided by the SAM this creates the list of technology characteristics used to define the option in POM • Costs (CAPEX, FOM, VOM) • Type (flexible profile, fixed profile, storage) • Modes of operation • Input / output products by mode	Data by technology option / resource product, by timeperiod (different characteristics by vintage), by cluster	A number of parameters could be subject to Monte Carlo simulation – particularly Costs and Efficiency for future vintages	Covers embedded generation/CHP, other heat network sources, network storage (electricity, heat), etc	*Existing* data will have to be compiled in conjunction with LA and DNOs *Future* data from various ETI sources (e.g. UK	N/A

			• Efficiency by mode / (losses for storage) • Lifetime (economic, technical, construction period) • Annualisation factor • Max availability (for flexible profile plant) • Load factor (for fixed profile plant) Separately this also need to cover resource availability by cluster				ESME database or specific technology programmes)	
3	Building archetype option parameters	Secondary	Combined with SAM ID for number of each existing archetype in each cluster to create all archetype characteristics needed in POM • Viable 'archetype conversions' pairs • Costs of converting one archetype into another • Archetype final energy demand profiles (e.g. gas, electricity for heat, electricity for EVs, electricity for other, network heat, biomass, etc)	Data by archetype cluster, by timeperiod Demand profile(s) for each half hour in each characteristic day	Relevant Monte Carlo parameters provided in relation to archetype conversion costs, final energy demand profiles	All relevant archetypes	HOM	# of remaining archetypes used to define area are subject to simplification in HOM Fixed choice of limited demand profiles to simplify optimisation problem rather than full dispatch of intra-building heat device / storage in POM
4	Archetype demand diversity scalars	Primary	Used to adjust HOM base demand profiles for diversity effects on the network to avoid overestimating peak demand, separate scalars for • Heat-related electricity demands • EV electricity demand • Other electricity demand • Heat network demand	By archetype, by characteristic day timeslice	N/A		Existing engineering estimates (e.g. WPD, Ramboll) Analysis via Monte Carlo simulation of demand profiles in HOM Survey data	
5	Network	Secondary	All data necessary to define network options	Data by network	Relevant Monte	Electricity, heat,	NAM	Cluster definition

	reinforcement and new build options		within POM • Costs (CAPEX, FOM, VOM) • Lifetime (economic, technical, construction period) Options already contain data on existing network features	option, by timeperiod (different characteristics by vintage), by cluster and by intercluster connection	Carlo parameters provided in relation to network costs, based on uncertainty of key component costs described in section 7.3.4	gas, hydrogen		naturally loses spatial information within it
6	National Pathway boundary conditions	Primary	To help ensure pathway design at local level aligned with national level pathway. Range of possible conditions, but is likely to contain at a minimum • Carbon shadow price • Cost, availability and carbon intensity of transmission connected electricity and heat at boundary of local area • Cost of increasing available peak supply at boundary • Fossil fuel and biomass prices • EV ownership • Potentially also relative cost trends by timeperiod for key technologies	By timeperiod By timeslice, by cluster where relevant	Based on national pathway simulation results	Local Area	ESME	More limited resolution (spatial, temporal, technology aggregation) means some intermediate processing is needed to translate direct ESME values into the EnergyPath™ Design Tool boundary conditions Need to ensure consistency of national level inputs (e.g. fuel prices) which generate a given set of national outputs (e.g. carbon shadow price)
7	Other user defined scenario assumptions	Primary	This includes other necessary scenario assumptions not already defined through data in other modules passed to POM. In particular • Local area industry final energy demands over time (by fuel, profile) • Key pathway constraints such as design standards ('peak winter'), deployment rate constraints reflecting supply chain limitations for the area as a whole, etc • Correlation factors for Monte Carlo	By technology option, by timeperiod, by cluster, by timeslice as relevant	Some constraint values could in principle be simulated Industry energy demands could be simulated	Various	Various – ETI, LA, etc	N/a

			simulation					

8.5. Key optimisation components

8.5.1. Objective function

As concluded as part of the EnergyPath™ Design Tool scoping workshops the core objective for the creation of a pathway design is to minimise total energy system 'economic resource costs'. The core cost components of the POM objective function are

- Option costs (e.g. building archetype retrofits, network reinforcements, etc) covering CAPEX, FOM (Fixed Operating and Maintenance), VOM (Variable Operating and Maintenance) costs.
- Resource costs for fossil fuels, biomass, and energy inputs that occur at the boundary of the local area (e.g. national transmission system electricity), this needs to consider the cost of both the electricity supply as well as expanding capacity for peak supply at the boundary.
- Carbon price (note that POM will also include the ability to cap absolute emissions)
- Placeholders will also be created for other 'societal costs' which can be monetised and linked to a particular driver within the POM, so that they can be factored explicitly into the least cost optimisation. For example, air pollutant emissions for NOx and SO2 could be estimated from underlying fuel use and UK Government estimates of monetised damage costs applied.

The impact on individual agent incentives or wider distributional impacts, and by extension the consideration of profit or consumer/producer surplus maximising behaviour, is not considered directly within the primary objective. *This means that taxes, subsidies and other distributional transfers are not included.* By extension the cost of land is not included where this represents a straight transfer from buyer to seller. However, it may be useful at a later stage to value the opportunity cost from the perspective of 'society within the local area' of converting land from one use to another – e.g. the conversion of productive farmland versus a brownfield site to an energy centre. This could be undertaken within the modelling framework by adding a location specific 'resource cost' for development. The values themselves are difficult to estimate and it becomes important that such costs are applied appropriately across the local area so as not to bias the cost optimisation.

The objective function will consider the minimisation of costs across the NPV (Net Present Value) of pathway costs from the base year until 2050, rather than trying to minimise the costs in each 'spot year snapshot'. As part of this process the CAPEX will need to be annualised (with varying discount rates by option[39]) to avoid pathway boundary issues such as incurring the full CAPEX cost of investment made in 2050 even though the investment will last beyond this point.

[39] This could be used to indirectly test consumer preference with respect to technology choice, e.g. domestic building retrofit options could have a much higher discount rate than non-domestic options or network upgrades (particularly as the latter are operated under pre-defined rate of return business models)

In addition, an overarching societal discount rate will need to be applied reflecting the preference for delayed costs. The benefit of minimising the total pathway costs is that this better captures the trade-offs around timing of investment such as early oversizing versus incremental upgrade. This representation will be enhanced further if it is also possible to represent economies of scale as discussed in section 8.3.3.

8.5.2. Key decision variables

The decision variables represent the choices that POM is able to make as part of determining the least cost pathway design given a set of input assumptions, these are summarised in the table below.

Table 8-3 Summary of key POM decision variables

Item	Timeperiod	Location	Timeslice*
Building archetype conversion and potential operation	Choice in each timeperiod for converting one archetype into another (at given cost and subject to valid conversions)	Separate choice of archetype conversion in each cluster	For some archetypes there will be the choice of >1 'operational' profile to reflect a number of dispatch states for operating flexible heating/storage devices. To simplify the optimization problem the choice is effectively a profile choice for the whole characteristic day as opposed to a unique 'dispatch' decision for each timeslice within each characteristic day
Networks build and operation	Choice in each timeperiod for reinforcing or deploying new network features for various energy vectors	Different choice of network option by cluster as well as options between clusters.	Additional decision variables to track the effective supply/demand balancing of network energy flows between clusters, for each relevant timeslice (for electricity and heat this covers both characteristic days and within day, but for gas only daily balancing is necessary)
Resource use	Choice of quantity of resource to be used. The degree of timeslicing will be dependent on the product – e.g. biomass, waste, geothermal and coal only need to consider supply/demand balancing at the annual level	Resource use may be cluster dependent in some cases – e.g. for biomass or geothermal, whereas others such as coal choice is at aggregate local area level only	Gas supply/demand balancing is considered at the daily level
Boundary heat / electricity	Choice in each timeperiod for use of resources given available capacity Choice in each timeperiod to reinforce boundary to access more peak demand and annual supply	Choice in clusters that are effectively connected at the boundary, network flows then ensure supply/demand balancing between clusters to point of demand	Choice for each timeslice within each characteristic day
Other network embedded generation / storage technologies for e.g. electricity	Choice in each timeperiod for build of technology	Choice within each cluster	Choice of dispatch by timeslice for flexible technology (e.g. CHP / storage)

and heat			Fixed profile dispatch for others (e.g. embedded wind / solar)

A number of additional complexities exist with respect to the deployment of a particular option, be it building archetype conversion, network or other technologies:

- For some options (e.g. network build or embedded generation) the variables will also be ***vintaged*** by the year of deployment. For example, CHP plants may be built in 2015 and 2020 and both will still be operational in 2025 given operating lifetimes. The 2015 and 2020 vintages of the plant will likely have different characteristics (e.g. the new plant may be more efficient) and hence the choice of how to operate them in 2025 may be different depending on the vintage.

- In addition, for some option build variables (such as large scale network upgrades) there may also be a construction period effect, whereby if development takes >1 than one year to complete some costs are incurred from the start of construction, but the operating ability of the option (e.g. to flow energy through the network as part of supply / demand balancing) may not be available instantaneously. When combined with the optimisation across the full pathway, this will help to understand better the real timing of investment, particularly where network reinforcement needs to precede additional load.

- Economies of scale may be achievable, by undertaking more deployment up-front versus incremental, modular expansion, for example, in the development of a heat network or in multi-street retrofits versus house-by-house. As described in section 8.7 the proof-of-concept activity has demonstrated the ability to represent this within an optimisation framework. It is then necessary to separate the components of the installation costs, which are fixed per unit (e.g. the direct material cost of a pipe) versus those which may exhibit economies of scale such as the labour costs of installation.

As discussed in the table above a number of the variables associated with the supply / demand balancing are timesliced at different levels (annual, characteristic day, within day) to simplify the problem optimisation problem. For example, it is not necessary to understand the supply / demand balance of coal at finer than an annual level as we can reasonably assume that stock-piling and supply chains are sufficiently developed to ensure that supply is available where it is needed. However, given the issues and cost associated with large-scale electricity storage and the need to more realistically account for peak demand as the key driver of capacity requirements it is necessary to track the supply / demand balancing of this product at the most granular within day level for each characteristic day. The number of periods represented within day in the POM is flexible as discussed in section 8.6.3.

Finally, it is important to note that the more focused scope of the EnergyPath™ Design Tool means that a number of more 'traditional choices' found in national level energy system models are not appropriate or covered. All of these factors will still need to be replaced by exogenous input assumptions informed by national pathway modelling from ESME or other sources (with variations tested via sensitivity) as they still indirectly impact on the local area energy system pathway. Factors not reflected by explicit decision variables include:

- Choice over the deployment of different vehicle types (e.g. H2, EVs) in the local area as this should be informed by developments at the national pathway level. Separate EV charging profiles would be layered into the building archetype electricity demand profiles subject to external assumptions about the shape of the profile, ownership of EVs by cluster, diversity impacts in charging profiles

- Deployment of new large scale transmission connected electricity generation as this is outside the boundary of the network system modelled (up to 33kV as outlined in section 7.3.1). However, as indicated in the table above, the choice to use this external supply (at a given cost) and expand the implied peak capacity of supply (at a given cost) at the boundary need to be factored into the decision variables, so that this can be traded off against alternative options within the local area such as efficiency measures, fuel switching or more embedded generation

- Choices around industry energy use such as fuel switching, building of new boilers or on-site generation within the boundary of the industrial site. Within the initial development of the EnergyPath™ Design Tool it is suggested that the net demand (or potential supply of electricity / waste heat) assumptions at the industrial site boundary are exogenous. For future development it would then be possible to introduce further elements of choice, for example in terms of heat supply options for building-grade heat within the industrial site boundary

8.5.3. Key constraints and 'design standards'

Constraints restrict the possible combinations of variables (options) that can be explored by POM to find a least-cost solution to mimic real-world issues. These can broadly be divided into core constraints (which must apply at all times to provide a sensible energy system representation) and optional constraints, which further restrict the possible solution (and be extension increase the costs of the pathway design):

- Core constraints

    - Supply must equal demand in all timeperiods, timeslices and clusters. This implicitly guarantees a feasible working energy system under normal operating conditions and by extension that household comfort requirements are met by the pathway design[40]. In addition, this supply / demand balancing requirement ensures that sufficient network capacity is developed to cover *net* peak demand (i.e. peak demand net of embedded supply within a cluster)

    - Restrictions on the maximum (or optionally forced minimum) operating availability of options such as CHP or network components – i.e. this limits the maximum output or flow that can be produce or accommodated, respectively

- Optional constraints

[40] As the final energy demand profiles constructed in the HOM, which must be supplied appropriately in the POM, are themselves based on operation of heating devices, which adequately meet the required comfort levels in the different building archetypes.

- Restrictions on building archetype / network / other technology deployment. This could reflect a maximum absolute limit (including zero) or rate of deployment in cluster X or the total local authority area by timeperiod. In addition build rate constraints could reflect 's-curve' type deployment whereby the rate of expansion is dependent on what has been built to date (i.e. to mimic the development of a supply chain)

- In addition these constraints could be further refined by targeting them at an individual option or a group of option (i.e. POM has the freedom to decide which options to build within the group, but up to the total group limit). This could be used to reflect a supply chain on the maximum number of local area insulation retrofits that can be undertaken, but without restricting where this retrofits are undertaken.

- It is also possible to invert the above constraint to force a minimum absolute limit or rate of deployment, for example, to calibrate deployment in the near term in response to a non-modelled policy incentive

- Restrictions on the maximum (or forced min) use of resources such as biomass in biomass in cluster X in timeperiod Y, as the alternative is to give them unlimited availability at a given cost

- Restrictions on maximum release of certain products e.g. to reflect a CO2 emissions cap

The proposed approach to building POM and the platform proposed (see Design Architecture (D3) deliverable) means that adding constraints is a relatively flexible process. The broad nature of the underlying formulations (e.g. a maximum level of deployment or a maximum rate of deployment) and the flexibility to target at one or a group of options, means that a wide range of real world issues can be represented such as area specific planning constraints (meaning that no build is possible) or delays in development.

A natural extension of the optional constraint formulation is the creation of '*design standards*' that the pathway design must adhere to. In other words, the POM must find the lowest cost system design for a given set of pathways that satisfies both the core and optional constraints, including the design standards.

It is possible for the user to create various design standards, but a key one is around the security of supply or resilience standard that we are designing the pathway to meet. For local energy systems in the UK it is the key ability of the system to deal with an extended extreme cold weather spell (e.g. a 1-in-20 winter rather than the long-run average), this could reflect a single day or an extended cold spell maintained for a number of days. Heating capacity must be sufficient to meet peak demand along with the wider network features and primary supply sources, particularly where heating is electrified.

In creating the design standard constraint it is necessary to consider what the

- Key energy system features are that need to meet the design standard. For example if peak demand is just considered in capacity terms it is possible to represent this by a simple proxy of peak demand scalar relative to typical winter conditions versus the available capacity. Whereas to understand both capacity adequacy and operational requirements on the peak day and potentially over a number of consecutive days[41] the POM would need to understand the supply/demand balance position across this period. This could be tackled by creating another 'fully resolved' characteristic day which represents a 1-in-20 winter day, alongside the typical winter characteristic day and adjusting the options, which may not be able to provide continued supply (e.g. storage) over an extended cold weather spell.

- The definition of how stringent a design standard is. For example, does it represent a P90 or P95 (i.e. 1-in-20) cold weather spell in terms of external temperature, and should further 'worse case' assumptions be added (e.g. no diversity benefits on the electricity system which further raises peak demand). The user will be able to change the stringency of the design standard.

8.6. Key POM process steps

8.6.1. Scenario data pre-processing [POM-001]

The first process step within the POM is to generate the detailed scenario and option data necessary to undertake the optimisation, using the inputs outlined in section 8.4. Two key points at this stage are:

- To create all the data that is needed, the POM must in some cases combine data from multiple sources. For example, the SAM provides information on what archetypes currently exist in each cluster. This needs to be combined with data from the HOM to understand what the current demand profiles for this archetype are, and what the possible options (including costs and demand implications) are for converting this starting archetype into another

- The SAM / NAM have already undertaken the process of aggregating data to the cluster level, but the remaining data which varies by timeperiod (e.g. the cost of options which may decline over time) and timeslice data related to energy demand profiles is still highly disaggregated (annual, half hour for each characteristic day respectively). This disaggregated data is then subject to pre-processing in the POM as discussed in the following sections.

8.6.2. Cluster-level energy demand diversity [POM-002]

The HOM will create *base* final energy demand profile(s) for each individual archetype in isolation (covering gas, network hot water for district heating, and electricity). Each cluster seen by the POM will then contain a number of existing archetypes and options for converting these into different archetypes, with implications for the resulting demand profiles (e.g. if a building with a gas boiler is then converted into one with a heat pump).

[41] E.g. to also understand the value of deploying further load shifting to reduce peak electricity demand whilst maintaining the overall annual level of electrified heating

However, simply adding the *base* demand profiles for the same energy vector will overestimate peak demand in the cluster and as a consequence the necessary network reinforcement (or new build) costs, which is particularly problematic in the case of electricity.

In reality there are natural diversity effects at a local level, for example, in terms of electricity demand on a street level feeder, which affect the overall level of peak demand seen on the feeder. Adjustments need to be made to the base archetype profiles in POM to account for diversity, however, it is important to separate out a number of different demand profile elements as they will be subject to different diversity adjustments:

1. Heat-related electricity consumption as variation is driven strongly by external temperature requirements and hence there is potentially more coincidence of demand on cold days

2. EV related consumption, as the extent of ownership and grouping of ownership in a cluster is a key exogenous assumption, and it will be important to easily adapt the diversity factor for this

3. Other electricity consumption covering appliances, lighting, etc which will again have some level of diversity, and is most closely related to historically observed diversity factors

There are two main options for implementing this

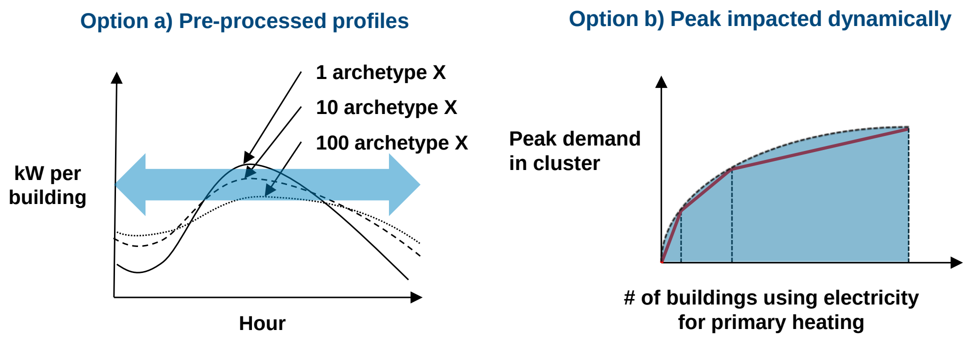
- The simpler (**Option a** in Figure 8-4) option is to pre-process the base profiles based on a 'stretch factor' that relates to the number of archetypes in the cluster – i.e. a cluster with 100 buildings would see more smoothing *per building* than a cluster with fewer buildings. Separate factors would be used for each of the elements above and the diversity scaling factors themselves could vary by archetype and by timeslice (e.g. less diversity benefit on a very cold peak winter day compared to a characteristic winter).

    - The downside to this option is that for the heat-related electricity component it makes an ex-ante assumption about the number of buildings that will be converting to (or from) electrical heating over time, whereas this is a dynamic choice in the POM. This may overestimate the diversity benefit for heat-related electricity consumption, but not for the other element as the assumptions are known ex-ante.

- A more accurate approach (**Option b**) is to link dynamically the diversity adjustment for the number of archetypes using electricity as a primary heating device. The *incremental* effect on peak demand in the cluster per additional building using electricity for heating would decline reflecting the diversity effect more accurately.

    - This option relies on a MIP formulation of the optimisation problem (to create a piecewise linear curve), which is being explored as part of the POM proof of concept work (see section 8.7)

Figure 8-4 Options for capturing cluster-level diversity effects



8.6.3. Problem simplification [POM-003]

The ability to easily flex the complexity of the POM optimisation, to help improve performance, is a key design feature for the EnergyPath™ Design Tool. There are number of possible routes to simplification both within the POM and as part of the data processing in other modules which feeds into the POM. The range of routes is important as it is difficult to tell ex-ante which areas of detail are most important for understanding the pathway design (and may indeed vary by area):

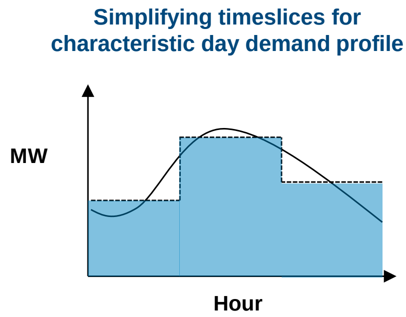
- As a general point the POM aims to retain parallel LP and MIP structures (via the use of rounded relaxation) as an LP problem is considerably quicker to solve than a MIP[42]
- POM pre-processing simplification
  - *Timeslicing* - building archetype energy demand profiles sent from the HOM are initially timesliced at their most granular level by typical characteristic day (e.g. average season, peak winter) and half-hourly within day. It will be possible to flexibly aggregate both the number of characteristic days (e.g. months to seasons) and / or the number of time periods within day (e.g. see Figure 8-5 below)
  - *Timeperiods* – in a similar manner to timeslices the POM receives timeperiod data at its most granular annual level initially (e.g. the cost of archetype conversion or network options). It will then be possible to aggregate this data to a smaller number of pathway steps e.g. 5/10-yearly periods or annually in the near-term with 5-year steps thereafter
- SAM / NAM pre-processing simplification – reducing the number of clusters
  - The number of clusters used to represent a local area is flexible and is defined at the SAM/NAM stage before being passed to the POM
- HOM / NAM pre-processing simplification – reducing the number of POM options

[42] As part of the longer term development of EnergyPath™ other approaches can be used to improve MIP performance, such as carefully bounding the range of possible integer variables or providing a partial starting solution from a previous problem to improve the solving time.

– As described in section 5.5.2 part of the HOM process involves simplifying the number of archetypes used to represent a local area (both existing and the options for conversion). This process is flexible and hence further reducing the number of archetypes would by extension simplify the optimisation problem in the POM

– In a similar manner the NAM would be able to aggregate similar network reinforcement options within a cluster where they are broadly similar in terms of cost and headroom provided

Figure 8-5 Example of POM timeslice pre-processing simplification



To further improve performance there are potentially other ways that the energy system can be more elegantly represented within the optimisation problem, without necessarily sacrificing detail about the trade-offs between options.

