Programme Area: Cross Cutting Projects

Project: UK Energy Systems Model

Title: Functional Definition of ETI Energy Model

Abstract:

This report was produced at the outset of the ESME project. Written at the time of the build of the first prototype energy model, it gives a functional definition for a complete model to meet the key requirements identified by ETI. The functional definition addresses model functionality, architecture, data requirements and the implementation path. Many aspects of this functional definition were later implemented in ESME via development of the original prototype model, however not all, so this report is not a factual record of the functionality of later versions of ESME. See the modelling paper dated April 2014 instead for a record of actual model functionality.

Context:

This publication has been produced as part of the work on the ETI's internationally peer reviewed energy system modelling environment (ESME) - a national energy system design and planning capability that helps to identify key areas for ETI investments. ESME is also used by UK Government to underpin and inform energy policy.

Disclaimer:
The Energy Technologies Institute is making this document available to use under the Energy Technologies Institute Open Licence for Materials. Please refer to the Energy Technologies Institute website for the terms and conditions of this licence. The Information is licensed 'as is' and the Energy Technologies Institute excludes all representations, warranties, obligations and liabilities in relation to the Information to the maximum extent permitted by law. The Energy Technologies Institute is not liable for any errors or omissions in the Information and shall not be liable for any loss, injury or damage of any kind caused by its use. This exclusion of liability includes, but is not limited to, any direct, indirect, special, incidental, consequential, punitive, or exemplary damages in each case such as loss of revenue, data, anticipated profits, and lost business. The Energy Technologies Institute does not guarantee the continued supply of the Information. Notwithstanding any statement to the contrary contained on the face of this document, the Energy Technologies Institute confirms that the authors of the document have consented to its publication by the Energy Technologies Institute.

FINAL DRAFT

Prepared For:

Energy Technologies Institute

Holywell Building
Holywell Way
Loughborough

LE11 3UZ

UK Energy System Model

Functional Definition

FINAL DRAFT

[Note final version may use the ETI report template – await guidance]

Prepared By:

Simon Ede

CRA International

99 Bishopsgate

London EC2M 3XD, United Kingdom

Date: 12 June 2009

CRA Project No. D14104-00

DISCLAIMER

TABLE OF CONTENTS

1. INTRODUCTION

1.1. BACKGROUND INFORMATION

The ETI aims to accelerate the development and commercial deployment of low carbon energy technologies to address climate and energy security. In particular, the objectives of the ETI are to support:

- Reducing greenhouse gas emissions (UK targets - 15% renewable energy by 2020, 80% GHG reduction by 2050);
- Accelerating development and deployment of affordable low carbon technology and service solutions;
- Increasing security of energy supply in conjunction with GHG mitigation; and
- Increasing the level and capacity of the low carbon skills pool, both in the UK and internationally.

As part of this remit, the ETI is developing a model of the UK Energy System that will support these objectives. In particular, an energy system model, combined with outlook scenarios for energy demand through 2050 and technology roadmaps, will comprise a core evidence base for informing the ETI's technology strategy and investment plans.

1.2. THE OBJECTIVES FOR THE MODEL

A key objective of the ETI's UK Energy System Model ("ESM") is to analyse the likely contribution of energy technologies in meeting 2050 greenhouse gas emissions targets. Outputs from the model, such as anticipated levels of deployment and utilisation for different technologies, will inform the ETI's strategy process by identifying, for example, high impact technologies that are resilient under a range of demand scenarios.

Box 1 summarises the key question that should be addressed by the model.

Box 1: Key question for the ETI Energy System Model

"What is the most likely least-cost portfolio of technologies to meet energy service end-use demand across the buildings, industry and transport sectors in 2050, that:-

- Meets the requirement to reduce UK emissions of greenhouse gases by 80%;
- Ensures UK energy security of supply;
- Recognises energy resource and supply chain constraints;
- Is technically feasible, that is consistent with the appropriate operation of energy technologies and networks; and
- Reflects interim operating and emissions constraints in the period to 2050".

1.3. HIGH-LEVEL MODELLING CRITERIA

The ETI has identified a number of key requirements for the model functions, data and technology architecture that are summarised as follows:

- The model should represent end-use energy service and energy-related supply and demand decisions resulting from the use of residential, commercial and public buildings, industrial production, and the transportation of people and goods;
- The model should reflect that supplies and demand for energy are geographically dispersed in many cases and so energy will be transported between locations;
- The physical infrastructure of the energy system should be included; the model will be technology rich and include energy networks and storage;
- The model should account for important seasonal and diurnal variations in demand as well as the capacity for certain technologies to affect those patterns (e.g. through energy storage or shifting demand between time periods);
- While the representation of the energy system is necessarily simplified, it is important that the model produces a reasonable estimation of its operation. As such, operating constraints of both technologies and infrastructure should be included;
- The model should constrain its solution to reflect greenhouse gas and other pollutant emissions limits, and security of supply (or resilience) criteria;
- The model will have a key focus on 2050 energy system designs, or "blueprints", but will also define its evolution through time, for example 2010-2050;
- Importantly, as there is great uncertainty around the input values for 2050, the model should incorporate probabilistic analysis for key variables;

- The model should be scalable and provide interfaces to other ETI models that are planned or under development, and an existing GIS package.

1.4. CONTEXT WITHIN UK ENERGY MODELLING LANDSCAPE

Energy and climate modelling is well established internationally in the public and private sector analyses:

- Macroeconomic and specialist models
- Addressing supply and demand side questions

The ETI will leverage this experience in developing analytical tools. The analysis set out in this document is based on a comprehensive review of this modelling landscape. There are three elements (when combined) that make the ETI modelling objective distinctive[1] in the context of the UK modelling landscape:

- A focus on 2050 and assessment of what is possible rather than forecasting forward to consider what is likely;
- Consideration the implications of regional diversity of supply and demand and the physical infrastructure to facilitate regional interactions; and
- Investigation of the impact of uncertainty inherent in technology development.

1.5. CONTEXT WITHIN THE ETI MODELLING ENVIRONMENT

The ESM will operate within a suite of analytical tools that are being employed by the ETI Strategy Directorate. As such, the interface between the models and software will be a key consideration in their development and utilisation (see Section 6.1.4). The key existing and planned developments are:

- **Technology cost model** (under development): This will represent detailed deployment and operating characteristics, including cost behaviour, of the various conversion, storage, transmission, demand reduction and end use technologies. Outputs from the cost model will form the basis for the technology profiles used by the ESM. Some of the results from the cost model may be imported directly into the ESM for direct use or for further manipulation.
- **Energy network model** (planned): This will model the capacity and behaviour of UK energy transmission networks at a reasonable level of detail. The model will facilitate the study of the existing networks and future network infrastructure including extensions of the electricity to offshore renewables. The model outputs will inform the data inputs to, and solutions from, the ESM, which will have a simplified treatment of energy networks, viewing each network as a series of direct connections between regional nodes.

---

[1] The UKERC concluded in August 2008 that spatial and stochastic analysis were some of the lesser developed modelling capabilities already existing in the UK energy modelling space.

- **ETI Member and 3rd party models**: It is expected that input assumptions for the ESM will be informed by models and data sources made available by the ETI Members, or obtained from published sources.
- **ArcView GIS:** The ETI has selected ArcView GIS as its spatial representation tool and this will be a key platform for storing detailed data that will form inputs to the ESM and for displaying results from the ESM i.e. technology deployment.

1.6. CONTEXT WITH ETI SCENARIOS

The ETI has, in parallel, developed a set of 2050 scenarios which will enable it to both frame the analysis of potential technology investments and also develop datasets for use in the ESM. The demand data within the model will be formulated on the basis of the scenarios that have been developed. In the future, it may be possible other elements of the overall model dataset will also be driven by the demand scenarios. The ESM, in addition to being a tool to understand the implications of the scenarios will also be a tool used to understand the importance of key variables and further develop the scenarios.

1.7. WORK CARRIED OUT SO FAR

The ETI has developed a comprehensive prototype to help it determine its requirements for the full model development. The prototype is intended as a working “desktop” model for use within the Strategy Directorate and has a single interface to the ArcView GIS. It covers all sector demands, technologies, spatial and temporal requirements but its level of detail is necessarily limited and its technical architecture has limited scalability.

A fuller description of the prototype model and how it differs from the model specified in this document is included in Appendix E .

As part of the development phase for the next version of the model, the experience garnered from the prototype will be used to update and clarify this document.

1.8. PURPOSE AND STRUCTURE OF THIS DOCUMENT

The purpose of this document is to set out the development requirements for a full ESM development that might inform an in-house development team or form the basis for a request for proposals to modelling software houses.

The document is organised as follows:

- **Section 2** sets out the objectives for the model;
- **Sections 3 and 4** describe the basic representation of the energy sector required in the model. Section 3 summarises the processes required to be modelled and Section 4 described the components of the system to be included;
- **Sections 5 and 6** describe the data structure and the system architecture envisaged for the model; and

- **Section 7** describes the potential implementation path for the next phase of model development. This includes a method for interpreting the experience with the prototype and a plan for developing the next version of the model.

The appendices to this document include more detailed information on the representation of the energy sector that is required, the structure of the data, a description of the prototype, and a glossary of terms.

2. MODEL OBJECTIVES

2.1. INTRODUCTION

The ESM should assess the most likely least-cost portfolios of technologies to satisfy end-use energy service demands resulting from the use of buildings, industrial production, and the transportation of people and goods in 2050.

The first part of this section describes the required solution to be determined by the model and the second part identifies the constraints.

2.2. OBJECTIVE FUNCTION

The model should select an optimal mix of technologies to satisfy end-use energy service demands which, when deployed, minimises the aggregate annualised cost in 2050.

Where cost is defined as the sum of:

- Aggregate investment cost and fixed cost for all technologies deployed;
- Aggregate non-fuel variable operating and maintenance costs of all deployed technologies;
- Aggregate variable resource/fuel costs of all deployed technologies; and
- Net costs of importing resources or energy carriers.

And where:

- The scope of the technology choice is within the UK (taking into account imports and exports of electricity and fuels);
- Technology includes energy conversion, networks, energy storage, and end-use energy service technologies;
- Where cost is “private cost”, ignoring all “externalities” created by the use of a technology but not borne by the technology user itself; and
- Supply technologies and demand for energy may be geographically dispersed at different “nodes”.

2.2.1. Constraints to the objective function

The solution to the above problem is subject to a number of constraints which may apply differently in different locations within the UK, and in different time periods:

- End-use energy service demand is satisfied for interim time points (through 2050)
- Total unabated greenhouse gas emissions should be less than its given target for a time period;
- All technologies should operate at all times within their capacity limits and within any specified operating profile constraints;
- Total deployed capacity of any technology at any time should be less than or equal to the allowed capacity for that time period;
- The total amount of any technology built in a single time period should be less than or equal to the allowable build-rate for that technology in that period;
- The production and consumption of any "product" (end use, energy carrier, or emission) should balance at every location after taking into account transfers of product between locations, injections and withdrawals from storage, imports and exports and losses incurred through conversion, storage, transportation;
- Consumption of any resource, e.g., gas or wind, in any period may not exceed its supply in any period (after accounting for inter-period storage); and
- Deployment of certain technologies is greater than or equal to a level determined as necessary to ensure safe system operation (examples might include requirements for a certain level of installed electricity capacity to meet peak demand at specific locations in the UK or a requirement to have a varied mix of technologies or maintain a certain level of stocks in storage technologies).

2.3. ADDITIONAL CONSIDERATIONS

A number of different formulations for the model have been considered. The degree to which the proposed formulation is satisfactory will be judged after the experience of using the prototype. Some issues to be addressed include:

- The above formulation leads to a solution for a final year (such as 2050) and ensures that the path to that final year is feasible and meets constraints on meeting demand and emissions. The ETI may also wish to extend the cost minimising solution to include all years (e.g., 2010-50). This could be done through creating a weighted formula for 2050 and intervening years that would allow the user to assign relative importance between the end-solution for 2050 and the 2010-50 pathway;
- While the prime consideration will be for 2050, the model should be sufficiently flexible to allow for any year to be considered as the solution year. The ETI may wish to use the model to review the sensitivity of the results to the final model year selected; and

- The time horizon of the analysis should also be flexible. For example, the above formulation considers a 2010-50 time period but this does not preclude the need to run longer or shorter analyses.

3. SECTOR DESCRIPTIONS

3.1. INTRODUCTION

The subject matter scope of the model has been widely discussed. The desired coverage is, therefore, relatively fixed. In this section, described at a high level are the underlying energy system processes that need to be represented within the ESM. As set out in Section 2, the primary objective is to satisfy end-use demands. The model, however, should also include the conversion of primary energy resources into energy carriers (such as electricity) to enable end-use technologies to meet these demands.

Appendix A contains a more detailed discussion of the representation desired in each of the energy sectors based on a review of other models and experience of the prototype.

3.2. SECTOR COVERAGE

Figure 1 shows a high level representation of the processes that should be included in the ESM. It shows the generation of end-use demands and the interaction with energy service technologies as well as the creation and supply of energy to those technologies.

The model will consider:

- The satisfaction of end-use energy service demands through end-use energy service technologies;
- The provision of energy to end-use energy service technologies through conversion technologies;
- The storage of energy; and
- The transportation of energy when supply and demand and not at the same location.

3.3. END-USE SECTORS

The model should include representation of three core end-use sectors:

- Buildings (residential and service sectors) and end-use demands;
- Industrial demands; and
- Private and commercial transportation demands.

The rest of this section summarise the key components of each of these sectors that need to be modelled.

3.3.1. Buildings and End-Use Demands

Residential Buildings

The ESM should model supply and demand for residential buildings by the UK population. The model should include both the existing stock of buildings and new build options. Table 1 shows the main features of this sector to be modelled. (Appendix A includes a more detailed description of the sector).

Table 1: Residential Buildings – Summary of Model Requirements

Model Process	Representation required in ESM	MODELLING COMMENTS			
		General comments	***Differences by location?***	***Differences through time?***	***Probabilistic?***
Demand for residential buildings *e.g., demand for high density dwellings*	• Based on population size and characteristics (housing occupancy levels) • Based on economic characteristics of population	• Demand should be calculated by the model based on exogenous inputs concerning population size/characteristics and economic characteristics • The relationship between housing demand will be a scenario input that will have annual representations for every location modelled	• Demand defined at nodal level	•Annual demand variation •No within year /day variation	•No, demand will be set through exogenous scenarios
Supply of residential buildings *e.g., semi-detached houses, flats, detached houses*	• There should be an initial stock of residential buildings with assumptions about building lifetimes • There should be a set of possible new house-build options including the ability to transform an existing house through renovation	•It should be possible to constrain the selection of housing types to given criteria about the population and its occupancy characteristics	• Supply options defined at nodal level	•Annual •No within year/day	•No probabilistic assessment of housing stock

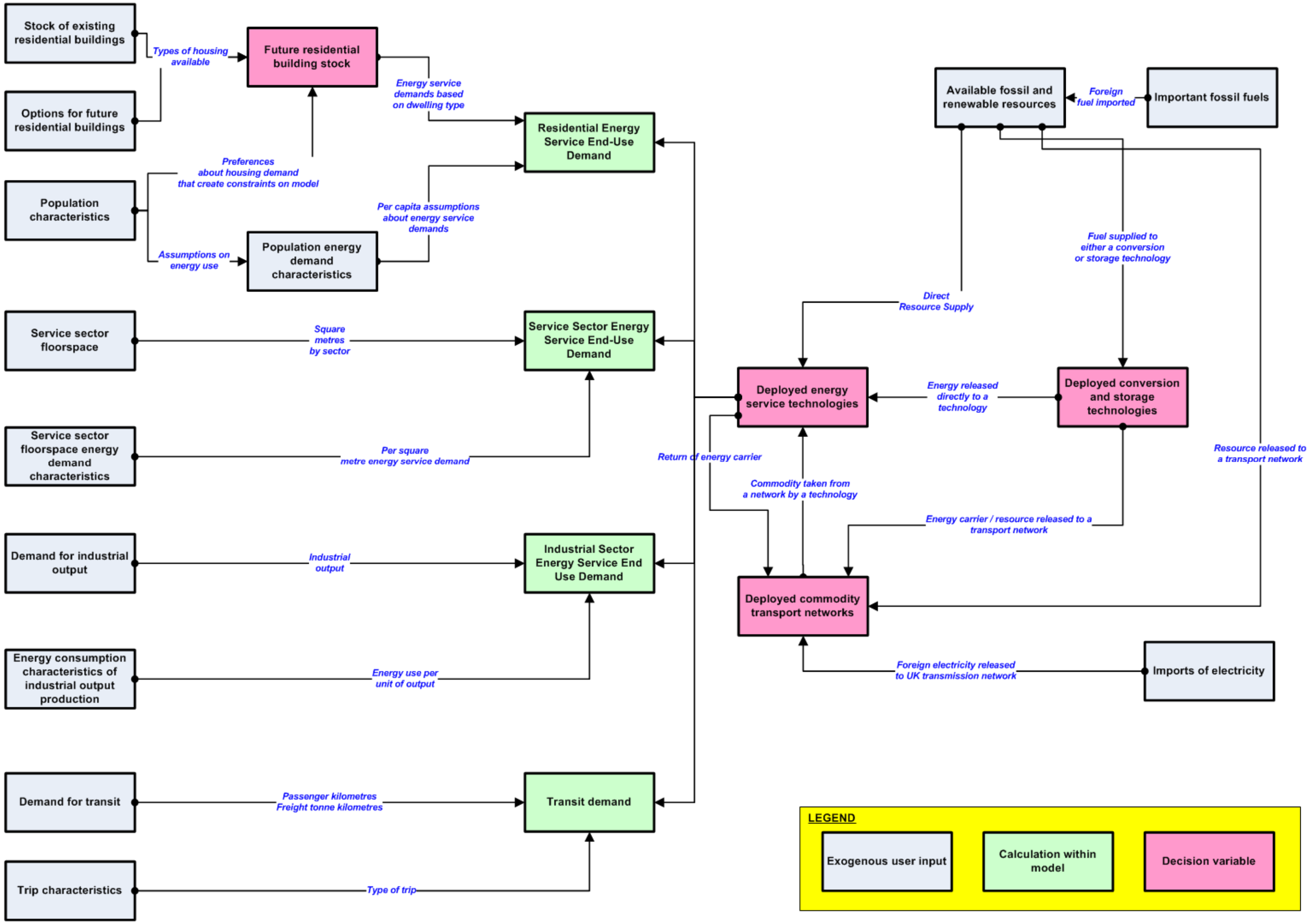


Figure 1: High level map of energy system to be modelled

Residential End-Use Energy Service Demand

The ESM will model the supply and demand for energy end-use services in the Residential Sector. The model will assess the selection of energy service technologies to meet demands which will be linked to either population characteristics and/or the housing choices of the population (described in the previous section). Table 2 summarises the main processes of this sector to be modelled. (Appendix A includes a more detailed description of the sector).

Table 2: Residential End Use Energy Services – Summary of Model Requirements

Model Process	Representation required in ESM	MODELLING COMMENTS			
		General comments	*Differences by location?*	*Differences through time?*	*Probabilistic?*
Demand for residential end-use energy services *e.g., heating, hot water, cooking, appliances, refrigeration, lighting*	• Based on population size and characteristics (housing occupancy levels) • Based on economic characteristics of population • Based on the selection of housing by the population • Specified as fundamental demand (e.g., cooling hours) rather than resulting final energy demand (e.g., refrigerator consumed X kWh)	• Demand should be calculated by the model based on exogenous inputs concerning population size/characteristics and economic characteristics and modelled choices of housing type • The relationship between end-use energy service demand and population and housing characteristics will be a scenario input that will have annual representations for every location modelled	• Demand to be defined by node	• Annual demand variation • Seasonal demand variation • Diurnal demand variation	• No, demand will be set through exogenous scenarios
Supply of end-use energy services *e.g., boilers, refrigerators*	• There will be an initial stock of technologies with lifetime assumptions • There will be a portfolio of technologies with a range of possible cost and capacity assumptions to meet demand • Demand management technologies (e.g., insulation) to be included	• Technologies will have a set out output products which can be used to satisfy specific end-use demands • Demand reduction will be modelled as a technology rather than a negative demand • Technologies should be feasible for the housing type that they are servicing – e.g., different set of possible technologies for a block of flats from a detached house	• Supply options to be defined by node	Technologies will be defined on the following bases: • Annual • Seasonal • Diurnal Consistent with demand specification	• Key probabilistic variables will be cost and availability of technology

Service Sector End Use Demands

The ESM should represent the supply of and demand for energy end-use services in the Service Sector. The Service Sector comprises of public sector and commercial end-use demands. The model will not represent the selection of buildings within this sector, unlike the residential sector, though demands will be related to assumptions made on the floorspace characteristics of the sector. The amount of floorspace consumed by the Service Sector should be an exogenous assumption in the model. The model should assess the selection of energy service technologies to meet these demands. Table 3 summarises the main Service Sector processes that should be incorporated in the model. (Appendix A includes a more detailed description of the sector).

Table 3: Service Sector End Use Energy Services – Summary of Model Requirements

Model Process	**Representation required in ESM**	**MODELLING COMMENTS**			
		General comments	***Differences by location?***	***Differences through time?***	***Probabilistic?***
Demand for service sector end-use energy services *e.g., heating, hot water, cooking, appliances, refrigeration, lighting*	• Based on economic characteristics of the sector such as gross value added or employment levels • Broken down by category of service, e.g., school vs. hotel • Based on floorspace consumed by the sector • Specified as fundamental demand (e.g., cooling hours) rather than resulting final energy demand (e.g., refrigerator consumed X kWh)	• Demand should be calculated by the model based on exogenous inputs concerning economic characteristics of the sector (i.e. GVA) and assumptions about the floorspace consumed or the number of employees • The relationship between end-use energy service demand and economic characteristics and floorspace will be a scenario input that will have annual representations for every location modelled	• Demand will be specified by location	• Annual demand variation • Seasonal demand variation • Diurnal demand variation	• No
Supply of end-use energy services *e.g., boilers, refrigerators*	• There will be an initial stock of technologies with lifetime assumptions • There will be a portfolio of technologies with a range of possible cost and capacity assumptions to meet demand	• Technologies will have a set out output products which can be used to satisfy specific end-use demands	• Supply of technologies will be specified by node	Technologies will be defined on the following bases: • Annual • Seasonal • Diurnal • Consistent with demand	• Key probabilistic variables will be cost and availability of technology

3.3.2. Industrial sector

The industrial sector (excluding power generation) should be modelled in less detail than other end-use sectors given the lower GHG emissions profile. Specific technologies to meet industrial process requirements should not be included. The model should instead include both representations of overall demand for industrial output and consequent demands for electricity, steam and fuel, and include associated outputs of greenhouse gases. Table 4 summarises the processes to be included. (Appendix A includes a more detailed description of the sector.)

