

Department of Architecture
and Built Environment

University of Nottingham

Analysis of strategies to decarbonise the UK housing stock

Gustavo A Sousa

Supervisors:
Prof. Darren Robinson
Dr. Benjamin Jones
Dr. Parham Mirzaei

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

August, 2017

"Essentially, all models are wrong, but some are useful"

– George E. P. Box

"Occasionally some apparently simple phenomena will require complex theories and apparently complex phenomena will require simple theories"

– Bruce Edmonds

Abstract

The purpose of this *Thesis* is to revise the adoption of strategies for both new and existing housing towards the reduction of carbon emissions associated to energy demand, particularly by identifying promising technical improvements, as well as by capturing potential incentives to adopt these improvements. The case of housing is particularly interesting because although it is straightforward to categorise each building, perhaps by its physical properties, it can be considered unique because of the quasi-random behaviour of their occupants which can be determined as collective (e.g. peer pressure influencing the installation of solar panels), circumstantial (e.g. local incentives or sales in efficient devices), biological (e.g. occupants needs according to age and health conditions), or cultural (e.g. habits and patterns). It is also interesting because domestic buildings—or *dwellings*—represent a starting point in which our decisions to improve energy demand are taken, specially because these are biased by the exercise of rights and privileges, which do not necessarily comply with inhabitant's comfort and satisfaction.

The study of energy performance of dwellings implies an accurate and appropriate representation of the building envelope—to faithfully evaluate energy flow pathways; of occupant behaviour—to capture activities and responses to adapt indoor conditions; and the characterisation of appliances and systems—to determine the magnitude of the energy demand. The representation of these components in housing energy models is obviously challenging, because of the multiple stimuli affecting the performance of each dwelling. To date, occupant behaviour has been poorly represented in housing energy models due to their unpredictable nature, as well as their high level of variability, which is found in practice; however, although its consideration conveys complexity, capturing behaviour is essential in the formulation of energy performance models, and more importantly, in the estimation of changes to the stock energy demand in the long term.

An exhaustive representation of the interplay among envelope, occupants and systems, is crucial towards the achievement of reliable modelling techniques and, consequently, the deployment of robust tools for testing decarbonisation strategies in the domestic sector. To abstract this interplay, housing energy models define their resolution (or granularity) in terms of the chosen units of observation, such as individuals, households, communities or regions; and their level of sophistication in terms of their ability to abstract and model energy flow pathways, evolving elements, adaptive behaviours, adoption of policies, and diffusion of technologies. As a result, the scope of models to study energy demand is vast and varied.

In the UK, the domestic sector has represented one third of the end-use energy demand, with approximately 80% of this attributable to space heating and hot water, in which most of this energy is supplied by fossil fuels. Reducing heating energy demand in UK dwellings has been the subject of much systematic investigation, and it is thus the most obvious objective towards the decarbonisation of the housing stock. Nevertheless, to achieve this, the primary solution might not necessarily be the *truncation* of such an energy demand (i.e. strong policies intended to cut consumption patterns, as well as technological improvements to improve energy

conservation in dwellings), but the *transition* to a more efficient energy performance of the stock (i.e. gradual modification of the consumption line, and changes in the way people use energy in their homes). To successfully elucidate the decarbonisation pathways in the UK agenda, it is vitally important to understand the relevance and usability of strategies designed to reduce energy demand, by improving modelling tools, by enhancing sources of information, and by incorporating decision-making processes.

This *Thesis* begins by describing the composition of the housing stock in the UK. It initially identifies the relevance of the domestic sector in the carbon footprint at national level. It also shows how interesting—yet complicated—can the study of domestic sector be, due to the combination of multiple elements that, together, give life and shape to the housing stock in the UK. Once the relevance of the domestic sector at national scale is identified, a summary of interventions is provided to define the context of the stock, as well as a description of models supporting these interventions. It then continues by reviewing these energy models, including the analysis of bases, assumptions, data sources and limitations involved in their development. This review aims to set the basis of a modelling platform, which considers the virtues of existing housing energy models, and focuses on those areas where the projects lack or are very general.

The *Thesis* then describes the development of an open-source platform for energy simulation, and its subsequent application in the development and testing of strategies to reduce energy demand in dwellings. The platform summarises the steps taken for its systematic development, which consist of the optimisation of a database to generate archetypes, the conformation of richer archetypes with more useful variables, the generation of volumetric archetypes, the dynamic simulation of archetypes and, finally, the systematic evaluation of results. Under this protocol, the foundations for developing scenarios are established in this work.

As a result, a set of preliminary scenarios are developed to measure the effect of very specific (and optimistic) changes in the variables that compose the archetypes. Then, a theoretical framework is provided to create a different set of scenarios, in which additional stimuli to represent processes of decision-making are included. Despite limited, the platform enables the possibility to include the consideration of multiple stimuli domains affecting these processes, such as social, educational, regulatory, financial and technological. Such a consideration is applied in an exploratory prototype to estimate energy-related interventions.

In summary, the development of a simulation platform provides the basis for improving the evaluation of variables in both models and sources of information. By highlighting and applying open-source and modularity, as foundations to the platform, not only the validation of algorithms is improved, but also the development of an aggregated data pool is boosted. To this end, attention is given to the data that have been used to describe the stock, and the way they have been modelled, identifying the most relevant attributes and influential elements that have defined the *archetypes* of the national housing stock. In this way, the platform develops to describe the carbon intensity transition with an emphasis on the underlying drivers that define the composition of the residential sector.

This *Thesis* summarises the relevant elements that must be considered to improve robustness in the development of strategies to improve energy conservation and to reduce energy demand in dwellings. It also identifies those points that should be considered promptly in the future, and concludes with the vitality of this project for the participation of multiple disciplines and the use of new paradigms of work that today are paramount.

Keywords : residential stock, decarbonisation,
dynamic simulation, decision-making,
modularity, open-source, policy support

Acknowledgements

I would like to express my thanks and gratitude to my supervisors. Thank you Prof Darren Robinson for your insightful guidance, motivation, understanding, and patience from the commencement of the research. Thank you Dr Benjamin Jones for your continued support, mentorship, and valuable questions throughout the research. Thank you Dr Parham Mirzaei for your guidance and constant encouragement. Thank you all for your warm and involved way of supervising, and your interesting contributions to my work during all these years.
A particular thanks goes to Dr Peer-Olaf Siebers for giving me so valuable advice and providing me with such detailed criticism.
I owe thanks to Dr Jason Palmer for all the time taken to read my work, and to patiently provide me with feedback and questions on my project.
Thank you Kunpeng Wang for providing me with such a valuable tool.
I have had the pleasure of being around the people in the Lenton Hurst rooms, from which I have learned much. You have taught me a lot.
There are many others who have contributed to this fascinating experience, to whom I owe a debt of gratitude. I still believe that the most useful advice may come from the person you expect the least, if you listen.
Thanks to my family for being there all the time.
Thanks to my wonderful wife for her love and support throughout.
Gracias a todos.