One example is the conversion of an archetype into another, at a given cost and with a given impact on demand profiles (reflecting the impact of different insulation or heating systems) in the new archetype. The default approach is for the POM to track all possible combinations of valid archetype conversion options simultaneously[43], but this is quite a complex problem structure and which expands quickly with more archetypes (as the problem is also compounded by the number of clusters and timeperiods).

It some cases it may be possible to extract a feature of the archetype, which is effectively *independent* of the underlying archetype *at the cluster level*. For example, by default the option to have solar PV on a building, or not, would virtually double the number of archetypes. However, at the cluster level the effect of solar PV is to net off the demand / supply balance within the cluster and does not directly impact on the individual building archetype operation. As a result, the decision to build solar PV in a cluster can be separated from the individual archetypes, reducing their total number of possible archetypes that have to be considered simultaneously and simplifying the POM problem structure. Note that the number and type of archetypes within a cluster would still be used to bound the maximum amount of solar PV that can be deployed. In a similar manner, the impact of an EV is effectively independent of the building archetype.

[43] Analogous to the *retrofit* code structure in ESME

8.6.4. National level boundary conditions [POM-004]

To assess the optimal pathway design for the local area, over the pathway to 2050, it is important that relevant national level pathway conditions are consistent with those assumed in the POM.

A fundamental assumption is that the national level conditions are unaffected by decisions made by an individual local area, analogous to a "price-taker" rather than a "price-maker". Over time as the EnergyPath™ Design Tool is used with multiple local authorities, insights from more accurate pathway design at a local could be used to refine the national level pathway (e.g. the costs and feasibility of district heating). An ongoing process would then be established such that insights from both the national and local area pathways are used to refine each other.

It is currently envisaged that the UK ESME model would provide this information. ESME can undertake 5-yearly pathway optimisation steps from a 2010 base year to 2050 and is geographically disaggregated to the 12 political regions of the UK. Its within year timeslicing is also currently limited to 3 characteristic days (peak winter, typical winter and summer) and 5 within-day periods. For some boundary condition data some additional interpretation is likely to be needed to bridge the gap between the resolution at the national pathway level and that needed by the POM within the EnergyPath™ Design Tool - e.g. 5-yearly to annual timeperiods (e.g. via linear interpolation) or moving from a regional spatial level (such as Scotland or London) to a specific Local Authority area within the region.

A range of national boundary conditions can be defined to inform the local area pathway design, from those which still provide all the original degrees of freedom in the POM, to those which tightly constrain certain aspects of the pathway. For example:

- The most flexible boundary conditions reflect price impacts, particularly the overarching fossil fuel prices consistent with the national pathway solution, as well as the shadow prices for carbon and other energy supply at the physical boundary of the local area (which are an output of the pathway) such as for electricity, heat and biomass[44]
  - For input supply at the boundary there is a need to consider not only the price, but the cost of expanding the availability of supply, particularly peak capacity. For example, the POM should also be able to see a proxy for the cost of expanding peak electricity capacity based on the national marginal costs of expanding the transmission system and generation in the relevant region
  - The shadow price at the boundary also provides a proxy for the value of export of surplus embedded generation within the local area energy system back onto the national system
  - The other characteristics of energy supply at the boundary of the local area also need to be considered such as the carbon intensity of transmission connected electricity

[44] I.e. it cannot automatically be assumed that biomass is used in the local area it is produced, as the value of this limited resource must be considered within the context of national decarbonisation options such as CCS

- Similarly, the cost of specific technology options at the national level should be mirrored to some extent at the national level. For example, the potential reduction in ASHP costs
    - However, a similar interpretation is needed to bridge the gap between the simpler technology representation at the national level, such as the cost of an ASHP, versus the wide variety of costs for different ASHP sub-types applied to different archetypes. It is likely to be more appropriate to use the trend in costs at the national level rather than the absolute level for the boundary condition
- The most stringent boundary conditions involve forcing the deployment of particular options within the local area in line with the national pathway. For example, if 50% of all households in the national region have ASHPs this forced deployment could be imposed.

There are a wide range of national pathway conditions that could be imposed at the local level, particularly at the more stringent level. The overarching design of key constraints in the POM (see section 8.5.3) is such that the user will be able to easily implement these depending on the operation of the tool.

As part of the overarching process of the tool it is proposed that more flexible boundary conditions are the starting point; as the primary purpose of the EnergyPath™ Design Tool is to understand how pathway design changes with a more detailed representation of the local geography, rather than constraining very specific outcomes of the national pathway, which may not be appropriate.

More stringent boundary conditions should be tested as part of wider sensitivity analysis, and changing national boundary conditions (such as the level of transmission connected electricity decarbonisation) should form a key part of resilience testing of the local area pathway.

In the initial development stages it is not proposed that UK ESME is hard-linked to the EnergyPath™ Design Tool, but is used manually to define the required inputs. The two models could then be soft-linked (i.e. semi-automating the process of passing data between the models), with formal hard-linking a long-term option. This would make it easier to automatically impose the boundary conditions from multiple probabilistic simulations (e.g. 100+) at the national level into an equivalent number of probabilistic simulations in the POM, whilst ensuring consistency of boundary conditions in each simulation.

8.6.5. Additional user-defined inputs [POM-005]

Here the user would define all other scenario inputs which have not already been entered or have come from other modules. These include:

- Local area industry final energy demands over time (by fuel, profile)
- Key pathway constraints such as design standards (‘peak winter’), deployment rate constraints reflecting supply chain limitations for the area as a whole, etc
- Correlation factors for Monte Carlo simulation

8.6.6. Monte Carlo input simulation [POM-006]

The process of sampling the Monte Carlo inputs in POM would need significant flexibility with respect to how it is configured, for example over:

- Which parameters to simulate - such as demands, technology costs, technology performance, technology deployment constraints, resource availability, etc. Any non-simulated parameters would use their default deterministic values in the optimisation
- The distributions used – such as normal or triangular (generally used by default when there is limited knowledge of the uncertainty surrounding a parameter as it is easier conceptually to define minimum / mode / maximum values)
- The parameters used to define the uncertainty distributions in each time period, as in most cases the spread of uncertainty increases to their zenith over the medium term (e.g. the next 10-20 years), but beyond this does not increase further.
- The correlation factors assumed.

Uncertainty in underlying data inputs, the flows and uncertainty propagation across the EnergyPath™ Design Tool and correlation of sampled inputs are discussed further in section 9.

8.6.7. Optimisation process [POM-007]

At this stage the final optimisation process is effectively automated and would undertake a repeated set of optimisation for all *sets* of input data generated by the Monte Carlo simulation process.

8.7. Proof of concept activity

The proof of concept work undertaken within Stage 1 for the POM is comprised of two main elements, both of which re-use the core of the ESME modelling code:

- Adding in MIP functionality to demonstrate the feasibility of a number of ways to enhance the structural representation of the local energy system (as discussed in section 8.3.3) within POM, including
  - *Lumpy investment* in new options as a more accurate reflection of the true cost of certain options (e.g. a CHP plant)
  - *Discrete interdependent choices* via conditional constraints, i.e. the availability, or not, of one option (or a change in its characteristics) is dependent on first deploying another option. As discussed in section 7.5.4 this would help to represent the option for meshing networks together as the network cost functions in each cluster are different before and after meshing.
  - *Piece-wise linear curves*. The benefits of this are two-fold. First it allows the representation of non-linear functions without resorting to a non-linear optimisation formulation. Secondly, it allows the representation of concave functions which directly or indirectly represent a marginally decreasing rate of cost for the optimiser objective function (e.g. representing economies of scale in

deployment or incrementally smaller levels of peak demand impact due to diversity effects), which are not possible in cost minimising LP formulation.

- High-level performance testing (both with / without MIP options) using an 'illustrative' EnergyPath^{TM} Design Tool-type problem and varying its complexity due to the number of
  - Clusters
  - Timeperiods
  - Timeslices

The work is described in more detail in the Appendix, section 13.3. All of the MIP functionality is feasible, however, it is recommended to keep parallel LP/MIP structures where possible given the longer solving times with MIP.

8.8. Output validation - detailed planning / analysis functionality

To make the POM optimisation problem tractable, detailed analysis of the existing and future energy system is first undertaken in the HOM / NAM / SAM, but this representation is then simplified for the POM primarily through spatial and temporal aggregation.

Whilst the design of the EnergyPath^{TM} Design Tool will ensure that the level of 'detail' seen by the POM is easily flexed, it is important to be able to undertake validation of the POM outputs via two different routes:

- The first is simply to provide the POM with more detailed inputs prior to the optimisation and explore how the solution changes compared to a more aggregated set of inputs. However, more detail will naturally limit the extent to which differences materialise under a wide range of scenarios.
- The second is to take the more aggregated cluster level results from the POM pathway and test their 'feasibility' and 'appropriateness' via the more detailed processes used in the NAM and the HOM. Where there is a divergence between the two sets of results, the inputs sent to the POM could be refined to better calibrate a more suitable outcome from the POM itself, or alternatively additional constraints could be added to the POM to drive the more appropriate outcome[45]

More specifically, in the:

- NAM, the network solutions chosen for electricity and heat at a cluster and inter-cluster level from the POM could be tested within the more detailed spatial load flow models to:

[45] In a similar manner to the way the UK ESME model contains a number of more detailed structures for analysing building heating and peak supply / demand balancing. These structures are 'turned off' by default to enable faster solving time, but the insights from solutions using the more detailed structures are calibrated into the default setup via custom constraints.

- Ensure the solution is feasible - i.e. the demand conditions do not violate the network constraints[46]. If this does occur it may imply an underestimate in reinforcement in the POM and this could be remedied (for example) by adjusting the level of peak demand that triggers reinforcement in the final network options that are sent to the POM

- Re-examine and refine the more detailed network deployment options, which would be chosen at the NAM-level as part of a least cost-solution given the boundary conditions of increasing load over time from the POM solution. Where this implied a significantly different set of network deployment choices than that seen in the POM; the options, or costs of the options, passed to the POM could be refined.

- Finally if the detailed network analysis in the NAM shows the solution is feasible and the wider POM analysis does not need to be updated, the information used for the individual project briefs for network options could draw more directly on the more detailed cost and other information from the NAM

- In a similar manner, the POM makes decisions about the operation of heating devices within an archetype in an indirect manner, via the selection from a small number of demand profiles options for each characteristic day. These profiles are created in the HOM and represent a subset of pre-defined possible choices for heating systems with optionality (e.g. hybrids or systems with storage). A limited subset of choices is necessary to reduce the complexity of the POM optimisation and these choices are pre-defined in the absence of knowing key drivers that are only resolved in the POM (e.g. the implied electricity price).

  - By passing these key drivers back into the dynamic energy simulation component of the HOM it would be possible to understand how close the archetype operation is to the profiles chosen in the POM. If these are significantly different, additional profiles could be created in the HOM and passed to the POM for a further iteration of the pathway optimisation, as this may indirectly wider pathway choices such as the level of peak demand and network reinforcement.

For Stage 2 it is proposed that this is a largely manual, expert-user processing step. Over the longer term, it would be possible to automate the way the results from the POM are passed back to the SAM/NAM to simplify the validation process. However, it is unlikely to be appropriate to automate any subsequent refinements to the POM inputs as this will be a result of expert user insight as opposed to an automated calibration.

[46] In a similar manner to the way the ESME solution for the electricity system is tested in a PLEXOS hourly resolution dispatch model to better understand whether the system operation is feasible from a security of supply and flexibility perspective.

9. Modelling approach to uncertainty

9.1. Overview

A key design requirement for the EnergyPath™ Design Tool is to allow the user to efficiently tackle 'uncertainty' associated with the inputs and interpretation of the outputs. This 'uncertainty' can be grouped into two broad categories:

- The more technical elements of tackling uncertainty within the tool itself *which is the purpose of this section*; e.g. what is the modelling approach to dealing with uncertainty in inputs and associated design issues, what are the most material input parameters which exhibit uncertainty, what other elements of the modelling approach introduce uncertainty and how, etc

- The overarching use of the tool itself (and the features within it) as part of a business process that aids real-world decision making, in light of the inherent uncertainty of the problem we are trying to tackle – i.e. a strategic pathway design for the next 30+ years. *This is discussed further in section 4*

The main technical approaches to tackling uncertainty in the tool will be two-fold

- Scenario analysis (i.e. manually changing one or more discrete input parameters) to enable what-if testing or resilience analysis. The benefit of this approach is that it allows the user to easily understand the directional impact of changing parameters (within the context of other fixed inputs)

- Monte Carlo simulation of key input parameters. The benefit of this approach is that it enables the user to more efficiently assess a wider range of parameter uncertainty on pathway design; both in terms of the number of states for an individual parameter (e.g. 100s of samples from a triangular distribution of fuel prices) and across multiple, and potentially correlated parameters (e.g. 100s of simulated sets of coal, gas, oil and biomass prices)

The types of uncertainty within the tool can be categorised broadly as follows:

- Data driven uncertainty:

  - Uncertainty in input data that can only sensibly be parameterised in a discrete sense (e.g. low, medium, high values) and which lends itself more to specific scenario values or discrete sensitivity testing. Examples might include the level of industry that could exist in LA area over a 30 year pathway, which is highly uncertain, but has significant implications for wider energy system design

  - Uncertainty in input data which can sensibly be parameterised as a continuous variable (e.g. normally distributed with a mean of X and standard deviation of Y) and which lends itself more to Monte Carlo simulation. Examples may include the distribution of technology costs[47]

- Modelling driven uncertainty (i.e. that driven by the approach or use of the data as opposed to uncertainty in the data itself).
  - Examples of this include the need to aggregate the spatial information from more detailed zone- to cluster-level. This means that spatial information on exactly where options are deployed within the cluster is lost for the purposes of decision making.
  - Mitigation strategies to overcome this modelling introduced uncertainty are important and in this example are two-fold; the first is to use assumptions which minimise any distortion in decision making as this spatial information is lost[48], the second is to have the ability to easily flex the level of spatial granularity and re-run the POM to understand how this impacts the solution.

9.2. Summary of module-specific uncertainty

The tables below summarise the uncertainty introduced by both data and modelling approach in the different parts of the EnergyPath™ Design Tool and how this proposed to be treated and mitigated, respectively.

[47] Note that when information on the distribution of a parameter is limited, a default approach is to parameterise the min, mode, max of a triangular distribution. However, for some parameters, such as the industry example above, it may still be better to understand the impact of changing this factor via discrete scenario testing rather than including this within a wider set of simulated inputs; as it then becomes difficult to unpick what is the key driver of the results

[48] For example, as outlined in section 0 the approach for aggregating load in a cluster for the purposes of assessing network reinforcements

Table 9-1 Key data driven uncertainty

Module	Uncertain parameters	Propagation across tool	Proposed treatment
HOM	Costs for converting one archetype into another driven by uncertainty in component costs for insulation, heating systems, etc	• Output passed to POM directly for use in pathway optimisation	• HOM would pass low / medium / high upgrade costs for each archetype pair in each timeperiod, which could be used by POM for Monte Carlo sampling. • POM would potentially need to make further adjustments to option cost profiles in future timeperiods in line with national pathway boundary conditions. • Component costs across archetypes using similar technology would have to be correlated (i.e. all archetypes using ASHP exhibit same shift in costs)
HOM	Archetype final energy demands and profiles. This is driven by underlying uncertainty around a number of components: • Heat related – External temperature profiles – Comfort requirements as day / occupancy levels and patterns – Building component performance (e.g. efficiency of heating devices) • EV demand and profiles driven by – Number of EVs – Charging profile • Other lighting / appliances – Driven by usage patterns	• Output passed to SAM (used indirectly as part of calibration of archetype matching using historic demand by area see 6.5.1) • Output passed to NAM to help determine likely maximum demand range to be tested by network modelling • Output passed to POM for direct use in pathway optimisation • Output passed to POM to help define cluster diversity factors	The treatment of uncertainty is different for each underlying component used to create the final demand profile for each archetype • Heat related – Uncertainty in external temperatures is captured by the creation of typical day and edge-case or peak days and so these are treated deterministically in the POM, with no additional Monte Carlo simulation (however, the number of characteristic days, definition of edge cases, is flexible) – In the creation of the main archetype demand profiles sent to POM temperature uncertainty is tackled through the creation characteristic days (average and edge cases), central behaviour profiles are also used, and the efficacy of building components is tackled by the use of low/central/high cases. I.e. all are tackled via deterministic sensitivities and there would be three sets of deterministic profiles from the HOM for each archetype. The POM would then use this information to define distributions to simulate changes in the efficacy of building component, but this would not embed uncertainty changes to the profile driven by household energy using behaviour. To account for both behaviour and efficacy uncertainty, the original

			base profile in the HOM related to behaviour would need to be updated and the above process repeated – To estimate cluster diversity parameters for the POM the HOM *could* undertake Monte Carlo simulation of changes to the final archetype profiles driven by changes in householder demand behaviour (both comfort and appliance/lighting use), but fixing the central deterministic values for efficacy of measures • For EVs uncertainty in charging profiles would be reflected by exogenous assumptions, whereas the number of EVs deployed would be consistent with the simulated National Pathway outputs i.e. scaling the number of EVs present in the local area • Uncertainty in other lighting / appliances – Shape could be simulated (although by default this is an exogenously specified profile as discussed in 5.5.1.3) as part of creating diversity scalars for use in the POM, however wider changes to the shape of demand from these items in the POM would likely be via deterministic testing.
NAM	Costs of network reinforcement or new build options. As discussed in section 7.3.4 the different parameters that lead to the uncertainty in cost will be treated differently, but can be grouped broadly as: • Base component costs of equipment and installation • General site context (urban, rural, etc) • Specific ground conditions (e.g. excavation difficulty)	• Output passed to POM directly for use in pathway optimisation	• The NAM would pass low / medium / high upgrade costs for each network option based on the base component costs, which could be used by the POM for Monte Carlo sampling – e.g. within a triangular distribution • On top of the simulated costs, location specific scalars could be used to represent the further markup/down of costs related to general site context and specific ground conditions • Base component costs across similar network options would have to be correlated
POM	Cluster diversity demand scalars separated by • Heat-related electricity demands • EV electricity demand • Other electricity demand • Heat network demand	• For heat-related electricity demands one potential approach to creating the estimates is to use Monte Carlo analysis of final energy demand profiles in the HOM to inform the diversity scalars. • This approach would indirectly embed the uncertainty factors considered as part of this Monte Carlo simulation into the diversity factors	• These will likely be explored by discrete scenario testing, however, the option to undertake Monte Carlo simulation alongside other outputs could still be included • Note that EV electricity demand would be affected indirectly via simulated outputs from the National pathway boundary conditions on the number of EVs deployed, which would scale the number of EVs present in the local area
POM	Other non-building / non-network energy system data for new options – e.g. CHP, other district heat sources,		• POM would contain cost data necessary to define distribution

	network storage or other network embedded generation. The uncertainty is most likely to cover cost components (CAPEX, FOM, VOM), efficiency and availability or load factor		for Monte Carlo sampling • Where simulated national pathway outputs are not independent of the data in a specific EnergyPath™ simulation further adjustments may need to be made with the national pathway likely to take priority. E.g. in terms of the relative evolution of technology costs in each future timeperiod
POM	National pathway boundary conditions. This can be divided into • Carbon shadow price • Variable costs, availability and carbon intensity of transmission connected electricity and heat at boundary of local area • Cost of increasing available peak supply at boundary • Fossil fuel and biomass prices • EV ownership • Potentially also relative cost trends by timeperiod for key technologies	• Taken from ESME outputs and used directly in POM	• ESME outputs are themselves a set of simulation results reflecting uncertainty. It will be important to ensure that – The number of simulations of national pathway outputs to be used is ≥ the number of simulations to be undertaken in EnergyPath™ – Where simulated national pathway outputs are not independent of the data in a specific EnergyPath™ simulation further adjustments may need to be made e.g. the national pathway takes priority or its values are reflected via correlation factors when the POM specific values are simulated. E.g. in terms of the relative evolution of technology costs in each future timeperiod
POM	Other user defined scenario assumptions – e.g. location and level of industry in each time period	• Used in POM pathway optimisation	• These will likely be explored by discrete scenario testing, however, the option to undertake Monte Carlo simulation alongside other outputs could still be included

Table 9-2 Key approach driven uncertainty

Module	Uncertain parameters	Propagation across tool	Mitigation strategy
HOM	Number of building archetypes used to describe area after filtering process (see section 5.5.2), affects the accuracy of the representation	• Final 'menu' of archetypes passed to SAM for archetype matching on a spatial basis • Also effectively passed to POM represents the options available for upgrading buildings by cluster	Archetype filtering process is defined to flexible such that user can relatively easily tailor the number of final archetype and add specific user-defined archetypes to the final set. This more granular information is then automatically propagated through to the rest of the EnergyPath™ Design Tool via the standard mechanisms.
HOM	Limited number of final energy demand profile options created for each archetype (to minimize computational issues in HOM itself and POM) with respect to • Characteristic days modelled (e.g. only 4 typical seasonal days + edge cases) rather than fully 365 days + edge cases • And limited number of possible profiles within each characteristic day where building devices exhibit optionality (e.g. due to hybrid devices, micro-CHP, integration with storage, etc)	• Effectively represents the options available to POM to alter the S/D balancing in a cluster (on top of the broader option to convert/upgrade an archetype) by altering the operation of the devices within the building.	Two key mitigation options • The process to create characteristic days and operation options (leading to new optional demand profiles) is flexible and hence more can be created in the HOM to provide the POM with a wider set of choices / trade-offs – e.g. more variations of load shifting with storage • Secondly, via the detailed analysis loop outlined in section 8.8 the user could test/validate the extent to which the available operating profiles are appropriate once the wider scenario conditions are established (e.g. once electricity 'shadow prices' are available from the POM scenario), which could inform the creation or refinement of additional profile options
SAM	Synthetic network generation (see 6.5.3) where real data is limited initially.	• Network topology data is passed to the NAM and will impact load flow modelling and the generation of reinforcement options used in the POM	The base process for generating artificial network topology is flexible in that it can accommodate actual data (even if only partially available) and can be calibrated to real data at a more aggregate level, which is likely to be available in most cases. E.g. total number of customers connected to a substation (even if it is not known exactly which final feeder each customer is connected to). The base process for generating an artificial topology has also been discussed with our DNO partner WPD and is judged to be the most appropriate approach in the absence of real world information.
SAM	Building archetype matching (see section 6.5.1) – uncertainty in understanding what the best match is for each unique building in the area given limited data	• Archetype matching affects options seen by POM within each cluster as part of pathway optimisation	Two key mitigation options: • Process is flexible to accommodate better data as it becomes available at different levels of granularity and improving matching process • Validation exercise to compare estimated annual demand for electricity and gas from matching of the existing stock with historic spatial estimates. Key issue for calibration exercise is that this is compound problem of uncertainty in HOM energy

			demand data (as described in table Table 9-1) and archetype allocation data
NAM	Accuracy of the derived cost functions for zones, clusters and inter-clusters • This includes how accurately network reinforcement functions are collated; • It also includes how accurately additional load is assumed to be distributed within a cluster;	• Network cost functions used by POM within each cluster as part of pathway optimisation	As described in section 7.5.4 there exists a spectrum of possible options for representing the accuracy of the network cost functions, we are proposing to implement the penultimate option 4b, and subject to the proof of concept work the most sophisticated option 5 if possible. In addition the two key mitigation options described below also apply in the case of the accuracy of the cost functions.
NAM / SAM	Cluster definition, both number and shape see sections 6.5.6 and 7.5.5), the driving principles of which are to: • Create the maximum number of clusters which is computational feasible in the POM to improve granularity • Create clusters which contain network reinforcement / new build options which are as representative as possible for all energy vectors so as not to unduly impact the cost-optimisation process in the POM	• Central to defining the network cost functions which are used in POM	Two key mitigation options: • The core cluster generation process aims to balance the representation of heat and electricity networks in particular. However, the process of defining the clusters is meant to be flexible such that the user can both automatically generate different cluster definitions (e.g. subject to different threshold and network constraints) as well as manually impose cluster definitions, and the final cluster data will automatically be processed in the required format for the POM • Secondly, via the detailed analysis loop outlined in section 8.8 the user could test/validate the extent to which the final (more aggregate) network solutions outlined in the POM, are valid when re-considered from the more detailed network analysis perspective in the NAM. This can be used to help refine the network options and / or cluster definitions
POM	Data simplification steps (see section 8.6.3) to minimize the computational complexity of the POM optimisation • Timeslice aggregation • Timeperiod aggregation • LP / MIP structure	• Central to the POM's ability to accurately trade-off different energy system design options	It is difficult to say ex-ante where granularity matters most for the defining the energy system pathways. Therefore the process of data simplification is designed to be flexible, such that the user can easily change granularity in one area and automatically reprocess the data in the necessary format for the POM

In the case of data driven uncertainty, the vast majority of this is propagated through the tool in what is effectively a 1-step process to the POM. A range of uncertainty on the inputs in preceding modules (such as the costs for archetype conversions in the HOM or network options in the SAM) is either passed directly or captured in the intermediate outputs and passed to POM for use in Monte Carlo simulation. By default the POM would use the central/mean values in the deterministic case.