Table 4: Industrial Sector End Use Energy Services – Summary of Model Requirements

Model Component	Representation required in ESM	MODELLING COMMENTS			
		General comments	*Differences by location?*	*Differences through time?*	*Probabilistic?*
Demand Industrial Output e.g., demand for Iron and Steel **Demand for Final Energy** *e.g., heating, hot water, cooking, appliances, refrigeration, lighting*	• Demand for output to be based on economic characteristics of the sector such as gross value added • Demand for final energy (e.g., electricity) to be specified by exogenously determined relationship to either total output or sector GVA • Demands broken down by industry category, e.g., Iron and Steel, Chemical	• Demand should be calculated by the model based on exogenous inputs concerning economic characteristics of the sector (i.e. GVA) and assumptions about the floorspace consumed or the number of employees • The relationship between end-use energy service demand and economic characteristics and floor space will be a scenario input that will have annual representations for every location modelled	• Demand will be specified by location	• Annual demand variation • Seasonal demand variation • Diurnal demand variation	• No
Supply of output *e.g., Iron Smelters*	• There will be one generic industrial technology per industry sub-category	• The generic technology will input a specified amount of fuel, steam and electricity and produce output to meet demand	• Yes, vary by node	• Annual • Seasonal • Diurnal	• Industrial production will not be part of the probabilistic analysis

3.3.3. Transportation

The ESM should represent the supply of and demand for private individual, commercial and public transportation. The model should assess the selection of transportation technologies to meet these demands. Table 5 summarises the main Transport Sector processes of this sector to be modelled.

Table 5: Transportation – Summary of Model Requirements

Model Component	Representation required in ESM	MODELLING COMMENTS General comments	 Will there be differences by location?	 Will there be differences through time?	 Will analysis be probabilistic?
Demand for private and commercial transportation *e.g. passenger or freight kilometres*	• Demand for transportation to be based on economic characteristics of population, service sector and Industrial GVA • Demand will be for transportation kilometres rather than mode of transportation (e.g., demand for 100 passenger kilometres and not demand for 100 road passenger kilometres) • Demands broken down at least between passenger and freight transportation	• Demand will be independent of mode of transportation • Demand for certain freight transportation will be related to the selection of power generation technologies (e.g., coal power will require more coal freight transportation)	• Demand will be specified by location – assumed that all trips emanate from a particular region – no explicit representation of trips between locations	• Annual demand variation • Not seasonally or diurnally specified	• No
Supply of transportation *e.g., road, rail, plane, marine*	• Covers road, rail, plane and marine travel in the UK • There will be an initial stock of transportation technologies with lifetime assumptions • There will be a portfolio of transportation technologies with a range of possible cost and capacity assumptions to meet demand	• It should be possible to constrain technology selection to limit the ability to switch between modes of transport	• Yes, vary by node	• Annual • Not seasonally or diurnally specified	• Key probabilistic variables will be cost and availability of transportation technology
Transport networks	• No explicit representation of transport networks • Transport networks may be represented as resources which are consumed when a passenger or freight tonne kilometre is produced	• No modelling of point to point transportation • Modelled as a resource, will have a cost and capacity associated with it • As a resource it will be an input product for transportation technologies with a capacity and a cost	• Yes, by node	• Annual • Not seasonally or diurnally specified	• No – impact of infrastructure could be assessed through scenario
Other transport infrastructure	• Filling stations, electricity charging points, hydrogen filling stations should be included with an initial supply level and options for future deployment	• Fuel delivery networks can be considered as an intermediate technology	• Yes by node	• Annual • Not seasonally or diurnally specified	• No – impact of infrastructure could be assessed through scenario

3.4. C ONVERSION, STORAGE AND ENERGY CARRIER TRANSMISSION

The ESM should represent the supply of final energy to energy service and transportation technologies. As a result, the ESM will include a representation of the conversion of fuels and renewable resources into electricity, heat and hydrogen. The model should also include the supply of resources, such as natural gas, to end use technologies. Because supply and demand are geographically dispersed, the model will include a representation of networks for electricity and gas. The model should include the capture, transportation and storage of resources, such as natural gas, and pollutants such as CO2. Table 6 provides a summary of the representation required in the model (This is described in more detail in Appendix A).

Table 6: Conversion, storage and transmission – Summary of Model Requirements

Model Component	Representation required in ESM	MODELLING COMMENTS			
		General comments	*Differences by location?*	*Differences through time?*	*Probabilistic?*
Conversion	• Fossil fuel, nuclear, and renewable generation • Refining • Security of supply requirements to apply to at least capacity and flexibility of electricity generation as well as possibly concentration of supply sources (HHI index) and on gas storage (days of stocks)	• Modelled in tranches of capacity where appropriate – e.g., wind	• Yes	• Annual • Seasonal • Diurbal	• Key probabilistic variables will be cost and availability of transportation technology
Storage	• Gas – large scale only – no representation of linepack • Electricity – pumped storage, batteries • Carbon	• If the model is not run chronologically, the representation of storage may be limited to shifting energy carrier between time periods with no representation of stocks • CCS will be a specific technology option that can be added to a conversion technology	• Yes	• Storage will be defined as having either a seasonal or diurnal storage capability (unless model is run chronologically)	• Key probabilistic variables will be cost and availability of transportation technology
Transmission	• High voltage and trunkline between nodes for gas and electricity • Pipelines for carbon • Multiple options to transport energy carrier if available – e.g., DC or AC • No representation of networks within a node	• Decision variable as to how much energy carrier can be transferred from node to node rather than explicit technology	• Yes	• Annual only	• Transmission outages / availability not
Distribution	• At least a basic representation of the requirement for distribution capacity within a geographic location	• Modelled as a technology, will have a cost and capacity associated with it	• Yes by node	• Annual • Not seasonally or diurnally specified	• No – impact of infrastructure could be assessed through scenario

4. REQUIRED MODEL FUNCTIONALITY

4.1. I NTRODUCTION

Section 3 considered the subject matter of the model. This section sets out the practical implementation of the subject matter and describes the main elements in the model, which are:

- Resources;
- Technologies;
- End-use energy service demands;
- Products (including Energy carriers)
- Emissions

- Spatial elements; and
- Temporal elements.

4.2. RESOURCES

There will be a number of primary renewable and fossil energy resources included within the model. These untransformed energy resources are used as inputs to technology processes that deliver either an energy carrier (such as electricity) or an end-use service (such as heat or transportation kilometres).

Inputs to the model will include values by time period and by region for:

- Availability for consumption that may in part be determined by the overall capacity of the resource;
- Constraints that availability of the resource has upon the installed capacity of conversion technology; and
- Cost per unit of consumption.

Resources might be broken down into blocks (or tranches) of capacity, each with an associated cost, availability and capacity. This would reflect the varying costs of different sources of the same resource.

Imported energy carrier products can also be considered as resources, since they are not produced by the technologies within the UK energy system. Hence, the model will take, as inputs, the availability and unit price for imported fossil fuels and electricity.

4.3. TECHNOLOGIES

There will be different classes of technologies:

  - "***Conversion***" technologies which convert either primary energy resources or energy carriers into energy carriers and which may have other non-energy by-products, including CO2;
  - "***End-Use***" technologies which convert either primary energy resources or energy carriers into end-use services (these include demand management technologies);
  - "***Storage***" technologies which can store through time either primary energy resources or energy carriers;
  - "***Network***" technologies which can transport energy carriers or primary energy resources; and
  - "***Buildings***" which can accommodate either private individuals or organisations.

- Technologies have a specific "configuration" that determines at any point in time or for any location its:

    - Deployment cost;
    - Operating characteristics; and
    - Fixed and variable operating costs.

  An example configuration might relate to a power plant technology of a particular construction year developed by a particular manufacturer.

- Each technology should have an initial configuration but it may also have a number of possible alternative configurations. The configuration of a technology could change through time, for example by the addition of subsequent equipment. For example, an emissions control technology could be added to an existing power plant to change its output of emissions.

4.3.1. General technology configuration values

Technology configurations should have a number of measures or attributes. Those measures which are common to all technologies will include for any location and time period:

- A maximum installable capacity;
- A maximum cumulative installable capacity;
- A forced outage rate (except for Buildings) or % of time unavailable due to unplanned incidents;
- A maintenance profile which describes when maintenance occurs;
- A decommissioning cost;
- An annual degradation coefficient by which intensity is reduced to reflect the end of technical lifetimes;
- An investment required to increase capacity by one unit (which reflects a sectoral or economy-wide discount rate);
- An investment required to convert technology between configurations;
- A fixed and variable operation and maintenance cost profile;
- A minimum and maximum operating range (in relation to capacity).

Particular technology types may have additional measures. These are listed in the following subsections.

4.3.2. Conversion technologies

Conversion technologies will have a number of attributes that should be reflected in the model. For a unit of intensity (or installation) of conversion technology with a specific configuration there will be:

- An input of energy carrier required;

- An input of primary energy resource required;
- An output of energy carrier;
- A release of emissions;
- A fraction of emissions that could be captured from the technology (i.e., by a technology such as CCS).

4.3.3. End-use technologies

End-use technologies should have a number of attributes that should be reflected in the model. For a unit of intensity (or installation) of end-use technology with a specific configuration there will be:

- An input of energy carrier required;
- An output of end-use service (to meet end-use demand);
- A release of emissions;
- An input of primary energy resource required;
- A fraction of emissions that could be captured onsite.

Demand management technologies should be represented slightly differently in the model. For each unit of intensity (or capacity), for a given configuration, there will be:

- An output of end-use service (equivalent to the amount of reduction in the end-use demand which would otherwise be met by the appropriate end-use technology).

4.3.4. Storage technologies

Storage technologies should have a number of attributes that should be reflected in the model. For a unit of intensity (or installation) of storage technology with a specific configuration there will be:

- A feasible set of storable primary energy resources, energy carriers, or emissions;
- A leakage coefficient to reflect loss of inventory for unit of time;
- A maximum injection rate per unit of time;
- A maximum inventory of resource or energy carrier per unit of time (applicable only if the model is run in a chronological manner);
- A maximum withdrawal rate per unit of time;
- A release of emissions per unit of injection of energy carrier or primary energy resource;
- A release of emissions per unit of withdrawal of energy carrier or primary energy resource;
- An input of energy carrier required per unit of injection;

- An input of energy carrier required per unit of withdrawal; and
- A fraction of emissions that could be captured onsite.

4.3.5. Network technologies

Network technologies will have a number of attributes that should be reflected in the model. For a unit of intensity (or installation) of network technology with a specific configuration there will be between each node per unit of time:

- A feasible set of transportable primary energy resources, energy carriers, or emissions;
- Annual degradation coefficient of transfer limit from node to node for resource, carrier or emission;
- Input of energy carrier required per unit of flow of resource, carrier or emission from node to node in a unit of time;
- Investments required to create one additional unit of transfer capability from node to node for energy carrier in a unit of time;
- O&M expenses per unit of flow of resource, carrier, or emission from node to node in unit of time; and
- Release of emission per unit of flow of energy carrier from node to node in a unit of time.

4.3.6. Buildings

Buildings represented in the model should have a number of attributes that should be reflected in the model. For a unit of intensity (or installation) of buildings with a specific configuration there will be:

- A maximum occupancy level;
- An input of end-use demands required for an occupancy level; and
- A capacity to add any specific technology to meet an end-use demand (e.g., ability to install insulation in a building).

4.3.7. Linked conversion technologies

The model should include intermediate technologies that are linked or associated with conversion technologies. For example, there will be a coal-fired generation technology. It may be possible to install carbon capture technology for a coal power plant and this will capture a fraction of the onsite CO2 emissions. There will be an associated transport network and storage capability. The choice to deploy CCS is separate from the choice to install a power plant but the two technologies are linked. An example of this is shown in Figure 2 below.

Figure 2: Associated technologies – example representation of CCS technology

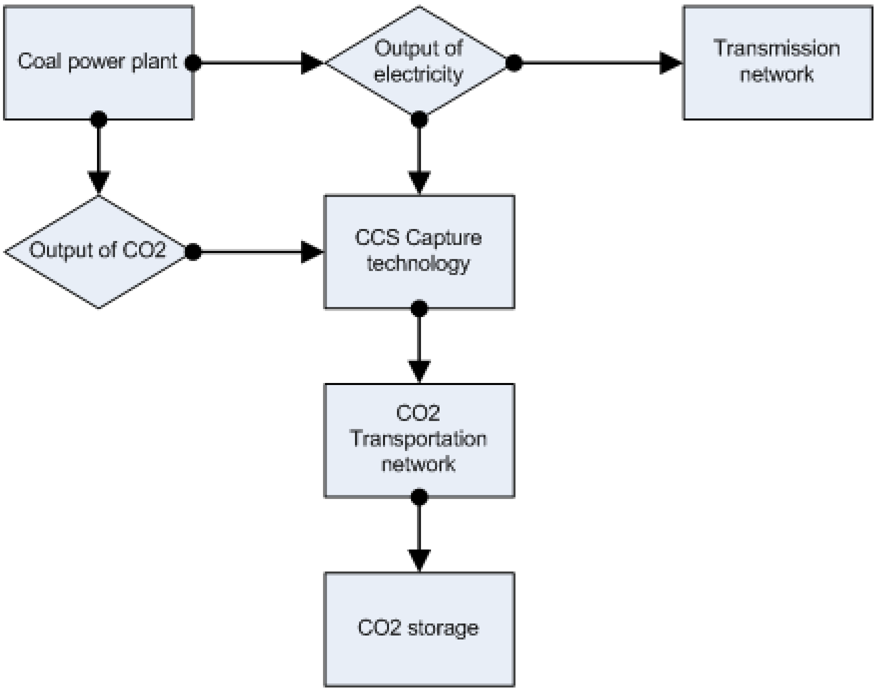


Table 7 provides some of the main examples of the linked technologies that will need to be incorporated in the model.

Table 7: Examples of Linked Conversion Technologies

Category	Technology	Link
Hydrogen	Large scale hydrogen production	Electricity generation technology
Hydrogen	Liquefaction	Large and small scale hydrogen production
Electricity generation	Carbon Capture and Storage	Electricity generation technology
Electricity generation	Emissions control technology - e.g., FGD technology	Electricity generation technology
Electricity generation	Refurbishment	Electricity generation technology

The addition of an associated technology can be thought of as a change in the configuration of the main technology. If an associated technology is installed then it may also change the operating characteristics and costs of the original technology. For example, the efficiency of conversion may change or the lifetime of a power plant may be extended. These associated technologies may also consume some of the output of the original technology. In the case of CCS this is a desirable property. The CCS technology consumes the CO2 output of the generation technology. It also consumes some of the electrical output to operate and it may affect the operating performance of the power plant. In some cases the associated technology could be part of the initial configuration of the base technology but the model should be able to handle the transition between configurations of a technology should an associated technology be added. This will be important to reflect, in particular, the choice to retrofit existing power plants with CCS.

4.3.8. Technology tranches

For some technologies there will be an increasing cost of deployment and/or a reduction in efficiency or productivity of the technology as total capacity of the technology increases. For example, the productivity of wind turbines may decrease in a specific location as more turbines are built and the most advantageous turbine sites (i.e. those with the best wind profiles) are filled. It may necessary for some technologies to represent total available capacity in a series of blocks or tranches. Each tranche would have a specified cost and operating profile. The model should be able maintain a link between those tranches, particularly since certain variables for technologies will be subject to probabilistic analysis and so they should be correlated.

4.3.9. Technology learning curves

Technology cost may decrease and capacity to install may increase through time as a result of experience of developing and using technologies. The model should include provision of this learning curve effect. For some technologies this cost effect may also vary with the scale of deployment of the technology within the UK. For others, the scale of UK deployment will not have a great effect on overall cost of the technology. The model should also enable the cost of the technology to be endogenous to the scale of deployment if desired, and if the final solution method in the model allows for non-linear optimisation. The decision whether to employ a non-linear optimisation method will only be considered after reviewing the experience with the linear methods included in the prototype. Until that point, learning curves will remain exogenous.

4.4. END-USE ENERGY SERVICE DEMANDS

There will be a number of end-use energy demands that will be specified on a nodal basis for each time period modelled.

End-use energy service demands are for services, rather than for energy itself. They are the final deliverable products of the energy system model. For example, lighting is an end-use energy product of an end-use technology (light bulb). The electricity required for the light bulb is an energy carrier product.

End-use service demands will either be a direct input to the model or the result of a pre-processing step. For example, residential end-use energy service demand will relate to the modelled residential housing choice. Transport demand, however, will be a direct input. Appendix A describes the calculation of end-use demands in each of the modelled sectors.

End-use demands will be exogenous to the model and will be perfectly inelastic. Demand will not be responsive to cost of energy carrier products calculated within the model.

4.5. ENERGY SERVICE END-USE PRODUCTS

Energy service end-use products are those products which satisfy an end-use energy service demand (e.g., lighting). There may be, in some instances, multiple end-use products that can satisfy and end-use demand.

4.6. ENERGY CARRIER PRODUCTS

Energy carrier products are those products which are produced by and may also be consumed by conversion technologies. An example is electricity. They are consumed by end-use technologies in order to output and end-use product (e.g., lighting).

4.7. EMISSIONS PRODUCTS

Emissions products are non-energy outputs of technologies. They include all greenhouse gases. The model, however, should be capable of extending to any pollutant (such as NOx and SOx). The concept of pollutant should be generic rather than specific to CO2.

The model is not required to represent non-emission by-products, such as ash from biomass and coal-fired power generation.

4.8. SPATIAL ELEMENTS

The ESM is intended as a UK focused model although the design should be such that it is easily portable to other geographies. The model should reflect geographic regions or “nodes” within the UK. The ESM will include both onshore and offshore geographic regions. Geography, as a dimension, will be used to specify not only the location of demand and supply but also the cost and operating characteristics of technologies. Constraints defined in the model could also include a geographic dimension.

The number of desired nodes is yet to be determined. The table below shows the nodes included in the current version of the prototype. There should be no constraint on the number of nodes; it should be a user choice based on the level of data detail available and desired running times for the model.

Table 8: Spatial dimensions to be included in the model

Category	Examples Nodes	Category	Example Nodes
Onshore	North East North West Yorkshire and the Humber East Midlands West Midlands East London South East South West Scotland Northern Ireland Wales	***Offshore***	Dogger Bank – Offshore wind / wave Norfolk – Offshore wind / wave Irish Sea – Offshore wind North Sea – Carbon storage Pentland – Tidal Southwest – Wave

For the ESM, it is likely that the nodes will be defined such that they do not overlap (consistent with real geographic boundaries) but it should be feasible within the model to define geography differently for different demand and supply features.

4.9. TEMPORAL ELEMENTS

The ETI does not anticipate the requirement for a chronologically operated model due to computational intensity and the time required to complete model runs. An acceptable approach would be to break time down into categories, or “time slices” and to model 2010-50 as a collection of time slices. A time slice might be a season or period within the day such as an off-peak hour.

The model will solve for a final year but that final year should be a user choice. It may also be desirable to solve for a block (for example 5 years) of time.

Instead of modelling, therefore, all 8760 hours within a year the model can simply by considering groups of time, as shown in the table below. Using this table, an example timeslice could be “2050 winter weekday offpeak”.

Table 9: Examples of possible time dimension categories

Category	Examples (not comprehensive)
Whole period	• Years • Groups of years
Within year	• Seasonal • Quarterly • Monthly • Weekly • Weekday / Weekend • Daily
Within day	• Peak • Off Peak • Shoulder

4.10. PROBABILITY AND CORRELATION

It should be possible to assign a probability distribution to any variable within the model. As a practical matter only a sub-set of variables will likely be given a distribution. Section 3 sets out the main areas where this is desired. The most common probabilistic variables will be the cost and capacity of technologies. As certain technologies will be linked it is also necessary that their probability distributions are linked either through a variance-covariance matrix or (more likely) through indices and correlation to indices.

Monte-Carlo draws will be performed on 2050 variable values. There will not be a representation of any stochastic pathway to that variable value. For example, if a technology has a range of possible outcomes for capacity between 0 and 10 units in 2050, the model will select based on the assumed distribution a value, for example 5 units. It will not make any assumptions about the implications for capacity in the period 2010-50. The model will only assess, given build rate for the technology, whether it is feasible for get to 5 units by 2050 from the specified starting stock in 2010.

The distributions of variables in the model should be a user input. It should be possible to define both continuous and discrete probabilities. The selection of appropriate distributions and correlations between variables will be a significant data challenge for the ETI but will be processed outside of the model technology.

5. DATA MODEL

This section describes the data that should be handled by the ESM.

5.1. ENTITIES AND ENTITY RELATIONSHIPS

The following Entity Relationship Diagram shows the “real world” entities of the UK energy system and the relationships between them. The entities shown relate to Section 4. For simplicity, the diagram ignores time-related dimensions, which are fundamental to the ESM, but which are included in Appendix C “Data model bus matrix”.

5.2. ATTRIBUTES AND MEASURES

The diagram below shows the main physical entities, but not their attributes. Many of the attributes that we are interested in are “measures” of various types that relate to end-use service demand, technology efficiency factors, availability factors, deployment and operating costs, etc., and various constraint values by which we wish to condition the model’s search for solutions.

In the simplified ERD diagram, the various measures, relating to a particular technology, installed in a particular region, are implied in the relationship between technology and region: e.g., “installed in / has installed capacity of”. This is shown as a many-to-many relationship, because each technology may be installed in many regions and each region may have an installed capacity for many distinct technologies.