Nottingham, 21 August 2017

G. S.

Contents

Abstract **iii**

Acknowledgements **v**

List of Figures **xi**

List of Tables **xiii**

List of Abbreviations **xv**

Publications **xix**

1 Introduction **1**
1.1 Context of Research 1
1.2 Research Questions, Aim and Objectives 3
1.3 Thesis outline 6

I Literature Review **9**

2 The British Housing Stock **11**
2.1 Composition and Evolution of the British Housing Stock 11
2.1.1 Structure of the Stock 11
2.1.2 Energy Demand and Environmental Contribution 19
2.2 Policies, Programmes and Assessments 23
2.3 Sources of Information 27
2.3.1 Scope and Quantity 27
2.3.2 Reductions and Imputations: Common Assumptions in Stock Data 30
2.4 Typologies of the Housing Stock 32
2.5 Conclusion of *the British Housing Stock* 41

3 Housing Stock Energy Models **43**
3.1 Modelling Rationale 43
3.1.1 Modelling Domains in Building Energy Models 46
3.2 Development of Housing Stock Energy Models in the UK 48
3.2.1 Classification and Use of Energy Models 48
3.2.2 Overview of Housing Stock Energy Models in the UK 52
3.3 The Conformation of Housing Stock Energy Models 65
3.3.1 Building Geometry and Fabric 65
3.3.2 Systems Components and Inter-Relationships 67
3.3.3 Occupant Behaviour in Housing Stock Energy Models 70
3.4 Evaluation of the Housing Stock Energy Model 72
3.4.1 Evaluation of Functionality 72

3.4.2 Anomalies and Omissions . . . 77
3.4.3 Future Functionality . . . 78
3.5 Conclusion of *Housing Stock Energy Models* . . . 79

II Development and Evaluation 81

4 Composition and Development of the Simulation Platform: EnHub 83
4.1 Simulation Platform Architecture . . . 84
4.1.1 Housing Stock Energy Hub (EnHub) Rationale . . . 84
4.1.2 Abstraction and Design . . . 87
4.1.3 `EnergyPlus` Modules . . . 94
4.1.4 Enriched Volumetric Generation . . . 97
4.2 Additional Functionalities of the Platform . . . 108
4.2.1 Verification, Validation and Calibration of the EnHub Platform . . . 108
4.2.2 Sensitivity Analysis and Uncertainties . . . 108
4.2.3 High Performance Computing and Cloud Computing . . . 113
4.3 Housing Stock Energy Simulation . . . 115
4.3.1 Similar Energy Performance Studies . . . 117
4.3.2 Energy Performance Indicators . . . 119
4.4 Conclusion of *Composition and Development of the Simulation Platform* . . . 120

5 Workflow for Efficient Housing Stock Representation 123
5.1 Methods: Statistical Analysis and Engineering Models . . . 124
5.1.1 Step a_1 Mining . . . 125
5.1.2 Step a_2 Reduction (or Sampling) . . . 136
5.1.3 Step b_1 Statistical Reduction . . . 142
5.1.4 Step b_2 Enrichment and Calibration . . . 143
5.1.5 Step $a-b_3$ Construction of Archetypes . . . 146
5.1.6 Step $a-b_4$ Dynamic Simulation . . . 147
5.1.7 Step $a-b_5$ Iterative Analysis . . . 149
5.2 Scalability of the Platform . . . 149
5.3 Conclusion of *Workflow for Efficient Housing Stock Representation* . . . 150

III Application 151

6 Analysis of Strategies to Decarbonise the British Housing Stock 153
6.1 Definition of Scenarios . . . 154
6.2 Evaluation of Scenarios . . . 157
6.2.1 Results and Findings . . . 157
6.2.2 Decarbonisation Strategies and Implications . . . 162
6.3 Further Evaluation: Household Decision-Making . . . 162
6.3.1 Triggers of Decision-Making . . . 162
6.3.2 Agency in Housing Stock Energy Models . . . 164
6.3.3 Network Environment . . . 169
6.3.4 Agent-based Uptakes . . . 171
6.3.5 Agent-based Outcomes, Discussion and Further Directions . . . 174
6.4 Conclusion of *Analysis of Strategies to Decarbonise the British Housing Stock* . . . 178

IV Closure **181**

7 Summary and General Conclusions **183**
7.1 Summary and Research Questions 183
7.1.1 Results and Findings 185
7.1.2 Description of the Possible and Probable Scenarios 186
7.2 Discussions and Reflections 187
7.3 Contribution to Knowledge 190
7.4 Directions for Future Research 191

A Standards and Regulations **193**

B Statistical Concepts **197**

C Modelling Process: Composition and Simulation **203**

D Unified Modelling Language **209**

E Glossary **219**

References **223**

List of Figures

1.1 Abstraction of components shaping domestic energy demand 3
1.2 Housing stock energy framework: bottom-up and top-down approaches 7

2.1 Housing stock growth comparison 12
2.2 UK population growth: people and households 13
2.3 Evolution of the UK housing stock 14
2.4 Registered average temperatures in the domestic sector 14
2.5 Ownership of heating systems among the UK housing stock 15
2.6 Evolution of household appliances 16
2.7 UK energy flow chart 2011 in million tonnes of oil equivalent 17
2.8 UK energy demand by sector, mix and services 18
2.9 Change in domestic final energy demand by fuel supply and demand services per year 20
2.10 Fuel supply by demand services in 2011 EPE: a zoom into Fig 2.7 domestic end-use 20
2.11 Historic domestic energy demand per year 21
2.12 Historic carbon emissions in the UK 22
2.13 Development of policies, programmes and assessments in the UK 23
2.14 Diagrammatic representation of the Milton Keynes projects 24
2.15 Simplified weighting process of nationwide surveys 30
2.16 Building typologies in the British housing stock 33
2.17 Building epochs in the British housing stock 34
2.18 Composition of dwellings by type and age of construction 35
2.19 Energy intensity by typologies 35
2.20 Energy performance indicators: type and age 36
2.21 Carbon emissions CDF by type and age of construction 37
2.22 Differential energy intensity CDF by heating systems 37
2.23 Mean annual temperature with degree days map overlay 39

3.1 Abstraction: how reality is seen through a model 43
3.2 Usability and reproducibility of multi-layered models 44
3.3 Balance between both granularity and sophistication on modelling techniques . . 44
3.4 Uncertainty expectancy by level of granularity 45
3.5 Classification of energy performance modelling: top-down and bottom-up 49
3.6 Classification of analytic approaches 50
3.7 Top-down and bottom-up approaches in domestic energy modelling 50
3.8 Classification of domestic energy modelling approaches 50
3.9 Evolution of Housing Stock Energy Models (HSEMs) in the UK 51
3.10 Diagrammatic representation of HSEMs 57
3.11 Energy flow pathways 64
3.12 Schematic illustration of the energy balance principle used by the BREDEM . . 65
3.13 Typical heat loss in dwellings 67
3.14 Simplified energy flow pathways in dwellings 68

4.1 Cuboid abstraction: representation of contiguous blocks . . . 87
4.2 Volumetric representation of dwellings . . . 88
4.3 EnHub architecture of the modelling platform . . . 93
4.4 `EnergyPlus` program schematic . . . 94
4.5 Heat balance control volume diagram . . . 96
4.6 EHS regions with EPW reference . . . 100
4.7 Average 24-hour electricity use profile for owner-occupied homes . . . 100
4.8 Example of associated EnergyPlus Input Data Files profiles . . . 101
4.9 Example of a volumetric cuboid for simulation . . . 104
4.10 An engineer's vision of UA, SA . . . 110
4.11 Examples of the SA module: local and global . . . 111
4.12 Examples of the SA module: screening method . . . 112
4.13 Simulation Flowchart: coupling `EnergyPlus` and `R` . . . 115
4.14 Distribution of energy demand by typologies . . . 117
4.15 Relation between energy demand and indoor temperature by typologies . . . 118
4.16 Relation between gas energy demand and external temperature for one archetype 120

5.1 Conceptual strategies to reduce the housing stock population . . . 124
5.2 The EnHub structure for the analysis of scenarios . . . 125
5.3 Linear model relationships . . . 126
5.4 FPCA visualisation . . . 127
5.5 Conditioned LHS iterations . . . 139
5.6 Comparison between the EHS and the reduced dataset . . . 141
5.7 Comparison between the EHS and the synthetic dataset . . . 145
5.8 Simulation and analysis flowchart: development of scenarios . . . 148

6.1 Summarised energy demand reduction . . . 158
6.2 Reduction of energy by typologies (stock indicator) . . . 158
6.3 Typological comparison in energy demand by tenure . . . 159
6.4 Variation in energy intensity and indoor temperature . . . 159
6.5 Estimated heat energy intensity . . . 160
6.6 Estimated carbon emissions . . . 160
6.7 Agent state machine diagram . . . 166
6.8 Structure of an agent-based model meta-diagram . . . 166
6.9 Weight of attributes for the four strategies present in the stock of dwellings . . . 167
6.10 Adoption-decision model . . . 168
6.11 Map of households represented as agents . . . 173
6.12 Differential adoption by strategies, in the emulated ABM interventions . . . 175
6.13 Differential adoption by weights, in the emulated ABM interventions . . . 176
6.14 Uptakes in the ABM emulated interventions . . . 177