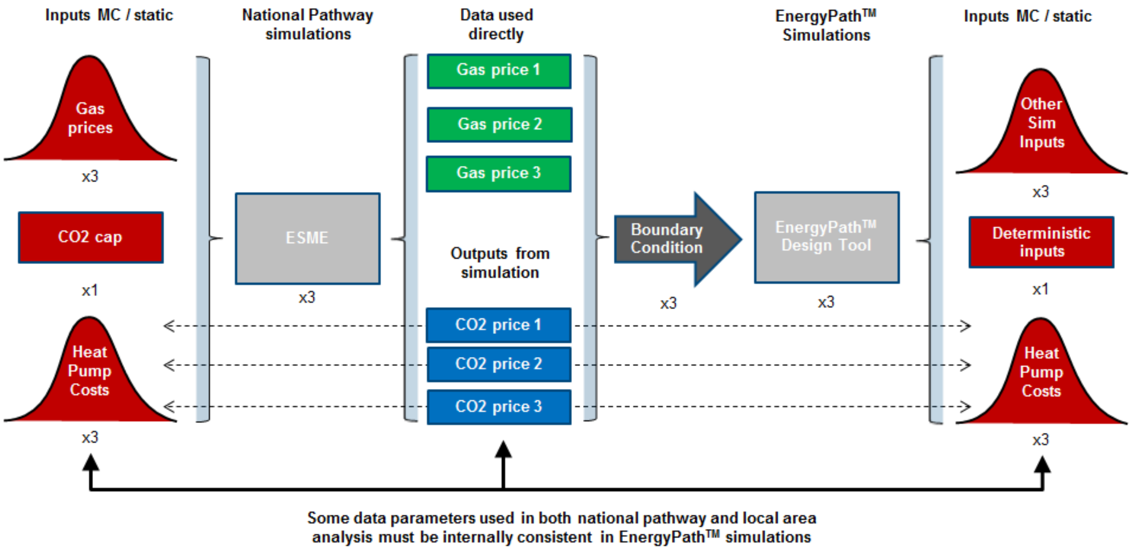
There are two more complicated areas of uncertainty propagation across the EnergyPath™ Design Tool.

The first is the application of simulated ESME outputs for reflecting uncertainty in national pathway boundary conditions can interact with the POM input data in two ways:

- The ESME input or output data may be used directly to reflect uncertainty – e.g. fossil fuel prices, or the number of EVs in the local area in each POM simulation could be scaled by the national level values, respectively.

- The ESME input data may interact with and potentially supersede other simulated parameters in the POM. For example, an ESME output may have considered the evolution of simulated costs for certain heating devices. These costs may also be simulated within the POM (i.e. the range of archetype costs). It would, for example, be inconsistent to have a single POM simulation using national boundary conditions (carbon price, electricity, etc) which reflect a limited decline in future heat pump costs, whereas archetype costs in that same simulation reflect a significant future decline in heat pump costs. The simplest option is to assume that the national pathway trends take precedence.

Figure 9-1 Illustration of consistency in national pathway inputs/outputs with the POM inputs



The second area, is the 2-stage uncertainty propagation from potential Monte Carlo simulation of some elements in the HOM (e.g. used to help create cluster demand diversity scalars for the POM) followed by additional simulation of other archetype-related parameters in the POM.

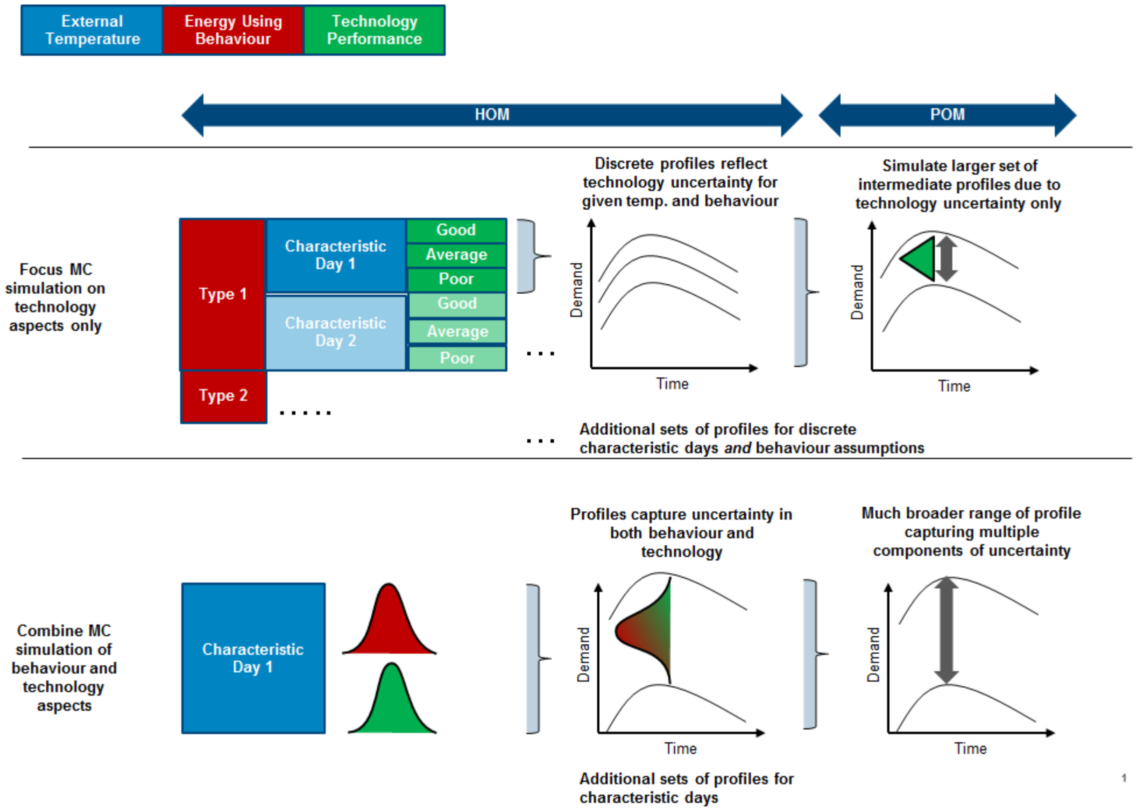
The key issue here relates to the heat-related final demand uncertainty, which is driven by uncertainty in three main drivers: **a)** external temperature, **b)** household behaviour (comfort requirements, occupancy/usage throughout day, etc) and **c)** efficacy of physical building components (insulation and heating device performance). It is proposed to focus the use of Monte Carlo on different aspects in the HOM versus the POM to minimise computational issues in the HOM and make it easier for the end-user to understand uncertainty propagation.

In the creation of the main archetype demand profiles sent to POM, temperature uncertainty is tackled through the creation characteristic days. A fixed set of behaviour assumptions (e.g. Type 1 in Figure 9-2 below) are used in each case, and the efficacy of building components is tackled by the use of low/central/high cases. I.e. all are tackled via deterministic sensitivities and there would be *three deterministic profiles from the HOM for each archetype on each characteristic day.*

The POM would then use this information to define distributions to simulate changes in the efficacy of building technology component, but this would ***not*** embed uncertainty changes to the profile driven by household behaviour. To account for *both* behavior and efficacy uncertainty, the original base behaviour profile in the HOM would need to be updated (e.g. Type 2) and the above process repeated to re-create the final POM data.

By contrast, combining the uncertainty in both behavior and efficacy of building technology performance to create the initial demand profiles would lead to a much spread of resulting profiles in the POM. Whilst this would technically be possible such a wide spread in profile outcomes is likely to make it difficult for the user to understand what is driving the final POM solutions.

Figure 9-2 Illustration of combining versus separating key drivers of building final energy demand profile uncertainty as part of Monte Carlo Simulation



Separately from the above, to better estimate cluster diversity parameters for the POM, the HOM could in principle undertake Monte Carlo simulation of changes to the final archetype profiles[49] driven by changes in householder behaviour, but ***fixing*** the central *deterministic* values for efficacy of measures. I.e. given the significant variation and uncertainty in demand due to consumer demand behaviour it is likely to be easier to test this more deterministically and only simulate technology uncertainty in the POM. As opposed to simulating the combined impact of behaviour and technology uncertainty in ***both*** the diversity scalars and the final archetype demand profiles in POM and trying to ensure the two are sensibly correlated. *The simplifying assumption here is that the diversity scalars move proportionally with changes to demand profiles caused by building component efficacy.*

[49] By using available survey / sample data on the range of key behaviour (e.g. desired temperature and occupancy/usage patterns) and technology uncertainty (e.g. efficiency) to estimate the distributions and correlation between key inputs parameters. These would be used to generate multiple sets of inputs for the dynamic simulation of each building archetype on different characteristic days. Multiple simulations would be run for an individual building and then the results of these combined (accounting for the correlation in sets of inputs) with different numbers and types of building archetypes to explore the combined profile effect. The final combinations would help to better understand the potential diversity effects in a localised area heating is electrified using different technologies. Aside from limited available survey data, the key difficulty is the number of dynamic simulations that would need to be run to generate a set of profiles under different behavioural and other conditions for each archetype.

9.3. Detailed design issues

As outlined in section 8.6.6 the approach to Monte Carlo simulation in the POM would need significant user flexibility with respect to how it is configured in term of

- Which parameters to simulate - such as demands, technology costs, technology performance, technology deployment constraints, resource availability, etc. Any non-simulated parameters would use their default deterministic values in the optimisation
- The distributions used – such as normal or triangular (generally used by default when there is limited knowledge of the uncertainty surrounding a parameter as it is easier conceptually to define minimum / mode / maximum values)
- The parameters used to define the uncertainty distributions in each time period, as in many cases the spread of uncertainty increases to a zenith over the medium term (e.g. the next 10-20 years), but beyond this does not increase further[50].

A key complexity involved with the simulation of input parameters is accounting for potential correlation in parameters between:

- Similar technology options – e.g. ensuring costs for ASHP variants which share similar core components, but which are associated with different building archetypes move in a similar direction between simulations
- The *same* technology option in different timeperiods – e.g. ensuring ASHP costs in 2015, 2020, 2025 move in a similar direction in the same simulation and subject to an overarching trend – i.e. if costs exhibit a declining trend over the pathway the costs in a subsequent timeperiod can only be ≤ that in the previous period.
- *Similar* technology options in different time periods – as above but ensuring that all ASHP variant costs associated with different archetypes all move in a similar manner across timeperiods in the same simulation
- National level boundary conditions – where the underlying boundary simulation data could be comparable to elements being simulated directly in POM. E.g. the carbon price boundary condition is itself based on certain technology costs assumptions such as ASHPs, which may be being simulated in ESME.
  - Hence in a single POM simulation using a carbon price boundary condition which itself reflects very high ASHP costs, it would not be consistent to use a POM simulated parameter reflecting very low ASHP costs; *unless* it can be assumed that the LA area is not a 'price taker' for this factor and it is valid that the evolution of heat pump costs at the national level could be significantly different from those at local level.

[50] This is in contrast to the approach taken in ESME whereby only the data parameters in 2050 are simulated and the deterministic profile of data values from 2010 to 2050 are scaled up or down to meet the simulated 2050 value. I.e. uncertainty always becomes greater over time and it is not possible to simulate and increasing range of uncertainty to e.g. 2020/2030 and a fixed range of uncertainty beyond this point.

- This does not imply that both ESME and the EnergyPath^{TM} Design Tool are being run and iterated in parallel, but that where a carbon shadow price output from ESME is being used as a boundary conditions, the associated parameters which generated that value in ESME initially should be consistent with the input conditions in the tool

This requires the creation of complex correlation matrices (more so than the approach currently used in ESME). As a result it is *not* proposed at this stage to allow flexibility in the simulation of different uncertainty parameters for each cluster in the local area (although this could be included in future). This is primarily because there is limited rationale for a key input (e.g. the cost of a technology or supply chain constraints) to have a radically different uncertainty distribution by virtue of being a small geographic distance away; but it avoids further complexity in the specification of correlation factors.

10. Summary of open functional considerations

Within Stage 2 there are a number of open functional considerations that will be explored and finalised early in Stage 2 through direct prototyping and experimentation. These do not reflect gaps in the proposed approach, but either the selection of preferred approach from multiple options that have been outlined in previous sections, or options for increasing the sophistication of the core approach. The section also highlights some potential longer term considerations for further development

Household Options Module

The main option functional considerations relate to the building archetype reduction process outlined in section 5.5.2. In particular, determining whether the catalogue of results will be a sparse grid with interpolation, or a dense grid with nearest-neighbour selection of best matching output, as this will depend on the computational performance of the process.

Beyond Stage 2 the key consideration is the extent to non-domestic buildings could and should be modelled in the event that significantly better data is available in future or within individual local authorities. As described in Section 5.5.4, the energy demand of the non-domestic stock is currently estimated using a statistically inferred benchmark kWh/m^2 approach, although a detailed bottom-up ISO13790 model (SBEM) does exist for UK non-domestic buildings.

The trigger-point for reconsidering this approach would be the availability of data, in a readily accessible format, used to produce Energy Performance Certificates / Display Energy Certificates (DEC). Depending on the level of detail that is recorded and held on central registers, it may therefore be possible to model these buildings in packages like SBEM from this information, and use SBEM to investigate the impact of retrofit from a bottom-up engineering-type approach. A further consideration would also be whether to build an in-house engine for non-domestic ISO13790 calculations (a non-trivial task), or to apply for permission from DCLG to use SBEM for commercial purposes and package it with other parts of the EnergyPath^{TM} Design Tool.

Spatial Analysis Module

For Stage 2 there are a number of possible refinements to the network topology synthesis (described in section 6.5.3) including an automated methodology that accurately represents the feeder branching for real networks that have been developed organically over time, where partial rather than a full set of network topology data exists. This may be based on maximum streets per feeder, minimum overlap of unique routes per feeder or average feeders per distribution substation.

Network analysis module

There are a number of open functional considerations that need to be finalised in Stage 2

- The choice of electricity network metrics for use in spatial clustering in the SAM. Possible options include: (i) the area under the derived cost function; (ii) the electricity network reinforcement cost at potential peak load; (iii) the derived cost function gradient, or others. This is explained in Section 7.5.6 in greater detail.

- The combination of threshold metrics that can be used in isolation or combination as part of the 'heat-led' clustering approach. These include a simple or detailed levelised heat network cost metric[51], the total size of the cluster (e.g. in terms of number of buildings), electricity network topology constraints, a comparable electricity network cost metric.

Pathway Optimisation Module

The key open functional consideration for Stage 2 is the extent to which MIP functionality (which has been demonstrated to be feasible) will be used to better represent the energy system, as outlined in section 8.7. In particular, extent to which piece-wise linear curves are used to represent economies of scale and demand diversity effects and conditional, interdependent options to capture network meshing interactions.

[51] Regardless of whether the simple metric is used within the automated cluster boundary definition process, once the final cluster boundaries are set the detailed analysis process is always used to define the final set of network options which are passed to the POM.

11. Concluding remarks

The intent of this document has been to describe the required purpose, behaviour, data and components of the EnergyPath^{TM} Design Tool to a sufficient level of detail, that modelling experts are subsequently able to develop the tool against. The functional specification should be read in conjunction with the other key Stage 1 deliverables, in particular the Design and Data Architecture (deliverables D2 and D3, respectively) to understand the overall picture for the design and implementation of the tool. In addition, within the Stage 2 proposal deliverable (D5) we have a proposed an implementation pathway under which the tool could be developed.

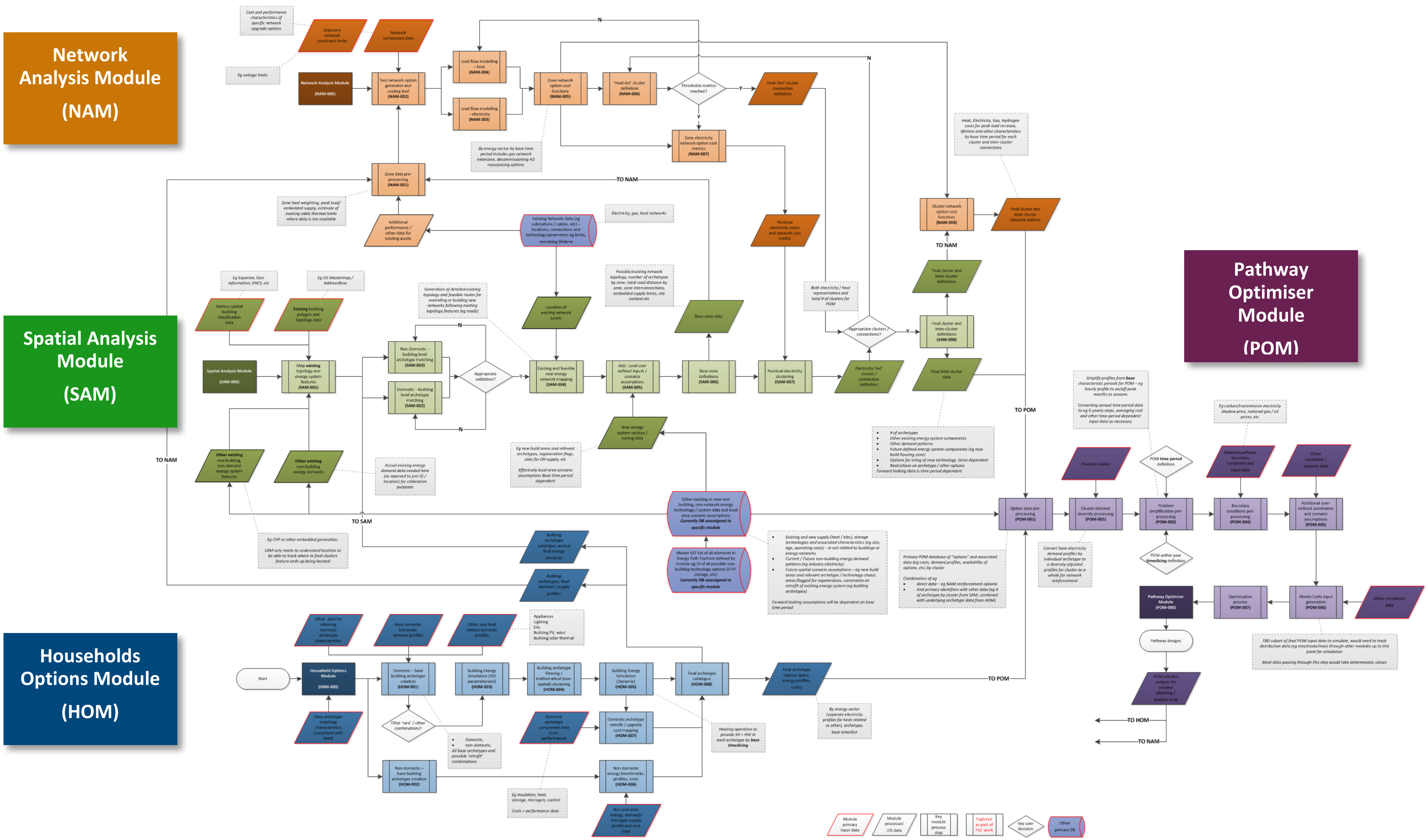
As part of creating this document we have also undertaken a number of discrete 'proof-of-concept' pieces of work (described in further detail in the Appendices), to prove that key logic steps are feasible and to gain a better understanding of the range of 3rd party software / tools, that could be integrated to enable the tool to more efficiently deliver its objectives.

The key functional and design challenge for the EnergyPath^{TM} Design Tool, which has been considered throughout the proposed approach, is the tight integration of 4 conceptual areas (buildings, energy networks of different vectors, granular spatial analysis and pathway optimisation) which are each highly complex and detailed in their own right. Integrating all of these effectively is essential to enable the tool to deliver its primary objective of helping to inform strategic energy planning decisions under uncertainty in real local areas.