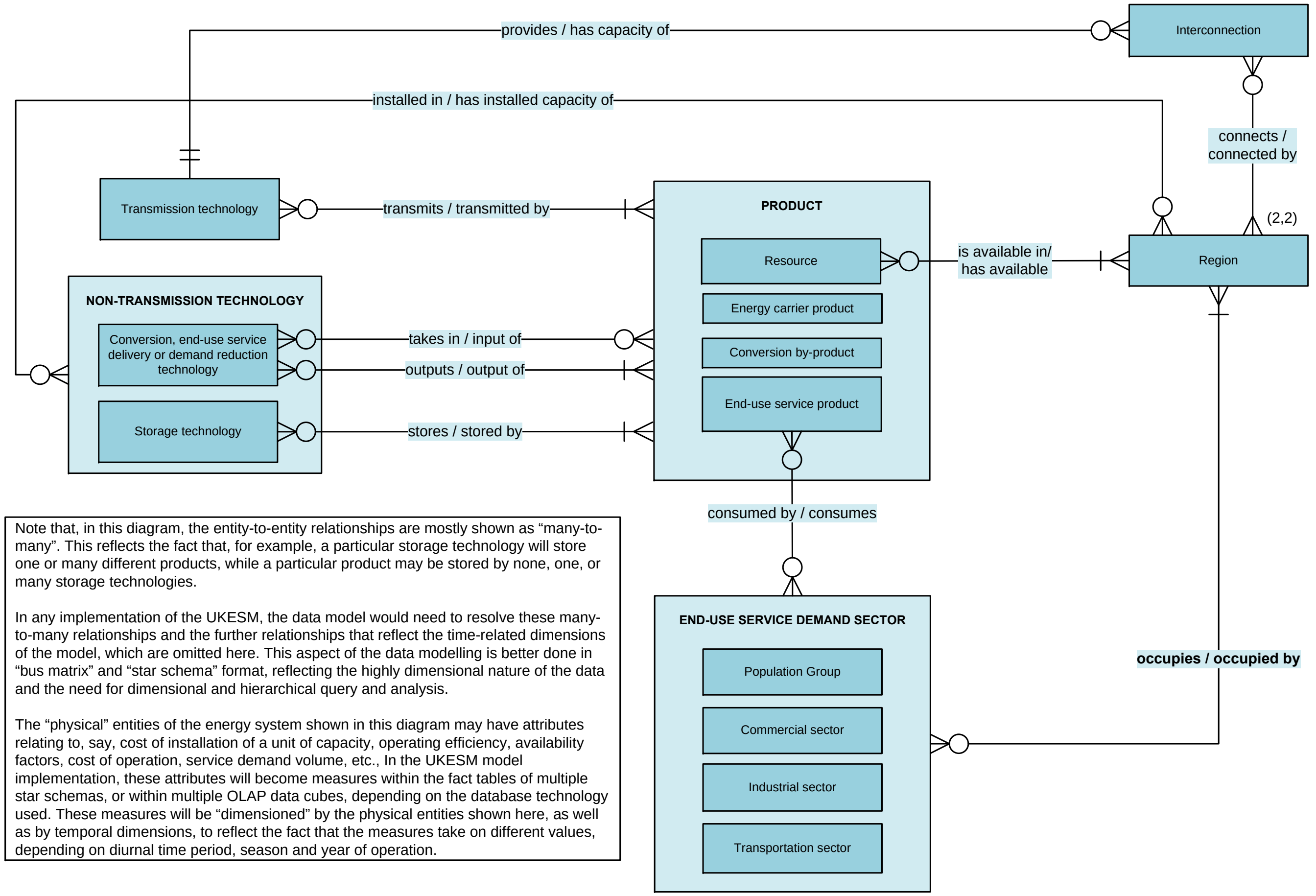
There will be a group of measures, relating to the installation of technology capacity in a region, including:

- Cost of building a unit of capacity;
- Limit to the rate at which capacity can be built;
- Limit to the maximum capacity that can be built.

In the data model for the ESM, these measures are not attributes of the technology entity, since they may take on many different values, depending on the region in which the technology capacity is built (for example a unit of onshore wind farm capacity may cost different amounts depending on the region in which it is installed). The values may also depend on the year in which the technology is built, since we wish to reflect in the ESM the impact of the learning curve for each technology, whereby the cost of building a unit of capacity may reduce over time, as the technology improves. They may also vary through time as a technology ages.

In fact, most of the measures which the ESM needs both as input data and as calculated values, are “dimensioned” in this way and the data model is therefore most usefully represented as a dimensional data model.

Figure 3: Physical Entity Relationship Diagram



Note that, in this diagram, the entity-to-entity relationships are mostly shown as “many-to-many”. This reflects the fact that, for example, a particular storage technology will store one or many different products, while a particular product may be stored by none, one, or many storage technologies.

In any implementation of the UKESM, the data model would need to resolve these many-to-many relationships and the further relationships that reflect the time-related dimensions of the model, which are omitted here. This aspect of the data modelling is better done in “bus matrix” and “star schema” format, reflecting the highly dimensional nature of the data and the need for dimensional and hierarchical query and analysis.

The “physical” entities of the energy system shown in this diagram may have attributes relating to, say, cost of installation of a unit of capacity, operating efficiency, availability factors, cost of operation, service demand volume, etc., In the UKESM model implementation, these attributes will become measures within the fact tables of multiple star schemas, or within multiple OLAP data cubes, depending on the database technology used. These measures will be “dimensioned” by the physical entities shown here, as well as by temporal dimensions, to reflect the fact that the measures take on different values, depending on diurnal time period, season and year of operation.

The “Bus Matrix” shown in Appendix C indicates the measures which need to be held within the ESM and how they may be dimensioned (and relate back to the descriptions in Section 4). These measures are shown as rows in the matrix. The columns represent the dimensions. A particular measure has multiple values, for each valid intersection of the dimensions that apply to it (denoted by an x). For example, the cost of building a unit of capacity has values by technology, by build year and by region.

Note that the dimension columns in the bus matrix correspond to the main entities shown in the Entity Relationship Diagram. Also, the many-to-many relationships between the entities correspond, broadly-speaking, to where measures are dimensioned by those multiple entities.

The Entity Relationship Diagram, Figure 4, combines the real world entity mapping with the concept of measure groups.

5.3. DIMENSION HIERARCHIES

The dimensions of the Bus Matrix relate to certain entity types which are the dimensions of the various measures within the ESM. Each of the entity types may contain one or more hierarchies, which indicate how entities may be “rolled up” to higher levels of the hierarchy. This relates to the functional requirement to be able to run the model at different levels of granularity.

For example, each individual technology may belong to a technology sub-group, e.g. wind power, and each technology group may, in turn, roll up to a higher level grouping, for example renewable energy.

The number of levels and the granularity of these hierarchies will be decided by the ETI modelling team, following experience with the extended prototype. The following section sets out some broad principles.

5.3.1. Data groupings/hierarchies

Granularity of data included within the model will be determined by both the level of data available to the ETI and the extent to which extra detail has a material effect upon results.

There may be groups of technologies, resources, and end-use energy service demands that the ETI will either wish to model as a group and/or report as a group (in addition to individually). Other elements within the model such as nodes and time slices may also have natural groupings.

As a result there will be a hierarchy within the data. The table below shows some examples of the hierarchies that could exist within the model.

Figure 4: Entity Relationship Diagram with Measure Groups

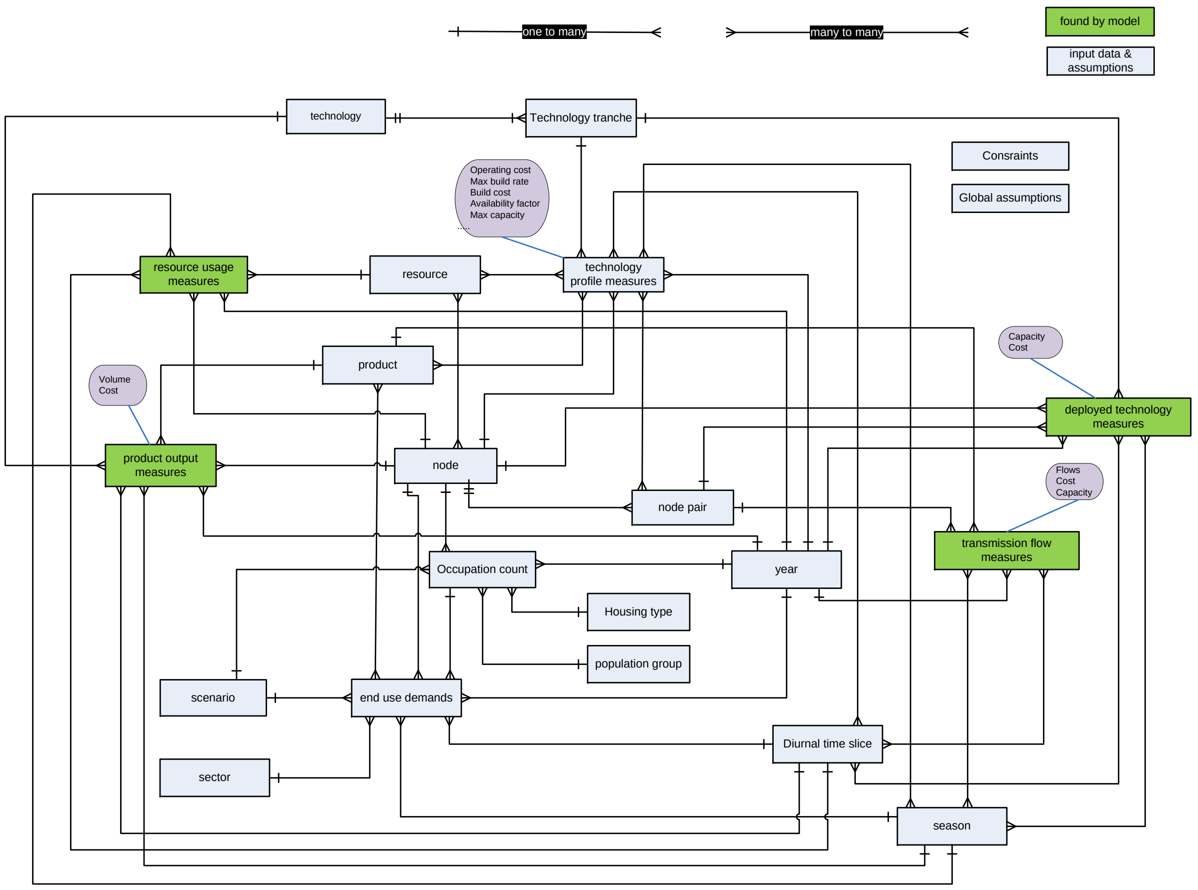


Table 10: Examples of data hierarchies

Group	Elements
Regions	Nodes Node pairs
Fossil fuel generation	Coal plant Gas plant
Gas Plant	Combined Cycle Open Cycle Combustion Turbine
Road Vehicles	Car Bus
Car	Hybrid Gasoline

The data structure should allow the user to both model at different level of aggregations and allow the user to report at an individual element level and grouped level as desired. This may be achieved by having the capability to store multiple sets of data or by having aggregation rules for the data.

5.4. SCENARIOS AND THE DATA MODEL

At a very general level, the ETI will maintain scenarios which will drive the inputted values for certain variables in the model. Currently, the ETI is developing scenarios which will affect mainly demand-side variables. It is also possible that these scenarios be extended to cover global variables in the model or technology input variables. The ETI may also develop variants or sensitivities of the scenarios which test changes in specific variables. The data warehouse (and model) should be able to accommodate the need for multiple data sets which reflect different scenarios and scenario sensitivities and that any variable could be part of a scenario.

The scenario, therefore, is also a dimension within the data model.

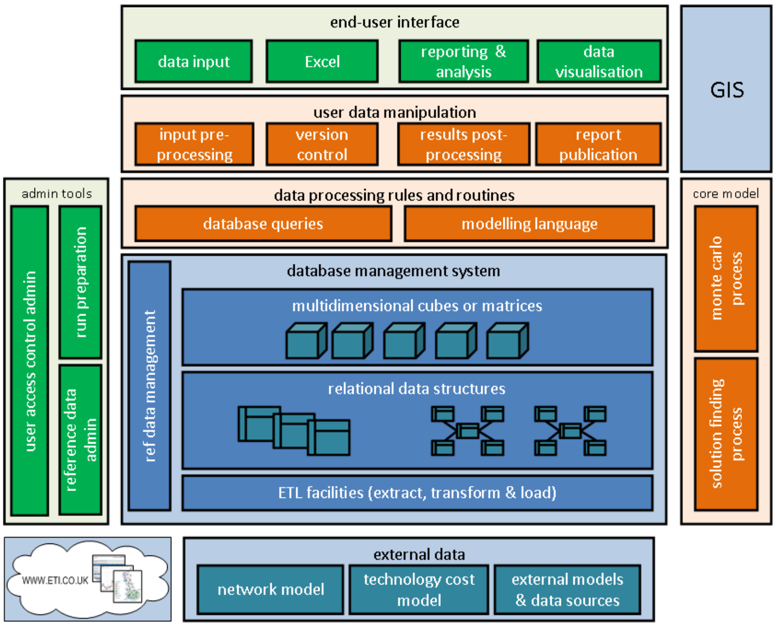
6. MODEL ARCHITECTURE

6.1. MODELLING SYSTEM FUNCTIONAL ARCHITECTURE

6.1.1. Summary

This section outlines the functional components and architecture of the ESM modelling system. The ESM is defined here as comprising all the components of the modelling system, not just the core solution finding model. Figure 5 sets out this envisaged architecture in a single picture.

Figure 5: High Level Functional Architecture of the ESM



The core model is embedded within a wider set of systems components that handle a range of functions, namely:

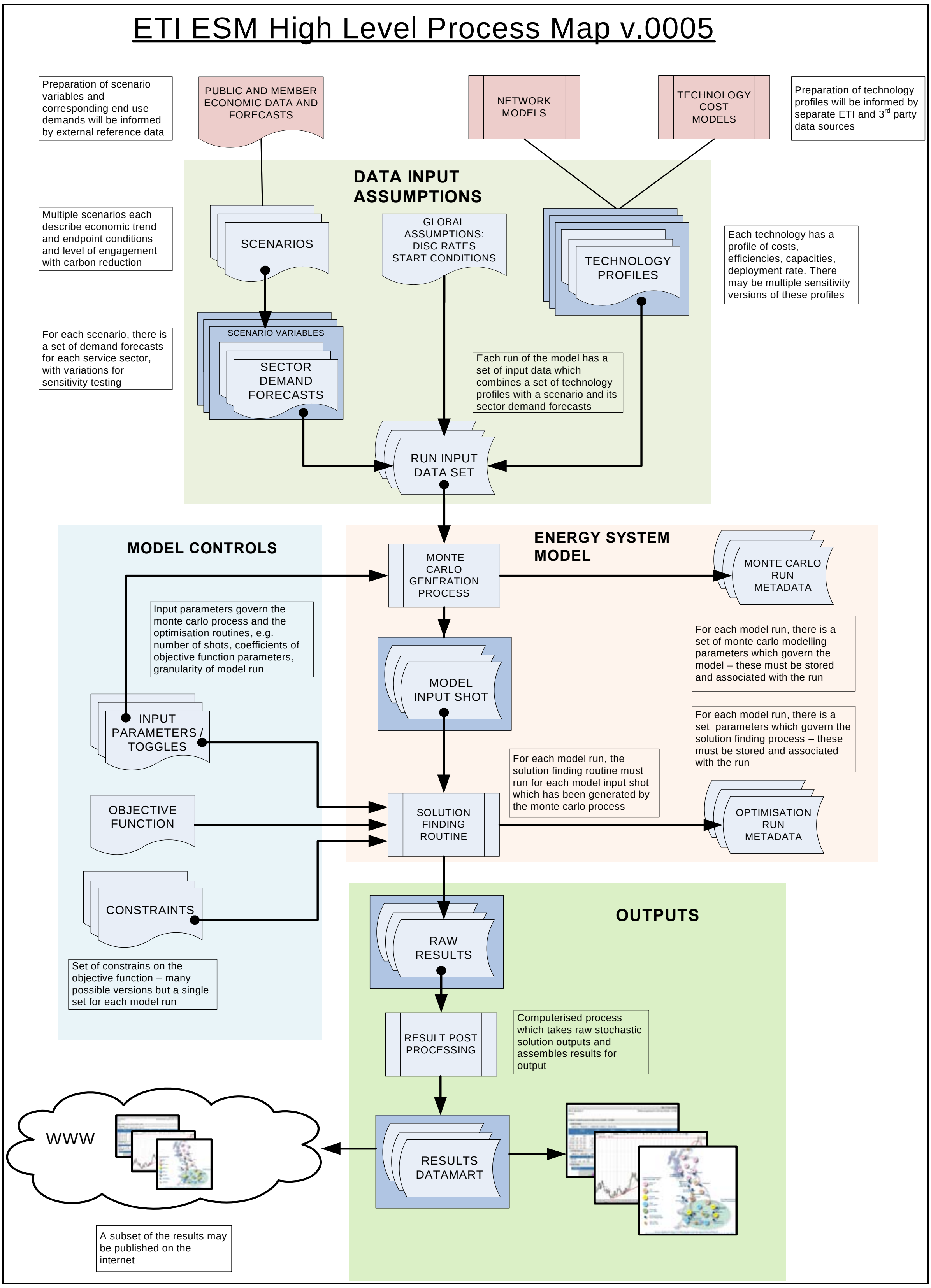
- Database management system, for the storage of all input, output and working datasets
- External models, data sources and publishing media
- ETL (Extract Transform and Load): to handle data interfaces to other models, data sources and publishing media
- Data input and data preparation facilities
- Data pre-processing routines
- Run preparation and run control
- Monte Carlo and solution finding modelling processes
- Results post-processing routines
- Analysis and reporting facilities
- GIS reporting
- External publishing
- Run automation

- Security management

The physical implementation of the system, in terms of the choice of specific software tools and the technical architecture, will be decided in the systems design phase. However, the functional and data needs of the modelling system suggest that the software tools will need to have certain characteristics.

The process chart in Figure 6 describes how these components will interact:.

Figure 6: High Level Model Process Map



ETI ESM High Level Process Map v.0005

Preparation of scenario variables and corresponding end use demands will be informed by external reference data

Preparation of technology profiles will be informed by separate ETI and 3rd party data sources

Multiple scenarios each describe economic trend and endpoint conditions and level of engagement with carbon reduction

Each technology has a profile of costs, efficiencies, capacities, deployment rate. There may be multiple sensitivity versions of these profiles

For each scenario, there is a set of demand forecasts for each service sector, with variations for sensitivity testing

Each run of the model has a set of input data which combines a set of technology profiles with a scenario and its sector demand forecasts

Input parameters govern the monte carlo process and the optimisation routines, e.g. number of shots, coefficients of objective function parameters, granularity of model run

For each model run, there is a set of monte carlo modelling parameters which govern the model – these must be stored and associated with the run

For each model run, there is a set parameters which govern the solution finding process – these must be stored and associated with the run

For each model run, the solution finding routine must run for each model input shot which has been generated by the monte carlo process

Set of constrains on the objective function – many possible versions but a single set for each model run

Computerised process which takes raw stochastic solution outputs and assembles results for output

A subset of the results may be published on the internet

6.1.2. Database management system

A database management system will be a key component of the modeling system. As this is a modelling and analytical system, the majority of data to be held is multi-dimensional in structure. This would suggest that the database management system used should be capable of holding data in a multi-dimensional or "cube" form, since this structure is conducive to rapid processing and analysis.

The modelling system will also need to hold some data in relational format, for example for tracking the compilation of input datasets and tracking and documentation of individual runs of the model. The system will also contain textual data, including commentaries on input data and interpretation of model results. Thus, the database management system will need to be hold data in both relational and multidimensional form, in order to facilitate input and output in the appropriate form.

The most intensive processing within the model will be the solution finding process, which will calculate the optimal technology deployment mix, based on the model's goal and constraints. The solution finding process will need to be run across multiple input data sets, generated by a Monte Carlo process. The solution finding process will, most likely, be implemented using a specialist mathematical modelling environment.

Hence, an option for the database management system would be a relational database management system, where the quantitative data is held in multi-dimensional tables. These would be accessed by the mathematical solution finding process, which would write its results back to the database for subsequent analysis and reporting.

6.1.3. Data required in the modelling system

There are several sets of data which will need to be managed within the modelling system. The structures and definitions of the main data elements within the model are set out Section 4.10. Appendix C describes these data sets in detail.

6.1.4. External models and other data sources

Some of the input data for the ESM will be derived from other models, run by the ETI and possibly by ETI Member organisations. These are described in Section 1.5.

Extract Transform and Load (ETL)

Since the ESM needs to bring in external data and may also need to export results to other systems and interested 3rd parties, there is a need for some extract, transform and load (ETL) capabilities. This will allow, for example, external data to be extracted from the ETI's other models, validated, summarized and transformed, as appropriate, into the particular formats needed by the ESM.

Depending on the regularity of updates the ETL facility may be required to automate the data interfaces with, for example, the technology cost model.

6.1.5. Data input and data preparation

The ESM should provide facilities for users to key in, generate and manipulate model input values, covering the different groups of data:

- Scenarios
- End-use demand forecasts
- Technology profiles (including probability distributions where applicable for variables)
- Global variables

The system will need to provide some of the following functions:

- The ability to key data into standard templates for each of the data groups (this should include the ability to assign standard probability distributions to variables such as Normal, Lognormal, Triangular, Uniform when desired).
- Since many of the input measures will be time series of values, the system should allow the user to generate time series semi-automatically, for example: growing at a given rate from a starting value; interpolating between given start and end values; growing according to a given template growth curve, or based on the growth pattern of another time series.
- Copying individual values, time series or multiple selected time series, from an existing input dataset.
- Generating a new version of an entire data group, for example a technology profile, from an existing version, which can then be amended.
- Importing data values, time series or multiple time series, from linked or external spreadsheets.
- Manipulating other data which has been imported from external models and other data sources, where this has not been automatically managed by the ETL functions (because the manipulation requirement is ad hoc in nature).

6.1.6. Data pre-processing

Since input data may be obtained in a format which is different to that which is needed by the model, additional forms of transformation may be needed. For example, there may be processes which will generate domestic appliance end-use service demand, based on the input of population data.

Some data may need to be allocated across dimensional elements and down through the hierarchies in some dimensions. For example, population growth trend data may be available at regional level, but may need to be reallocated down to sub-region level, using assumptions of population ratios by sub-region, perhaps reflecting a particular scenario’s expected future movements in population concentrations between regions.

Although some transformation processes may be quite fixed in nature, and could be carried out by the ETL process, other processes will require some form of modelling language, which can be programmed by the modelling analysts, rather than depending on deeper technical programming skills.

6.2. RUN PREPARATION AND RUN CONTROL

Having input, imported, or otherwise prepared, all of the input data for a model run, the system needs to facilitate the preparation of specific model runs. The following functions will be required:

- Creating an input dataset, by explicitly selecting the relevant versions of each data group.
- Specifying the level of granularity at which the model should run: for example, some runs of the model may run with data specified at regional node level, whereas more detailed runs may be required with sub-region node analysis; or for example some runs of the model may be run at a technology group level rather than at specific technology level.
- Creating an input dataset, by copying and modifying an existing input dataset: for example one run may be required which is identical to a previous run, except for a different version of the demand forecasts for one or more of the end-use services. This is particularly useful for sensitivity testing, where many runs may be done with variations of a number of input assumptions.
- Selecting relevant model parameters to be used in the run. These may include:
    - The version of the core model algorithms to be used
    - The number of shots to be generated by the Monte Carlo process
    - The on/off toggles for particular constraints within the solution finding process
    - The values of specific constraint constants
    - The co-efficient values of the objective function
    - Parameter values for heuristic algorithms which may be used in solution finding
- Input of relevant narrative for the model run, explaining its purpose and context and any specific additional comments on the selection of data group versions and model parameters.

Since many model runs may use much of the same data, it would be ideal if the system avoided duplication of data from previous. Thus, a model run would be defined as references to the relevant versions of each data group used, so that only one copy of each version need be held and to be re-used by multiple model runs.

Depending upon the intensity of usage of the ESM, there may be a requirement for automation and scheduling of model runs. Thus, the modeller may wish to prepare several model runs which will include the running of the core modelling processes, together with the preparation of selected reports and analyses. These multiple run jobs could then be scheduled to occur automatically, in a predetermined sequence, perhaps overnight, while the analyst is absent or otherwise engaged.

6.2.1. Version creation and model runs for non-specialist users

Although most modelling runs will be prepared and run by the core modelling team at ETI, it is also envisaged that other less specialist users should be able to run the model with alternative versions of input data.