List of Tables

2.1 UK countries population in 2011: people, households and dwellings 13
2.2 Five-year budgets to revise the decarbonisation process in the UK 22
2.3 UK regional classification in 2011 28
2.4 Summary of surveys and datasets 31
2.5 Energy summary estimates by age and type of construction 38
2.6 Key Statistics of the 2011 UK Housing Stock 40

3.1 Positive and negative attributes of housing energy modelling approaches 73
3.2 Evaluation of models according to the level of transparency as a function of accuracy, computability, reproducibility, sensitivity, adaptability and usability . . 76

4.1 Stock typologies represented by cuboids 89
4.2 Climate data summary for the EHS regions 99
4.3 Building properties for England and Wales 103
4.4 Fifteen high-level heating system types, used in the modelling process 106
4.5 `EnergyPlus` typology-specific components with UML reference 107
4.6 Energy performance summary by age and type of construction, using dynamic simulation 116

5.1 GLM rating on BREDEM 129
5.2 Reduction variables 137
5.3 Kruskal-Wallis Test (KWt) summary 140

6.1 Opted scenarios for perfect and conditional uptakes 154
6.2 Scenarios results: the potential of interventions 157
6.3 Network topologies: different configurations based on their connectedness 170

7.1 Improvements for Housing Stock Energy Models 187

List of Abbreviations

ABM Agent Based Modelling

ADEPT Annual Delivered Energy, Price and Temperature

AHOME Agent Home Owner Model of Energy

ANOVA ANalysis Of VAriances

ASHRAE American Society of Heating, Refrigerating, and Air-Conditioning Engineers

BRE Building Research Establishment

BREDEM Building Research Establishment Domestic Energy Model

BREEAM Building Research Establishment Environmental Assessment Methodology

BREHOMES Building Research Establishment's Housing Model for Energy Studies

CCC Committee on Climate Change

CCP Climate Change Programme

CDEM Community Domestic Energy Model

CDF Cumulative Distribution Function

CHM Cambridge Housing Model

CIBSE Chartered Institution Building Services Engineers

CLI Command Language Interpreter

CTF Conduction Transfer Function

DCLG Department for Communities and Local Government

DDM Domestic Dwelling Model

DECADE Domestic Equipment and Carbon Dioxide Emissions

DeCarb Domestic Energy and Carbon Model

DECC Department of Energy and Climate Change

DECM Domestic Energy and Carbon Model

DECoRuM Domestic Energy, Carbon Counting and Carbon Reduction model

DEFF Domestic Energy Fact File

DEFRA Department for Environment, Food and Rural Affairs

DEMScot Domestic Energy Model for Scotland

DHW Domestic Hot Water

DOMVENT Domestic Ventilation Model

DUKES Digest of UK Energy Statistics

ECO Energy Company Obligation

ECUK Energy Consumption in the UK

EDEM (Energy Systems Research Unit) Domestic Energy Model

EDM Energy Demand Model

EEP Energy and Environment Prediction

EFUS Energy Follow Up Survey

eHS equivalised Household Size

EHS English Housing Survey

EnHub Housing Stock Energy Hub

EPC Energy Performance Certificate

EPE Equivalent Primary Energy

EPW EnergyPlus Weather

EST Energy Saving Trust

E3MG Energy Economy Environment Model at the Global Level

EUI Energy Use Intensity

FIT Feed-In Tariffs

FMI Functional Mock-up Interface

FPCA Focused Principal Component Analysis

GHG Greenhouse Gases

GIS Geographical Information System

GLM Generalised Linear Model

GOR Government Offices for the Regions

GUI Graphical User Interface

HEED Home Energy Efficiency Database

HIDEEM Health Impact of Domestic Energy Efficiency Measures

HPC High Performance Computing

HSEM Housing Stock Energy Model

HVAC Heating, Ventilation and Air Conditioning

IEA International Energy Agency

idf EnergyPlus Input Data Files

KWt Kruskal-Wallis Test

LA Local Authority

LARA Local Area Resource Analysis

LHS Latin Hypercube Sampling

MARKAL Market Allocation

MCMC Markov-chain Monte Carlo

MDM Cambridge Multisectoral Dynamic Model of the British Economy

MKECI Milton Keynes Energy Cost Index

MTP Market Transformation Programme

NEED National Energy Efficiency Data-framework

NHER National Home Energy Rating Scheme

NHM National Housing Model

No-MASS Nottingham Multi-Agent Stochastic Simulation

NIHCS Northern Ireland House Condition Survey

NUTS Nomenclature of Territorial Units for Statistics

OAC Output Area Classification

OAT One-at-a-time

ONS Office for National Statistics

OOM Object-Oriented Modelling

OS Operating System

PAF Postcode Address File

PCA Principal Component Analysis

RdSAP Reduced Standard Assessment Procedure

RHI Renewable Heat Incentive

SA Sensitivity Analysis

SAP Standard Assessment Procedure

SDLT Stamp Duty Land Tax

SHCS Scottish House Condition Survey

SOA Super Output Area

STEP Seasonal Temperature Energy Price

SUSDEM Stochastic Urban Scale - Domestic Energy Model

TARBASE Technology Assessment for Radically improving the Built Asset base Domestic Model

TPB theory of planned behaviour

TRY Test Reference Year

UA Uncertainty Analysis

UK United Kingdom of Great Britain and Northern Ireland

UKDCM UK Carbon Domestic Model

UML Unified Modelling Language

UNCED United Nations Conference on Environment and Development

UoN-HPC University of Nottingham High Performance Computing Facility

Eng England

Sct Scotland

Wal Wales

NI Northern Ireland

List of Publications

Journal Papers

+ G. Sousa, B.M. Jones, P.A. Mirzaei, D. Robinson, A review and critique of UK housing stock energy models, modelling approaches and data sources (2017), Energy and Buildings, 157, 66–80.

+ Decision-makers in the UK housing stock decarbonisation (2017), Energy Policy - In preparation

+ G. Sousa, B.M. Jones, P.A. Mirzaei, D. Robinson, An open-source simulation platform to support the decarbonisation of the UK housing stock (2017), Energy and Buildings - In review

Conference Proceedings

+ G. Sousa, B.M. Jones, P.A. Mirzaei, D. Robinson, Pivotal Platform of the Housing Stock Energy Models in the UK to evaluate decarbonisation scenarios (2017), Building Simulation 2017.

1 Introduction

1.1 Context of Research

The domestic sector and its contribution to the national energy demand

The UK housing stock is responsible for 27% of national energy demand attributable to 4 key services: 60% space heating, 20% Domestic Hot Water (DHW), 17% lighting and appliances, and 3% cooking [60]. Satisfying an energy demand generally implies the emission of pollutants to the environment when combustion is involved, but is dependent on the fuel properties and the processes required to deliver that energy to the more than 27 million dwellings that comprise the UK housing stock. Gas is the most common supply of domestic energy (68%), which in the UK became prominent in the 1990s when electricity generation switched from using coal to natural gas. Electricity is also a prominent supply (24%) and is generated by nuclear, wind, and hydro-power plants; other fuels include petroleum, coal, coke and breeze, and with minor share solid fuels such as bio-energy and waste [61]. Around two-fifths of the emissions attributed to housing demand were caused by generation of energy in power stations and the remainder by end-use emissions [58]. This end-use is shaped by the needs of individual households, which in turn are a function of their socio-demographic characteristics, associated activities and context of each dwelling.

The housing stock is mainly composed of houses (80%), however, apartments are increasing at a substantial rate [179]. This change is due to population growth, but also—and perhaps most likely—due to the adoption of lifestyles and collective aspirations, that have fragmented the composition of households. More than half of the existing stock was built before the 1970s, when building energy regulations became part of the agenda in the UK [196]. Nevertheless, whilst more recent dwellings have improved in their energy performance, changes on lifestyles have counteracted this trend [203]. For instance, even though fabric insulation has improved the heat transfer in dwellings (i.e. less losses), the chosen temperatures of comfort have increased heating energy demand. Likewise, although the efficiency of household appliances has increased during the last years, its energy share has grown due to the rebound effect that increases their ownership and so their usage intensity [51, 176, 180].