12. Appendix – Full logic / process flow diagram

The diagram below shows the full logic / process flow diagram for the tool. This should be viewed in conjunction with the (D2) Design Architecture which highlights how these flows are organised from a technical design perspective – e.g. how databases are assigned and coordinated across the tool.

Figure 12-1 Full logic / process flow diagram



13. Appendix - Proof of concept work

13.1. SAM

13.1.1. Building archetype analysis

13.1.1.1. Overview

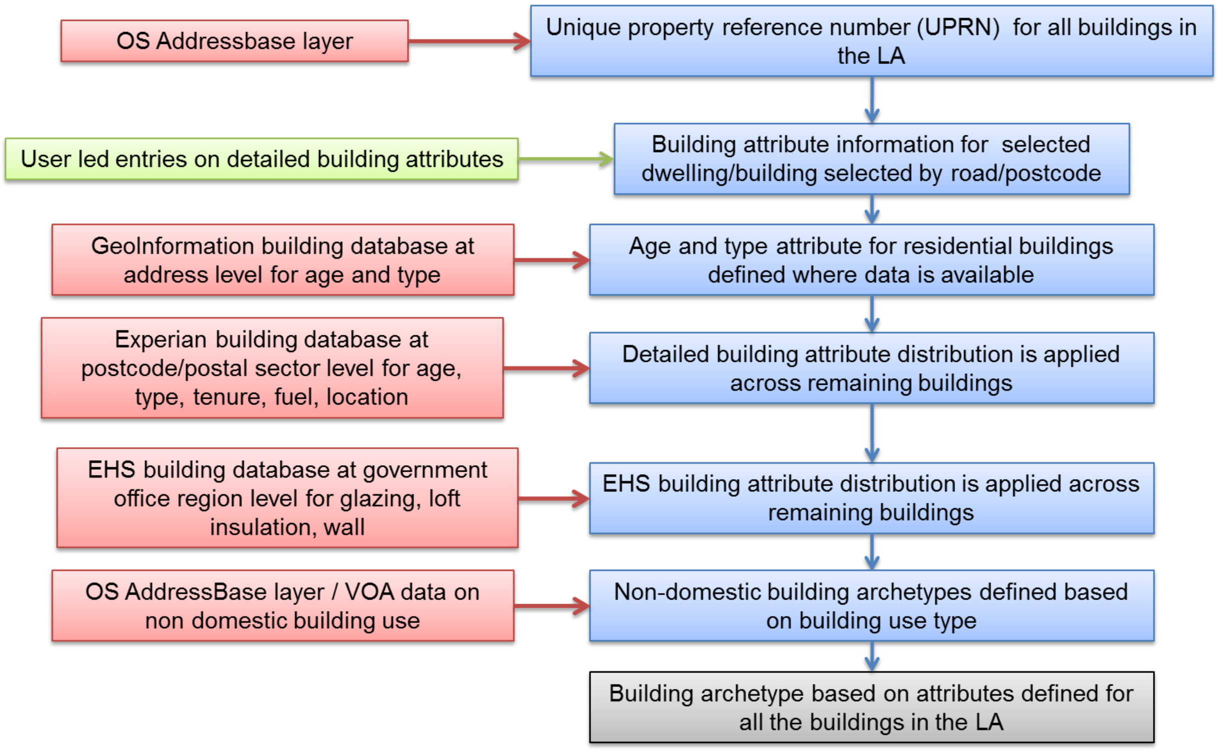
The spatial analysis module is responsible for developing a building archetype database for the area under consideration. The proof of concept work involves developing and testing methodology for enriching basic building-level information available, to include the more detailed building characteristics required for archetype assignment. This involves the following key steps:

1. Loading a core database based on the Ordnance Survey (OS) AddressBase layer that contains data at address level

2. Allowing the user to edit and enter any building level data known by the local authority

3. Combining building location data with other databases containing building attribute data at various geographic levels for the remaining buildings

Figure 13-1 Schematic for building archetype analysis methodology



13.1.1.2. Description of methodology

A semi-automated excel based tool has been developed to test the proposed methodology. This tool relies on merging the building location data and their physical attribute characteristics from several external datasets. This involves the following steps:

Step 1: Loading building type and location

AddressBase GIS layer is used to populate a building database containing UPRN, dwelling / premise type, postcode, street name. This provides the basic spatial and building use information. For the prototyping model, sample data on Exeter region, available from the OS website, has been used. The main data fields in this dataset are shown below:

Figure 13-2 Sample data on the building stock in Exeter

OBJECT	uprn	toid	udprn	buildingNumb	throughfareName	postcod
5240	1.0004E+11	osgb1000002274361562	8773047	3	CLEVELAND STREET	EX4 1BB
5241	1.0004E+11	osgb1000002274361526	8773048	30	CLEVELAND STREET	EX4 1BB
5242	1.0004E+11	osgb1000002274361525	8773049	31	CLEVELAND STREET	EX4 1BB
5243	1.0004E+11	osgb1000002274361524	8773050	32	CLEVELAND STREET	EX4 1BB
5244	1.0004E+11	osgb1000002274361584	8773051	33	CLEVELAND STREET	EX4 1BB
5245	1.0004E+11	osgb1000002274361523	8773052	34	CLEVELAND STREET	EX4 1BB
5246	1.0004E+11	osgb1000002274361522	8773053	35	CLEVELAND STREET	EX4 1BB
5247	1.0004E+11	osgb1000002274361521	8773054	36	CLEVELAND STREET	EX4 1BB
5248	1.0004E+11	osgb1000002274361561	8773056	4	CLEVELAND STREET	EX4 1BB
5249	1.0004E+11	osgb1000002274361560	8773057	5	CLEVELAND STREET	EX4 1BB
5250	1.0004E+11	osgb1000002274361559	8773058	6	CLEVELAND STREET	EX4 1BB
5251	1.0004E+11	osgb1000002274361558	8773059	7	CLEVELAND STREET	EX4 1BB
5252	1.0004E+11	osgb1000002274361557	8773060	8	CLEVELAND STREET	EX4 1BB
5253	1.0004E+11	osgb1000002274361556	8773061	9	CLEVELAND STREET	EX4 1BB
6009	1.0004E+11	osgb1000002274361510	8773034		CLEVELAND STREET	EX4 1BB

Step 2: User led data entry for known buildings

At this stage the user has the option to enter detailed building attribute characteristics for known buildings. The buildings can be selected based on a street name or postcode as shown below:

Figure 13-3 User selects option to define buildings for a street or postcode

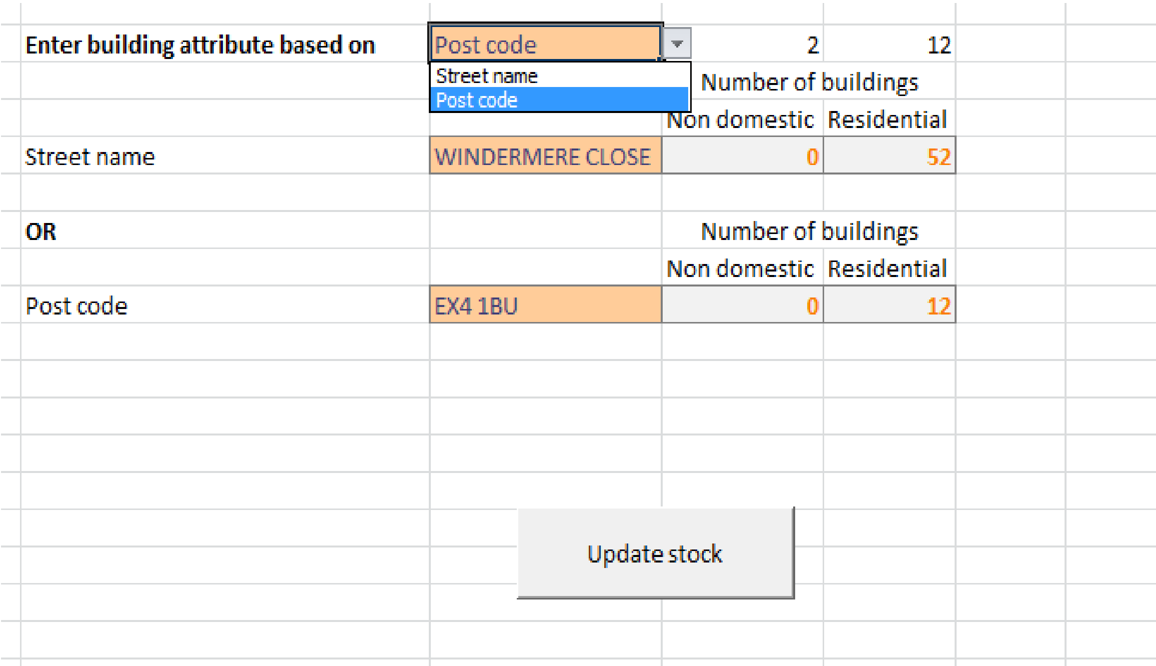


Figure 13-4 User enters the detailed attribute characteristics for all known buildings

1	7	8	12	Building attributes							
OBJECTID	buildingNumber	throughfareName	postcode	Age	Size	Wall	Tenure	Heating fuel	Glazing	Loft Insulation	Location
7117	1	CORNWALL STREET	EX4 1BU	Old	Detached	CWI	Owner occupied	Gas	Single	51-100 mm	Urban
7118	10	CORNWALL STREET	EX4 1BU	Old	Detached	CWI	Owner occupied	Gas	Single	51-100 mm	Urban
7119	11	CORNWALL STREET	EX4 1BU	Old	Detached	CWI	Owner occupied	Gas	Double	51-100 mm	Urban
7120	12	CORNWALL STREET	EX4 1BU	Old	Detached	CWI	Owner occupied	Gas	Double	101-150 mm	Urban
7121	2	CORNWALL STREET	EX4 1BU	Old	Detached	CWU	Privately rented	Gas	Double	101-150 mm	Urban
7122	3	CORNWALL STREET	EX4 1BU	Old	Detached	CWU	Privately rented	Gas	Double	200+ mm	Urban
7123	4	CORNWALL STREET	EX4 1BU	Old	Detached	CWU	Privately rented	Gas	Single	200+ mm	Urban
7124	5	CORNWALL STREET	EX4 1BU	Old	Detached	CWU	Privately rented	Gas	Single	200+ mm	Urban
7125	6	CORNWALL STREET	EX4 1BU								
7126	7	CORNWALL STREET	EX4 1BU								
7127	8	CORNWALL STREET	EX4 1BU								
7128	9	CORNWALL STREET	EX4 1BU								

This procedure is repeated for all the streets or post codes where building attribute data is available.

Step 3: GeoInformation data is used for age and type

GeoInformation provides an external dataset which has actual data on the age and type for domestic buildings. This dataset provides around 70% coverage of UK, mainly focused in the urban areas. Therefore, where available, address level data is used to populate age and type building attributes for dwellings. This is done by matching the UDPRN for each property.

Figure 13-5 Illustrative example of GeoInformation data used

udprn	Residence Type	Property Age
8775919	Semi-detached	1946-1954
8775920	Semi-detached	1946-1954
8775921	Semi-detached	1946-1954
8775922	Semi-detached	1946-1954
8775923	Semi-detached	1955-1979
8775924	Semi-detached	1955-1979
8775925	Semi-detached	1955-1979
8777446	Terraced	1955-1979
8777447	Terraced	1955-1979
8777448	Terraced	1955-1979
8777449	Terraced	1955-1979
8777450	Terraced	1955-1979
8777451	Terraced	1955-1979
8777452	Terraced	1955-1979
8768874	Terraced	1920-1945
8768875	Terraced	1920-1945
8768876	Terraced	1920-1945
8768877	Terraced	1920-1945
8768878	Terraced	1920-1945
8768879	Terraced	1920-1945

Step 4: Experian data is used for additional building attributes

Experian ConsumerView provides aggregated data at street or postal sector level. This is used to populate building age and type for buildings not already covered by GeoInformation or defined by the user. It also contains additional attribute data on tenure, location and fuel type. This data is smeared proportionally across building stock in the area for which the data is available, taking into account any user-entered.

Figure 13-6 Illustrative data on attributes available from Experian at aggregated level

Postcode	Postal Sector	Mains Gas Flag	Rural Urban Description	Residence Type	Tenure	Property Age
EX4 1BB	EX41	Y	Mid Urban	Semi-detached	Privately rented	1946-1954
EX4 1BB	EX41	Y	Mid Urban	Semi-detached	Privately rented	1946-1954
EX4 1BB	EX41	Y	Mid Urban	Semi-detached	Privately rented	1946-1954
EX4 1BB	EX41	Y	Mid Urban	Semi-detached	Privately rented	1946-1954
EX4 1BB	EX41	N	Mid Urban	Semi-detached	Council/housing association	1955-1979
EX4 1BB	EX41	N	Mid Urban	Semi-detached	Council/housing association	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Semi-detached	Council/housing association	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Owner occupied	1955-1979
EX4 1BB	EX41	N	Mid Urban	Terraced	Privately rented	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Privately rented	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Privately rented	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Privately rented	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Privately rented	1955-1979
EX4 1BB	EX41	Y	Mid Urban	Terraced	Privately rented	1955-1979

Step 5: EHS, LiW and SHCS data at GOR level is used

EHS, LiW and SHCS provide data on wall, window and loft insulation, aggregated at a GOR level. The GOR-level distribution of wall, glazing and loft insulation combination versus known age, type, tenure, location and fuel is applied to the buildings, taking into account any data already entered by the user.

Figure 13-7 Distribution of stock across GOR based on combination of all attribute characteristics

Tenure	Size	Age	Location	Fuel	Wall	Window	Loft insulation	North East	North West
Privately rented	Flat	Recent	Rural	Gas	CWI	single	0-50 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	single	51-100 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	single	101-150 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	single	151-200 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	single	200+ mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	double	0-50 mm	0.00%	0.05%
Privately rented	Flat	Recent	Rural	Gas	CWI	double	51-100 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	double	101-150 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWI	double	151-200 mm	0.00%	0.09%
Privately rented	Flat	Recent	Rural	Gas	CWI	double	200+ mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	single	0-50 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	single	51-100 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	single	101-150 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	single	151-200 mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	single	200+ mm	0.00%	0.00%
Privately rented	Flat	Recent	Rural	Gas	CWU	double	0-50 mm	0.00%	0.07%
Privately rented	Flat	Recent	Rural	Gas	CWU	double	51-100 mm	0.00%	0.05%

Step 6: Final building archetypes are defined and matched to HOM

The archetype definition for individual building depends on the physical attribute characteristics. With an enriched address-level building characteristics database, the attribute characteristics are used to match each building to known building archetypes as defined in the Household Options Module.

Figure 13-8 Archetype definition based on the attribute characteristics

				Building attributes							
buildingNumber	throughfareName	postcode	Final archetype ID	Fuel	Location	Size	Tenure	Age	Wall	Glazing	Loft insulation
117	OKEHAMPTON ROAD	EX4 1ER	16	Gas	Rural	Flat	Social	Recent	CWU	double	0-50 mm
119	OKEHAMPTON ROAD	EX4 1ER	21	Gas	Rural	Flat	Social	Recent	SW	single	0-50 mm
1	NAPIER TERRACE	EX4 3EZ	21	Gas	Rural	Flat	Social	Recent	SW	single	0-50 mm
1	ELDERTREE GARDENS	EX4 4DE	22	Gas	Rural	Flat	Social	Recent	SW	single	51-100 mm
37	LONGBROOK STREET	EX4 6AW	22	Gas	Rural	Flat	Social	Recent	SW	single	51-100 mm
0	BLACKBOY ROAD	EX4 6ST	26	Gas	Rural	Flat	Social	Recent	SW	double	0-50 mm
37	PRIORY ROAD	EX4 7AP	28	Gas	Rural	Flat	Social	Recent	SW	double	101-150 mm
39	PRIORY ROAD	EX4 7AP	29	Gas	Rural	Flat	Social	Recent	SW	double	151-200 mm
1	WELLSWOOD GARDENS	EX4 1RH	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
10	WELLSWOOD GARDENS	EX4 1RH	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
30	LOWER NORTH STREET	EX4 3EU	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
0	NEW BRIDGE STREET	EX4 3JW	51	Gas	Rural	Flat	Social	Old	SW	single	0-50 mm
11	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm
17	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm
18	RICHMOND ROAD	EX4 4JA	52	Gas	Rural	Flat	Social	Old	SW	single	51-100 mm

Thus the building analysis module is able to define the archetypal classification for the whole building stock in the local authority.

13.1.2. Network analysis module

13.1.2.1. Overview

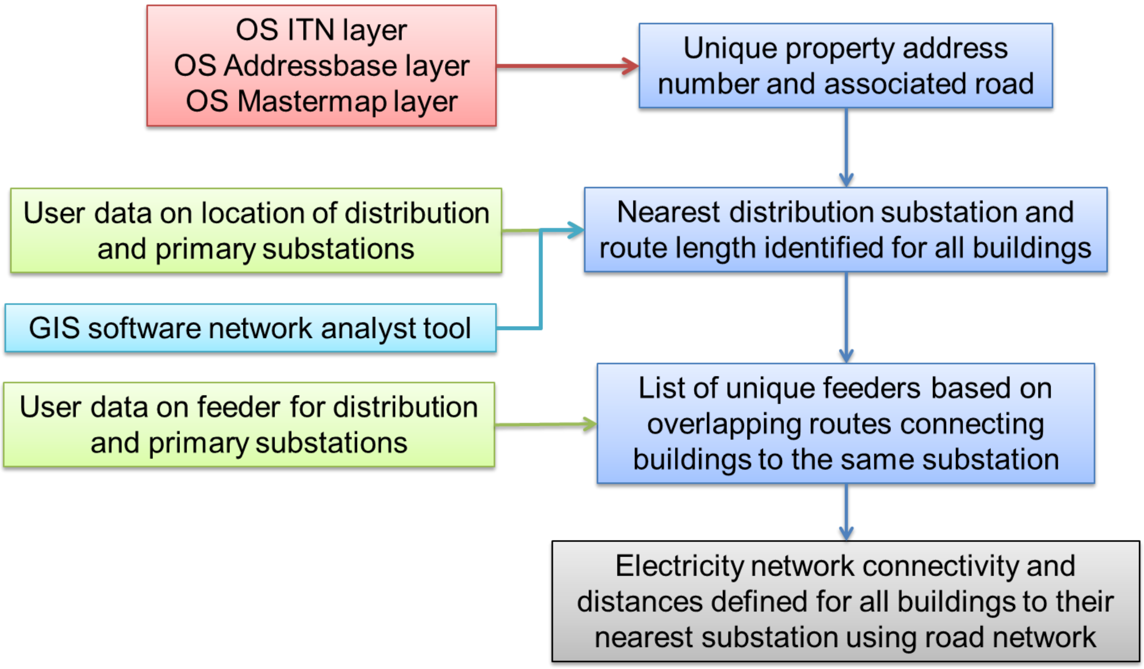
The spatial analysis module is responsible for synthesising a representation of the local electricity distribution network for the local authority under consideration. This process also results in outputs used to develop heat networks within NAM. The proof of concept work involves developing and testing methodology for synthesising the network to determine the connectivity of the buildings to their nearest substations and determine the distances. Key steps of this proof of concept work include:

1. Information from the Ordnance Survey (OS) AddressBase, MasterMap and Integrated Transportation Network (ITN) layers is loaded into a Graphical Information System (GIS) package (ArcGIS in this case)

2. This is combined with Distribution Network Operator (DNO) data on the location and number of feeders for local distribution/primary substations

3. GIS network analysis software algorithms and an Excel-based tool are then used to synthesise a representation of the local electricity distribution network, determine feeder lengths and distances for these buildings to their nearest distribution and primary substations

Figure 13-9 Schematic for network synthesis methodology



13.1.2.2. Description of methodology

ArcGIS is used to perform the initial network analysis while an Excel-based tool has been developed to process the initial network-related GIS outputs to test the proposed methodology. This involves calculating final outputs on unique feeders, distances of buildings from their nearest connected distribution and primary substation. This involves the following steps:

Step 1 and 2: Loading GIS layers with building and road location

AddressBase GIS layer, obtained from ordinance survey, is used to populate building database containing unique property reference number (UPRN) for the local authority under consideration. This contains data on the location of individual buildings. Next, the ITN GIS layer from ordinance survey is used to populate the roads and their end nodes. This provides the relevant data to develop a localised road network that provides the routes to be used to develop an electricity network. For the prototyping model, sample data from Exeter (AddressBase and ITN layers from the OS website) has been used in ArcGIS as shown below:

Figure 13-10 Sample Exeter data showing buildings, roads and road nodes

Step 3: Calculating distances from building to adjacent roads

The spatial analysis functionality in ArcGIS is used to assign each building to the adjacent road and calculate the distance to the nearest road. This data is embedded in the database of the layer containing building location (Addressbase).