Hence, it is important that the data input and model run preparation facilities can be set up as semi-automatic scripted processes. Thus the non-specialist user should be able to:

- Select an existing model run as the basis for creating a new run variant
- Change a selected range of input data values, representing the “what if?” variant that they wish to explore
- Run the model or, if the model run is likely to take a significant amount of processing time, submit the run for later off-line processing
- View the output results, including analyses of differences to the original variant, once the mode has been run

6.2.2. Monte Carlo "shot" generation and solution-finding modeling processes

These two processes form the core modeling process and are interdependent. The Monte Carlo process generates multiple shots of input data for those input assumptions which have ranges of possible values. The solution-finding process finds a technology deployment solution for each of these shots, based on the given goal and set of constraints.

Choice of solution finding approach

The formulation of the solution finding process is covered separately in Section 2. However, an important aspect of the ESM architecture is to allow for alternative solution finding tools to be used.

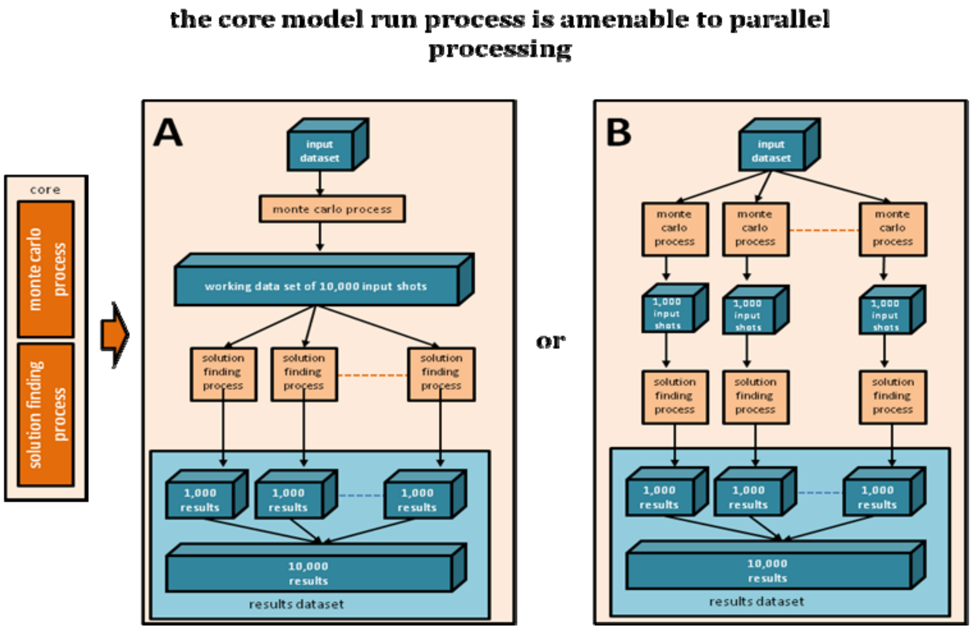
The current prototype uses the AIMMS software package to implement an LP solution. An early version of the prototype used a heuristic approach to solve for the problem. The ETI currently prefers the final solution of the model to implement both an LP and heuristic approach. The experience gained from using the prototype will help the ETI determine a prioritization of the approaches. It is important, therefore, that the modelling system does not become dependent upon one particular LP or heuristic solution tool. In particular, all of the ESM’s data content should reside in a separate database management system, rather than in other modules. This will mean that the input and data analyses will not be affected, if and when the LP tool is changed, other than in amending any ETL process which is required to present the input dataset to the solution finding tool.

Processing and run times

Run time should be configurable to some degree within the model. A full probabilistic run of the model should not take longer than overnight.

Depending upon the level of granularity of the model, the efficiency of the solution finding process and the number of Monte Carlo shots generated, some model runs may be very intensive in terms of data processing requirements. Hence, it may be necessary to consider some form of parallel processing, in order to reduce elapsed run times.

Figure 7: High Level Model Process Map



Since the solution for each input data shot can be found independently of other shots, it may make sense to solve these in parallel on a multiprocessor computer. For example, if a model run requires 8,000 Monte Carlo shots, the run could be split into eight independent sub-runs, each of 1,000 shots, each with a different seed value for the pseudo-random number generator.

These eight runs could then be processed independently, in parallel, which should be significantly faster than running the original 8,000 shot run, with each shot being solved in series. The multiple run output results could then be recombined, in order to compile the full results dataset. This combined result would be equally as valid as the single run. The multiple seed values would need to be stored with the run dataset, in case the run needed to be identically reconstructed.

6.2.3. Post-processing of results

Depending upon the way in which the chosen Monte Carlo and solution finding tools handle input and output of data, it may be necessary to carry out post-model processing of the data.

Data processing

An example of post-processing would be the reallocation of output values for resource and product usage and output products, in particular CO2, to the original end-use services demand.

The modeling system should provide a means of automating such processes, possibly using the ETL component or, preferably, using the modeling language to set up scripted routines.

Secondary simulation analysis

For some analyses it may be interesting to understand the robustness of results to certain input assumptions. It should be possible to perform secondary risk analysis in which the output data set from the model run is further analyzed and where the technology portfolio defined in the model results is now an input rather than a decision variable. Some example questions might be:

- Do the model results hold if a technology is eliminated?
- What is the sensitivity of results to a change in cost or emissions target?
- What is the sensitivity of results to a change in fuel costs or a disruption to fuel supply?
- Do the results change if technology build rates/lifetimes are made stochastic?

6.3. ANALYSIS AND REPORTING

6.3.1. Summary reporting requirements

The modelling system will need provide the ability for the users to create various analytical reports from model run results. In addition to tabular reporting of the values of input and results data, the modelling system should provide a rich capability for data visualisation, through charting.

Both tabular reporting and charts should provide the typical end-user interaction facilities, such as drilling down through data hierarchies to expose greater detail, and drilling across - from the analysis across one dimension to another dimension. For example, an analyst may wish to drill down to the CO2 output attributed to a particular end-use demand sector and then drill down further to show the regional node analysis of the output values.

The analysis and reporting component should enable the automation of the production of standard pre-defined sets of reports, so that these can be run as part of the modelling runs, without user intervention. Appendix E describes the reports in more detail.

6.3.2. GIS reporting

The ESM should represent some results via a GIS system to reflect the spatial nature of the data. Examples of GIS visualisations of input and output data may include:

- Map zones, coloured and enclosed by isotachs (lines of same wind speed), isohels (same sunshine hours), isochimes (same average winter temperatures and isobaths (same average water depth), isorithms (same population densities). Such maps, together with their data, will help to determine some of the input variables and input data assumptions for geographic nodes

- Zonal maps to represent output results from the model, in terms of energy supply and end-use service demand

- Where data input and output is best represented by bar chart or time series charts, then the GIS should position charts for the data at geographic node level, in the correct position on a UK map

Map positions relevant to particular data values should be held within the main database. For example, the model user may wish to specify the position for particular tranches of wind turbine deployment within each region.

It is important that the GIS should be able to read data directly from the database management system. Alternatively, data could be transferred from the modelling system database into the GIS via an ETL process which would take care of any reformatting and reorganisation of the data, to suit the requirements of the GIS component.

6.3.3. Ad hoc enquiry/analysis

In addition to the regular set of reports which may be produced by model runs, the modelling system should provide the end user with the ability to produce ad hoc analyses from the underlying data.

It is essential that end users can access the modelling system database using Excel. This means that the table and cube data structures within the database should be directly meaningful to the Excel end user.

It is unlikely that Excel, alone, will provide sufficiently flexible analysis and charting capabilities to satisfy end-user requirements. Hence it is important that the other reporting and visualisation tools selected will be amenable to ad hoc use by as wide an audience as possible, without the need for specialist technical skills.

6.3.4. Delivery /publishing within ETI

Beyond the immediate model users, the results of the ESM will be of interest to a wider audience within the ETI. Hence, the reporting tools used within the modelling system should be capable of publishing reports and analyses for that audience. It would be preferable to provide a number of facilities for doing this:

- Publication of sets of pre-run or “canned” reports, for direct access by end users

- Publication of reports within a ETI portal
- Email of key reports in Excel or PDF form to a pre-set distribution list

The advantage of the web-portal approach is the opportunity to provide richer commentary and association of results with other less structured data, which may provide useful context to a particular set of results. In addition, a portal presentation would provide a means to collect commentary from end users, which can then be perused by other users and the modelling analysts.

6.3.5. Delivery/publishing to 3rd parties/internet

Taking the portal idea to the next stage, it is ETI's intention to publish certain results of the ESM to external audiences and potentially for general public access.

It is still undecided what aspects of the ESM will be published, but some or all of the following may be considered:

- Documentation of the model function and data structure
- Published scenarios and run results, including input and output data sets
- Facility to run the model on-line, with third parties submitting and storing their own input data versions and model results
- Facility to download and run the model and selected input datasets offline
- Facility to store third party data versions on-line
- Facility for 3rd parties to enter commentary on the ETI's model, results and conclusions
- Facility for open online discussion of ETI model, results and conclusions

6.3.6. Data processing/modelling language

Although Excel may be used for some input, data manipulation and reporting activities with the ESM model, it is likely that a more powerful multidimensional modelling language will be needed. Much of the pre-processing of input data and post-processing of results can be regarded as the manipulation of multidimensional cube or matrix structures.

Such multidimensional processing can be achieved in several different ways:

- Direct manipulation in the multidimensional database itself – for example using the DMX language in Microsoft SQL Server. This approach would only be suitable for processing which is unchanging, since it usually needs more specialist IT skills than some of the more mathematical or commercial programming languages
- Using the rules language of multidimensional database products, such as Cognos TM1, or Oracle Essbase

- Using a multi-dimensional modelling language within a mathematical modelling tool, such as AIMMS or MATLAB, reading from and writing data back to the core ESM database

6.3.7. Run automation

The modelling system should have a means of scripting and scheduling individual processes, so that any appropriate part of the modelling process can be automated. For example, the end user may wish to schedule a number of sensitivity runs for the model, whereby the pre-processing, Monte Carlo generation and solution finding, together with post-processing and production of reports and analyses may all be run unattended and, for example, scheduled for overnight running.

6.3.8. Security component

The security component of the ESM modelling system should enable the system's administrators to restrict access to data and functions, selectively, for particular uses or groups of users.

The security system should be able to operate across the multiple components of the modelling system, so that user rights and user sign-on are based on a single set of user identities and their various permissions.

Ideally, the security system should also be capable of linking automatically with the ETI's wider user security management system, to avoid the need for multiple sign-on and password management.

7. IMPLEMENTATION PATH

7.1. REVIEW OF THE PROTOTYPE

The prototype will be used by the ETI for a period of time, before development of the full model described in this document. The experience gained during use of the prototype will inform a review and revision of this functional definition and the subsequent development of the full model.

The ETI may conclude that the prototype is an exact or close match with requirements and that little or no further development is required. It is also conceivable that use of the prototype may lead to the need for a radically different model formulation.

However, it is an assumption of this section of the Functional Definition that the most likely conclusions from the prototype will be that there is a need for extension and refinement of the model, rather than a radical reformulation, and that the full model is likely to be accessed rather more widely than the prototype.

Hence, this section outlines an iterative and incremental approach to further development of the full model, from the basis of the extended prototype.

The first development step, beyond the extended prototype, will be a thorough assessment of the prototype, in the light of several weeks or months of usage.

7.1.1. An assessment of the future scalability

The ETI will need to consider, in the light of experience with the prototype, what is the appropriate scalability that is required in the full model.

On one hand, the model may need to handle greater detail of the components of the energy system, for example: a finer breakdown of geographic regions, or a larger number of variations of technology types, to reflect the subtleties of their performance characteristics and costs. On the other hand, the prototype may highlight practical limits to this level of granularity due to limits in the availability of valid assumption data or due to the run times of the model, when handling very large volumes of data. There will certainly be a level of detail beyond which the greater precision of finely granulated modelling will be nullified by the spurious nature of the estimates for model input data.

The prototype will yield direct experience of the run times for the solution finding and Monte Carlo aspects of the model. Both features could lead to relatively long run times. A pragmatic limit might be that any model run should be able to complete overnight. Any longer run times would suggest that it would be better to compromise on the level of detail, on the number of input variables which are subject to probability distributions, or on the number of Monte Carlo shots to be generated.

However, it will also be important to consider that that performance of the full model could be made many times faster than the prototype, on a like-for-like basis, if the database were migrated to a server-based architecture, using, for example, Microsoft SQL Server 2008, rather than Access. This would have implications for cost and development effort and should only be done if there is a real need.

With regard to performance, it is likely that much faster model run times could be obtained if some parallel processing of the Monte Carlo shots were introduced. This is because the solution for each shot is an independent process which can be carried out in parallel with solving other shots. There are also techniques for using parallel processing to reduce the run times of large linear program solvers. Such an architectural change may incur extra cost and development effort, but would provide greater flexibility for more demanding model runs.

The other main aspect of scalability will be the number of people, in the ETI and beyond, who may wish to have access to the model and how much concurrent activity there may be from multiple users.

If the model will only ever be used by one or two specialist analysts in the ETI Strategy Directorate, then a modelling system that operates on single workstation PCs will be sufficient. If, however, experience of the prototype indicates that a wider audience would benefit from direct use of the modelling system, then a server-based architecture for the full model should be considered, in order to give concurrent access to the components of the modelling system.

7.1.2. End users: groups and functional requirements

The identification of the end users for the full model will also affect the degree of automation required in the full modelling system. A small specialist team will gain great familiarity with the operation of the mode, and require relatively little automation and guidance from the user interface. However, within a wider and less expert user community, some will need more automation and guidance. Moreover, with a wider group of users, the integrity of the modelling system will need greater protection from inadvertent error and damage to the model.

7.1.3. Publishing requirements

A key aspect of the full model which would have a significant bearing on its future development will be external publishing.

If the ETI Strategy Directorate will control the external publishing of selected results from the model, then very little additional reporting functionality will be required, beyond the prototype. The modelling team will be able to prepare reports, charts and commentaries on selected model run results, off-line, and then publish these, in a controlled manner, through an intranet portal, shared folders, or through email distribution.

If, on the other hand, ETI were to conclude that the model would be made available for use beyond the strategy team and even to third parties, then additional functionality would be required to deliver and control such access to the modelling system. This would also require a greater level of ongoing administrative effort.

7.1.4. Model solution formulation and analysis

The experience gained from the use of the extended prototype will provide a basis for reviewing the functionality defined in this functional definition. In particular, the model formulation may need to change to reflect a revised set of objectives and constraints which the ETI may wish to model.

Use of the prototype is likely to give rise to many additional requirements for reporting and analysis of model results.

7.2. DEVELOPMENT OF THE FULL MODEL

7.2.1. Introduction

It is impossible at the time of writing to anticipate exactly how the full model will be developed, since the functionality will be so impacted by the experience with the prototype. However, on the assumption that the prototype is generally successful, it seems likely that an incremental approach should be considered. The technology used for the prototype could, largely, be upgraded incrementally, without having to rebuild.

7.2.2. Upgrade from Access to Microsoft SQL Server

The upgrade of the database from Access to SQL Server would be the most fundamental step to increase the scalability of the modelling system, from prototype to full modelling system. Even if the full model were not to require multiuser access, the move to SQL server could still be worthwhile, in order to improve performance and robustness of the application.

It would be possible to upgrade to other database products, such as Oracle. However, given that SQL Server is already used within the ETI and that it offers a more cost-effective and simpler migration process from Access, SQL Server would be the preferred upgrade. SQL Server 2005 is a very well-proven product, while SQL Server 2008 is likely to be mature and robust by the time the prototype is due to be reviewed and this later version appears to provide significant performance improvements over SQL Server 2005**.**

If, as suggested, an incremental approach is taken, the prototype could be migrated to the SQL Server platform, while retaining other functionality of the prototype. This would allow the modelling team to continue with modelling activity, with a relatively short interruption and with fewer requirements for data changes and retesting of the model. Further development of other aspects of the full model could then proceed on this new platform.

One of the advantages of SQL Server is its rich functionality for manipulating data. Hence, the move to SQL Server will provide opportunities to automate the extraction of data from source systems such as the GIS database or from other ETI models.

A further benefit of a move to SQL Server would be the richer capability for controlling user access to specific data. This could be a mandatory requirement, if it is decided to make the modelling system available to a wider audience.

7.2.3. Extend model formulation

Following the upgrade to the database, the core modelling functions could then be extended and refined, incrementally, rather than starting a new development from scratch. Thus, the linear programme formulation could be adapted to reflect any changes in the objective function or constraints suggested by the review of the prototype.

The easiest change to the model could be an extension to a finer level of granularity of region, time, end-use service sector or technology. It is important, however, that the data structures of the full model are designed in such a way that it is possible to add lower levels of detail within, say, the hierarchy of technology types, without further configuration or programming of the modelling system itself.

Although extending the data structures may be relatively simple, an important consideration, in increasing the data granularity, would be how to facilitate the use of the model at different levels of detail. For example, the modelling analyst may wish to carry out model runs where, say, end-use service demand is expressed at a relatively high level of detail, while many variants of a particular technology group are included to test sensitivity of the solution. Changing between low and high levels of detail is rarely trivial. The parameters for a higher level grouping of several technologies cannot be calculated as averages, for example, but need to be assessed carefully, to ensure that the parameters are properly representative. It is essential that the full model facilitates the user in selecting alternative levels of detail for specific model runs.

7.2.4. Redesign of data input and run control

In the prototype, much of the data entry, version control and run control is controlled through functionality programmed in Visual Basic. This approach provides a pragmatic mix of the familiarity of Excel with some automation of the modelling process. A potential disadvantage, however, is the dependence on the VB coding expertise, when future changes may be required, in addition to the potential risk of error and the need for retesting when making such changes.

If a more maintainable modelling system is needed, then there would be a case for considering a data collection, forecasting and workflow technology. Such functionality is provided in some CPM tools, such as Clarity 6. The advantage of such tools is that the modelling system administrator would have the ability to reconfigure end user data input forms, processes and reports, without recourse to the more error-prone process of Visual Basic programming.

7.2.5. Development process

Whether the above multi-step incremental development process is adopted, or a more fundamental redesign is needed, the development of the full model should follow a typical good practice systems development process. Key steps would include:

- **Design**
  Decide on technical architecture
  Revise FD data model design
  Process design

- **Development**
  Data model
  ETL, interfaces
  Pre and post-processing data manipulation
  Data entry, version control and run automation
  Core model: monte carlo and solution finding
  Reporting: regular, ad hoc and external publishing
  Administrative processes: reference data, archival, access control

- **Testing**
  Functional testing
  Sensitivity testing
  Calibration of the model (check back to prototype)

- **Data transfer from prototype**
- **User training and process documentation**

7.3. TECHNOLOGY ARCHITECTURE

7.3.1. Change in technology between prototype and full model

Appendix F.4 describes the main changes in the technology chosen to deliver the prototype and the more likely end-point for later versions of the model.

7.3.2. Modularity

A key consideration in the design of the full model should be modularity. For example, it may be that, during the lifetime of usage of the modelling system, the solution finding or Monte Carlo components may need to be changed for more effective alternatives. Hence, it is important that these components are configured within the overall architecture so that they can be relatively easily substituted.

8. CONCLUDING REMARKS

This purpose of this functional definition document has been to:

- Clarify ETI modelling requirements;
- Map out the high level modelling processes and data required
- Recommend modelling and computational approaches.
- Propose a model architecture including interfaces to other ETI/external models
- Proposed a model development path

As discussed in the last section, we would now expect the ETI to engage with the prototype and test the assumptions and directions proposed in this document.

APPENDIX A SECTORAL DESCRIPTIONS

The purpose of this Appendix is to discuss possible approaches to representing the subject matter scope within the model. Some elements of the final representation will depend on the experience gained with the prototype and importantly on the availability of meaningful and accurate data. The section, therefore, provides examples rather than definitive recommendations. In all cases a structure of representation is recommended but the level of detail has yet to be decided.

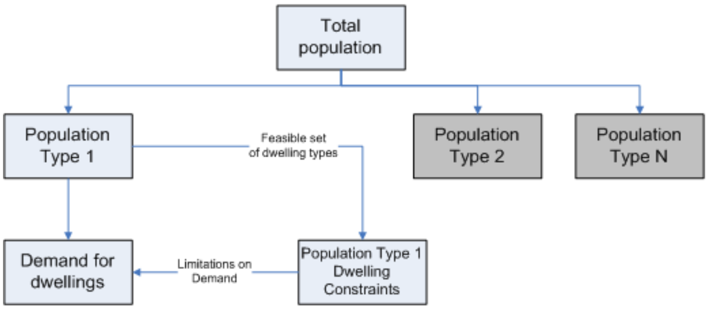
A.1 RESIDENTIAL BUILDINGS

The model should calculate the optimal mix of residential dwellings for each node and in each time period and given user inputted:

- Existing stock of dwellings;
- Population driven demand for new dwellings; and
- Feasible options for future dwellings given demand limitations for particular types of dwelling established by population group.

Figure 8 shows the formulation of such a demand structure.

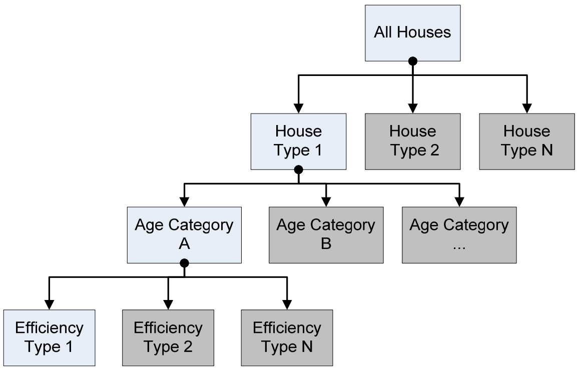
Figure 8: Example formulation for the demand for dwellings



For each node and time period, the model should accommodate a selection of the types of dwelling available to meet population demand. This will be reflected by a decision as to what types of new houses should be constructed and also the extent of renovation within the existing stock – for example, the sub-division of older (large) houses into flats.

Figure 9 shows a schematic representation of how the supply of dwellings might be represented.

Figure 9: Residential Dwelling Supply for each node and time period



The specification of the breakdown of dwelling types may vary by scenario or model run depending on either the data available or the purpose of the analysis. The Office of National Statistics uses some standard categories: detached house, semi detached house, terraced house, purpose built flat, converted flat and other. It may also be desirable to further sub divide property types to characterise by age (e.g., existing or new) and energy efficiency level (e.g., making assumptions about the amount of insulation).