The challenge of carbon intensity transition

Although it is possible to define repeating typologies of dwellings, each dwelling is nonetheless unique, as is its demand for energy, and the associated consequences for the environment. These energy demands are caused by the actions taken to improve health and well-being, by the activities of the inhabitants, and by the use they make of household appliances. It is crucial to develop frameworks able to fully encapsulate the energy flow pathways in dwellings; to characterise the factors influencing them; and to explain the emerging connections from such these activities taken place in dwellings.

The frameworks need to consider the quality of data and models describing the interplay of these factors; and they need to accurately capture the composition and evolution of the *dwellings*, despite their uniqueness. Hence, the representation of such a phenomena requires further efforts in multiple domains [36, 65]. This is a challenging task, but it is essential in the formulation of robust decarbonisation policies and transition strategies.

The role of occupant behaviour in domestic energy performance

The main purpose of a dwelling is to provide shelter from the weather, but it also provides security; visual, thermal, acoustic comfort; and privacy. Accordingly, indoor environment quality affects the physical and mental health of occupants.

The role of inhabitants, either as individuals entities or as collective units (i.e. units forming part of a community), becomes fundamental to the representation of the decisions taken to improve or deteriorate indoor conditions. Inhabitants directly shape the energy flow pathways in dwellings [145], yet these pathways are highly dependent on the physical properties of the buildings and the set of appliances contained within them (i.e. devices, goods and objects). Because of this, it is important to have a faithful understanding of the stimuli affecting inhabitants' decision-making; its representation is necessary to understand uncertainty in the analysis of strategies to reduce energy demand and, therefore, to decarbonise the housing stock.

A fit for purpose platform

To improve upon the representation of domestic energy performance, a paradigm shift is required in the way modelling strategies and data sources are designed, used and shared [164, 131]. Open modularity and active participation become two pillars in the development of this shift. They add flexibility, adaptability and resilience to frameworks and platforms describing domestic energy performance. Only in this way can triggers of decision-making be linked with modelling strategies that represent energy flow pathways, and changes in the composition of the housing stock be explicitly evaluated.

To study the processes of decarbonisation of the UK housing stock, a fit-for-purpose platform should be deployed. Based on these pillars of modularity and participation, a platform must thus be bound to draw together the underlying elements that affect the demands of energy in dwellings, by coupling triggers of decision-making with energy flow pathways, and by measuring uncertainties that arise from models' sophistications and data granularities [128, 74].

The purpose of this platform consists of parsimoniously evaluating properties of elements (occupants, fabric and appliances), connections between elements (technology usage, dynamic actions) and interactions between them. To date, most models describing housing energy performance and comfort are developed in *isolation*; however, to support the application of policies and programmes destined to reduce energy demand in the housing stock, models need to be prepared to process dynamic stimuli and evolving up-to-date information. In consequence, the analysis of energy-related intervention in the stock should be benefitted by the development of more accurate energy performance indicators, and ubiquitous data.

1.2 Research Questions, Aim and Objectives

The study of energy demands in dwellings has been addressed as a function of the physicality of dwellings (fabric and context), their occupancy (demographics and activities) and appliances (their ownership and efficiency). This abstraction, represented in Figure 1.1, is commonly evaluated, by Housing Stock Energy Models (HSEMs), in terms of the resulting energy demand and its associated carbon intensity.

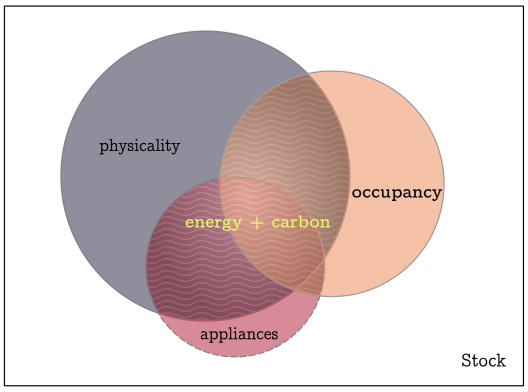


Figure 1.1: Abstraction of components shaping domestic energy demand

Ideally, a HSEM should be able to explicitly represent each component of this interplay, and how it is impacted upon by policy, market conditions and available technology. But some unavoidable limitations have historically narrowed their scope, such as levels of computation, resources to make surveys, resolution of the data, and sophistication of algorithms applied to these tools. Currently, HSEMs have partially addressed the aforementioned interplay, that leaves gaps in the way the stock has been described. For example, due to the composition of the end-use energy share in the stock, most HSEMs have improved upon the representation of heat gains on domestic buildings. This explains (and justifies) the numerous techniques developed to describe thermal phenomena; though, a number of simplifications regarding other contributors to the energy share remain unaddressed, which conveys a significant degree of uncertainty, and analytical bias, in the outcomes derived from these tools. Similarly, in order to represent household appliances and occupant behaviour, the modelling tools require a high level of granularity and fidelity in the sources of information, which is difficult—and sometimes impossible—to achieve, and so a considerable bias is introduced.

The core of this work consists of evaluating the components affecting energy demand in dwellings and, consequently, analysing strategies to reduce their associated carbon intensity. To this end, it is required to gain insights on the characterisation of each of these components, and on the description of algorithms employed to represent their relationships. Therefore, this Thesis addresses the following key research questions:

Q-1 Which **infrastructure parameters** are the most important for characterising dwelling typologies, and for modelling energy demand?

Q-2 How appropriate is the resolution of **domestic stock surveys** and datasets for the description of the UK housing stock and national carbon intensity?

Q-3 In what ways can the datasets be **optimised to analyse strategies** to reduce energy demand, and how well is the loss of information (or propagation of uncertainties), in derived analyses, tracked?

Q-4 Which are the main elements for the validation of **tools to estimate energy performance** and associated carbon emissions, and how can these be improved?

Q-5 How would it be possible to represent policy outcomes, as well as energy conservation measures, through different abstractions regarding **curtailment behaviours** (better practices) and **investment behaviours** (technological improvements)?

Q-6 How should the effectiveness of policies and **strategies**, addressing these behaviours, be validated?

This work will focus on the reconciliation of existing modelling tools, by adding emphasis on the sensitivity of both variables and algorithms, not only to better characterise physical components, but also to reconcile triggers of decision-making that influence occupants to use more or less energy in dwellings. Additionally—and as a direct consequence of this focus, the work will provide basis for testing further interactions as well as more detailed behaviours, where household-specific stimuli can be tested. This ability to represent dynamic interactions, defined as agency, will enable the evaluation of unique responses of householders and their decisions to improve energy performance.

In summary, this work aims to be a platform of models, where the virtues of existing HSEMs are encapsulated in a modular and standard way, to increase their scope, and to enhance the analysis of decarbonisation strategies. A formal definition of the aim of this work is then stated as:

> "To develop a modular and dynamic simulation platform of the UK housing stock, in order to gain a better understanding of inhabitants decision-making, from responses towards comfort satisfaction to investments into better performing environments. The simulation platform needs to emulate decarbonisation transition, to perform a methodological assessment of uncertainties, and to enhance disaggregated end-use level metrics."

To add clarity, the terms employed in the statement are explained below:

- **UK housing stock with its typology** - an in-depth review of the UK typology will minimise redundancy in the components describing the stock;
- **dynamic simulation platform** - by emulating indoor activities, interactions and responses to the external environment, the resolution of the evaluation can be improved. To do this, the platform operates as a linked mechanism, in which an Object-Oriented Modelling rationale is suitable to link algorithms and data to perform (and validate) the evaluation;
- **modular** - by encouraging the development of transparent and accessible components in modelling approaches, these can be straightforwardly validated and edited. In addition, a modular approach makes the components exchangeable and detachable within the platform;
- **methodological assessment of uncertainties** - a stable module that can test the sensitivity of the model and quantify uncertainties analysis is required;
- **disaggregated end-use level metrics** - by representing information at a disaggregated level, the decision-making triggers can be improved (e.g. more granular, updated and accurate) and dwelling-specific constrains affecting the decarbonisation transition can be identified;
- **inhabitants' decision-making** - by studying stimuli affecting decisions that influence energy demand in dwellings, improved algorithms can be developed. This is as an attempt to include a human-in-the-loop approach;
- **performing environments** - by studying specific metrics and by evaluating scenarios, the quality of tools designed to formulate and apply interventions in the housing stock is increased, and the energy performance in dwellings is better characterised;
- **comfort satisfaction** - by providing household specific metrics, the ability to predict particular changes increases, emulating improved conditions for occupants such as improved thermal conditions, and optimal air renovations; and
- **decarbonisation transition effectiveness** - by developing a modular platform, a large set of variables can be effectively tested to evaluate the potential of interventions to reduce carbon emissions, taking into consideration changes to components and occupant behaviours.