Figure 13-11 Data matching each building with the adjacent road, nearest node and the distances

Step 4: User entered DNO data on the location of substations

At this stage, the user is required to load a file into ArcGIS containing the DNO data on the location of distribution and primary substations. This is then displayed on top of the existing road network. For the prototype model, 9 distribution substation were used located equidistant within the road network grid as shown below:

Figure 13-12 Distribution substations used for the network synthesis

Step 5: ArcGIS network analysis

The network analyst tool in ArcGIS is used to connect the end nodes of each road to the closest substation along the road network. This results in detailed information about the routes used and also calculates the route lengths. The connectivity to the distribution substations can also be displayed via color coding as shown below:

Figure 13-13 Resulting network connectivity for the distribution substations

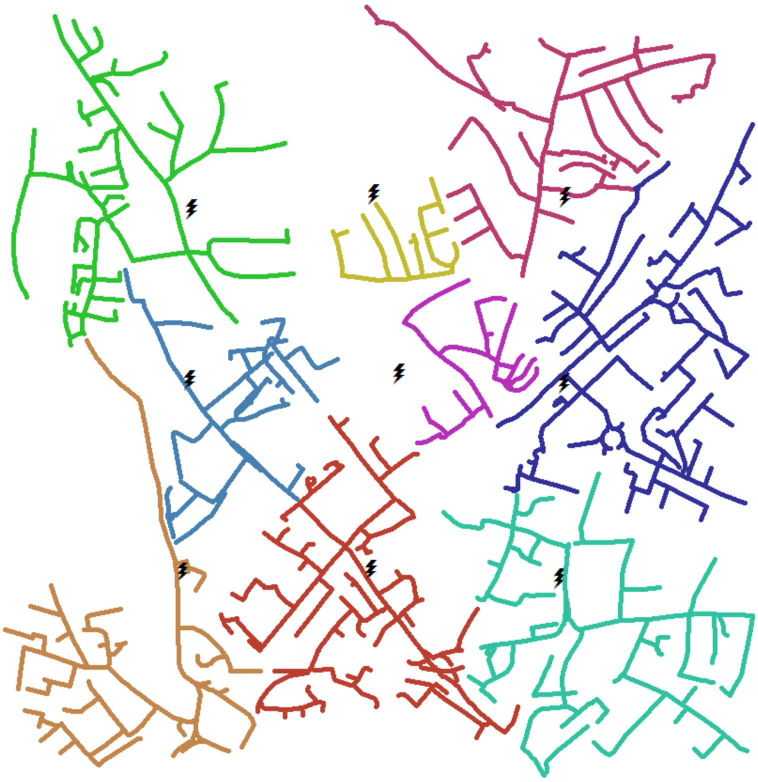


Figure 13-14 Detailed routing information generated by the network analysis tool

Directions (Closest Facility)

[–]	**Route: Location 2 - Location 1**	848.4 m
1:	Start at Location 2	
2:	Go northeast on 4000000025320915 toward 4000000025335843 / 4000000025335877	114.4 m
3:	Turn right on 4000000025335877	123.6 m
4:	Turn right on 4000000025335874	74.1 m
5:	Continue on 4000000025345889	14.3 m
6:	Continue on 4000000025335868	31.2 m
7:	Bear left on 4000000025320937	62.8 m
8:	Bear left on 4000000025345876	36.9 m
9:	Bear left on 4000000025305980	161.7 m
10:	Continue on 4000000025305981	171.5 m
11:	Turn right on 4000000025345897 and immediately turn right on 4000000025345901	49.6 m
12:	Continue on 4000000025351577	8.2 m
13:	Finish at Location 1	
	Driving distance: 848.4 m	

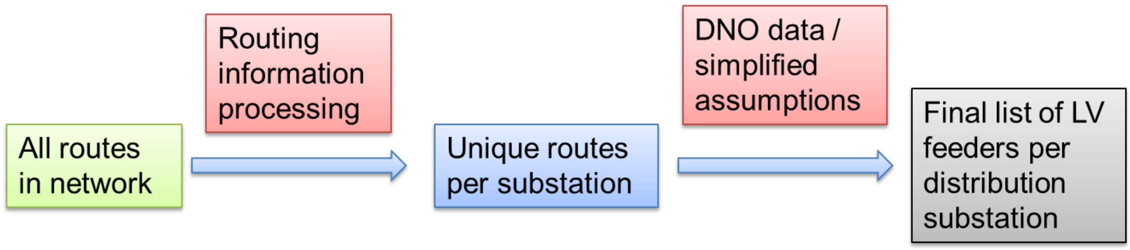
Step 6: List of unique feeders is derived

In order to perform the detailed electricity network upgrade cost calculations, a list of all LV feeders, their connectivity to the distribution substations and the distances of buildings connected to it are needed. This information is derived by processing the initial routing information generated by the network analysis. A semi-automated excel based tool has been developed to perform these processing steps.

All routes from each distribution substation to each connected road link end-node are first filtered to arrive at a list on unique routes i.e. routes that are not themselves part of a longer route. These unique routes connect the substation to the end-points of the distribution network. The unique routes are then associated with individual feeders. For the initial prototyping tool, a simple placeholder methodology has been used that associates all the unique routes in a quadrant, with the distribution substation at the centre, to be part of the same feeder. Thus a distribution substation could have up to 4 feeders.

In Phase 2, a methodology for this step will need to be developed in conjunction with DNOs – this could be based on DNO data, if available, or simplified assumptions e.g. average feeders per substation, minimum route overlap per feeder etc.

Figure 13-15 Processing steps to convert routing information into unique LV feeders



Step 7: Calculating distance to nearest distribution substation

Each LV feeder is associated with several unique routes. The detailed routing information, i.e. the roads used by each of these unique routes allows each building, based on its adjacent road, to be associated with the LV feeder connecting it to the nearest distribution substation. These relationships are used to associate individual buildings to LV feeders, the nearest substations and calculate the distance from each building to their connected substation. This procedure can then be repeated to connect distribution substations to the primary substations to calculate HV feeder connectivity and distances between each distribution substation and its nearest primary substation.

13.1.3. Cluster analysis

13.1.3.1. Overview

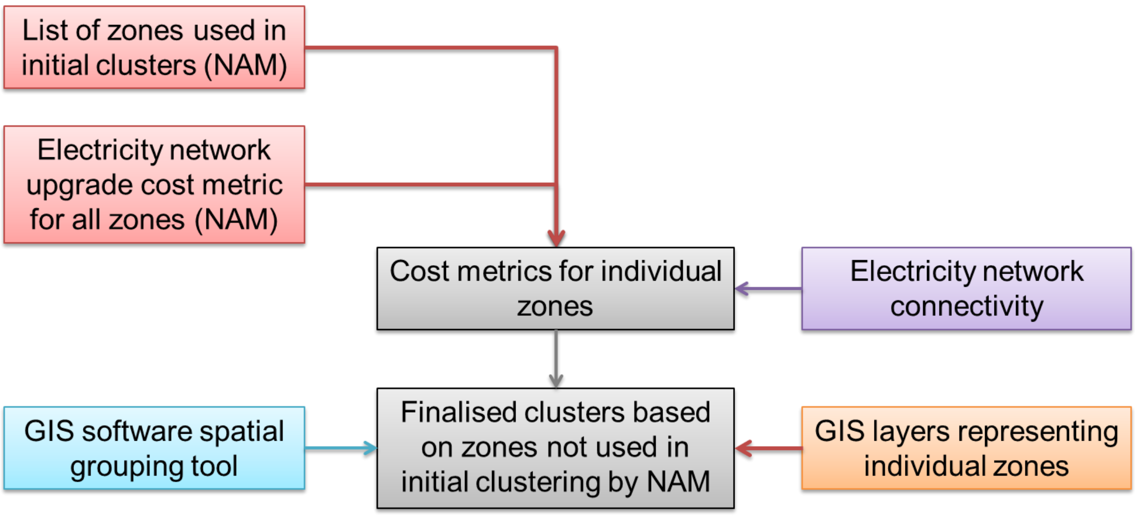
The Spatial Analysis Module is responsible for clustering zones that are not already used in the initial heat network based clusters in the NAM. These clusters consist of zones that are spatially contiguous and have homogenous electricity network upgrade costs. The key steps of this proof of concept work include:

1. Loading output layers from the network synthesis process within SAM into a Graphical Information System (GIS) package (ArcGIS in this case)

2. Using illustrative electricity network upgrade costs from NAM for individual geographic zones

3. Using GIS spatial clustering software algorithms and an Excel-based tool to cluster these zones

Figure 13-16 Schematic for the clustering methodology



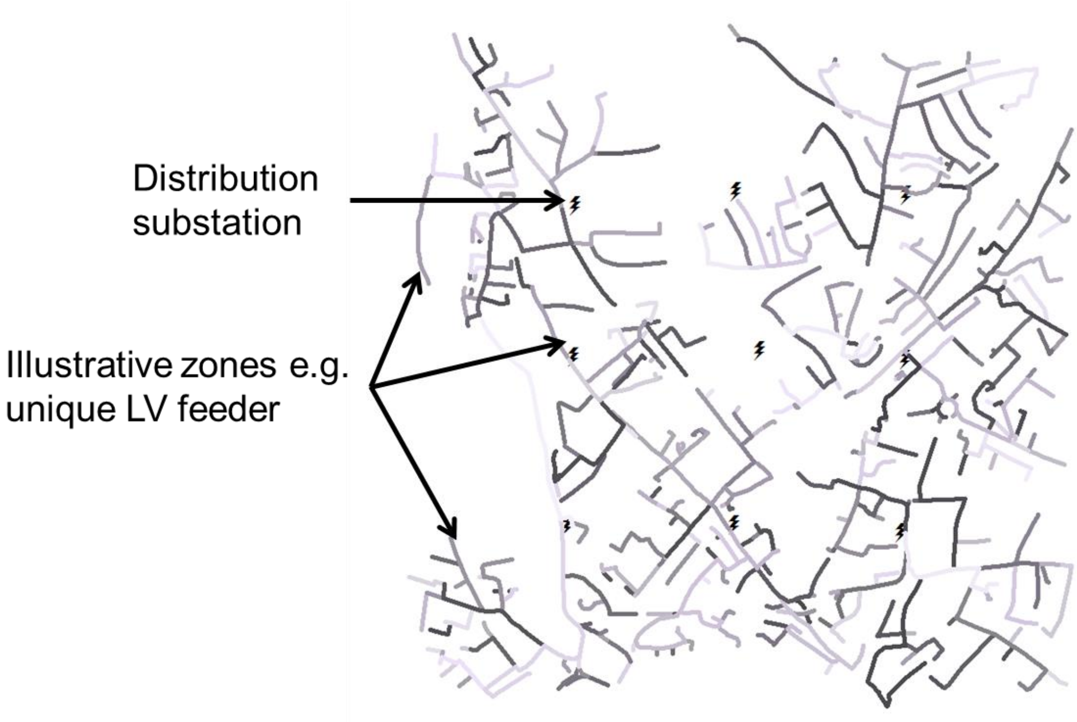
13.1.3.2. Description of methodology

Geographical information system software (ArcGIS) is used to perform the spatial clustering of the zones. The clustering algorithm creates spatially contiguous zones with homogenous clustering metric and also follows network connectivity constraints. The clustering metric is based on the NAM electricity network upgrade costs. This involves the following steps:

Step 1: Loading the synthesis electricity network

Outputs of the electricity network synthesis are loaded into ArcGIS to represent the spatial relationship between all the zones in the local authority. Sample data from Exeter is used for the prototyping.

Figure 13-17 Illustrative example of the zones in local authority



Step 2: Electricity network upgrade costs and use if zones in initial clusters is passed by NAM

An initial clustering of zones is performed in NAM based on the heat network costs. The NAM then passes information on the zones already clustered and the electricity network upgrade costs for the remaining zones. Placeholder data has been used for the network upgrade cost metric in the prototyping tool.

Figure 13-18 Data on zones used in initial clusters and network upgrade costs passed by NAM

Object_ID	TOID	Substation	Feeder	Electricity network cost metric	Used in initial clusters
1	4000000025288017	1	2	£ 68.73	1
2	4000000025305980	1	4	£ 73.33	1
3	4000000025305981	1	4	£ 72.72	1
4	4000000025306017	3	4	£ 83.95	0
5	4000000025320996	1	2	£ 71.25	1
6	4000000025320997	3	4	£ 71.28	0
7	4000000025335958	3	4	£ 75.97	0
8	4000000025335959	3	4	£ 87.62	0
9	4000000025345875	1	3	£ 73.96	1
10	4000000025345881	1	3	£ 97.66	1
11	4000000025345938	3	4	£ 69.57	0
12	4000000025345962	3	4	£ 79.59	0
13	4000000025345963	3	4	£ 75.80	0
14	4000000025345964	3	4	£ 67.41	0
15	4000000025351567	1	3	£ 69.63	1
16	4000000025384085	3	4	£ 80.89	0
17	4000000025447466	1	2	£ 77.26	1
18	4000000025447467	1	2	£ 90.14	1
19	4000000025447468	3	4	£ 64.65	0
20	4000000025295468	3	2	£ 67.00	0

Step 3: Cost metric is processed to ensure network connectivity constraints are followed

A key constraint on the clustering process is that the clusters cannot cross the boundaries of different network levels i.e. clusters must be a collection of whole building blocks, or part of a single large building block, but cannot be a collection of partial building blocks. Thus all LV feeders of a distribution substation are clustered first, then different distribution substations on the same HV feeder are clustered etc.

However, the clustering algorithm only works on a single cluster metric. The cost metric of LV feeder are therefore processed by adding to its value, based on the connectivity to distribution substation, HV feeder and primary substation. An illustrative example of this methodology is shown below:

Figure 13-19 Illustrative example of processing NAM cost metric into clustering metrics

Object_ID	Substation	Zone ID	Electricity network cost metric	Used in initial clusters	Clustering metric
1	1	12	£ 97.79	1	10.00
2	1	14	£ 76.84	1	10.00
3	1	14	£ 69.48	1	10.00
4	3	34	£ 96.63	0	30,096.63
5	1	12	£ 83.39	1	10.00
6	3	34	£ 60.22	0	30,060.22
7	3	34	£ 63.13	0	30,063.13
8	3	34	£ 88.12	0	30,088.12
9	1	13	£ 86.18	1	10.00
10	1	13	£ 92.52	1	10.00
11	3	34	£ 65.68	0	30,065.68
12	3	34	£ 86.21	0	30,086.21
13	3	34	£ 60.27	0	30,060.27
14	3	34	£ 85.29	0	30,085.29
15	1	13	£ 91.62	1	10.00
16	3	34	£ 90.63	0	30,090.63
17	1	12	£ 91.61	1	10.00
18	1	12	£ 71.17	1	10.00
19	3	34	£ 80.91	0	30,080.91
20	3	32	£ 73.29	0	30,073.29
21	9	91	£ 91.14	2	20.00
22	1	14	£ 91.72	1	10.00
23	4	43	£ 86.63	1	10.00
24	7	73	£ 92.44	0	70,092.44
25	7	73	£ 60.16	0	70,060.16
26	4	44	£ 84.64	1	10.00
27	8	83	£ 80.83	2	20.00
28	8	83	£ 81.95	2	20.00

Zones in the same initial heat network based clusters get the same metric

Zones connected to the same distribution substation have metric based on zone NAM cost value and substation ID

The generalised methodology is:

Clustering metric = NAM cost metric + Substation ID*10^4 + HV feeder ID*10^6 + Primary substation ID*10^8

Step 4: ArcGIS clustering algorithm is used to define additional clusters

The ArcGIS spatial grouping tool allows spatially contiguous clusters to be defined that have homogenous cluster metrics. This provides the final clusters that contain the zones not previously used in the initial clusters in NAM, as shown below:

Figure 13-20 Illustrative example of final clustering in ArcGIS

Thus, the final clusters are obtained whereby each zone is associated with its own cluster.

Figure 13-21 Illustrative outputs from clustering

Object_ID	Zone ID	Substation ID	Used in heat clusters	Cluster ID
1	12	1	1	3
2	13	1	1	3
3	14	1	1	3
4	21	2	1	3
5	22	2	1	3
6	23	2	1	3
7	24	2	1	3
8	31	3	0	2
9	32	3	0	2
10	33	3	0	2
11	34	3	0	2
12	41	4	1	3
13	42	4	1	3
14	43	4	1	3
15	44	4	1	3
16	51	5	0	2
17	53	5	0	2
18	61	6	0	2
19	63	6	0	2
20	64	6	0	2
21	71	7	0	1
22	72	7	0	1
23	73	7	0	1
24	74	7	0	1
25	81	8	1	4
26	82	8	1	4
27	83	8	1	4
28	84	8	1	4
29	91	9	1	4
30	92	9	1	4
31	94	9	1	4

Figure 13-22 Substations included in each cluster

Cluster ID	Cluster metric	Substation ID		
1	Electric	7		
2	Electric	3	5	6
3	Heat	1	2	4
4	Heat	8	9	

13.1.3.3. Conclusions

The ArcGIS clustering algorithms only allows spatial clustering based on a single metric. This single numeric value has to define:

3. The network upgrade cost as calculated in the NAM

4. The network connectivity constraints

This is achieved by processing the cost metric into a single meta-variable based on the network connectivity (e.g. the cost metric is adjusted to reflect the connection of zones to distribution substation, HV feeder and Primary substation) and use in initial clusters (all zones used in initial clusters get the same cost metric of their relevant cluster ID).

The result of the connectivity constraints is that they take precedence over the cost metric in determining final clusters, thereby reducing the number of possible clustering outcomes. Thus the initial clustering is determined by the network connectivity while the cost metric(s) from the NAM is/are a secondary clustering parameter. Thus all the distribution stations on a HV feeder need to be clustered before two HV feeders can be clustered together. The final clustering results in individual zones being grouped together in spatially contiguous clusters that have similar network upgrade cost metrics and obey network connectivity clustering rules.

13.2. NAM

The key tasks within the NAM proof of concept study can be summarised as follows:

- Test and help to select a suitable commercial network load flow tool that can provide the basis of the operational analysis required in NAM
  - A commercial network flow model considerably simplifies development and allows us to focus on interactions between the load flow model and the scenario costing tool in order to produce cost functions;
  - During the testing phase we have considered a number of criteria including: (i) performance; (ii) ability to handle GIS data sets; (iii) ability to model multiple energy vectors; (iv) ease of inputting and outputting data (particularly in Excel format); (v) support services provided by the software developer; (vi) cost
  - The two most suitable software tools that have been identified and tested are Siemens' PSS SINCAL and BCP Switzerland's NEPLAN.
  - Wider assessment of the two packages is described in deliverable (D2) Design Architecture
- We have carried out a range of simple load flow studies under different networks to understand what the most limiting network constraints are likely to be;
- Based on these load flow studies we have then run a simple internal process whereby a number of mock up test configurations and network reinforcement options were evaluated for a given area;
- Based on the above we have then created a set of representative network cost functions for the area in question in order to describe the obtained costs for the range of potential peak demands;

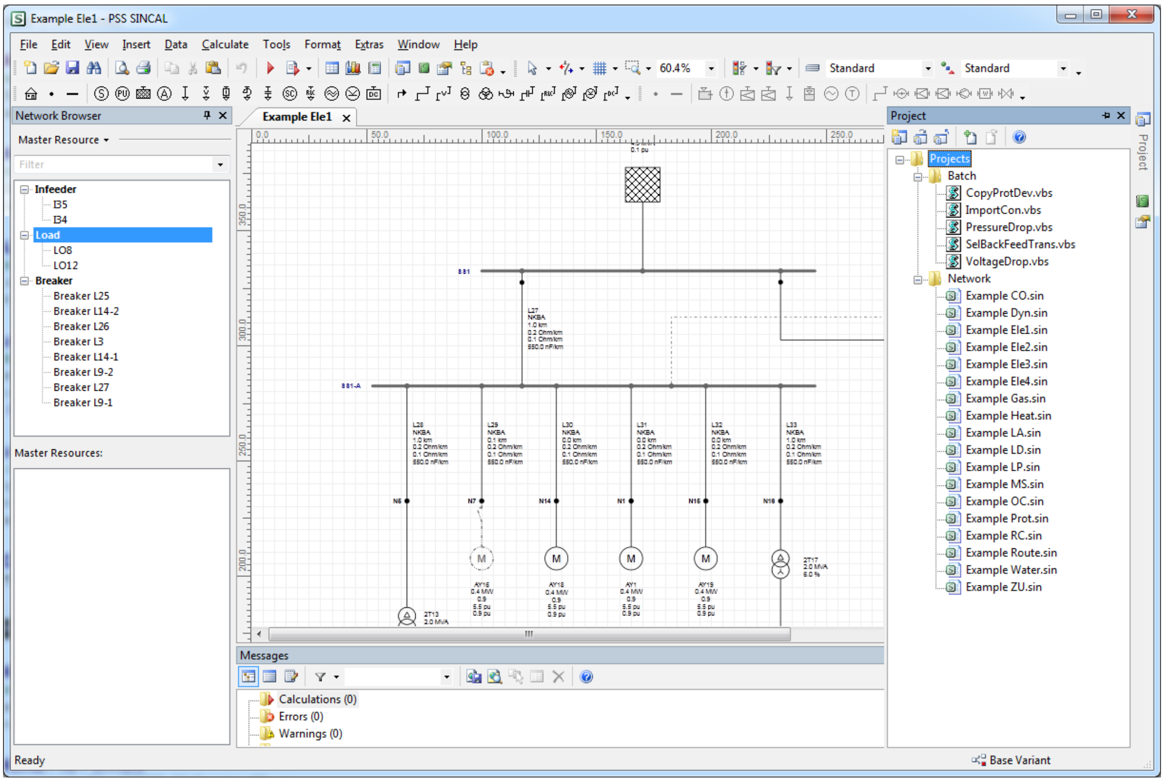
- Finally, we have also considered how the 'concentric-circles' methodology could work in practice by outlining the process by which 'heat-led' clusters can be created based on their heat network characteristics.

13.2.1. Selection of a suitable commercial network load flow tool

The first part of the proof of concept stage was to test PSS SINCAL and NEPLAN in terms of their suitability to provide the load flow modelling 'engine' for building up network cost and peak load pairs in the NAM.

SINCAL (SIemens Network CALculation) is a network flow modelling developed by SIEMENS as part of their Power System Simulator product suite. The software package is well developed and has an intuitive GUI as shown in Figure 13-23.

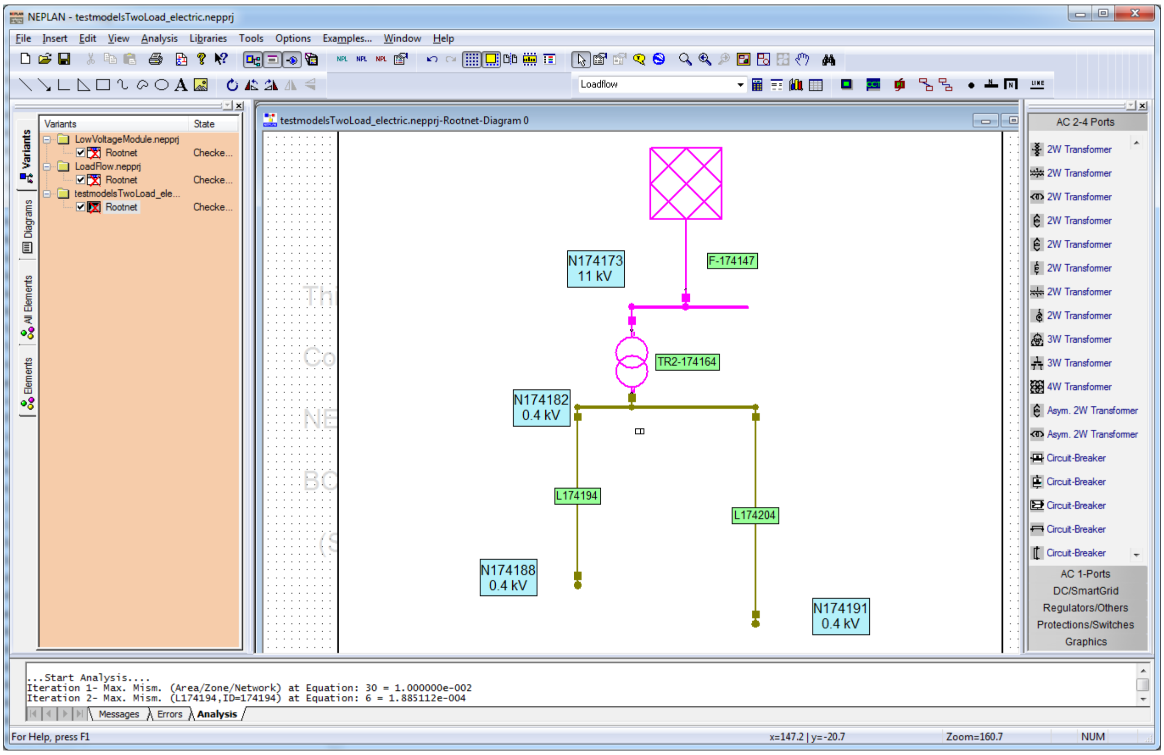
Figure 13-23 The SINCAL GUI



An alternative network flow modelling software package is NEPLAN, produced by the Swiss software developer BCP. The software package is similar to SINCAL, with modules to allow network flow modelling of electricity, heat, and gas networks, and an intuitive GUI as shown in Figure 13-24.