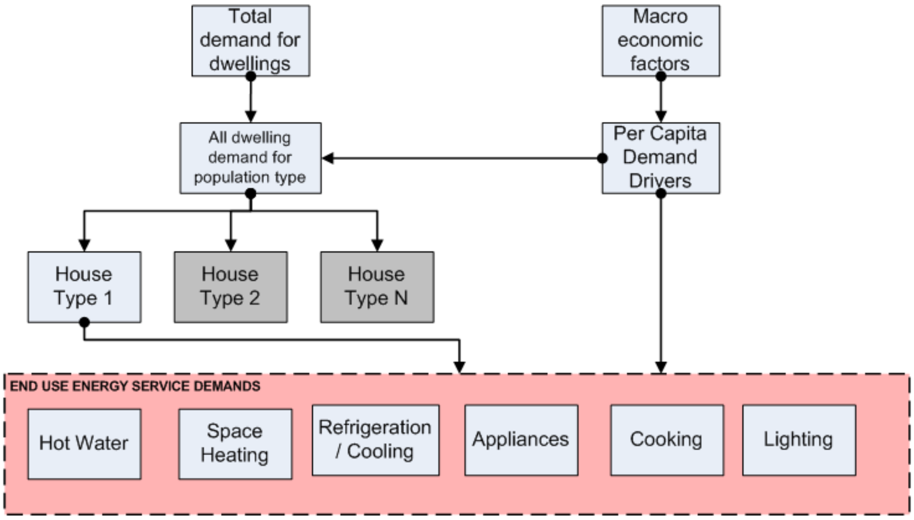
Appendix B.1 includes an example categorisation of housing stock that shows the potential level of detail required in the model.

A.2 RESIDENTIAL END-USE

End-use energy service demand should be modelled as a function of population, economic factors and dwelling occupancy. The demand for energy services, therefore, will be determined either by (a) the type of dwelling or (b) the number and type of people in the dwelling combined with a per capita estimate of energy use (based on economic drivers such as affluence) or (c) all of the above. These relationships will vary by location and through time as well as exhibiting seasonal and diurnal patterns. The choice of (a) – (c) should be a user choice and could vary between model runs.

End-Use demands may be variously specified but a common breakdown is for Hot Water, Space Heating, Refrigeration, Appliance, Cooking and Lighting. Figure 10 shows a possible formulation of energy service demand in light of the dwelling type and per capita input assumptions.

Figure 10: Possible Formulation for Energy Service Demand

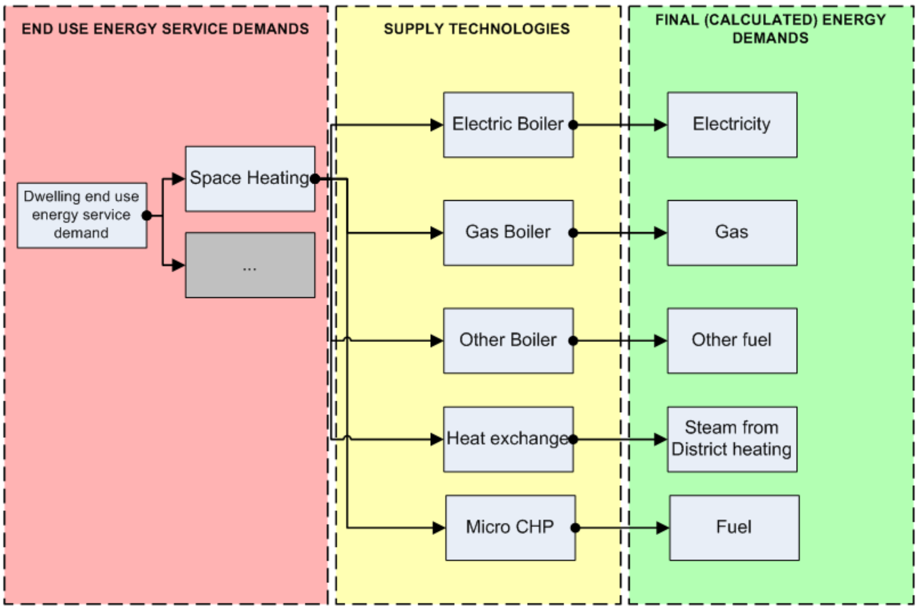


For each node and each time period and for each of the energy service demands there will be a set of technologies that can meet those demands. These technologies consume energy carriers such as electricity and output energy services. The model should reflect these multiple inputs and outputs. The output of multiple end-use services may be in fixed proportion or could be a model decision. For example a micro-CHP unit has a flexible output of steam and electricity and this could be varied. In the event that the ratio of outputs is flexible, then the model will need to reflect that the operating characteristics and costs may also vary with the outputs.

Appendix B.2 provides a provisional list of residential end-use technologies. It should be noted that demand management technologies such as smart metering and insulation will be included as technologies, and not as negative demand. Additionally, technologies may be constrained by the type of dwelling that they service. Different dwelling types will therefore have different sets of feasible technologies associated with them.

Figure 17 shows a possible representation of the relationship between energy service demand, energy technology and final energy demand resulting from that process. The example below is for heating in residential buildings but the approach should be replicated for any end use energy service demand and associated end use technologies.

Figure 11: Possible formulation for the representation of interaction of supply and demand for residential space heating (example)



In this example, demand for space heating could be supplied by a choice of technologies (gas boiler, electric boiler, etc). The supply technologies are simplified in this example and might be broken down further. Other technologies, such as lighting, do not map to the specific demand represented here and so are not included as an option. Given the choice of technology there is a consequent demand for a primary energy resource or energy carrier. The technologies may also have an output of pollutants.

A.3 SERVICE SECTOR END-USE

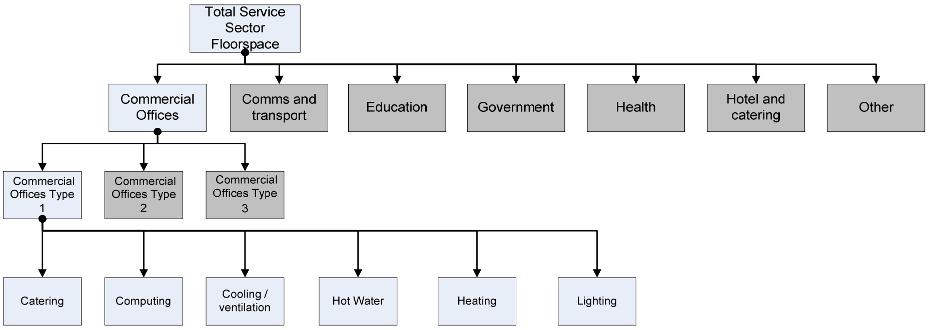
A number of service sectors may be represented in the model and they will be variously defined. Appendix B.3 provides an example list of sub-categories that could be varied based on the availability of relevant demand data.

The model should calculate the optimal mix of end-use technologies for each node and in each time period in each of those categories given user inputted:

- Mix of service sector floor space,
- Information about end-use energy service demand by square metre by floor space

End-Use demands may be variously specified but a common breakdown is for Hot Water, Space Heating, Refrigeration, Appliance, Cooking and Lighting which is similar to the residential sector. Figure 12 shows a possible formulation.

Figure 12: Possible Formulation for Energy Service Demand for Service Sector

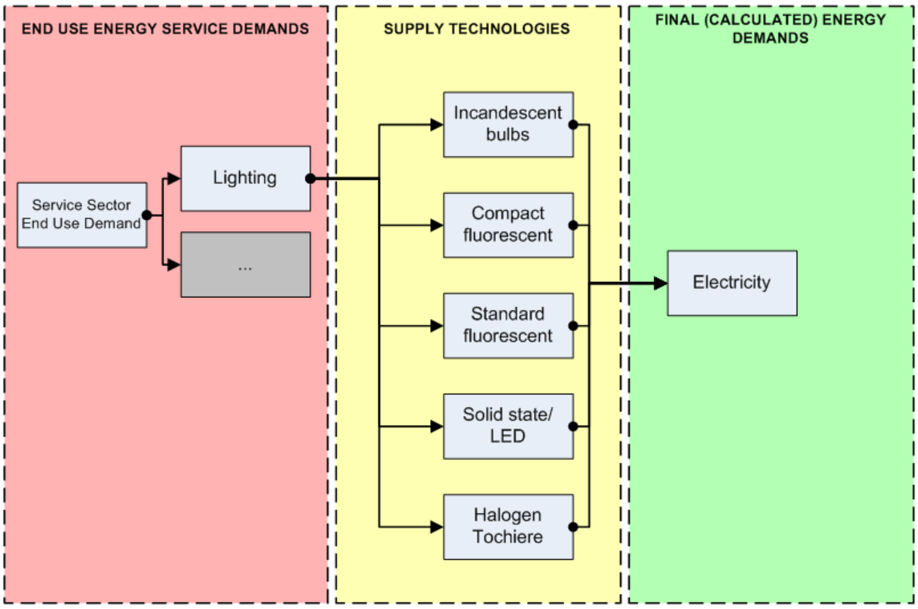


For each node and each time period and for each of the energy service demands there will be a set of technologies that can meet those demands. These technologies consume energy carriers such as electricity and output energy services. The model should reflect these multiple inputs and outputs. The output of multiple end-use services may be in fixed proportion or could be a model decision. For example a micro-CHP unit has a flexible output of steam and electricity and this could be varied. In the event that the ratio of outputs is flexible, then the model will need to reflect that the operating characteristics and costs may also vary with the outputs.

Appendix A.1 includes an example list of technologies applicable to the Service Sector.

Figure 13 shows a representation of the relationship between energy service demand and energy technology and final energy demand resulting from that process.

Figure 13: Possible formulation for the representation of interaction of supply and demand for service sector lighting



In this example, demand for lighting could be supplied by a choice of lighting technologies (incandescent bulb, LED, etc). The supply technologies are simplified in this example and could be broken down further. Other end-use technologies, such as boiler do not map to the specific demand represented here and so are not included as an option.

Given the choice of deployment levels for each technology there is a consequent demand for primary energy resource or energy carrier. These technologies may also have an output of pollutants.

A.4 I NDUSTRIAL SECTOR END-USE

The industrial sector (excluding power generation) will not be modelled in the same level of detail as other sectors to reflect its more limited CO2 impact and less opportunity for the ETI to impact industrial process improvement. As a result, the model should only be able to calculate the optimal way to satisfy final energy demands for electricity, fuel and steam, rather than end-use energy service demand. A number of industries may be represented in the model. Appendix B.4 provides an example list of industrial sub-categories to be modelled.

Each industry type will have an associated set of per unit of output final energy demands. There are several different ways of measuring industrial output. One approach to allow for comparability between different sectors is to use Gross Value Added (“GVA”). GVA measures the contribution to the economy of each individual producer, industry or sector in the UK and is broadly estimated by calculating the value of goods and services produced by an industry net of the value of inputs used in the production process. The Digest of UK Energy Statistics (“DUKES”) reports GVA figures and energy statistics in relation to GVA for UK industry.

In order to be able to account for GHG emissions and final energy demand consistently, the model should maintain at least a simplified representation of the industrial sector. A solution would be to represent (single or multiple) generic production facilities for each industrial type which produce an output equal to demand for that product and have a demand equal to the final energy demand for that industry. As such, one might consider these generic production facilities as end use technologies as modelled in the other sectors.

Figure 14 shows the industrial sector as it might be represented in the model.

Figure 14: Example formulation for the Industrial Sector

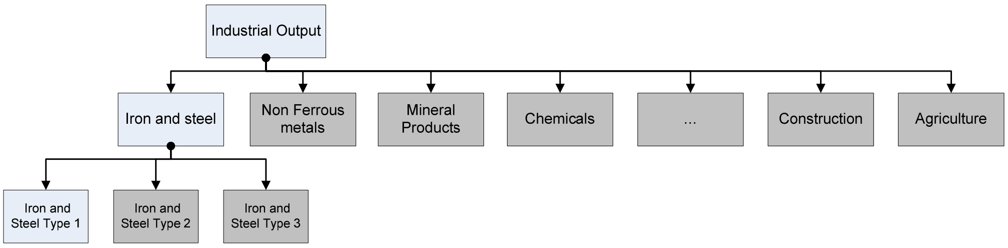
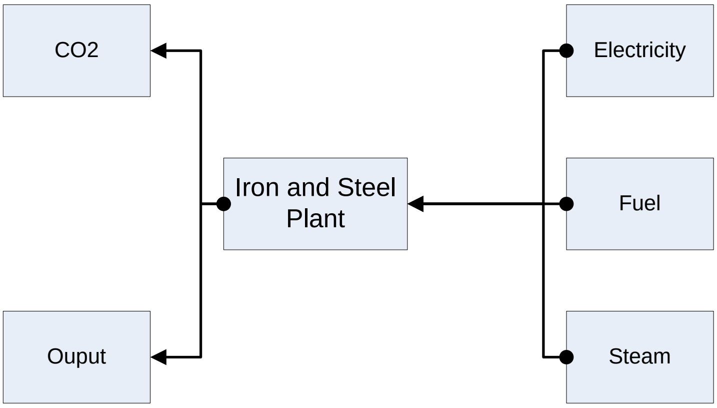


Figure 21 shows how a generic production facility might be represented in the model:

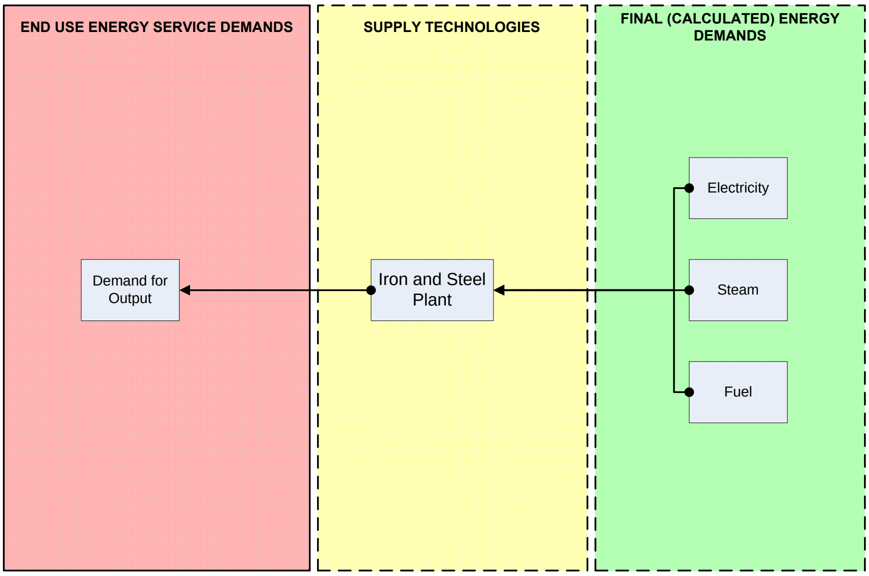
Figure 15: Representation of energy demand through generic industrial plant



The model shall choose sufficient size of generic production facilities to meet demand for output in each of the industrial sectors.

Figure 16 shows a possible representation of the relationship between demand for industrial output (akin to end-use energy service demand in other sectors), end use technologies and final energy demands.

Figure 16: Simplified formulation for the representation of interaction of supply and demand for the industrial sector



In this example, demand for iron and steel output is supplied by a generic iron and steel plant. Other technologies, such as a chemicals plant do not map to this specific demand and so are not included as an option. Given the choice of the amount of the technology there is a consequent demand for fuel. There will be an output of emissions

A.5 TRANSPORTATION

A.5.1 Transportation supply and demand

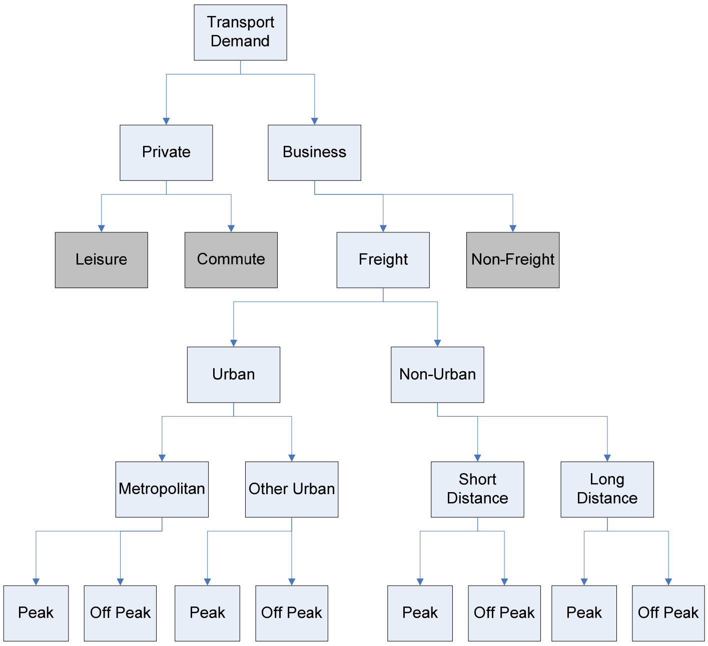
The model should calculate the optimal mix of transportation technologies for each node and in each time period and given user inputted:

- Overall demand for passenger and freight transportation,
- Information about the characteristics of the trips that will be made (for example, is it a long or short journey)

Transportation demand should be broken down into categories for which there will be a meaningful choice of technologies to satisfy it. The model should be flexible enough to accommodate various categorisations. One categorisation could be private journeys and business journeys. Private journeys could be specified in passenger kilometres (or miles) and business journeys could also be specified in passenger kilometres but if freight is moved as a result then will be specified in tonne kilometres (or miles). These categories would be broken down further. For example, private journeys might comprise leisure and commuting kilometres. For example, business journeys might be broken down into freight and non-freight journeys. Each of these journeys takes place in a set of circumstances that might, amongst other things, be characterised by the location and duration of the trip. For example, both private and business trips may take place in an urban or rural environment or be of a long or short duration.

As a result there will be a hierarchy of transportation demands combining purpose and type. The figure below shows a possible representation of such a hierarchy.

Figure 17: Example formulation for transportation demand



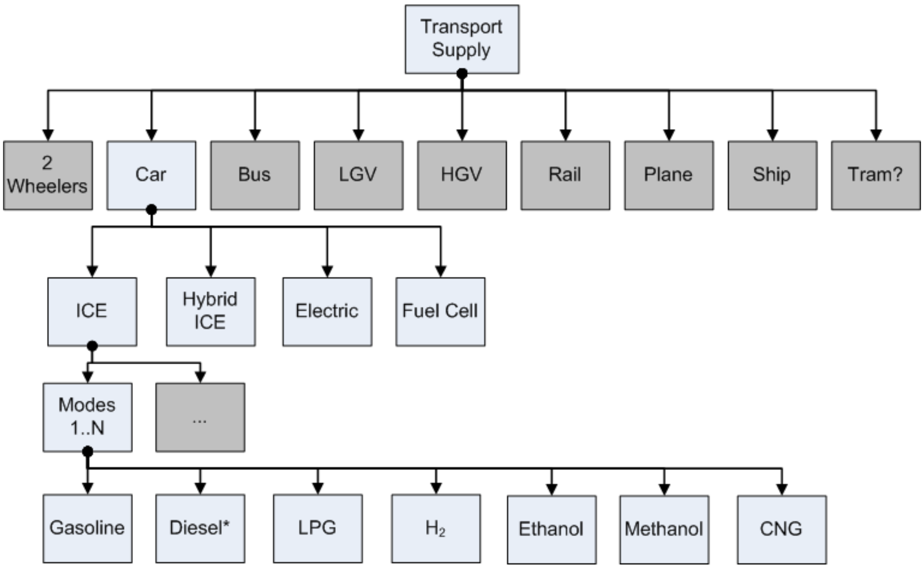
Transportation demand should not specified by mode of transport. As such, a “Peak Metropolitan Urban Freight Tonne Kilometre” makes no requirements as to the technologies which can be used to satisfy it, though that choice is unlikely to be constrained.

There will be a set of transportation technologies that can meet demand. These technologies can be categorised into modes of transport such as Road, Rail, Marine, and Air. Appendix B.5 contains an example list of transportation technologies. These categories may be broken down. For example, car may be broken down into diesel internal combustion, gasoline internal combustion, hybrid, etc. For example goods vehicles might be broken down into HGV and LGV.

Each energy service technology should be defined as being able to output certain types of transportation types. For example, an LGV may supply freight tonne kilometres but not passenger kilometres.

Figure 24 below shows a potential representation the hierarchy of transportation energy service technologies.

Figure 18: Transportation supply hierarchy with outputs example for HGV internal combustion engine supply freight kilometres



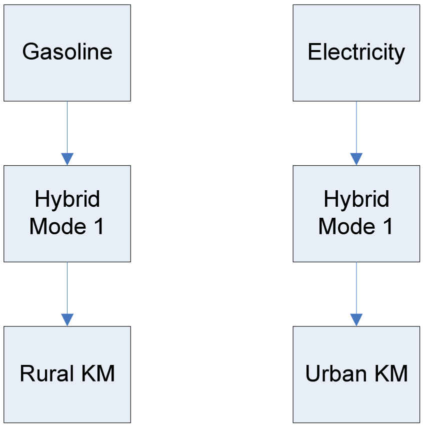
There will be constraints in the model to ensure that the sum of outputs from a particular technology across end use demands is less than or equal to the capacity to supply kilometres by that technology in aggregate.

It may also be necessary to specify a total number of vehicles required in order to ensure that vehicles have redundancy as observed in real life (e.g., cars are not shared so that they are used 100% of all hours). This may be done through a relationship to the number of households or to a per capita assumption. This factor will be a user input.

The model should be able to reflect that the possible output of multiple end-use trip types by a transportation technology may be in fixed proportion or could be a model decision. It is also possible that in the event that the ratio of outputs is flexible, then the model will need to reflect that the operating characteristics (including the type and amount of inputs required) and costs may also vary with the outputs.

Figure 28 shows a schematic representation of this.

Figure 19: Possible representation of modes for a hybrid car



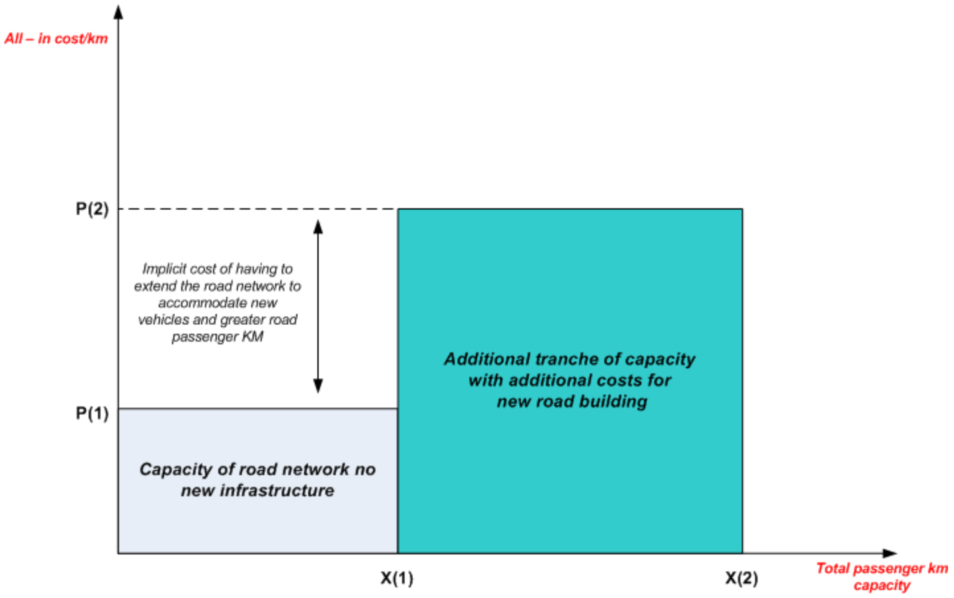
In this example, the hybrid car is capable of supplying either rural or urban kilometres. If it wishes to supply Rural KM it will run in mode 1 and input gasoline. If it wishes to supply urban KM then it will input electricity. The model will have a constraint that limits outputs from vehicles across mode types to an aggregate capacity for kilometres over a given time period.