More specifically, the objectives and potential contributions of this work are summarised as following:

1. To develop a HSEM platform focused on evaluation and validation of components (data and algorithms).
2. To report the usability of available data at UK national scale and to highlight where there are gaps and flaws.
3. To provide a set of energy and comfort indices to provide major parameters to support the analysis of strategies at stock level.

4. To evaluate uncertainties and their propagation in the study of strategies designed to minimise energy use in dwellings.

5. To test the uppermost potential of decarbonisation interventions (regarding occupant's behaviours, building fabric and building systems), identifying where efforts should be prioritised (which measures in which typologies).

6. To provide a basis for a prototype framework which enables to possibility to predict the (probable) uptake of decarbonisation interventions by incorporating household-specific conditions, as well as external stimuli affecting household decision-making; a suitable candidate is found in an agent-based modelling approach, which is able to represent actions and interactions of autonomous elements, named agents.

7. To highlight the usefulness of modularity by tackling a set of diverse tasks on the same platform (sensitivity, archetype modelling, stock analysis and agency).

1.3 Thesis outline

Figure 1.2 represents an abstraction of the factors influencing housing stock energy performance and strategies for addressing them. This diagram is helpful for improving, upon the holistic representation of elements when developing a platform to test the effectiveness of decarbonisation strategies in the housing stock. Here, different approaches can be studied from a top-down perspective, where data is mostly based on contextual fluctuations (fuel, sales, local authorities estimations, etc.), but also from a bottom-up perspective, where detailed information, regarding physical components, impacts on the stock performance. This abstraction helps to understand the relationships (and dependencies) between the social and physical variables. Both physical attributes and contextual factors define the context of the residential stock, as a link between the tangible elements of the stock (building shapes, materials, fuels and all physical elements) and the organisational components (tenure, regulation, taxation). The contextual factors are connected with the behavioural elements by a regulatory framework, where actions may be included from an organisational perspective. The physical attributes and behavioural elements are linked as investments, which can be perceived as the long term adaptive actions taken by the occupants on the buildings. At the core of the framework is sub-group focussed on the study of the utility and functionality of the platform. This sub-group is used to calibrate models in their performance (diagnostic) and to augment accuracy in their prediction (prognostic).

Following this rationale and in satisfaction of the aims articulated above, the first part of this work starts with a **literature review** of the housing stock infrastructure in the UK, considering its contribution to the national energy demand and description of typologies, as well as the policies and regulations that have been applied so far to influence energy use in the housing stock. Chapter 2 reviews the availability, granularity and usability of data to characterise the housing stock. Chapter 3 evaluates and critiques existing modelling techniques. This evaluation is used to re-define where efforts should be made to improve these techniques, and to set the basis for developing a modular platform to analyse the components shaping the domestic energy demand and the strategies of decarbonisation.

The second part of this work describes the **development and evaluation** of a new platform to analyse housing stock decarbonisation strategies. The platform, based on modularity and

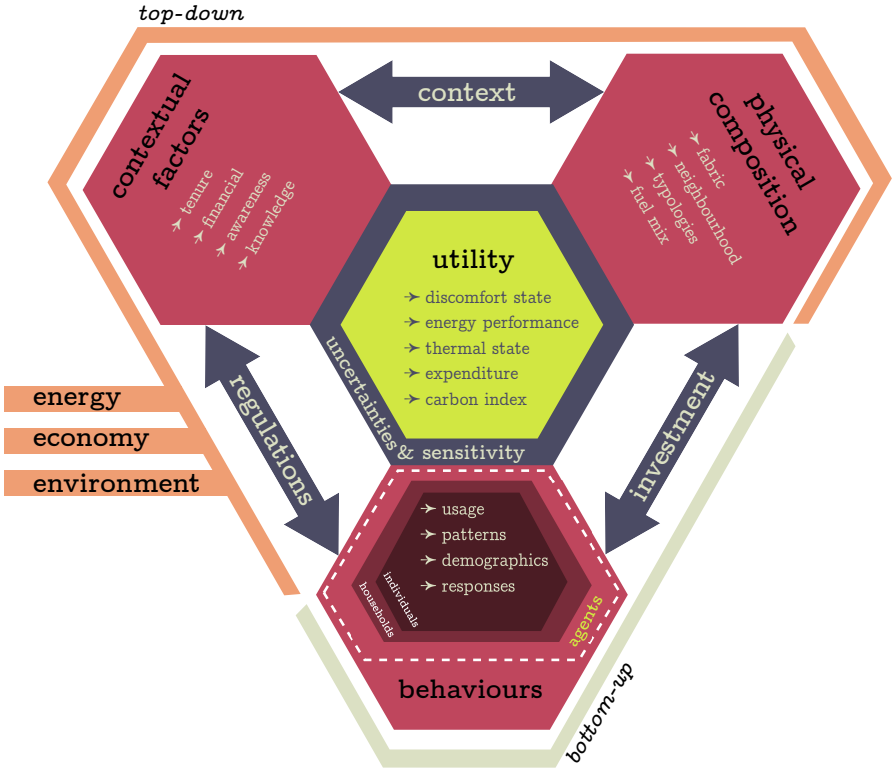


Figure 1.2: Housing stock energy framework: bottom-up and top-down approaches models. Four key elements: physical composition (tangible elements), contextual factors (organisational), behaviours (understanding) and utility (indicator of performance)

transparency, defines a standardised architecture to represent the properties of dwellings and households, and the algorithms employed to capture their relationships. A number of tests are described here to validate and improve the use of data and models. Chapters 4 and 5 describe the platform's ability to accurately track errors, to identify gaps in the modelling techniques, and to explore the utility of different techniques. By addressing this, the support for policies and programmes is increased considerably.

The third part describes the **application** of decarbonisation strategies to evaluate their potential to reduce stock energy demand. Chapter 6 describes and tests the potential of uppermost decarbonisation interventions in the stock, including those improvements related to decision-making. By including more specific information, some household constrains are included in the evaluation. This pragmatic consideration of constraints to uptakes is followed by the development of a conceptual framework for the agent-based modelling of households' decisions to modify their operational and investment behaviours in response to decarbonisation policies and strategies.

Chapter 7 concludes the thesis by highlighting some of the main findings of this research, and by elaborating on discussion points. It also reflects on future research in this area.

Literature Review Part I

2 The British Housing Stock

2.1 Composition and Evolution of the British Housing Stock

2.1.1 Structure of the Stock

As of 2011[1], the UK housing stock was comprised of 27.4 million dwellings (83.3% England, 5.1% Wales, 8.9% Scotland and 2.7% Northern Ireland; see Table 2.1). It has homogeneity in its physical features and can be divided into five groups: 28% terraced, 9% bungalow, 26% semi-detached, 17% detached, and 20% apartments [179] (see Figures 2.1:a and 2.3a). In the last forty years, the number of detached dwellings and apartments increased by 124% and 75%, respectively [179, 166], but this increase does not correlate well with the number of householders. The number of UK households is rising at around 1% per annum [14, 179], but the average household size of 2.2 people per household is gradually falling because of increased family fragmentation and reduced fertility rates (see Figures 2.1:b and 2.2). Concurrently, changes in the ownership and use of appliances and systems are causing fluctuations in dwelling energy demand [51, 176, 180], despite improvements in their efficiency[2]. The energy demand of some systems, especially those for heating, is correlated with the associated housing typologies [61, 229]. This is due to façade elements and construction technologies adopted in their period of construction.