In the following sections we shall first focus on performing key NAM functions using SINCAL, for electricity and heat. We shall than discuss performing the same functions within NEPLAN, and any differences we perceive between SINCAL and NEPLAN.

Figure 13-24 The NEPLAN GUI



13.2.2. Electrical load flow studies in SINCAL

The following features of SINCAL were tested as part of the proof of concept for electrical load flow modelling:

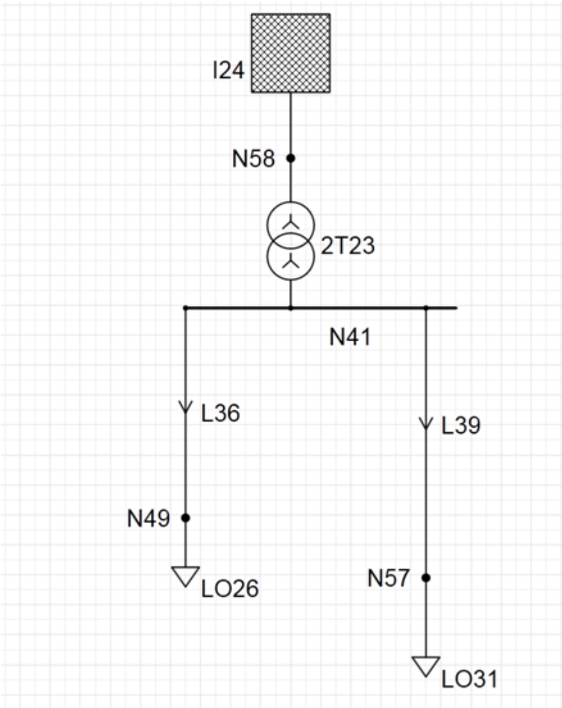
- Basic electricity distribution network construction using SINCAL's GUI;
- Importing properties for standard electricity distribution network components using SINCAL's GUI;
- Using Visual Basic (VB) script to execute SINCAL externally from command line;
- Using VB script to perform 'load scan', i.e. progressively increasing the load in a basic electricity distribution network in order to find the maximum allowable load;
- Building a curve of cost/peak load pairs for network reinforcement of a basic electricity distribution network.

These steps are described below in greater detail.

Basic network construction

The basic electricity distribution network that has been tested is shown below in Figure 13-25. For simplicity all components have been assumed to use three-phase connection interfaces although we recognise in reality that the majority of components in a LV distribution network use single-phase connections.

Figure 13-25 Basic electricity distribution network



The constructed distribution network consists of:

- One MV (11kV) infeeder (Line 'I24');
- One two-winding 11/0.4kV distribution transformer (Transformer '2T23');
- Two LV (0.4kV) feeder lines (Line 'L36' which is assumed to be 0.5km long and Line 'L29' which is assumed to be 0.8km long);
- Two LV load sources (Load 'LO26' and Load 'LO31');
- Four nodes connecting the above-mentioned network components (Node 'N41'[52], Node 'N49', Node 'N57' and Node 'N58').

Before the components can be added, first network 'levels' must be added, representing the different voltages in the network. In this simple example two levels are specified for a typical UK electricity distribution network, medium (11kV) and low (0.4kV) voltage. Nodes and elements may then be easily added using SINCAL's GUI toolbar.

52 It should be noted that Node 'N41' has been drawn as a 2D line, representing a busbar rather than a point node.

Importing properties for standard electricity distribution network components

To avoid the need for inputting all properties each time a network component is added, all components can be assigned to predefined 'standard types'. This is an important feature because limiting the total number of reinforcement options that are tested (see section 7.5) is critical to ensure a tractable optimization. In SINCAL, this could be achieved by limiting the number of 'standard types' that can be applied to each component when reinforcement occurs. In addition to keeping the problem tractable, this also reflects the level of standardisation that is typically seen in current electrical distribution networks in the UK.

In this proof of concept exercise we have tested reinforcement of the 0.4kV feeder lines, and the 11/0.4kV transformer. For simplicity we have only considered a small subset of potential network reinforcement options, which are consistent with the component data found in the ETI 2050 Energy Infrastructure Outlook cost database:

- For underground cables (Table 13-1) we have considered three cable sizes (95 mm², 185 mm² and 300 mm²) which are considered to be typical in urban distribution networks;
- For transformers (Table 13-2) we have only considered one standard size (500 kVA) which again is typical in urban distribution networks.

Table 13-1 Set of reinforcement options considered – underground cables

Material	Location	Voltage level (kV)	CSA (mm²)	Resistance (Ohm/km)	Reactance (Ohm/km)	Capacitance (nF/km)	Maximum current (kA)
Aluminium	Urban, buried	0.4	95	0.33	0.38	9.74	0.30
Aluminium	Urban, buried	0.4	185	0.17	0.37	10.00	0.55
Aluminium	Urban, buried	0.4	300	0.10	0.35	10.47	0.81

Table 13-2 Set of reinforcement options considered – distribution transformer

Location	High voltage (kV)	Low voltage (kV)	Maximum apparent power (MVA)
Urban	11	0.4	0.5

Using VB script to execute SINCAL externally, including performing a 'load scan' in a network

It is possible to automate the majority of the functionality of SINCAL by executing procedures outside the GUI using standard text commands. A number of languages can be used to provide this functionality via standard COM interfaces, including VBA and Windows Scripting Hosting.

Using SINCAL to quantify the maximum allowable load in a distribution network involves performing a number of load flow calculations at different levels of load, and assessing at what point network constraints (voltages, thermal limits etc.) are exceeded. This procedure is referred to here as a 'load scan'.

As part of the proof of concept for electrical load flow modelling we have developed a simple script that allows the user to perform a 'load scan' in the basic electricity distribution network from Figure 13-25 – the key steps are:

- Connect to SINCAL model;
- Execute steady state load flow calculation in the basic electricity distribution network;
- At different load points check the voltage drop at the remote end of the network as well as the electrical current through lines and transformer and compare these values with their maximum allowable limits to assess whether the network is operating within its limits;
- If within limits, increase load within network through simple multiplier on all load objects (real and reactive scaled equally) and repeat process;
- If any component out with limits, stop script and report final load levels and details of the limit that has been breached.

The developed script is shown in Figure 13-26 below. It uses some of the object oriented properties and procedures available from the SINCAL library and, once it is run, the load in the network is increased until one of the operational constraints is reached.

Figure 13-26 VB script for performing a network 'load scan' in SINCAL

```
' while loop - runs for a maximum of 2000 time, and while 5 operational
constraints are observed: voltage drop on load 1 & 2, max current in lines 1 & 2
and the transformer
Do While (iLoop < 2000) And (u_un1>VltDropPct) And (u_un2>VltDropPct) And
(Inb_Ele1 < 100) And (Inb_Ele2 < 100) And (Inb_Ele3 < 100)

        ' text output to separate outputs from each loop
        WScript.Echo vbCrLf & "-------- " & CStr( iLoop ) & " --------"

        ' We modify the load by adding a set % of load in each loop, for each
        both loads
        Call ModifyLoad( strLoad1, LoadObj1, StepSizePct )
        Call ModifyLoad( strLoad2, LoadObj2, StepSizePct )

        ' Start loadflow simulation
        SimulateObj.Start strLF
        If SimulateObj.StatusID <> siSimulationOK Then
                WScript.Echo "Load flow failed!"
                Exit Do
        End If

        ' Get all load flow results for node of both loads
        If LFNodeResult1 Is Nothing Then Set LFNodeResult1 = NodeObj1.Result(
        "LFNODERESULT", 0 )
        If LFNodeResult2 Is Nothing Then Set LFNodeResult2 = NodeObj2.Result(
        "LFNODERESULT", 0 )

        ' Get voltage drop results for each load
        If Not NodeObj1 Is Nothing And Not NodeObj2 Is Nothing Then
                u_un1 = LFNodeResult1.Item( "U_Un" )
                u_un2 = LFNodeResult2.Item( "U_Un" )

                ' Output voltage drop results to terminal
                WScript.Echo "Node1 U/Un = " & FormatNumber( u_un1 ) & "%"
                WScript.Echo "Node2 U/Un = " & FormatNumber( u_un2 ) & "%"
        End If



        ' Get all load flow results for each element (lines and transformer)
        if LFElementResult1 Is Nothing Then Set LFElementResult1 =
        ElementObj1.Result( "LFBRANCHRESULT", 1 )
        if LFElementResult2 Is Nothing Then Set LFElementResult2 =
        ElementObj2.Result( "LFBRANCHRESULT", 1 )
        if LFElementResult3 Is Nothing Then Set LFElementResult3 =
        ElementObj3.Result( "LFBRANCHRESULT", 1 )

        If Not LFElementResult1 is Nothing And Not LFElementResult2 is Nothing
        And Not LFElementResult3 is Nothing Then
```

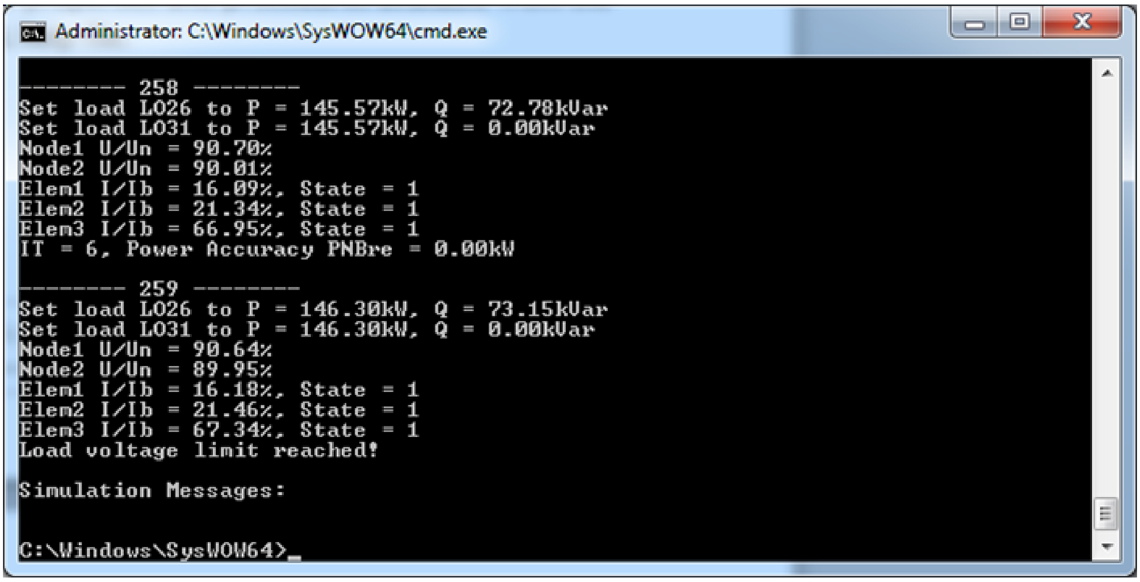
```
            ' Get operational state results (set by current limits of
            component)
            Dim LFElement1State, LFElement2State, LFElement3State

            LFElement1State = LFElementResult1.Item( "Flag_State" )
            LFElement2State = LFElementResult2.Item( "Flag_State" )
            LFElement3State = LFElementResult3.Item( "Flag_State" )

            ' Get current level as % of current limit
            Inb_Ele1 = LFElementResult1.Item( "Inb" )
            Inb_Ele2 = LFElementResult2.Item( "Inb" )
            Inb_Ele3 = LFElementResult3.Item( "Inb" )

            ' Output current and state results for each element
            WScript.Echo "Elem1 I/Ib = " & FormatNumber( Inb_Ele1 ) & "%,
            State = " & LFElement1State
            WScript.Echo "Elem2 I/Ib = " & FormatNumber( Inb_Ele2 ) & "%,
            State = " & LFElement2State
            WScript.Echo "Elem3 I/Ib = " & FormatNumber( Inb_Ele3 ) & "%,
            State = " & LFElement3State

            ' Output error if current limit reached and stop loop
            If LFElement1State = 2 Or LFElement2State = 2 Then
                  WScript.Echo "Element limit reached!"
                  Exit Do
            End If
      End If

      ' Output error if voltage drop limit reached and stop loop
      If (u_un1<=VltDropPct) Or (u_un2<=VltDropPct) then
            WScript.Echo "Load voltage limit reached!"
            Exit Do
       End If


    ' Display some global result information
    Call OutputLFAccurResult( SimulateNetworkDataSource )

' repeat loop if no error
    iLoop = iLoop + 1
Loop
```

Figure 13-27 below shows the output from performing a network 'load scan' using the developed script. It can be seen that assuming a total load of 292.6 kW in the network (evenly distributed between LO26 and LO31), the voltage on the line supplying the second load ("Node 2 V/Vn") has dropped below the limit of 90% of the nominal voltage level (400V). Similarly, the voltage on the line supplying the first load is near the limit (90.64%).

In terms of cable thermal limits, the electrical currents on the two LV feeders have been found to be well below their operational limits at that load point, with the first LV feeder operating at approximately only 16% of capacity and the second LV feeder operating at approximately 21% of capacity.

Finally, in terms of transformer thermal limits, the distribution transformer supplying the network has also been found to be operating well within its thermal limits (approximately 67%).

Figure 13-27 Output from performing a network 'load scan' in SINCAL



Building a curve of cost/peak load pairs for network reinforcement

Using the script outlined above, it is possible to assess the maximum allowable load for a range of different network reinforcement options in the basic electricity distribution network under consideration. For the proof of concept stage we have only considered a small subset of potential network reinforcement options:

- Replace an existing feeder line with a line with increased cable size – this would reduce voltage drop along the line and also increase its thermal limits;
- Add a new feeder line in order to increase headroom by splitting load onto multiple lines – this would reduce power flows across network lines, thus potentially alleviating voltage drop and cable thermal issues;
- Add a new distribution transformer in order to increase headroom by splitting load onto multiple transformers – this would reduce power flows across transformers, thus potentially alleviating voltage drop, cable and transformer thermal issues;

Table 13-3 shows a number of illustrative peak load/cost pairs for the basic distribution network under consideration.

- The first option (R #1) refers to the 'starting point' of the network where the load in Line 1 is supplied by one feeder line with a CSA of 95 mm², the load in Line 2 is also supplied by one feeder line with a CSA of 95 mm² and that there is also just one distribution transformer. Under this configuration, the maximum load that can be accommodated in the network is 98.2 kW, split evenly between P1 and P2. Any additional load after this point would result in the voltage drop limits across Line 2 to be exceeded.

- The second (R #2) and third (R #3) reinforcement options assume that Line 1 has now been replaced by a network line containing conductors with a CSA of 185 mm² and 300 mm² respectively, at an assumed additional cost of between £18,995/year to £21,585/year. However, no additional load can be accommodated in the network as the most limiting constraint remains voltage issues across Line 2.

- This is shown in the next four reinforcement options (R #4 to R #7) where it is assumed that Line 2 has been replaced by a network line containing conductors with a CSA of 185 mm² or 300 mm² and that Line 1 is supplied by conductors with a CSA of 95 mm², 185 mm² or 300 mm². It can be seen that the maximum load that can be accommodated in the network has now progressively increased between 104.2 kW to 174.2 kW, at an assumed additional cost of between £30,391/year to £56,120/year.

- It is also possible to add new lines in the network and increase headroom by splitting load onto multiple lines. This is explored under reinforcement options R #8 to R #16 – under option R #16, for example, there is a total of 447kW spread over 8 network lines, at a cost of £224,481/year.

- R#16 fails due to the thermal limits of single transformer. By adding a new transformer and spreading the load over 2 transformers, option R#17, a maximum load of 620 kW can now be supported, at a cost of £228,486/year.

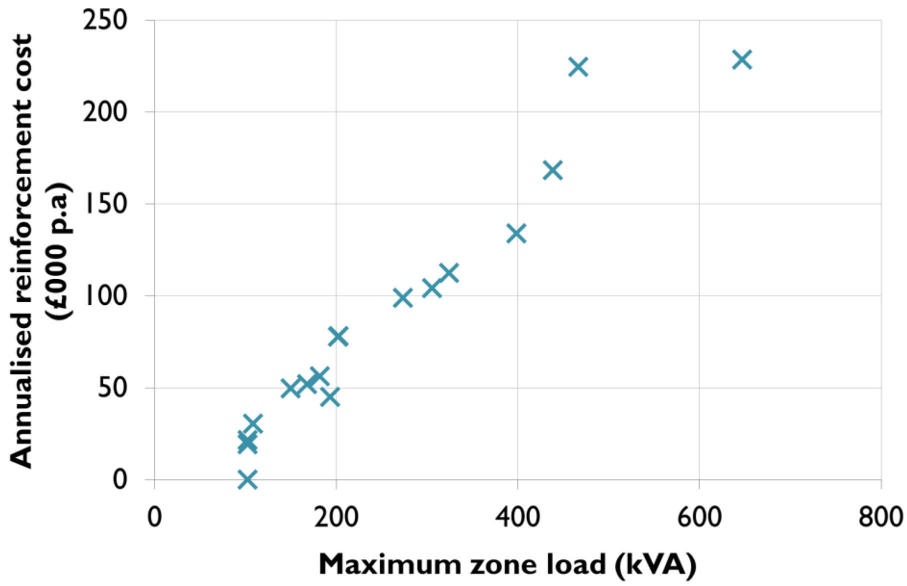
Table 13-3 Maximum allowable load / reinforcement cost pairs for the basic network

Option	Line 1 (mm²)	No. Lines	Line 2 (mm²)	No. Lines	No. Trans	Max load – zone (kW)	Max power - zone (kVA)	Limiting Element	Cost (£/year)
R #1	95	1	95	1	1	98.2	102.5	Line 2 V	0
R #2	185	1	95	1	1	98.2	102.5	Line 2 V	18,995
R #3	300	1	95	1	1	98.2	102.5	Line 2 V	21,585
R #4	95	1	185	1	1	104.2	108.8	Line 1 V	30,391
R #5	185	1	185	1	1	144.0	150.3	Line 1 V	49,386
R #6	300	1	185	1	1	160.8	167.9	Line 2 V	51,976
R #7	300	1	300	1	1	174.2	181.8	Line 1 V	56,120
R #8	95	2	95	2	1	185.1	193.3	Line 2 V	44,896
R #9	95	2	185	2	1	194.4	203.0	Line 1 V	78,050
R #10	185	2	185	2	1	262.0	273.5	Line 1 V	98,772
R #11	300	2	185	2	1	292.6	305.5	Line 2 V	103,952
R #12	300	2	300	1	1	193.4	201.9	Line 2 V	77,705
R #13	300	2	300	2	1	310.6	324.3	Line 1 V	112,241
R #14	300	3	300	2	1	382.0	398.8	Line 2 V	133,825
R #15	300	3	300	3	1	420.0	438.5	Line 1 V	168,361

R #16	300	4	300	4	1	447.0	466.7	Trans. 1 I	224,481
R #17	300	4	300	4	2	620.0	647.3	Line 1 V	228,486

These "Peak load / reinforcement cost" pairs are plotted in Figure 13-28, to show the range in costs for this simple network. These valid reinforcement options are the key output from the NAM, to be used in the POM when designing the optimal heat delivery pathway.

Figure 13-28 Building a curve of cost/peak load pairs for network reinforcement



13.2.1. Heat network load flow modelling in SINCAL

For heat networks the NAM must provide two functions:

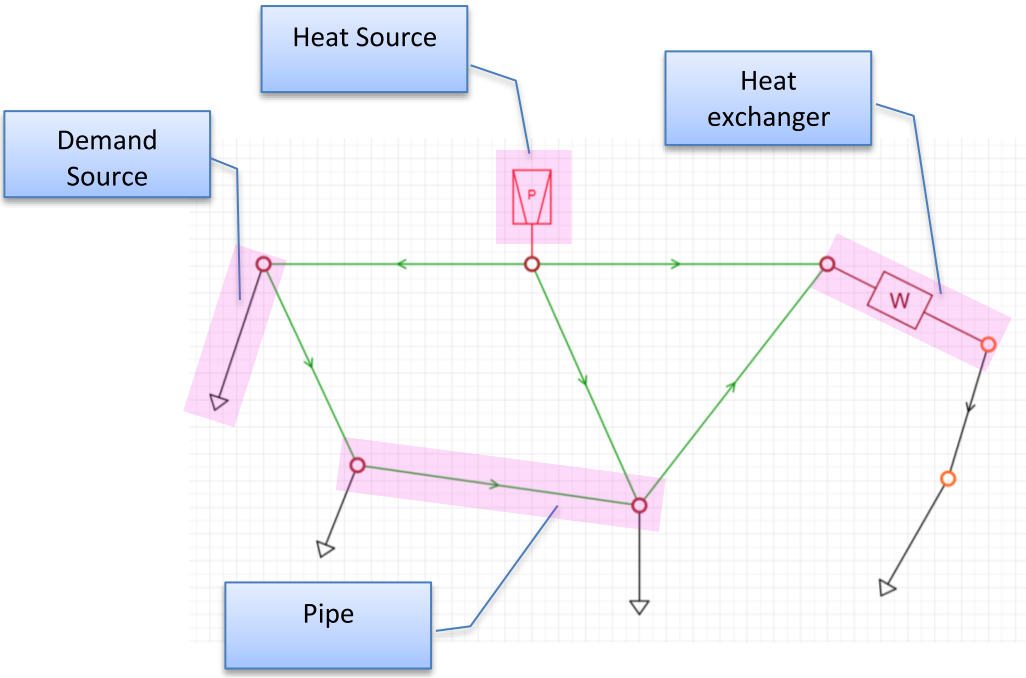
- *Network build cost estimates* – similar methodology to electricity networks, testing each new build configuration by progressively increasing load and checking for any breach of operational limits
- *Heat-led clustering* – defining the boundaries of clusters by assessing the cost of potential heat networks formed by connected zones

We have assessed the Heat networks functionality of SINCAL against the both functions, both are discussed below. We have characterised a heat network as having four main components:

- Heat source (could be large transmission pipe, or dedicated heat source such as CHP)
- Pipe network
- Heat exchangers and pumps
- Demand sources

Figure 13-29 shows a simplified heat network as represented in SINCAL, with each component type highlighted. This network is a simplified version of a real district heating network, and is supplied with SINCAL as an example. We shall use an extended version of this network to test the two key NAM functions described above.