A.5.2 Transport networks

Transport networks will not be modelled explicitly but cannot be ignored, particularly if modal switching is permitted in choice of transport. A suggested approach could be that transport networks are considered as a resource, within the context of the ESM. A required input for any transportation technology to delivers passenger kilometres or freight tonne kilometres would be transport network kilometres. Those network kilometres could be tranched into capacity and cost groups reflecting the link between additional network usage and infrastructure cost. This could be specified separately for private and freight travel and by node or globally depending on the data available to set these conditions.

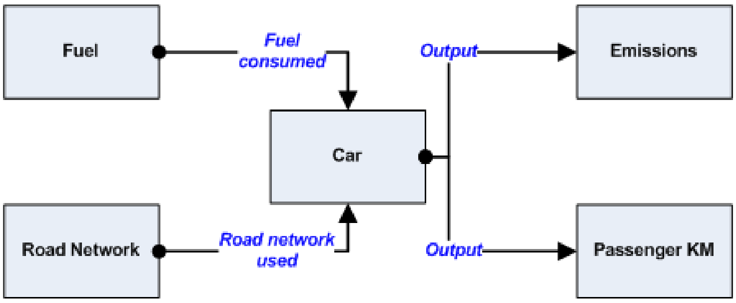
In the simplified example below, the road network has a capacity of X(1) passenger kilometres available at the cost of P(1). This does not include fuel cost. P(1) would be the input cost for each passenger kilometre travelled relating to the use of roads. For every kilometre above X(1) up to a maximum of X(2) the cost of a passenger kilometre using the road network. Variables X(1), P(1), X(2), P(2) are exogenously determined to the model. The difference P(2) – P(1) is the per km cost of extending the road network.

Figure 20: Road Transport Network modelled as a resource



As a result of this formulation, the figure below shows an input/output process map for an example car technology

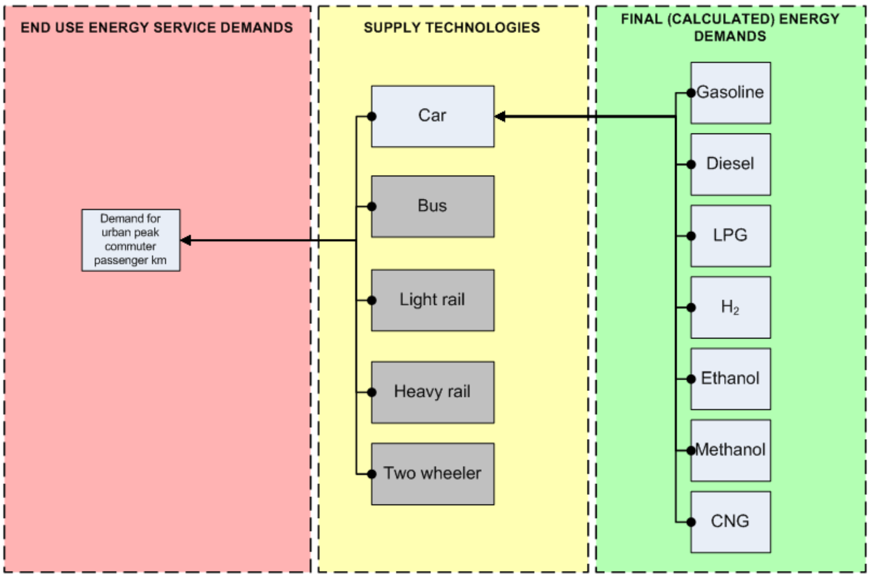
Figure 21: Input output process for car with network resource



The model would also be expected to incorporate the fact that vehicles require fuel delivery points. For fuels such as gasoline and diesel this network exists but for electricity and hydrogen they do not. It is important, therefore, that the costs of building and operating these distribution networks be included in the model. Such distribution networks could be handled as a linked or intermediate technology that needs exist for a certain amount of a transportation technology to exist (see Section 4.3.7).

Figure 28 shows a possible representation of the relationship between energy service demand and energy technology and energy demand resulting from that process.

Figure 22: Possible simplified formulation for the representation of interaction of supply and demand for the transportation sector



In this example, demand for urban peak commuter passenger kilometres could be supplied by a choice of technologies (car, bus,…etc). The supply technologies are very simplified in this example and could be broken down further. Other technologies, such as aeroplane do not map to the specific demand represented here and so are not included as an option. Given the choice of technology there is a consequent demand for primary energy resource and energy carrier.

A.5.3 Linkage to other sector results

In Section A.6 we describe the requirement to model the conversion sector. In particular the decision as to whether to deploy coal generating power plants would have an impact upon the level of freight transportation demand on the UK rail network. The model should incorporate this linkage.

A.6 C ONVERSION / STORAGE AND TRANSMISSION

The deployment and utilisation of end-use technologies to meet end-use demand generates a demand for energy carriers such as electricity and heat. Consequently, the model will require a representation of conversion technologies which input resources (or other energy carriers) and output energy carriers for consumption by other conversion technologies or by end-use technologies to satisfy end-use demand.

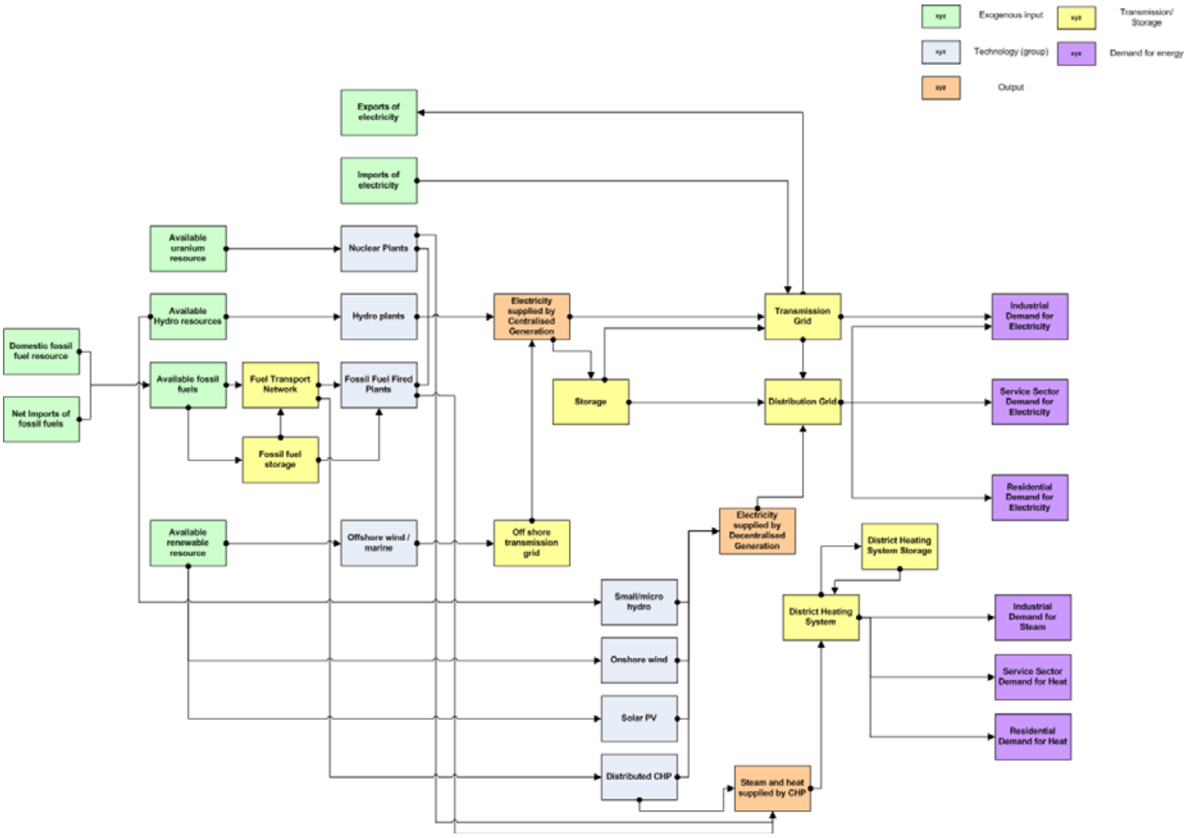
The main conversion technologies that will need to be included are power generation, hydrogen production and refining.

The energy produced by these technologies will need to be transported to the location of the demand for the energy or it will be stored for future use. Consequently the model will need to represent energy carrier transportation networks and storage technologies. In the following sections we describe the combinations of conversion, network and storage technologies and how they map together to satisfy demand for energy from end-use technologies.

A.6.1 Electricity and heat sector

The deployment and utilisation of end-use technologies in the buildings (residential and service), industrial and transport sectors generates a requirement for energy carriers such as electricity and heat. Figure 23 shows a high-level representation of this interaction with respect to the provision of electricity and heat.

Figure 23: High level representation of electricity and heat sector

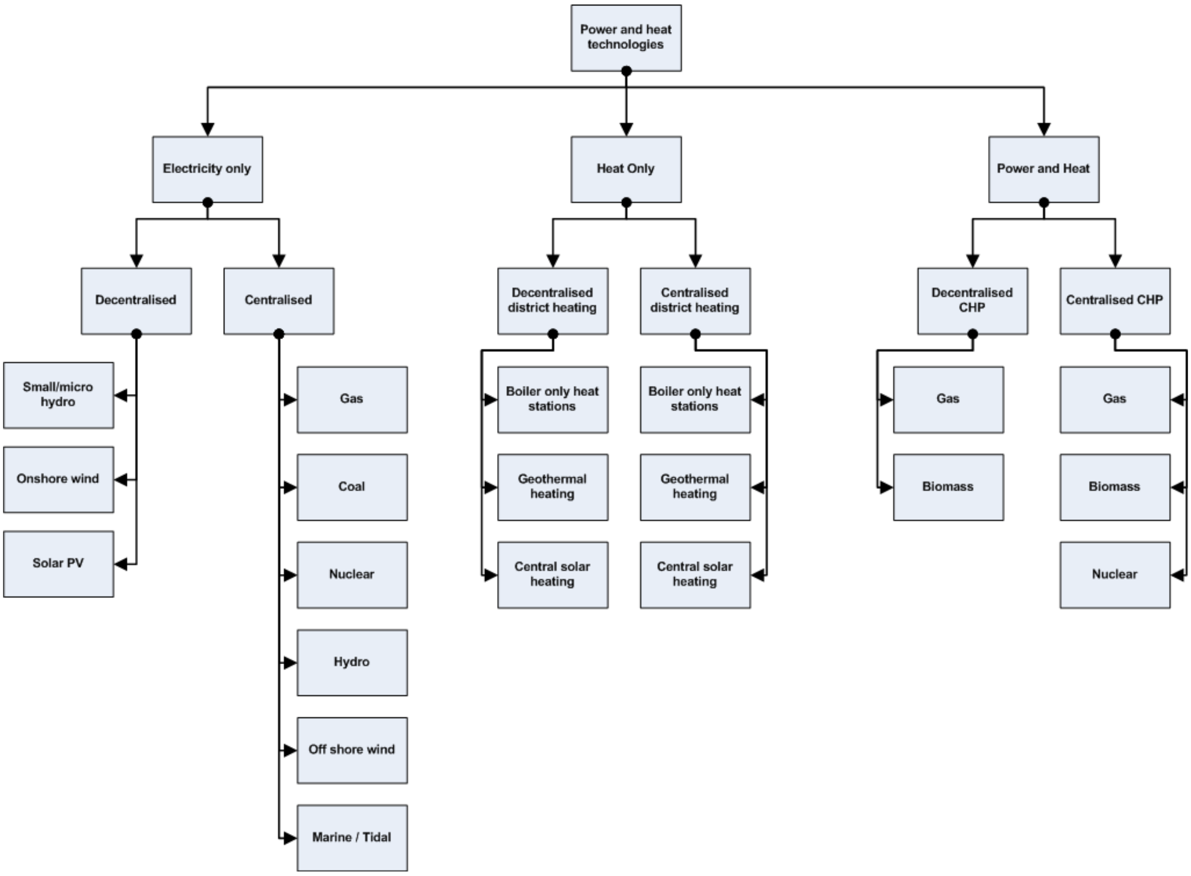


Inputs of fuel or resource (such as wind) are converted by generating technologies into either heat or electricity (depending on the technology). It is transmitted by network to the end-use demand or it is stored and used in a subsequent period. The model should, therefore, consider generation, storage and transmission technologies for electricity and heat.

A.6.2 Electricity generation and heat production technologies

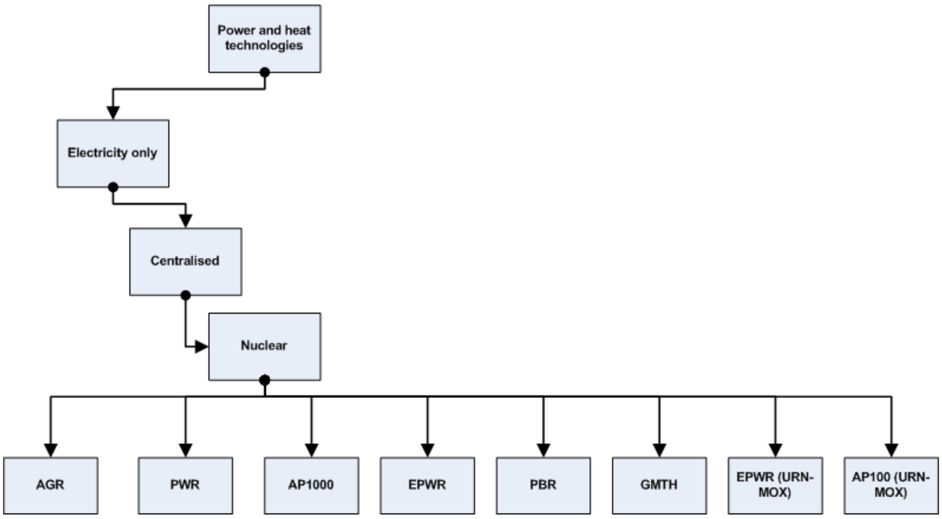
There will be multiple categories of electricity and heat production technologies. Categories and sub-categories may represent particular configurations of similar technologies. Figure 24 shows a possible hierarchy of electricity generation and heat production technologies.

Figure 24: Possible Simplified Electricity and Heat Technology Hierarchies



This is a high-level representation. For each technology, e.g., centralised gas generation, there are likely to be sub-categories, e.g., open cycle gas turbine and combined cycle gas turbine. Figure 25 shows an example of how a technology category might be broken down into sub-categories.

Figure 25: Possible technology hierarchy for Nuclear Generation



Some technologies may not have fixed input and output relationships. For example, a CHP facility could generate a number of different combinations of heat and power for a given amount of inputted fuel. This would need to be represented in the model. In the same manner that certain end-use technologies have modes of operation, one possible solution is for conversion technologies to have modes. A CHP plant could, for example, be entirely outputting power or gas or a pre-defined combination of power and heat.

A.6.3 Refining

Refining will be considered a conversion technology in that it takes in resources and produces energy carriers. The most obvious candidate will be the inclusion of oil refining, the output being refined oil products such as gasoline, diesel and fuel oil. The refining sector will also include refining of bio-fuels.

It may be argued that since the supply of oil products comes from the global market rather than from UK specific refineries, and that UK refineries also supply the global market. As such it will be difficult to determine the requirement for refineries in the UK based on UK demand for fuels. In the event that this relationship can be determined and will have a meaningful impact on the results from the model, then the refining sector should be included and the functionality should be available in the model to do so. It is also possible that selected refined products will be modelled and some may not.

A.6.4 Network technologies

Production of electricity, gas, heat, hydrogen and others are often geographically dispersed from the consumption by end-use technologies. As a result, the model should incorporate transportation networks for these energy carriers.

For the purposes of this model, only trunk line networks will be considered, for example only high voltage transmission of electricity will be represented. Network technology will be defined as having the capacity to move an energy carrier from one node to another.

There may be multiple possible network technologies for each energy carrier. Each network technology, however, should be able to carry only a single carrier. Figure 26 shows an example hierarchy for electricity transmission.

Figure 26: Electricity and Heat Network Technology Hierarchies

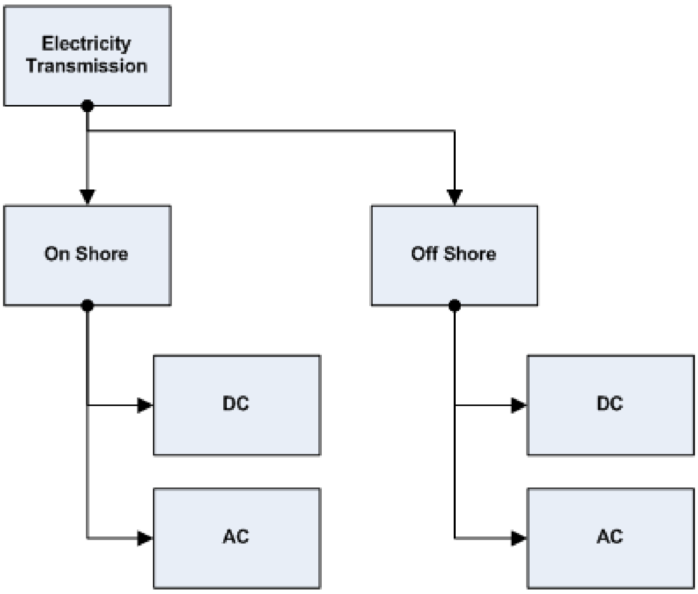
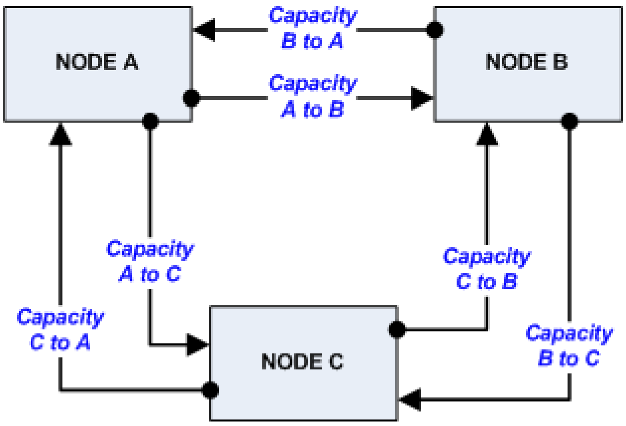


Figure 27 shows a simplified representation of the model process for inter-nodal transmission of electricity and heat. In this example, there are three nodes: A, B, and C. Transfer capacity between nodes is defined in each direction for each transmission technology. If, for example, we wish to model both AC and DC networks then there would be a separate representation for each technology with separate inter-nodal transfer limits.

Figure 27: Possible representation of the model process for a single transmission network technology



As a result, in this example there will be a 3x3 matrix of transfer capacities with 6 values. The translation of the Figure 27 in that matrix is shown in the

Table 11: Tabular representation of transfer capacities

Node from - to	A	B	C
A	-	Capacity A to B	Capacity A to C
B	Capacity B to A	-	Capacity B to C
C	Capacity C to A	Capacity C to kjB	-

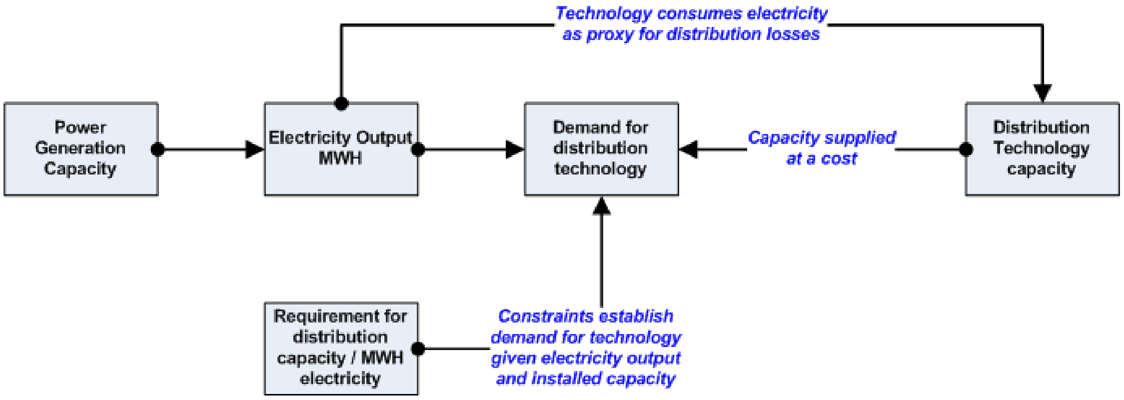
As distribution networks and local district heating networks will exist within a node and do not contribute to inter-nodal transport, local networks will be handled differently. The model should incorporate the notion that there is a required amount of local distribution capacity that is required for a level of electricity or heat production in that node. The model should capture distribution costs and losses and be able to link them to demand for electricity and correlate it to the installation of certain technology, e.g., distributed generation. It is not necessary to represent the flows between points within the region.

One possible formulation would be to establish the distribution network as a technology in the model that has a cost and a capacity. A constraint in the model formulation could be established such that the capacity of the distribution technology should not be less than some value related to the production of electricity or heat in that node.

Costs of the distribution network would then be recognised through the capital and operating costs of the distribution network technology and losses could be recognised by making the distribution network technology consume some amount of electricity related to its capacity (and therefore is indirectly linked back to electricity production/consumption in the node). Costs and availabilities could be modelled in tranches to simulate extensions and reinforcements to the network.

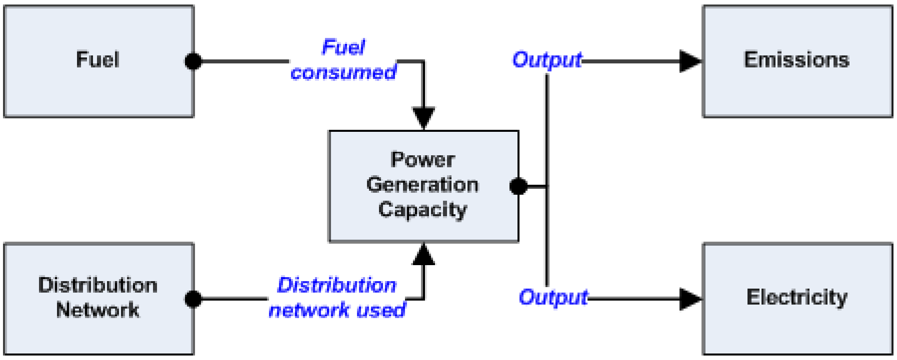
Figure 28 is a representation of this possible model approach for electricity but the same could be applied to the intra-nodal delivery of heat.