Six age bands are traditionally used to characterise the UK housing stock: pre 1919, inter-war, post-war 1950s, industrial 1970s, modern 1980s and post 2000 (see Figure 2.1:c). More than one-third of the stock date from the first two bands, in which most of the constructions employ solid masonry (see Figure 2.1:d) and the floor area is significantly larger than the average size of 92 sqm [166]. The housing stock increased by 45% between 1970 and 2011, with a majority of these dwellings built with cavity walls. Changes in construction methods during this period were, in general, intended to reduce the thermal transmittance of the envelope as well as uncontrolled air leakage. The registered age of construction provides thus an indication of the energy performance

[1] This document uses 2011 data as a reference because: (1) it is the same year of the last Census; (2) the CHM, a reference HSEM for national documents, also uses 2011; and (3) Super Output Areas, the sample units of many HSEMs, are constructed for this year.

[2] The same phenomenon was pointed out more than a hundred years ago by Jevons [121], although specifically for coal. This has recently been termed the rebound effect [207] and is a relevant phenomenon that affects the energy performance at stock level.

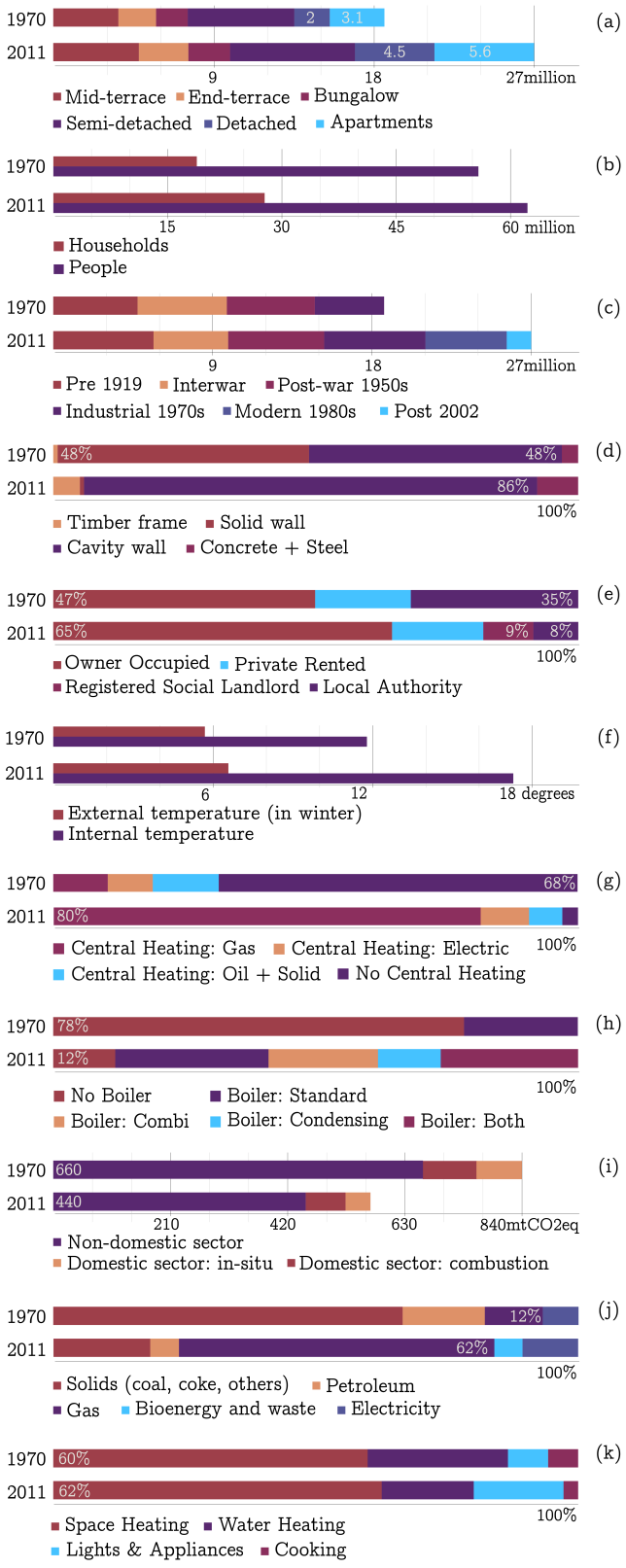


Figure 2.1: Housing stock growth comparison
Source: [61, 58]

Table 2.1: UK countries population in 2011: people, households and dwellings
Source: Office for National Statistics (ONS)–www.ons.gov.uk/

	People	Households	Dwellings
United Kingdom	62,055,838	27,651,734	27,580,884
Great Britain	60,267,499	26,903,499	26,832,836
England and Wales	55,071,113	24,429,618	24,359,880
England	52,059,931	23,044,097	22,976,066
Wales	3,011,182	1,385,521	1,383,814
Scotland	5,196,386	2,473,881	2,472,956
Northern Ireland	1,788,339	748,235	748,048

of the current housing stock, considering that energy conservation standards for dwellings were developed in response to the 1973 oil crisis and were maintained in the 1980s by the Housing Act [196, 220, 166].

In addition, the conformation of typologies within these bands is complemented by changes in the tenure composition. These changes reflected government (in particular monetarist) policies to encourage home ownership and the associated transfer of ownership of local authority housing [19]. Dwellings constructed during the 1950s and industrial periods were built in response to a restructuring society; they contained a range of previously uncommon amenities and services (such as community centres and shops), were built in large settlements of small properties under social housing tenure schemes, and relied heavily on the use of prefabricated components and lighter materials [196, 166]. During this period, an initiative to develop self-sufficient neighbourhoods was promoted, which increased social housing, satellite settlements, and development of apartments; but the result was not as successful as expected and a detour was made, which resulted in the decline of such housing schemes [196]. However, changes on the construction methods remained as part of the adopted building regulations. By 1970, 47% of properties were owner-occupied; by 2011, this tenure increased to 65% of the stock (see Figure 2.1:e and 2.3b) [179, 58]. As a result, not only these compositional changes affected services and energy usage, but also indoor conditions, although not always positively [203].

Since 1970, the mean external air temperature in winter has increased by around 1 K whereas the mean indoor air temperature has risen by around 6 K [61] (see Figures 2.1:f and 2.4). The increase in the external temperature can be associated with climatic change—also with heat island effects in urban areas, whereas the significant increase in the internal temperatures is attributable to improvements in fabric properties, the prevalence of central heating systems, an increase in

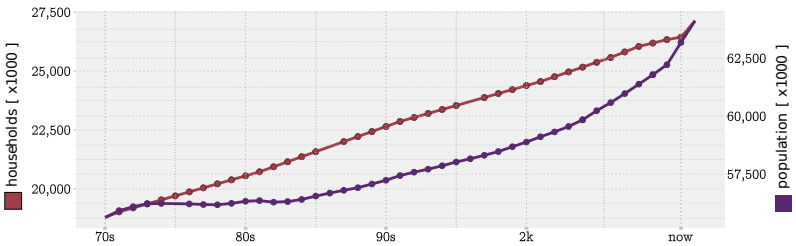
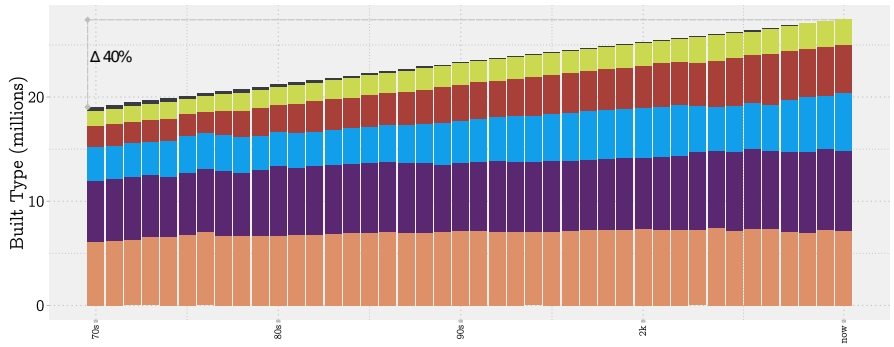
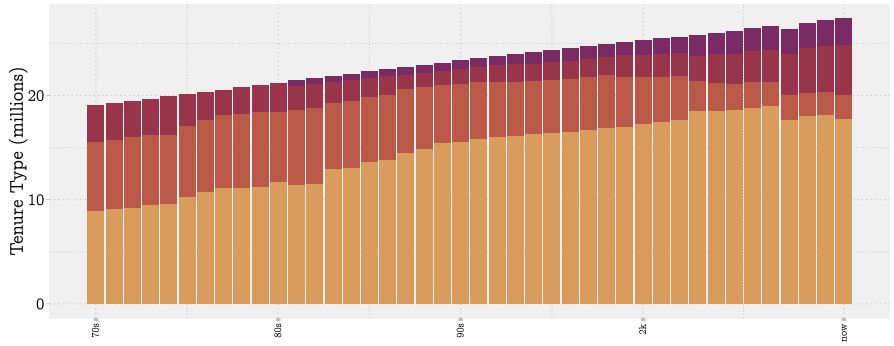