Figure 13-29 Example heat network



Network reinforcement cost estimates

Calculating peak load / cost estimates for all reinforcement options is conceptually very similar for heat networks and electricity networks. The methodology described in section 13.2.2 can be used for heat networks, in terms of progressively increasing the load on each demand source in the network until an operational limit is breached.

For heat networks operational limits are set by:

- Temperature
- Flow velocity in pipes
- Pressure in pipes
- Power output from heat exchangers and pumps

There is some operational flexibility in these parameters, for example increasing the temperature of heat flows to allow for reduced flow velocities. To fully test reinforcement options for heat networks, load flow modelling should be performed for a range of operational states, and the peak load/ cost information should be sent to the POM for each reinforcement configuration *and* operational state. The POM may then choose to build a network that operates at a higher temperature for example, incurring greater heat losses from the pipes but resulting in a higher peak capacity as a result of relieving peak flow velocity operational constraints.

In SINCAL the input heat source is typically defined as operating at a constant pressure and temperature. The flow velocity at each point in the network is therefore an output of the load flow modelling. To test for varying operational states, the input pressure and temperature may be varied iteratively and the peak load calculated in each state. This can be done using a two further "while loops" in a vbs script, scanning through a series of temperatures and pressures and calculating the load at which operational limits are met. These extra loops have the potential to increase the number of load flow calculations and POM optimization variables considerably, and so it is expected that only a limit set of typical operating temperatures and pressures would be used, to keep the problem tractable.

Of the four operational limits listed above, two are inputs (pressure and temperature) and can be set at suitable levels within practical limits. One is a simple output (flow velocity) and can be checked against a pre-defined operational limit. The final operational limit is on the power output from heat exchangers and pumps. In SINCAL these components are not rated by power; a heat exchanger, for example, is specified by the input and output temperatures, with power flow being an output from the load flow calculation. This means that the network reinforcement option for these components (i.e. size of heat exchanger and associated peak power flow) cannot be specified within SINCAL itself, and must be inferred ex-post using the output power flow of each component. This differs subtly from the electricity network reinforcement cost methodology, where all components can be fully specified within SINCAL.

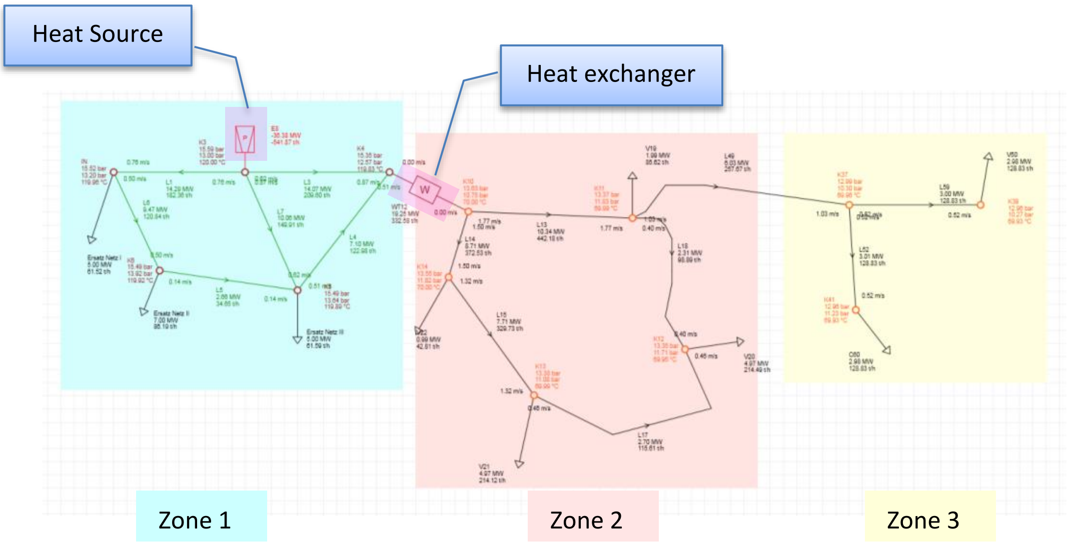
Despite the differences outlined above for heat networks (need to test different operating points, ex-post calculation of heat exchanger and pump limits) the general methodology for network reinforcement cost calculation is broadly similar for both heat and electricity, and we have confidence that using SINCAL and vbs scripts provides a suitable solution.

Heat-led clustering

The proposed methodology for clustering "zones" into larger building blocks is to define these clusters based on potential heat network topology in the first instance, as described in section 7.5.5. In this section we outline the process using a simple example.

Figure 13-30 shows a simple heat network, consisting of three distinct zones. Zone 1 is an industrial area, with high demand and a need for high temperature heat (120°C). Zones 2 and 3 represent a mixture of commercial and domestic demand sources, lower in volume and temperature (70°C). This network is an adaptation of the example heat network supplied within SINCAL.

Figure 13-30 Potential heat-led cluster, 3 zones



Using the "concentric circles" methodology, the cost of delivering load through a heat network must be calculated for progressively larger and large cluster sizes, until some metric indicates that a suitable boundary has been met. In the example described here we have expanded from Zone 1 into the other zones, calculating the levelised cost of heat delivery (associated with the network itself and not the source) at each stage. A boundary is set when the levelised cost of heat delivery is minimized. For this example we do not consider the additional boundary constraint, that heat-led clusters must not cross key building blocks of the existing electricity network topology.

Currently there are few district heat networks in the UK, and in zones with no current network the NAM must "build" a suitable heat network capable of supplying load to all demand sources. In this example we do not fully perform this step – we assume that the network skeleton has already been designed, routing pipes along existing roads for example. However, we do not assume that the pipe size has been specified. We make the simplifying assumptions that all pipes in the network are of the same size, and perform multiple load flow calculations using different pipe sizes to check for cost and if operational limits are violated.

The components used to cost the network are: heat source, pipes, and heat exchanger. Details of the options considered are given in Table 13-4 and Table 13-5.

Table 13-4 Standard pipes and associated costs

Type	Operating temperature	Diameter	Cost
-	℃	*mm*	*£/km/year*
120°C - 300mm	120	300	184,310
120°C - 150mm	120	150	170,253
70°C - 300mm	70	300	175,224
70°C - 150mm	70	150	89,595

Note: Costs taken from ETI database – Urban area, medium cost estimate

Table 13-5 Standard heat exchangers and associated costs

Type	Capacity	Cost
-	*MW*	£/year
120°C to 70°C	1.5	7,364
120°C to 70°C	18	22,653
120°C to 70°C	50	37,720

Note: Costs taken from ETI database – Urban area, medium cost estimate

The process followed in this heat-led-clustering example is as follows:

1. Begin with Zone 1 (zone of highest demand)
2. Assume heat source is situated in Zone 1, a co-firing CHP
3. Perform load flow calculation twice, first using 150mm diameter pipe throughout network, then with 300mm diameter pipe
4. Check all operational limits: temperature and pressure set as inputs, power flow constraint not relevant as no heat exchangers in Zone 1, but check peak velocity – assume 2m/s is maximum flow velocity
5. Discount networks with a pipe size that results in a breach in operational limits, i.e. peak velocity > 2m/s
6. Calculate cost of heat network for configurations satisfying operational limits, adding cost of all components and scaling by demand to give a £/kW/y figure for levelised cost of heat
7. Select network configuration with lowest levelised cost of heat network
8. Extend cluster to include Zone 2 as well as Zone 1
9. Add heat exchanger between Zones 1 and 2 to account for difference in temperature requirements in these zones

10. Repeat steps 3-5 above for new enlarged cluster

11. Check power flow through heat exchanger and choose the appropriate standard size (1.5MW, 18MW, 50MW)

12. Repeat steps 6-7

13. Enlarge cluster to include Zone 3

14. Repeat steps above until cost of network increases due to enlargement of cluster

Table 13-6 below shows the results of this heat-led clustering process. With a cluster containing only Zone 1, operational constraints are observed for networks of both pipe sizes (150mm and 300mm). In this case, the lower cost option is to use the smaller diameter pipe size of 150mm, at a total levelised cost for the network of 27 £/kW/year.

When Zone 2 is added to the cluster, the smaller pipe size is no longer suitable, due to flow velocities breaching the operational limit of 2m/s. It should be noted that this limit is somewhat arbitrary at this proof of concept stage, and further research is required to define an appropriate limit. Despite the need to use larger pipes, and the addition of a 18MW heat exchanger, the total cost of heat delivery is reduced slightly by adding Zone 2, to 26 £/kW/year, due to Zone 2 being at a lower temperature with associated lower cost pipes.

When the cluster is extended to contain Zone 3, the additional load requires a larger 50MW heat exchanger. This drives an increase in total network levelised costs, to 28 £/kW/year. At this point, and using just one very simple increasing cost threshold condition, the concentric circles method has found the boundary of the cluster, as the addition of additional zones has increased the cost of heat delivery. The final cluster containing Zones 1 & 2 and is the lowest cost cluster size.

Table 13-6 Heat-led clustering results

Zones	Cluster Load	Pipe diameter	Pipe length @ 120C	Pipe length @ 70C	Heat exchanger load	Heat exchanger size	Peak flow velocity	Flow limited reached?	Pipe Cost	Heat Exchanger cost	Total cost	Levelised Cost
-	*MW*	*mm*	*km*	*km*	*MW*	*MW*	*m/s*	-	*£ 000 / year*	*£ 000 / year*	*£ 000 / year*	*£/kW/year*
1	20	150	3.139	0	0	0	1.69	OK	534	0	534	27
1	20	300	3.139	0	0	0	0.422	OK	579	0	579	29
1 & 2	30	150	3.139	1.087	13.16	18	4.759	FAULT				
1 & 2	30	300	3.139	1.087	13.16	18	1.19	OK	769	23	792	26
1 & 2 & 3	36	150	3.139	2.327	19.25	50	7.097	FAULT				
1 & 2 & 3	36	300	3.139	2.327	19.25	50	1.775	OK	986	38	1024	28

13.2.2. Comparison of NEPLAN and SINCAL

We have tested NEPLAN, trying to perform the same functions tested in SINCAL, namely:

- Ability to set "standard types" for components
- Load flow modelling electricity
- Load flow modelling heat
- External Scripting

Ability to set "standard types" for components

NEPLAN allows "libraries" to be created that contain a predefined list of standard components. A new library was created with 4 standard components – 3 feeder lines and 1 transformer, using parameters in the ETI network cost database. Standard types were then applied to the network elements of a simple network.

Load flow modelling electricity

The same simple two load network tested previously within SINCAL was recreated in NEPLAN and a load flow study performed. Very similar results (<1% difference) for voltage drop and current flow were output from NEPLAN and SINCAL. When network components operate outside their limits, SINCAL gives better feedback of this from the GUI. NEPLAN and SINCAL both output a full results table with the operational state of each model element which can be queried.

Load flow modelling heat

Using the sample network file supplied with NEPLAN a load flow study was performed. The available network components are very similar to SINCAL, with pipes, heat sources, pumps and heat exchangers and similar defining parameters for each component. Changing operational states of the network (temperature, pressure) is fairly straightforward. While there is a caveat that we have not tested the same heat network in SINCAL and NEPLAN and so cannot compare results, the functionality seems very similar.

External Scripting

Within the Trial version of NEPLAN tested here there was no ability to use external scripts to call NEPLAN functions. However, the NEPLAN Programming Library (NPL) is a C/C++ API library containing functions to access most of the functionality of NEPLAN through a C/C++ program. The use of this library is untested, but from the documentation would appear to allow all of the features necessary for the load flow "wrapper" sub module - opening model, running simple load flow calculation, changing load, extracting results, etc. The use of C/C++ in NPL, rather than visual basic scripts as for SINCAL makes external execution somewhat more complicated using NEPLAN.

13.3. POM

This section outlines how the feasibility of key optimisation structures required within the EnergyPath^{TM} Design Tool have been tested using an adapted version of the ETI's EU Energy Systems Modelling Environment (ESME) framework.

The process has also undertaken a number of performance tests associated with the re-purposed model. The aim is to create problems of the magnitude that the EnergyPath^{TM} Design Tool could face to evaluate if they can reasonably be solved using the ESME framework. Furthermore, we will attempt to break down performance drop due to increasing key data items such as building archetypes, clusters, time slices and time periods as well as resulting from the newly added functions, in particularly as many of these are reliant on MIP (Mixed Integer Programming) functionality which can significantly increase the solving time.

In this testing exercise, it is important to note that we are not interested in the input data and solutions coming from the proof-of-concept model, only the feasibility of demonstrating the structure and the solver performance.

13.3.1. Context and objectives of proof of concept testing

The EnergyPath^{TM} Design Tool is aimed at representing energy systems in local areas. As such, it will require the representation of key features that are currently not present in ESME:

- **Detailed modelling of building archetype conversions (LP)**: representation of multiple building archetypes (e.g. apartment building, semi-detached house, etc.) with different levels of conversion/upgrade possibilities (including heating technology and insulation);
- **Lumpy investment (MIP)**: modelling discrete investment sizes (e.g. CHP unit or building retrofit) to more accurately represent the size and cost of the investment;
- **Discrete interdependent choices (IP)**: representing binary choices with consequences on the availability of a set of technologies (e.g. network reinforcement options with or without meshing). The decision to build a particular option could be mutually exclusive or alter the subset of remaining choices.
- **Piecewise linear approximation (LP)**: to provide more flexibility in defining functions (e.g. representing economies of scale or peak electricity demand due to diversity factors). Piecewise approximation is a way to model concave curves in a cost minimization problem (which would otherwise not be feasible as the highest marginal price step from the perspective of the objective function needs to be undertaken first) while keeping the formulation linear and solving times to a minimum.

We have also undertaken a set of performance tests using the models:

- With/without these added features;
- With various datasets so as to be able to allocate a performance impact to a particular dimension of the problem. The key dimensions considered here are:
  - **Clusters**: from 2 to 20 to 40;

- **Building archetypes**: 10 to 100;
- **Timeslices**: from 2x2 (characteristic day x within day time periods) to 5x5; and
- **Pathway timeperiods**: from 4 to 8.

ESME simplification

The ESME model contains a number of elements (data items and constraints) which add complexity to the model (e.g. transport and industry, flexibility reserve margin for electricity), but which are not relevant to the EnergyPath™ Design Tool formulation, so we removed this extraneous features for the purposes of this testing.

Buildings archetypes and retrofits

In order to better represent the variety of buildings present in local areas, we have modelled a series of buildings as combination of the two following notions:

- A building archetype represents a type of building (e.g. apartment building, semi-detached house, etc.); and
- A level of thermal efficiency is used to represent the thermal performance of the building, which in turn depends on its insulation and its heating technology (e.g. heat pump, gas boiler, district heating, etc.).

In our datasets, each building archetype is represented with 10 levels of performance. For each building archetype, it is possible to retrofit the building from any thermal performance level to any other thermal performance level (90 possible retrofits per archetype) as seen in Table 13-7. This approximates the type of problem within the EnergyPath™ Design Tool whereby one base archetype can be converted into many others and the conversion could represent insulation upgrades, a change in heating system or both.

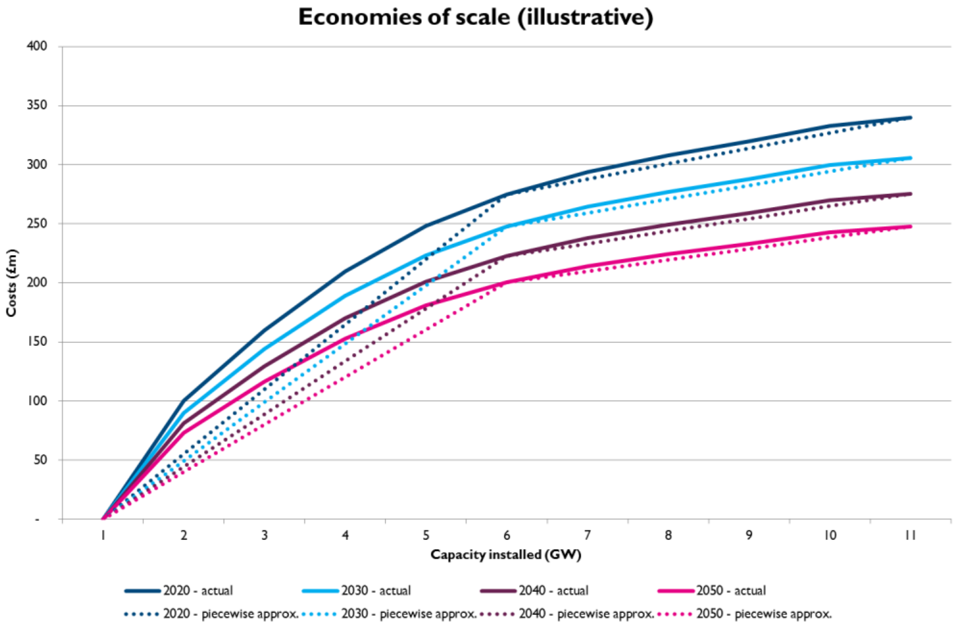
Table 13-7 Retrofit options for one building archetype

	Th1	Th2	Th3	Th4	Th5	Th6	Th7	Th8	Th9	Th10
Th1		✓	✓	✓	✓	✓	✓	✓	✓	✓
Th2	✓		✓	✓	✓	✓	✓	✓	✓	✓
Th3	✓	✓		✓	✓	✓	✓	✓	✓	✓
Th4	✓	✓	✓		✓	✓	✓	✓	✓	✓
Th5	✓	✓	✓	✓		✓	✓	✓	✓	✓
Th6	✓	✓	✓	✓	✓		✓	✓	✓	✓
Th7	✓	✓	✓	✓	✓	✓		✓	✓	✓
Th8	✓	✓	✓	✓	✓	✓	✓		✓	✓
Th9	✓	✓	✓	✓	✓	✓	✓	✓		✓
Th10	✓	✓	✓	✓	✓	✓	✓	✓	✓	

Piecewise linear approximation

Piecewise linear approximation can be used to represent concave curves in a cost minimization problem. Economies of scale (e.g. from reduced labour costs for solar PV when installing large quantities in one area simultaneously as described in Figure 13-31) as well as other issues (e.g. diversity effects when scaling of peak electricity demand with the number of buildings using electrical heating technologies) are examples of this.

Figure 13-31 Economies of scale approximate representation



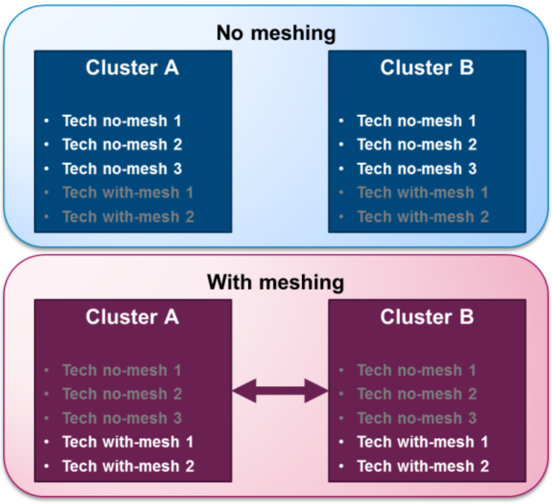
Lumpy investment

In the EnergyPath™ Design Tool, it is better to represent discrete investment decisions (e.g. CHP) as they happen in reality (i.e. this is a binary choice to invest or not) and reflect a more limited choice of investment options (e.g. CHP plant only available in 1, 3 or 10 MW sizes and not an idealised MW fraction (e.g. 6.945… MW) as would be constructed as part of a Linear Program representation.

Discrete interdependent choices

Conditional constraints can be put in place to represent discrete investment decisions that impact the availability of technologies (e.g. electricity network reinforcement options costs and availability would depend on whether or not distribution network of contiguous clusters are meshed). Figure 13-32 presents a mockup example of network reinforcement technologies available with and without meshing clusters A and B.

Figure 13-32 Discrete interdependent choices (meshing clusters A & B)



13.3.2. Additional features modelling in AIMMS

This section details how the new features (i.e. new building archetypes and retrofits, lumpy investment, piecewise linear approximation and interdependent binary choice) can be implemented using an adapted ESME framework.

Some of these features (i.e. building archetypes and retrofits, piecewise linear approximation[53]) are compatible with linear programming (LP) while others (i.e. lumpy investment and interdependent binary choices) involve a binary decision and require mixed integer programming (MIP). It is expected that LPs will solve faster than MIPs, since solving a MIP requires solving the associated LP and finding the combination of integers that approaches the objective function to the LP solution as closely as possible.

Buildings archetypes and retrofits (LP)

We have adapted the retrofit formulation already available in ESME to be able to mimic the conversion of a building archetype into another. In ESME this is limited to one-way technology improvements, but in the re-purposed model we are able to change back and forth across all possible combinations of retrofit – i.e. to mimic switching between archetypes with competing heating technologies.

We have prepared two databases with 100 and 1000 base archetypes (i.e. representing 900 and 9,000 possible combinations of archetypes available in any given timeperiod). We have reduced the amount of time slices to a minimum (2x2: 2 seasons crossed with 2 diurnal time slices) so the problem would solve as fast as possible and to limit data exchanged between MS SQL Server and AIMMS.

[53] Technically this requires binary integers, but is recognised as a special form of optimisation problem (Type 2 Special Ordered Set) that can be tackled very efficiently by a number of commercial LP solvers.