Figure 28: Possible representation of intra-nodal electricity distribution networks



An alternative formulation would be to represent distribution networks as a resource. A distribution network "resource" could be established for each node which has a MWh capacity and associated cost. The resource would be capped and have an associated cost. Higher cost resource capacity could be established to simulate necessary extensions or reinforcements to the network. Figure 29 shows this representation.

Figure 29: Alternative possible representation of distribution network

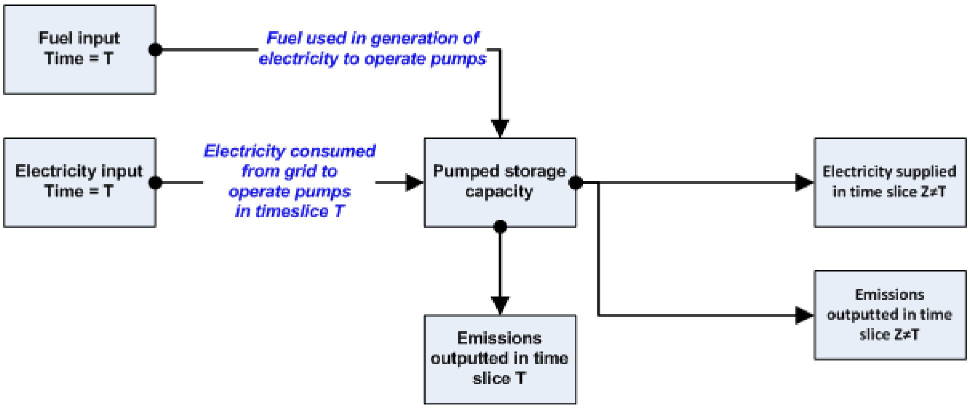


A.6.5 Electricity and heat storage technologies

The model should represent electricity and heat storage. Pumper storage is an aexample and this technology will consume electricity in the off-peak time period and deliver electricity in the subsequent peak period.

District heating networks sometimes include storage capability. This will consume heat in one time slice and deliver in the next time slice.

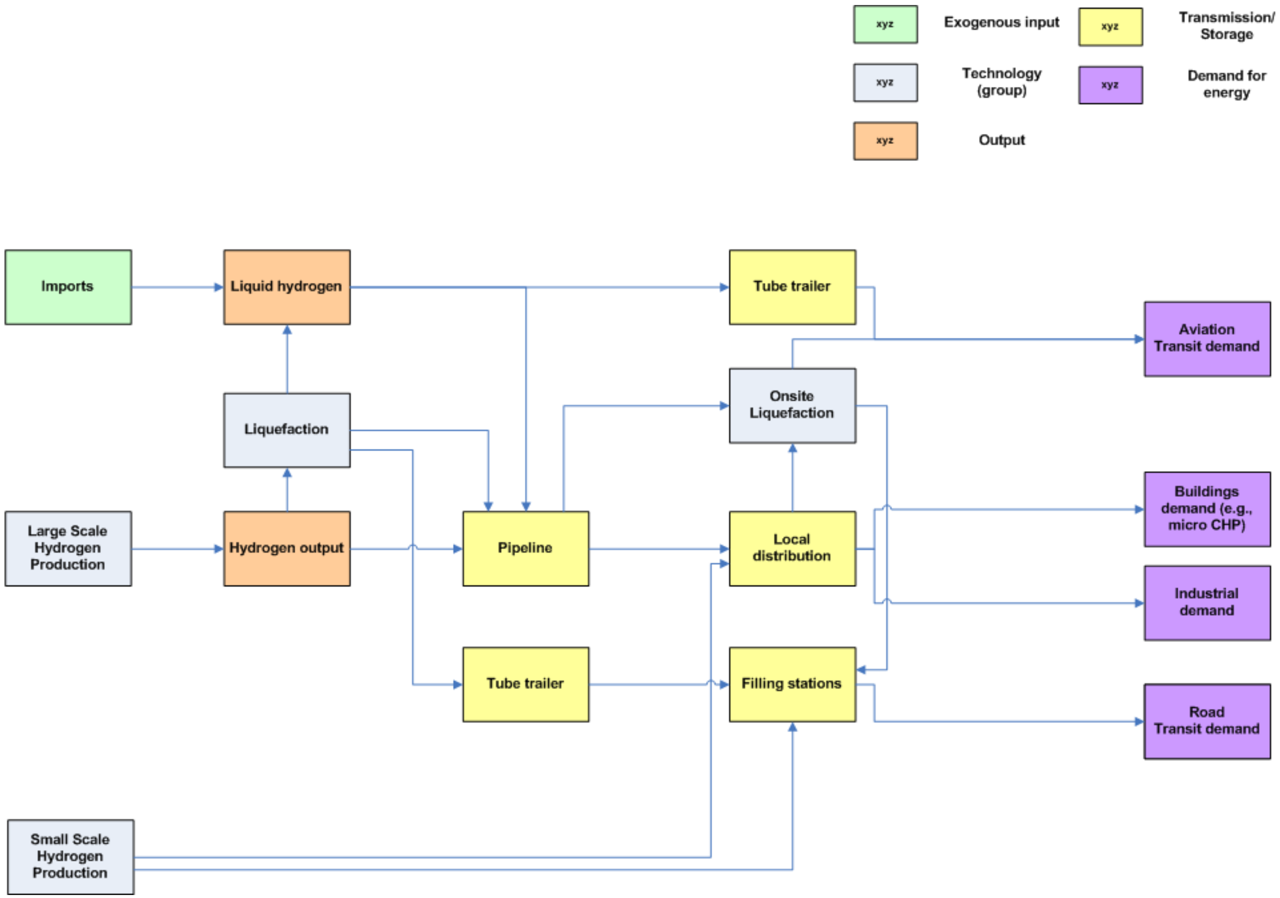
Figure 30: Possible representation model process pumped storage technology



A.6.6 Hydrogen

The deployment and utilisation of end-use technologies in the buildings, industrial and transport sector could generate a requirement for hydrogen. Figure 23 shows a high-level representation of this interaction with respect to the provision of electricity and heat.

Figure 31: High level representation of hydrogen sector

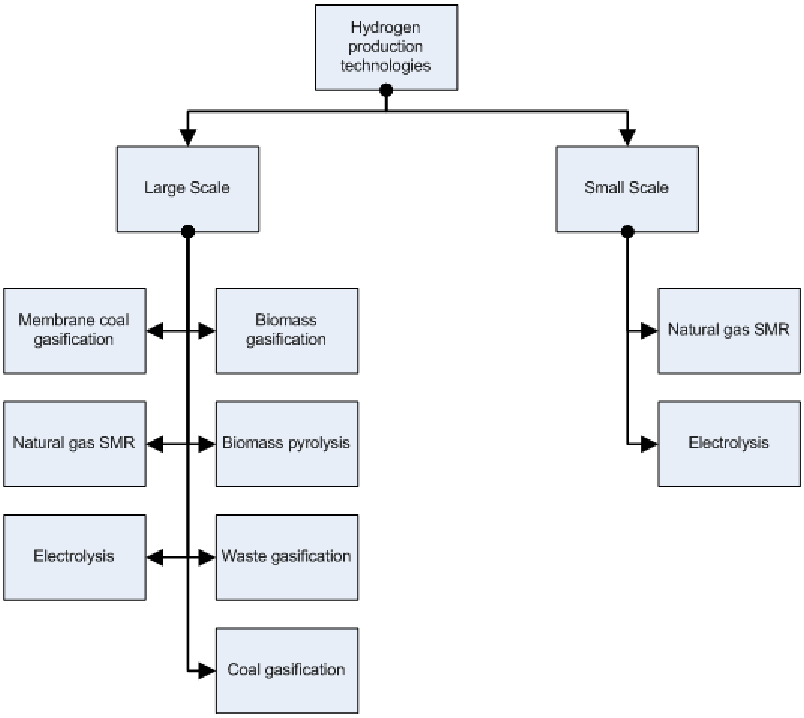


Hydrogen is produced by either small or large scale and can be transported in either liquid or gaseous state. It is transmitted by either pipeline or by tube trailer to its end-use demand location. Depending on the state of the hydrogen transported there may be receiving onsite liquefaction facilities. The model should, therefore, consider production, storage and transmission technologies for hydrogen.

Hydrogen production technologies

The major categories of hydrogen production technologies are steam methane reforming (SMR), gasification of coal, biomass and municipal wastes, and electrolysis. These technologies should be represented in the model. The two categories will be broken down into sub-categories.

Figure 32: Possible representation of hydrogen production technology hierarchy



Many of these technologies are connected to the generation of electricity and so their existence and operating characteristics should be contingent upon the existence of a base technology. Refer to Section 4.3.7 for a discussion on model requirements for linked technologies in the Conversion sector.

Hydrogen transportation and delivery

The model will need to accommodate multiple modes of transport for hydrogen:

- Tube trailer direct to end-use demand
- Tube trailer to filling station.

- Pipeline to a local distribution network or filling station
- Local distribution network to buildings and industrial end-use demands

Major pipelines for hydrogen should be modelled consistently with the other trunkline network technologies such as high voltage transmission. Local distribution networks and filling stations can be modelled consistently with electricity distribution (i.e., not explicitly a model of flows within a node but as a technology or resource to reflect the costs of distribution).

A.6.7 Security of supply

As set out in Section 2, the model should reflect constraints placed on the deployment and utilisation of technologies (and particularly electricity technologies) that ensure security of supply in the energy sector.

The key security of supply constraints that will be placed on the model are:

- Peak installed capacity margin by region at specified points of time; and
- Minimum levels of installed “flexible” generating capacity
- Maximum concentration levels (measured through an HHI) to prevent over-reliance on any single technology

While an element of these constraints will be exogenously determined and inputted into the model, the model should also enable dynamic security of supply constraints. For example, if a region has more wind generation, it may require more flexible generation capacity to support it or a higher level of installed capacity margin against peak demand.

In addition, the contribution to meeting peak installed capacity targets may be different between technologies. For example, wind may contribute a lower % of nameplate capacity toward meeting the margin requirement than a CCGT.

In addition it may be desirable to test the robustness of a model solution to variability in certain variables. This functionality is described in Section 6.2.3.

A.7 RESOURCES

There will be primary energy resources in the model. These may be employed by either end-use technologies (for example, a gas boiler consumes natural gas) or conversion technologies (for example, a CCGT consumes natural gas). Primary energy sources may be fossil or renewable.

As the resources will be considered an input in a technology they will need to be defined by node and have a potential for usage in a specified time period to reflect different availabilities at different times. They will also have a price for each node and specified time period.

If refined products are not modelled as part of the conversion technology representation and so are handled only as part of industrial demand then refined products such as diesel or biofuel will also be a resource with a nodal annual usage constraint and price.

Resources will either be domestic or imported. They may be differently specified – i.e., have different nodal availabilities or have different prices for each specified time period.

Figure 33 shows a proposed hierarchy of fossil fuels and Figure 34 shows the equivalent diagram for the renewable resources.

Figure 33: Possible representation of fossil fuel supply hierarchy

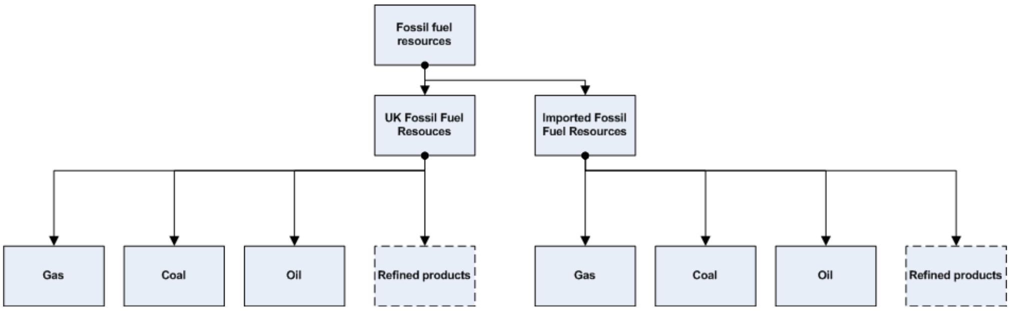
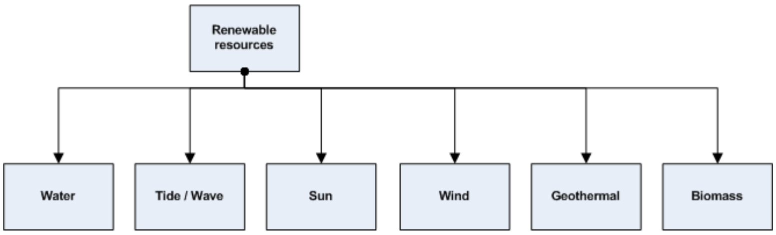


Figure 34: Possible representation of renewable fuel hierarchy



Whereas the availability of fossil fuels will limit only the consumption of the fuels by different end-use and conversion technologies, the availability of renewable resources may also limit the installation of conversion technologies. For example the amount of wind resource in the North Sea will establish a limit on both the installed capacity (GW) of wind generation and the yield. These limits will, again, be on a nodal and time period basis. These will be specified externally to the model in a resource profile.

APPENDIX B ADDITIONAL SECTORAL INFORMATION

This appendix provides examples of the detail likely to be represented in the model assumptions. It should be read in conjunction with Appendix A .

B.1 RESIDENTIAL BUILDINGS – ADDITIONAL DETAIL

The table below shows a potential categorisation of housing types created from the Office of National Statistics and the English National Housing Survey

Table 12: Possible dwelling types and sub-categories

Dwelling Types	Categorisations within Housing Types	
	Age Ranges	In Built Efficiencies
Detached house	*Pre 1919*	*High efficiency for age range*
Semi-detached house	*1919-1944*	*Standard efficiency for age range*
Terraced house	*1945-1964*	*Low efficiency for age range*
Purpose built flat	*1965-1980*	
Converted flat	*1981-2016*	
Other	*Post 2016*	

Source: Housing Types (ONS), Age Ranges (ENHS, CRA), In built efficiencies (CRA)

In this formulation there could be up to 6 (Housing Types) * 6 (Age groups) * 3 (Efficiency Groups) = **108 house types to be represented in each node and time period**.

B.2 RESIDENTIAL END-USE – ADDITIONAL DETAIL

Below, we provide a provisional list of technologies to meet residential end-use energy service demand and was used for the purposes of the prototype.

- **Building insulation (roof)**
- **Building insulation (cavity walls)**
- **Building insulation (solid walls)**
- **Lighting (Incandescent)**
- **Lighting (CFL)**
- **Lighting (LED)**
- **Conventional Gas boiler**
- **High Efficiency Gas Boiler**
- **Space Heating (Air Source Heat Pump**
- **Space Heating (Ground Source Heat Pump)**
- **Water Heating (Electrical)**
- **Macro CHP**
- **"Network hot water" Transmission**

B.3 SERVICE SECTOR – ADDITIONAL DETAIL

The list below shows some possible categories for service sector types and the scale measurement as taken from DUKES.

- **Service Sector**
- **Commercial Offices**
- **Communication and Transport**
- **Education**
- **Government**
- **Health**
- **Hotel and Catering**
- **Other**

The list below shows a provisional list of technologies to meet service sector end-use energy service demand and was used for the purposes of the prototype.

- **As per residential sector plus**
- **Non-Residential Appliances (Current Standard)**
- **Non-Residential Appliances (High Efficiency)**
- **Non-Residential Electric Cooking**
- **Non-Residential Gas Cooking**

B.4 INDUSTRIAL SECTOR – ADDITIONAL DETAIL

Below, we provide an example list of industries reported on by DUKES which could be used in the model for industrial categories.

- **Iron and steel**
- **Non-ferrous metals**
- **Mineral products**
- **Chemicals**
- **Mechanical engineering etc**
- **Electrical engineering etc**

- **Vehicle manufacturing**
- **Food, beverages**
- **Textiles, leather**
- **Paper, printing**
- **Construction**
- **Agriculture**
- **Other industries**

B.5 TRANSPORTATION – ADDITIONAL DETAIL

Below, we provide a provisional list of technologies to meet transportation demand and was used for the purposes of the prototype.

- **Car (Conventional Diesel)**
- **Car (High Efficiency/Hybrid)**
- **Car (Plug In Hybrid)**
- **Car (Battery Electric)**
- **Car (Hydrogen)**
- **Bus (Current Standard)**
- **Bus (High Efficiency)**
- **HGV (Current Standard)**
- **HGV (High Efficiency)**
- **LDV (Current Standard)**
- **LDV (High Efficiency)**
- **Rail (diesel)**
- **Rail (electric)**
- **Aviation**
- **Maritime Freight**

APPENDIX C MODEL DATA SETS

C.1 REFERENCE DATA

Reference data will specify the main categories, or dimensions, within the system and the members, attributes and hierarchical structures within these categories. The initial set of reference data structures is shown in Section 4.10.

The modelling system should provide a means of setting up and maintaining these reference data structures, which will change during the life of the model. It will be important to be able to manage the history of such changes, so that historical data can be retained in its original context and particular model runs can be recreated, by re-running against the relevant data structures.

C.1.1 Input data groups and versions

There are several types of input data required by the system:

- *Scenarios*
  - Scenario data will be keyed into the system and will comprise both quantitative data and free-form, textual information such as scenario narratives and references to external sources which the scenarios have drawn upon.
- *End-use service demand*
  - End use service demand will provide the annual demand estimates over the period to be modelled, i.e. until 2050, also the seasonal and daily profiles, in order to deal with peak levels of demand in the model.
- *Technology profiles*
  - This data group provides the costs and constraints for building capacity of each technology and the operating costs and rates of input and output.
- *Resource profiles*
  - This data group provides the costs and constraints on the use of an energy resource in the UK e.g. biomass.
- *Global model variables*
  - The model will run with a number of global variables, such as discount rates, start values for housing stock, installed technology capacities and existing resource stocks. These will be common to a number of scenarios and versions of service demand forecasts and technology profiles,

- *Model run parameters*
    - A number of run parameters will be specified for each complete run of the model, including the number of Monte Carlo shots to be generated, the coefficients used in the model's solution finding formulae and the values of model constraint variables, in particular for CO2.

The model database should store multiple versions of each of these groups of data. For example, there may be several versions of the values which define a particular technology's profile. In such a case, the different versions may either describe alternative assessments, or changing assessments over time, or "what if" variations for testing the sensitivity of model results to specific variations.

The model needs to reflect the uncertainty around the value of the input data values for technology profiles and, possibly, for end-use service demand values. For a particular run, the model user may wish to designate specific input measures to have a range of values. Therefore, the user will need to be able to select some inputs from the complete set, and specify both a range of values over which the measure's value may vary, together with a probability density function for the range of values.

These probability density functions will be used by the model's Monte Carlo process in order to generate multiple "shots" of input values. Each shot will then be solved by the solution finding process.

In addition to probability distributions for input measures, an input dataset will also have an associated correlation table. This table will define correlation coefficients between the input measures. As described above, this could be implemented through the use of indices and correlation of variables to those indices. These correlations are necessary to constrain the Monte Carlo process, so that generating input values are correspondingly correlated, rather than independently varying.

C.1.2 Input dataset

An input dataset will specify which selected version of each of the input data groups to use. This is to avoid data duplication and the need to change the same data in multiple copies. For example, a particular version of an input data group, say, for one particular technology profile, can be changed once, and then all model runs which use that version can be immediately re-run. Thus, two separate model runs might use one common version of end-user service demand, e.g. for the transportation sector, but each run may use a different version of the technology profile values for, e.g., hybrid cars.

C.1.3 Granularity of the model

Sections 3 and 4 indicate the potential scale of information that could reside in this model. Section 4.10 sets out how that data is structured. It will be important, however, for data granularity to flexible within the model. Hence, demand is specified in a hierarchy form to allow for different levels of granularity. See Appendix A for examples of the demand hierarchies in each sector.

It may be desirable in some model runs to aggregate at a low level of detail and in others to specify demand at a high level of detail. Similarly, technologies may be grouped and considered together in some model runs and in others considered as separate technologies. It is important, therefore, that the system be able either to store multiple datasets of different levels of detail (and manage the consistency between them) or provide a mechanism or set of rules by which to aggregate and disaggregate data.

C.1.4 Working datasets

The Monte Carlo process will generate multiple "shots" of input values. Each shot will contain a set of input measures values, where some of the measures have the same value for every shot, and where some measure values are generated by the Monte Carlo process. Many thousands of these Monte Carlo "shots" may need to be held, for post-model analysis, for each complete model run.

In addition to the Monte Carlo multi-shot process, the model may also be run in "single shot" mode, with an input data set which contains single values for each of the input variables. This can be regarded as merely a special case of the Monte Carlo approach, where only one shot is created and the upper and lower bounds are equal, for every input variable.

A working dataset will comprise the single or multiple input data shots, together with the global variables and the model parameter values used for a particular run. Included in the model parameter values will be the seed value(s) of the random number generator used for a particular run. This is needed in case a particular model run needs to be recreated, or to test the impact of changes to the model formulation.

C.1.5 Results datasets

The output of the solution-finding process will be a 'set of results' dataset. This dataset will contain one or many results, depending on whether the working dataset represents a single input shot or multiple shots generated by the Monte Carlo process. The solution result for each shot will quantify:

- The deployment and utilisation of technologies, over the operational timescale of the model run;
- The consumption of resources and the input and output of other products, including the products for end-use service consumption;
- The costs of resources and technology deployment and operation;
- Security of supply factors.

Where the Monte Carlo process has been used to generate multiple input shots, the results set will contain a corresponding number of result shots. The frequency distribution of the output values will reflect the probabilistically generated input values for each of the input shots.

C.1.6 Run data sets

Over time, the model will be run many times, to find solutions for different scenarios and their associated sets of input assumptions, as well as to explore optional model parameters and constraint values. Hence, it will be necessary to store and manage a large volume of input and output data.

Each run dataset will include the input data for a particular run, the multiple input shots, generated by the Monte Carlo process, the corresponding results dataset, and the model parameter values used for that run.

In many cases, in order to manage the total amount of data held in the system, the model user will choose to retain only the original input dataset and summary output data, dispensing with the multiple input shots and output results, resulting from the Monte Carlo process.

However, it may be necessary to recreate any particular model run, perhaps to carry out additional analysis, or to test the impact of changes to the model formulation, using the same input data values. Hence, the seed value(s) for the random number generator for each run should be stored as model parameter values, with the run dataset, so that the same working set of input values can be generated in the re-run.

C.1.7 Published data

Following many runs of the model, the model users may choose to select specific run datasets for publishing to a wider audience. This may be internally within the ETI, or to related parties, such as ETI members, or to non-related interested parties. Usually, publication will follow appropriate review and authorisation.

Published data will contain specific pre-prepared reports and analyses, together with commentary, appropriate to the intended audience and relevant to the specific conclusions from the particular model run, or group of runs, which were the basis of the analysis.

C.1.8 Workflow data

As well as the actual data content, each data group or dataset, as described above, should have attached data which defines its status and provenance. For example, the model users will wish to know who input, changed and authorized particular data values, when the data was entered into the system or subsequently changed, and by whom, what the original source may have been and what the publishing status may be.