Figure 2.2: UK population growth: people and households
Source: [47]



(a) Building type



(b) Tenure type

Figure 2.3: Evolution of the UK housing stock
Source: [48]

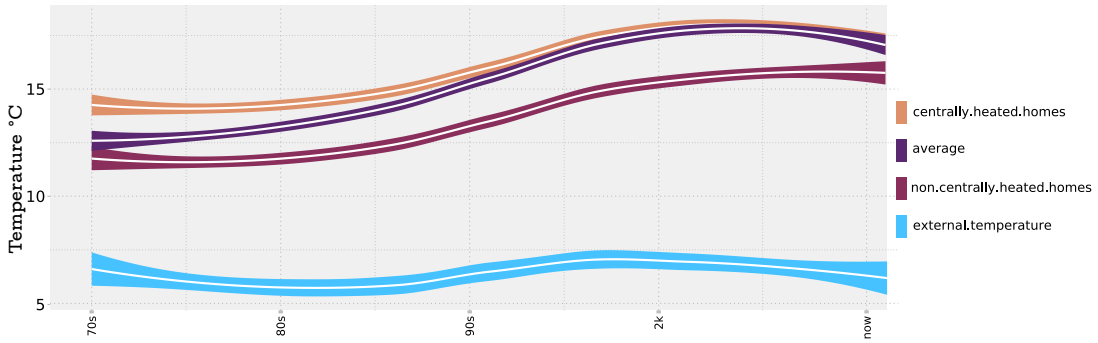
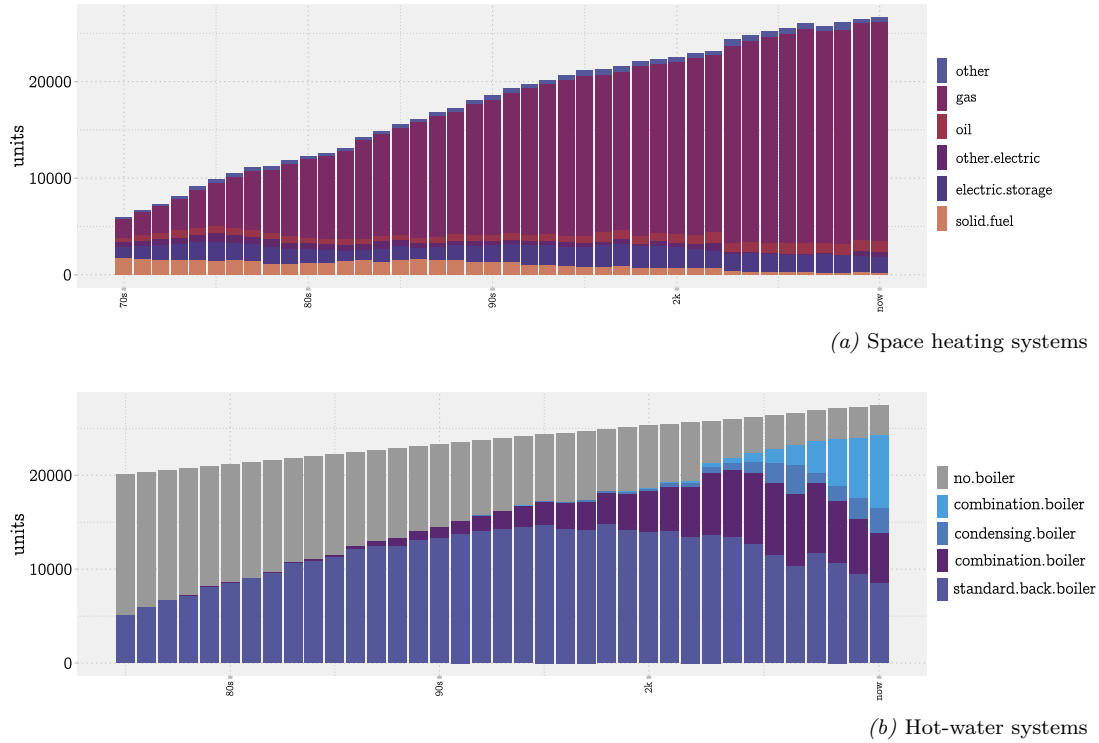


Figure 2.4: Registered average temperatures in the domestic sector
Source: [61]



(a) Space heating systems

(b) Hot-water systems

Figure 2.5: Ownership of heating systems among the UK housing stock *Source:* [47]

acceptable comfort temperatures [203], and more onerous regulations relating to the conservation of heat and power. Some of these changes have delivered improved indoor comfort and health co-benefits [85], but have not significantly reduced fuel bills nor lowered carbon emissions.

Domestic Space and Water Heating Systems

In 1970, 68% of dwellings had no central space-heating system. Gas-related systems were present in 10% of dwellings, solid fuel chimneys in 14%, and electrical storage in 8%. By 2011, gas accounted for 81% of space-heating, electrical storage for 6%, and solid fuels for 1%. Such an evolution is shown in Figure 2.5a. Other systems introduced by 2011 include electric ceiling heating, electric underfloor heating, communal heating, and combined heat and power units [60].

In 1970, 78% of the stock had no boiler at all, and only 22% had a standard back boiler. By 2011, back boilers comprised 31% of all water heating systems, condensing combination boilers 28%, combination-only boilers 19%, and condensing boilers 9%. The remaining 13% of dwellings had no boiler installed. Figure 2.5b show the increase in the ownership of boilers in the housing stock. Other less common water heating systems include dedicated boilers, electric immersion heaters and instantaneous heaters [70].

The combination of gas fuelled boilers with central heating systems are the most common system of their type found in UK dwellings. Therefore, highly disaggregated information about these systems should be included in datasets so that they can be modelled accurately by HSEMs.

Domestic Appliances

Since 1970, the ownership of household appliances and its aggregate energy demand has experienced continuous growth, having doubled in size [13, 175, 59, 179] (see Figure 2.6). Meanwhile, significant improvements have been made in appliance energy efficiency over the same period [174, 59]. Increased ownership is partly attributable to population growth, consumer-biased lifestyles and the rebound effect, where improvements in appliance efficiency have created a gap in dwelling energy demand that is filled by the acquisition of more appliances, or their more intense use [203, 228].