Lumpy investment (MIP)

Implementing lumpy investment in AIMMS requires creating integer variables to model:

- **Technology build:** Number of units of each technology built in a given time period ("Number_units_built" variable below);

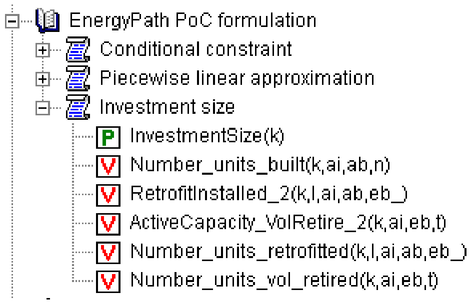
- **Technology retrofit:** Number of units of each technology retrofitted in a given time period ("Number_units_retrofited" variable below); and

- **Voluntary retirements:** Number of units of each technology voluntarily retired in a given time period ("Number_units_vol_retired" variable below).

These would be decision variables under the control of the optimization engine, which would force the use of MIP. These decision variables are coupled with an "InvestmentSize" parameter, defining the nameplate capacity for each technology. For this proof of concept testing, we chose to limit lumpy investment modelling to power plant. Figure 13-33 Formulation of lumpy investment in AIMMS

presents a tree view of the parameters and variables created to implement lumpy investment in AIMMS.

Figure 13-33 Formulation of lumpy investment in AIMMS



The installed capacity formulation is changed to reflect lumpy investment for power plant. Figure 13-34 shows the AIMMS code used to calculate the installed capacity of technologies: for power plant, we have swapped a continuous variable "NonAssoc_InstalledCapacity" for an integer one "Number_units_built".

Figure 13-34 Formulation of installed capacity using discrete investment sizes

Type	Variable
Identifier	InstalledCapacity
Index domain	(k,ai,ab,n) in NewTechInstalls_kibn
Text	Decision variable that represents quantity of new technology capacity k, installed at node i, in build year b, cost tranche n.
Range	free
Unit	TechCapacityUnitScaledCapacity(k)
Default	
Property	Inline
Priority	
Nonvar status	

```
if Problem_type_switch = 'LP' then
        NonAssoc_InstalledCapacity(k,ai,ab,n)
elseif Problem_type_switch = 'MIP' then
        if k in PowerPlants_k then
                if k in Piecewise_techs_k then
                        sum(l_s, Lambda(k, l_s, ab) * piecewise_approx_x(l_s, ab))
                else
                        InvestmentSize(k) * Number_units_built(k, ai, ab, n)
                endif
        else
                NonAssoc_InstalledCapacity(k,ai,ab,n)
        endif
endif;
```

Similarly, we replace the continuous “RetrofitInstalled” variable by the discrete one “Number_units_retrofitted” for power plant as shown in Figure 13-35. Voluntary retirement is handled the same way.

Figure 13-35 Formulation of retrofit capacity using discrete investment sizes

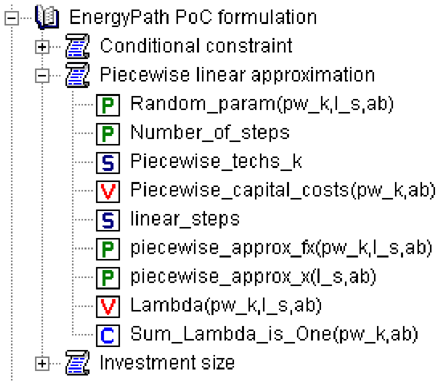
Type	Variable
Identifier	RetrofitInstalled_2
Index domain	(k,l,ai,ab,eb_) in FeasibleRetrofits_klibb
Text	Quantity of parent tech capacity k, retrofitted into technology l, at node i, fitted in year b, parent build year b_
Range	nonnegative
Unit	RetrofitUnitScaled(ai,k,l)
Default	
Property	
Priority	
Nonvar status	
Definition	see below

```
if Problem_type_switch = 'LP' then
        RetrofitInstalled(k,l,ai,ab,eb_)
elseif Problem_type_switch = 'MIP' then
        if k in PowerPlants_k then
                InvestmentSize(k) * Number_units_retrofitted(k, l, ai, ab, eb_)
        else
                RetrofitInstalled(k,l,ai,ab,eb_)
        endif
endif;
```

Piecewise approximation (LP/MIP)

Figure 13-36 presents a tree view of the sets, parameters, variables and constraints created to implement piecewise approximation.

Figure 13-36 Formulation of piecewise linear cost in AIMMS



In particular, we note:

- "Linear_steps" is the set used to define the number of steps in the piecewise linear curve;
- "Piecewise_approx_x" and "Piecewise_approx_fx" are used to define the piecewise linear curve interpolation points; and
- "Lambda" is the decision variables used as weights for the steps in the piecewise linear curve so as to effectively calculate values between the interpolation points. Figure 13-37 presents the formulation of piecewise approximated capital costs ("Piecewise_capital_costs") using the "Lambda" decision variables as weights to the "Piecewise_approx_fx" parameters.

Figure 13-37 Formulation of piecewise linear approximation of capital costs

Type	Variable
Identifier	Piecewise_capital_costs
Index domain	(pw_k, ab)
Text	
Range	nonnegative
Unit	£/year
Default	
Property	
Priority	
Nonvar status	
Definition	sum(l_s, Lambda(pw_k, l_s, ab) * piecewise_approx_fx(pw_k, l_s, ab))

In this proof of concept testing, we limited the piecewise approximation to capital costs for solar PV (in "Piecewise_tech_k" set). The capacity of solar PV can be either at one interpolation point or between two interpolation points, therefore the sum of all lambdas across linear steps must be equal to one (i.e. lambda can be viewed as weights). Furthermore, only two consecutive lambdas can be non-zeros by construction: this formulation has been codified as Specially Ordered Set #2 (SOS2) and can be recognized by a LP solver. Figure 13-38 shows the implementation of a SOS2 constraint on "Lambda" decision variables in AIMMS.

Figure 13-38 Formulation of SOS2 constraint

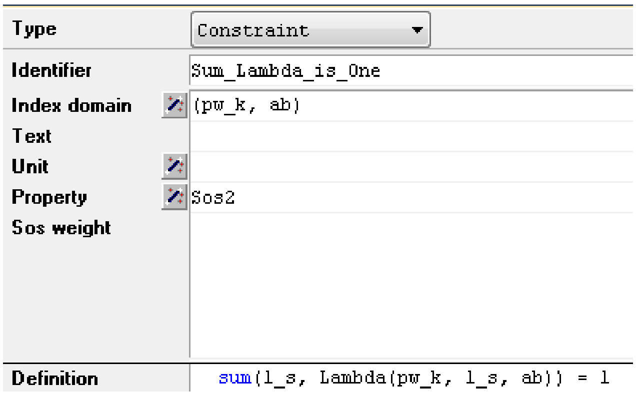
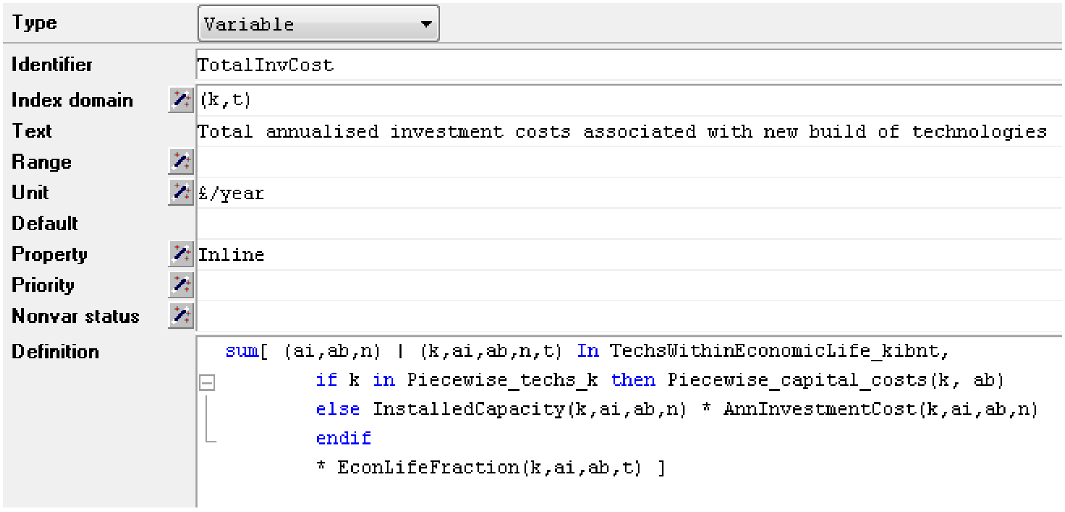


Figure 13-39 show how piecewise approximated capital costs have been integrated into the "TotalInvCost" inline variable and are then passed on to the objective function.

Figure 13-39 Formulation of piecewise capital costs in cost function



Discrete interdependent choices (MIP)

Modelling conditional constraints is required to represent different electricity network reinforcement options with and without meshing. This scheme allows the model to flip the availability of a series of technologies (e.g. reinforcement options) only if a particular technology (e.g. meshing) has been built. This piece is relatively complex to formulate as it requires defining:

- Sets of technologies that are available only if meshing is built (or not): "No_Meshing_k" and "WithMeshing_k" below;
- Binary decision variable to materialize whether the meshing technology is built "Mesh_On" or not "Mesh_Off". These decision variables are always opposite to one another, which we express in the "Mesh_On_Or_Off" constraint (see figure below);

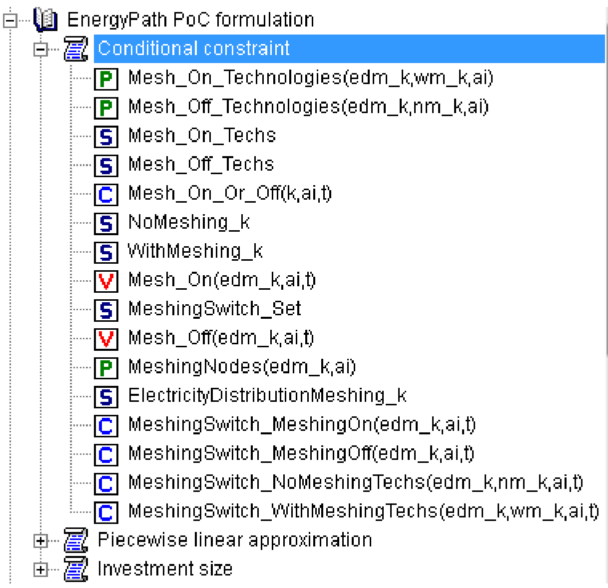
Figure 13-40 Constraint making sure the meshing technology is either on or off

Type	Constraint
Identifier	Mesh_On_Or_Off
Index domain	(k, ai, t) in MeshingSwitch_Set
Text	The meshing technology can be either ON or OFF
Unit	TechCapacityUnitScaledCapacity(k)
Property	
Definition	Mesh_On(k, ai, t) + Mesh_Off(k, ai, t) = 1 [TechCapacityUnitScaledCapacity(k)]

- Parameters to determine the feasible combinations of no-meshing and with-meshing technologies, meshing technologies and nodes: “Mesh_On_Technologies”, “Mesh_Off_Technologies”, “MeshingNodes”;
- Constraints to limit availability of technologies and force the build of meshing technology if necessary

The figure below shows the tree view of all items created in the AIMMS model to represent conditional constraints in the case of network meshing.

Figure 13-41 Formulation of conditional constraints in AIMMS



In order to turn constraints on and off, we use a very large number (10^10) such that when the binary variable is equal to zero, the right hand side (RHS) of the constraint is zero as well but when the binary variable is equal to one, the RHS is so large that the constraint can be ignored in practice. The figure below shows how we applied this principle to modelling the activation/deactivation of capacity of the meshing technology using the binary decision variable “Mesh_On”.

Figure 13-42 Activation/de-activation of the meshing technology

Type	Constraint
Identifier	MeshingSwitch_MeshingOn
Index domain	(edm_k, ai, t) in MeshingSwitch_Set
Text	If meshing technology is off then its active capacity is set to zero.
Unit	TechCapacityUnitScaledCapacity(edm_k)
Property	
Definition	sum(eb, ActiveCapacity(edm_k, ai, eb, t)) <= 1e10 * Mesh_On(edm_k, ai, t)

Figure 13-43 uses the same idea to force the system to build at least one unit of meshing technology when the "Mesh_On" decision variable is equal to one. This is used to include capital costs of meshing into the objective function.

Figure 13-43 Formulation of constraint to force the build of meshing when on

Type	Constraint
Identifier	MeshingSwitch_MeshingOff
Index domain	(edm_k, ai, t) in MeshingSwitch_Set
Text	This is to force the build of at least one unit of meshing technology if it is activated.
Unit	TechCapacityUnitScaledCapacity(edm_k)
Property	
Definition	sum(eb, ActiveCapacity(edm_k, ai, eb, t)) >= 1 * Mesh_On(edm_k, ai, t)

Figure 13-44 and Figure 13-45 present how the capacity of no-meshing technologies (resp. with-meshing technologies) are excluded from the solution when "Mesh_Off" (resp. "Mesh_On") decision variables is equal to zero.

Figure 13-44 Activation/de-activation of no-meshing technologies

Type	Constraint
Identifier	MeshingSwitch_NoMeshingTechs
Index domain	(edm_k, nm_k, ai, t) in Mesh_Off_Techs
Text	If meshing tech is ON, the no-mesh technologies are not allowed to be built.
Unit	TechCapacityUnitScaledCapacity(edm_k)
Property	
Definition	sum(eb, ActiveCapacity(nm_k, ai, eb, t)) <= 1e10 * Mesh_Off(edm_k, ai, t)

Figure 13-45 Activation/de-activation of with-meshing technologies

Type	Constraint
Identifier	MeshingSwitch_WithMeshingTechs
Index domain	(edm_k, wm_k, ai, t) in Mesh_On_Techs
Text	If meshing tech is OFF, the with-mesh technologies are not allowed to be built.
Unit	TechCapacityUnitScaledCapacity(edm_k)
Property	
Definition	sum(eb, ActiveCapacity(wm_k, ai, eb, t)) <= 1e10 * Mesh_On(edm_k, ai, t)

13.3.3. Test results

The second aim of the proof of concept exercise is to evaluate how performance (i.e. solving time) scales in different solving modes as well as when dataset is flexed in the following dimensions:

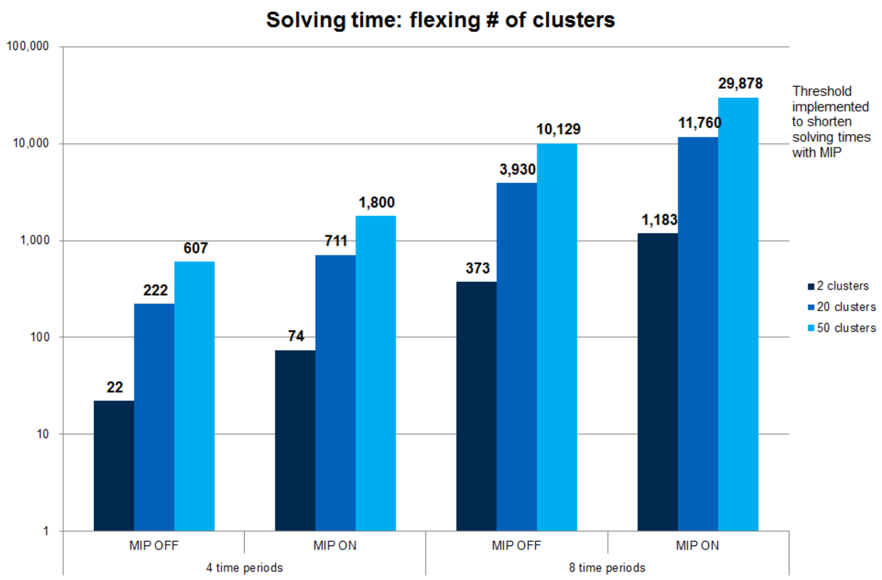
- **Clusters**: reflecting the spatial granularity of a local area
- **Buildings:** reflecting the number of different building archetypes that can be used to represent an area
- **Timeslices**: reflecting the number of characteristic days and within day periods that can be used to more accurately reflect supply/demand balancing
- The performance of the model when the additional functionalities described above are switched on is also monitored.

Whilst the absolute solving time will ultimately be critical within the context of Monte Carlo analysis the test was only undertaken on one specific hardware platform (which would have been considered a leading edge desktop machine approximately 3 years ago) and does not represent current state of the art performance. Hence the relative scaling of the solving time is of more relevant. In addition the simulations could be undertaken across multiple parallel computers (subject to licensing) to reduce the overall solving time for multiple simulations.

Clusters

Figure 13-46 represents how solving times (in seconds) increases when the number of clusters rises from 2 to 50. Solving times seem to scale linearly with the number of clusters. MIP run times are 3 to 5 times larger than LP solving times.

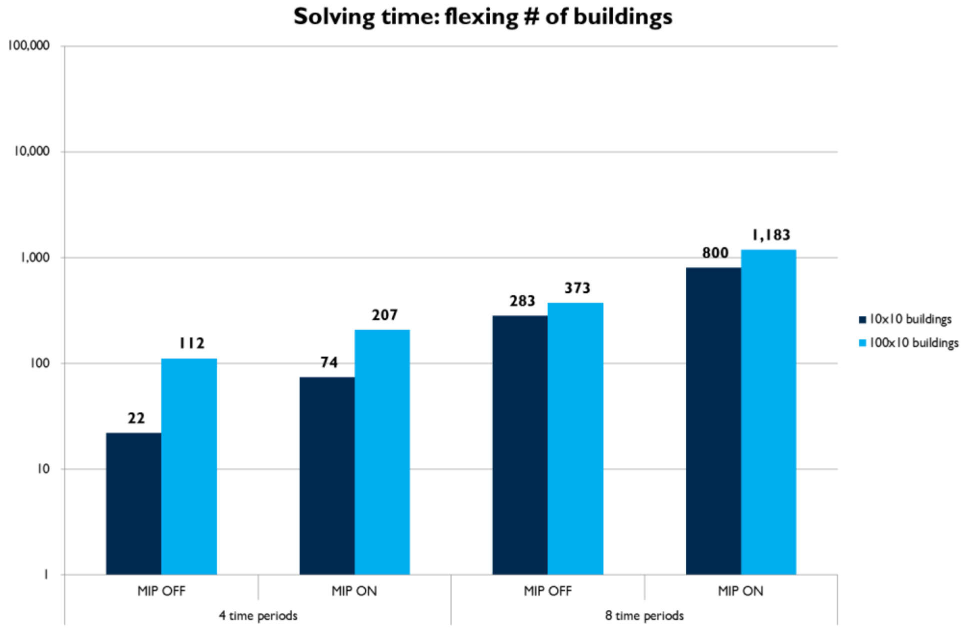
Figure 13-46 Flexing the number of clusters (log scale) – fixing 2x2 timeslices and 10 base buildings



Buildings

Figure 13-47 shows how solving times scales when the number of base buildings is increased from 10 (900 combinations) to 100 (9000 combinations), with (MIP ON) and without (MIP OFF) the additional functionalities. Data shows that turning MIP on significantly increases solving times (by a factor of 3 to 5).

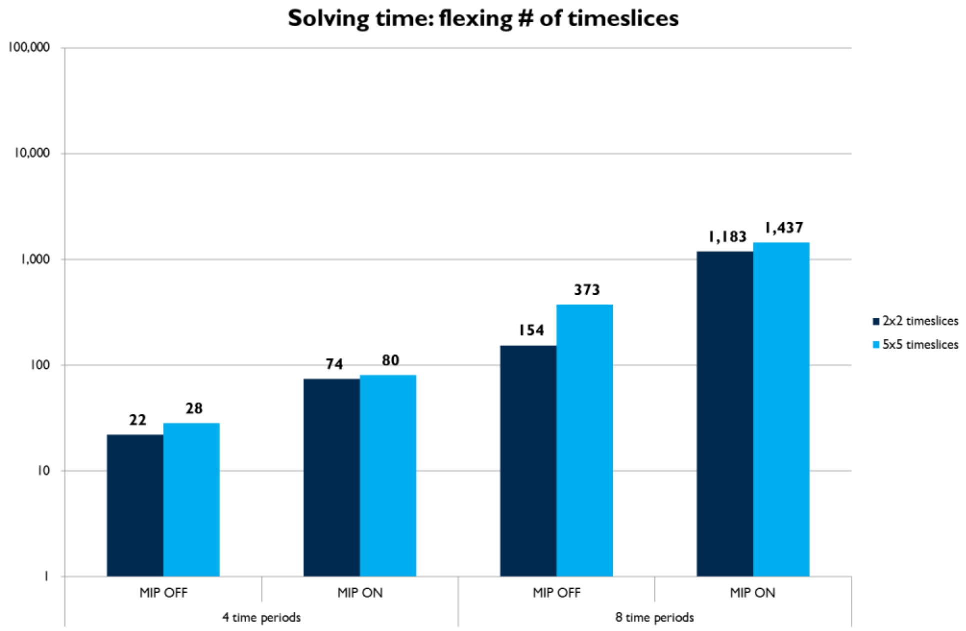
Figure 13-47 Flexing the number of buildings (log scale) - fixing 2 clusters and 2x2 timeslices



Time slices

Figure 13-48 shows how performance scales when the number of time slices is flexed from 4 (2x2: 2 seasons and 2 diurnal time slices) to 25 (5x5: 5 seasons including fully resolved peak and 5 diurnal time slices). Solving times of the MIP are 4 to 8 times as much as the LP's (scaling in line with time periods).

Figure 13-48 Flexing the number of time slices (log scale) – fixing 2 clusters and 10 base buildings



13.3.4. Conclusions for the EnergyPath^{TM} Design Tool

MIP solving times are significantly higher than the comparative LP problem, even when a threshold is implemented to stop the solver when it gets close enough to the solution. However, despite longer solving times, the problems remain feasible. Table 13-8 summarizes the conclusion of performance testing for each dimension explored.

The absolute and relative results clearly emphasise the importance of building in flexibility across the EnergyPath^{TM} Design Tool with respect to easily changing the temporal and spatial and granularity of the optimisation problem. It also highlights the importance of trying to maintain a parallel LP/MIP structure.

Table 13-8 Conclusions of performance testing

Dimension	Impact on solving times
Buildings	Reasonable scaling with moderate impact on performance
Timeslices	Poor scaling inducing large solving times
Clusters	Solving times are roughly in line with problem complexity
Timeperiods	Solving times are roughly in line with problem complexity