APPENDIX D DATA MODEL BUS MATRIX

ETI ENERGY SYSTEM MODEL DATA BUS MATRIX

Version 2a

UKESM DATA MODEL BUS MATRIX **DIMENSIONS**

Input measures	**Technology/technology tranche**	**Transmission technology**	**Resource**	**Product**	**Region**	**Population group**	**End-use service demand sector**	**Building type**	**Build year / vintage**	**Operating year**	**Diurnal period**	**Seasonal period**	**Scenario**
Scenario													
GDP / GVA (Per capita)					X	X				X			X
Population					X	X				X			X
End-user demand													
Demand for dwellings/floorspace					X	X		X		X			X
End-use product demand volume				X	X		X			X	X	X	X
Temporal fractions													
Diurnal time slice fraction				X	X		X			X	X	X	X
End use seasonal fraction				X	X		X			X		X	X
Resource Characteristics													
Maximum resource available			X		X					X		X	X
Price of resource			X		X					X		X	X
Constraints													
Emission target				X	X					X		X	X
Security of supply target				X	X					X			X
Building Deployment characteristics													
Starting stock / capacity					X			X	X				
Max new build quantity					X			X	X				
Unit capital cost					X			X	X				
Maximum deployment capacity					X			X		X			
Discount rate								X					
Building operating characteristics													
Annual fixed costs					X			X	X	X			
Variable O&M cost					X			X	X	X			
Technical Lifetime								X	X				
Economic Lifetime								X	X				
Technology Deployment Characteristics													
Technology unit capital cost	X				X				X				
Max technology new build quantity	X				X				X				
Maximum deployment capacity	X				X				X				
Starting stock / capacity	X				X				X				
Discount rate	X												
Technology Operating Characteristics													
Input / output quantities per unit capacity	X			X	X				X	X		X	
Variable O&M cost	X				X				X	X		X	
Maintenance rate	X				X				X	X		X	
Forced outage rate	X				X				X	X		X	
Technical Lifetime	X				X				X				
Economic Lifetime	X				X				X				
Annual fixed costs	X				X				X	X			
Capacity credit (intermittency)	X			X	X				X	X	X	X	
Flexibility credit index	X			X	X				X	X	X	X	
Transmission Technology Deployment Characteristics													
Technology unit capital cost		X			X				X				
Max technology new build quantity		X			X				X				
Maximum deployment capacity		X			X					X			
Starting stock / capacity		X			X				X				
Discount rate		X											
Transmission Technology Operating Characteristics													
Variable O&M cost		X			X				X	X		X	
Annual fixed costs		X			X				X	X			
Maintenance rate		X			X				X	X		X	
Forced outage rate		X			X				X	X		X	
Transmission loss factor		X			X				X	X			
Technical Lifetime		X			X				X				
Economic Lifetime		X			X				X				
Security of supply contribution factor		X		X	X				X	X	X	X	

ETI ENERGY SYSTEM MODEL DATA BUS MATRIX

Version 2a

UKESM DATA MODEL BUS MATRIX **DIMENSIONS**

Output measures	Technology/technology tranche	Transmission technology	Resource	Product	Region	Population group	End-use service demand sector	Building type	Build year / vintage	Operating year	Diurnal period	Seasonal period	Scenario
Technology Deployment Measures													
Deployed Capacity	X				X				X	X			X
Total cost of deployed capacity	X				X				X	X			X
Annualised total cost of deployed capacity	X				X				X	X			X
Deployed technology utilisation measures													
product volume consumed	X			X	X					X	X	X	X
Total product output	X			X	X					X	X	X	X
Load Factor	X				X					X	X	X	X
Realised VOM cost	X				X					X	X	X	X
Realised FOM cost	X				X					X			X
Transmission Technology Deployment Measures													
deployed Capacity		X		X	X				X	X			X
Total cost of deployed capacity		X			X				X	X			X
Annualised total cost of deployed capacity		X			X				X	X			X
Transmission Technology utilisation measures													
Total flows on network		X		X	X					X	X	X	X
Load Factor		X		X	X					X	X	X	X
Realised VOM cost		X		X	X					X	X	X	X
Realised FOM cost		X		X	X					X			X
Maximum deployment capacity		X		X						X	X	X	X
Resource usage Measures													
Aggregate resource usage			X		X					X	X	X	X
Aggregate resource cost			X		X					X	X	X	X
Product Output Measures													
Aggregate output cost	X			X	X					X	X	X	X
Aggregate output volume	X			X	X					X	X	X	X

APPENDIX E REPORTING REQUIREMENT EXAMPLES

E.1 DESIRED BASIC REPORTS

The model should report the results from each and every Monte Carlo shot. In addition it should be able to report a summary analysis of all the model shots. This section describes the basic output reports desired by the ETI. These are:

- For each model run, the amount of each technology deployed and its utilisation rate for each year modelled. Figure 35 and Figure 36 give examples.

Figure 35: Single Year Portfolio cost and utilisation output reports

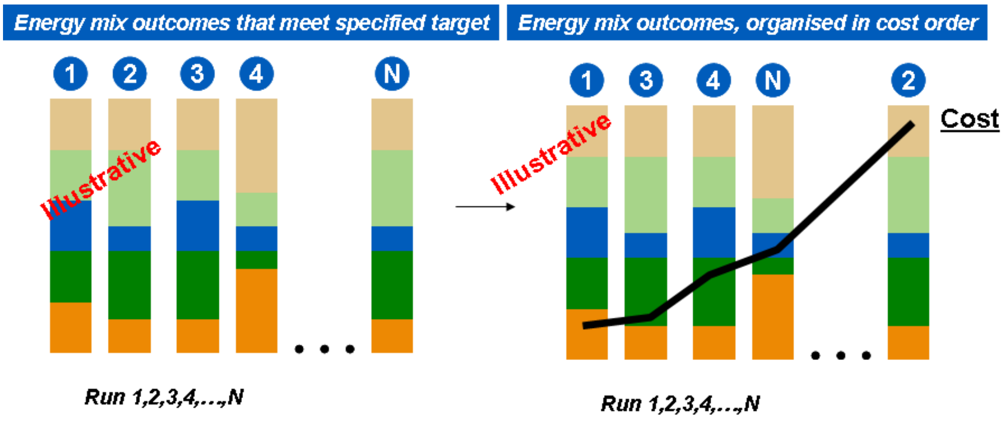
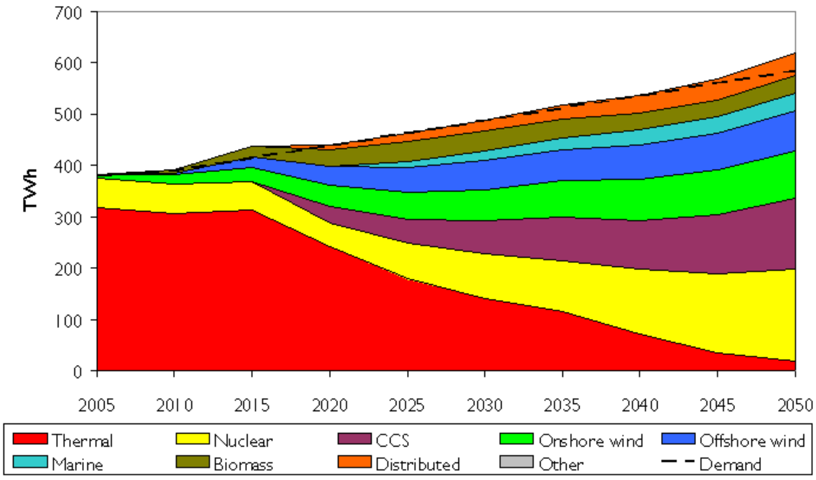
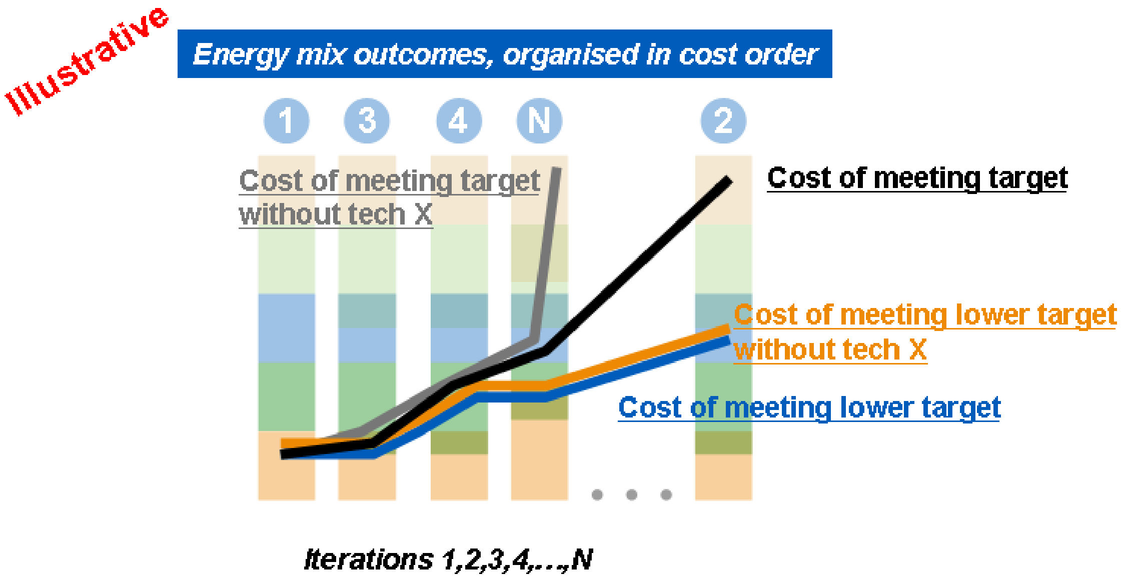


Figure 36: 2010 – 2050 Technology mix/utilisation output report



- Across scenarios, a comparison of the above metrics, as shown in Figure 37 below:

Figure 37: Scenario Results Comparison Report Variation



- Across all model runs, the ETI would like to view comparative statistics for the contribution to the energy mix – the probability of a technology being deployed and its utilisation if deployed. Figure 38 and Figure 39show a possible way of producing a report view on the contribution to energy mix.

Figure 38: Technology Energy Mix Contribution Report

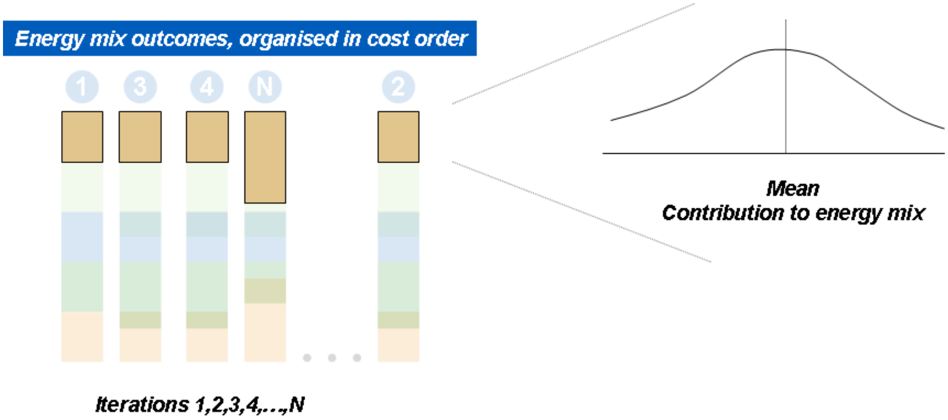
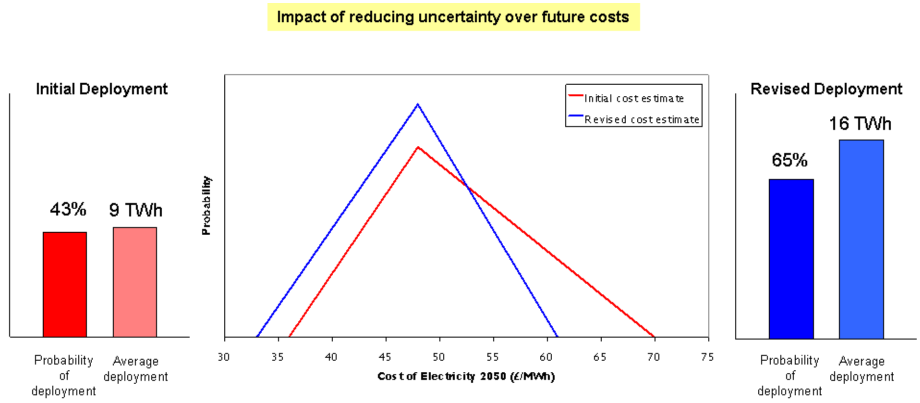
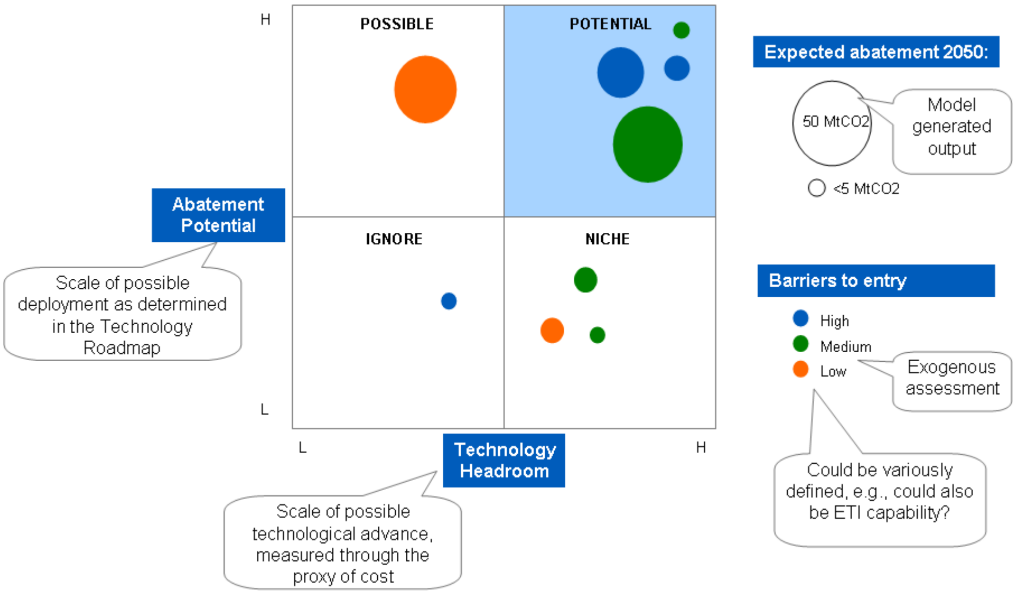


Figure 39: Deployment and utilisation report



- It would be useful to combine results information from the model with externally held information. For example, we would like to be able to create reports along the lines shown in Figure 40.

Figure 40: Example report combining model results with static data



APPENDIX F COMPARISON TO PROTOTYPE MODEL

F.1 INTRODUCTION

The prototype has been used as a means to test elements of both the model scope and processes. The prototype has been intended as a working "desktop" model which could easily and frequently be used within the ETI Strategy Directorate. Its scope includes all three sectors that will be included later versions of the model but the detail of its coverage has been necessarily restricted given the time constraints. As such the main difference in scope comes in depth.

Later versions of the model might be used by multiple (and possibly external) stakeholders to the ETI. If so, it may be desirable to move to a server-based solution but in nay case it will require robust audit and validation capabilities. In a number of scope areas the level of detail anticipated will be greater than for the prototype and so later versions of the model will need to be scalable, be able to incorporate significant volumes of data and allow for the user to vary and select the level of detail to be modelled.

This section sets out the main scope and functionality differences between the prototype and the full ambition that has been declared for later versions of the model.

F.2 DIFFERENCES IN SCOPE

The table below shows the key differences between the prototype and the plan for later versions of the model.

Table 13: Scope Differences

Scope	Main differences
Buildings / transport / industry-other	• Greater level of detail in both the number of demand types that could be modelled and also a greater breakdown of technology types • Later versions of the model may include refining as a conversion technology. The prototype includes refined fuels as a resource whose production is exogenously determined. Later versions will reflect the conversion of resources such as crude oil.
Networks and infrastructure	• The prototype only models the electricity network and CO2 transport. Later versions of the model will also permit the modelling of any energy carrier transportation network. These may include for example gas. • The prototype allows no technology choice in the transmission network. Later versions of the model will allow for multiple

	alternative network types – e.g., DC / AC networks for electricity • The prototype does not explicitly represent distribution networks within a node. Later versions of the model will incorporate such representations. For example, road networks will be modelled as a resource. Electricity distribution networks will be modelled as an intermediate technology. This will allow later versions of the model to reflect the impact that new technologies have upon the need for local networks.
Technology learning curves	• There is a desire for later versions of the model to allow for endogenous learning curves and to allow for learning curves to incorporate shifts instead of being exogenous and smooth. This will depend on the use of non-linear programming in later versions.
Seasons and diurnal effects	• Later versions of the model will incorporate possibly finer definitions of the time slices used.
Transportation modal switching	• The prototype defines transport demand by the mode – e.g., road, rail, air. It does not allow for switching between modes; it only allows for switching between transport types of the same mode. Later versions of the model will allow for switching between modes, e.g, between road and rail. When switching is permitted, however, the ability to switch will be constrained to avoid unrealistic changes in mode.
Security of supply	• The prototype covers requirements for regionally defined peak capacity margins for electricity and requirements for flexible generation (through a flexibility credit). Later versions of the model will take this further by making peak requirements dynamically dependent upon the type of power plants existing. • The model will also reflect the need to overdependence upon single technologies by allowing for constraints on the market share of specific technologies. • The model may also include constraints on technologies other than electricity generation.

F.3 DIFFERENCES IN FUNCTIONALITY

The table below shows the main differences in functionality of the model.

Table 14: Differences in functionality

Functionality	Main difference
Scalability	• The granularity of data is manually fixed within the prototype. While

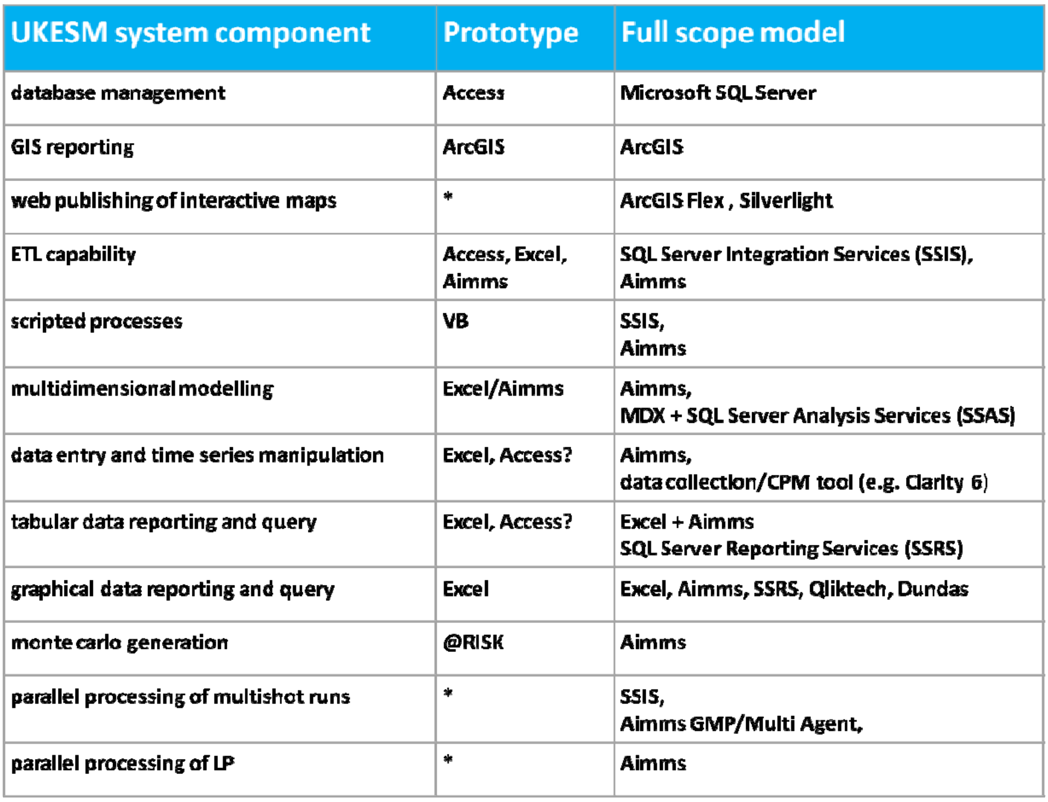
	it can be varied, the management of this process is manual and difficult. • Later versions of the model will allow the user to incorporate alternative levels of detail, e.g., nuclear or class of nuclear generator. The model will allow the user not only to report at different levels of aggregation but also select between running the model at different levels of detail.
Objective function	• The prototype solves for 2050 and ensures that the path to 2050 is feasible (but it does not necessarily have to be optimal). It does not, for instance, solve for all demand. The prototype does constrain on emissions limits and security of supply in the path to 2050. • Later versions of the model (depending on the experience of using the prototype) may solve the objective function for the whole 2010-50 time period. Later versions will also enable the model to ensure that all demands are met during the path to 2050.
Assumption detail	• Later versions of the model will have greater flexibility in the manner by which data is defined in the model. Data may be definable in more ways. For example, in the prototype operating costs are defined by the vintage of the technology. In later versions, the dimensions will be a user choice. For example, operating costs may be defined by both vintage of the technology and the operating age of the technology.
Solver	• The prototype solves using Linear Programming methods only. Later versions of the model may allow the user to select between approaches. For example, it may be desired in some circumstances to solve using a heuristic method.

F.4 DIFFERENCES IN ARCHITECTURE

As with the functionality of the full model, it is difficult at the time of writing to judge how significantly the technology platform may need to change from that used in the prototype.

On the assumption that the full model may need to be significantly more scalable, more automated and more widely accessible, than the prototype, then the following table indicates how the technology components may differ.

Table 15: Possible differences in technology choices for model platform



UKESM system component	Prototype	Full scope model
database management	Access	Microsoft SQL Server
GIS reporting	ArcGIS	ArcGIS
web publishing of interactive maps	*	ArcGIS Flex , Silverlight
ETL capability	Access, Excel, Aimms	SQL Server Integration Services (SSIS), Aimms
scripted processes	VB	SSIS, Aimms
multidimensional modelling	Excel/Aimms	Aimms, MDX + SQL Server Analysis Services (SSAS)
data entry and time series manipulation	Excel, Access?	Aimms, data collection/CPM tool (e.g. Clarity 6)
tabular data reporting and query	Excel, Access?	Excel + Aimms SQL Server Reporting Services (SSRS)
graphical data reporting and query	Excel	Excel, Aimms, SSRS, Qliktech, Dundas
monte carlo generation	@RISK	Aimms
parallel processing of multishot runs	*	SSIS, Aimms GMP/Multi Agent,
parallel processing of LP	*	Aimms