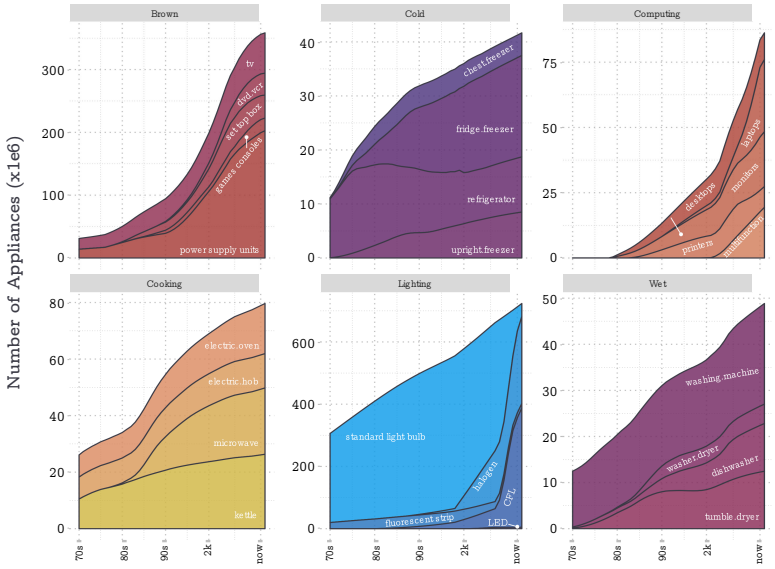


Figure 2.6: Evolution of household appliances
Source: [179]

Correspondingly, cold appliances (fridges, freezers) are responsible for 3.6% of overall domestic annual energy demand, wet appliances (washing machines, dishwashers) use 3.4%, lighting uses 2.9%, cooking (ovens, stoves, cookers) uses 3.1%, and brown electronics (TV, audio systems, PCs) use 4.9%. Cold and wet appliances may be associated with dwelling typologies because of their relationship with household size, and their overall energy demand can be described well by their energy efficiency labels. However, the energy demand for lighting, cooking, brown appliances, and to some extent wet appliances, is more related to users' activities and interactions [14, 230, 215], which is subject to stochasticity.

Characterising appliance energy demands is thus relevant for the analysis of drivers that affect decision-making. Nevertheless, the changes in the appliances ownership and related technologies are significantly volatile, and their prediction carries a wide range of uncertainties, due to the fact that some inclusions are not easy to predict. For instance, in the *40 % House* study [13], the expected evolution of appliances is significantly different than the current scenario; here, not only the introduction of more energy-efficient technologies affected the report's estimates, but also the typical ownership of appliances, where the number of small appliances increased considerably, modifying in consequence the energy intensity in the domestic sector [175].

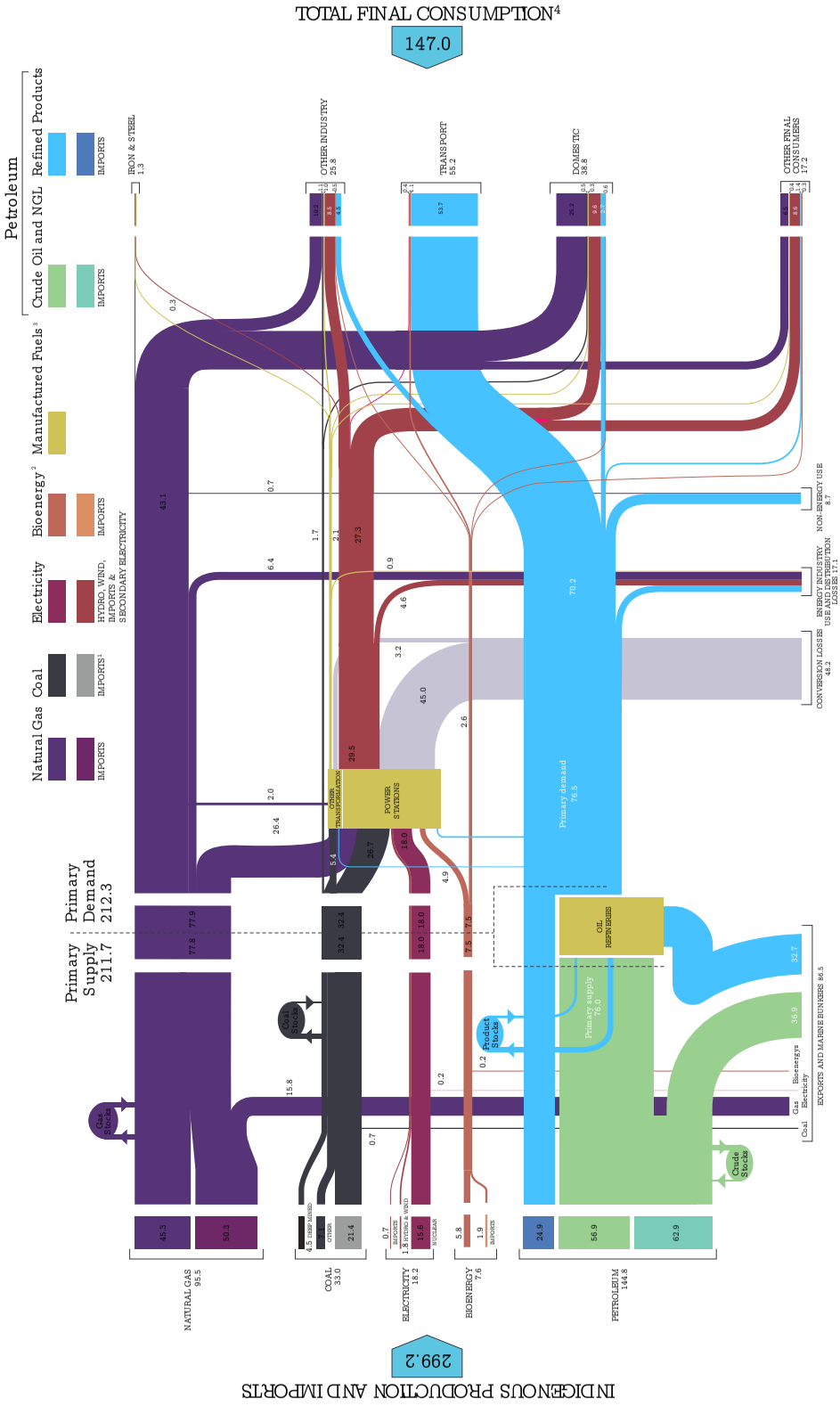
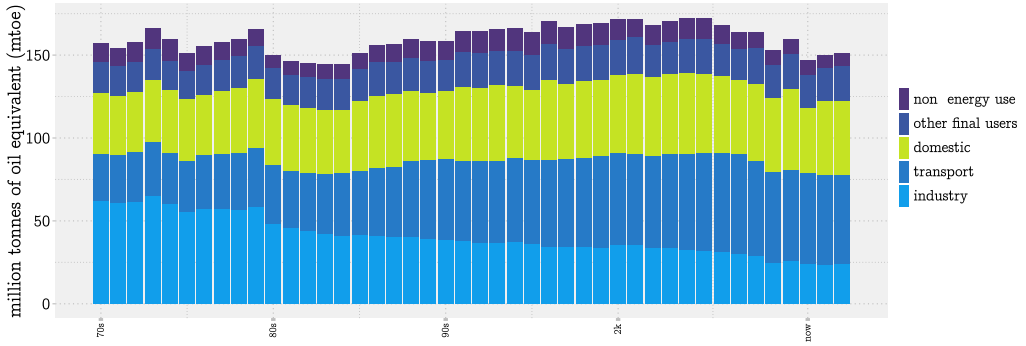
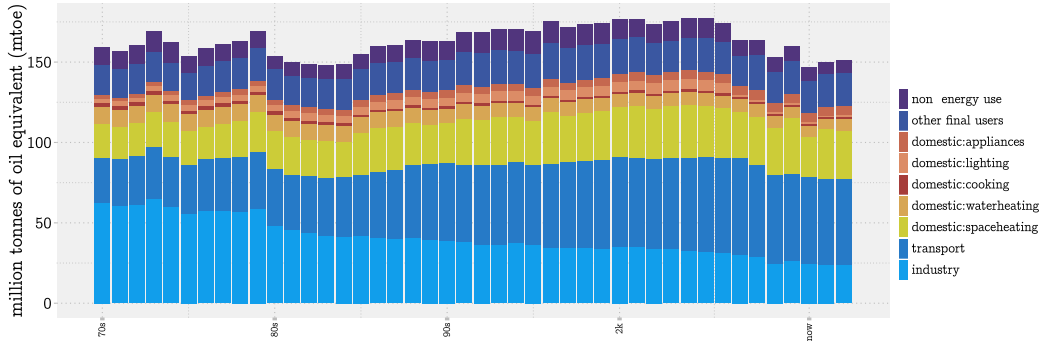


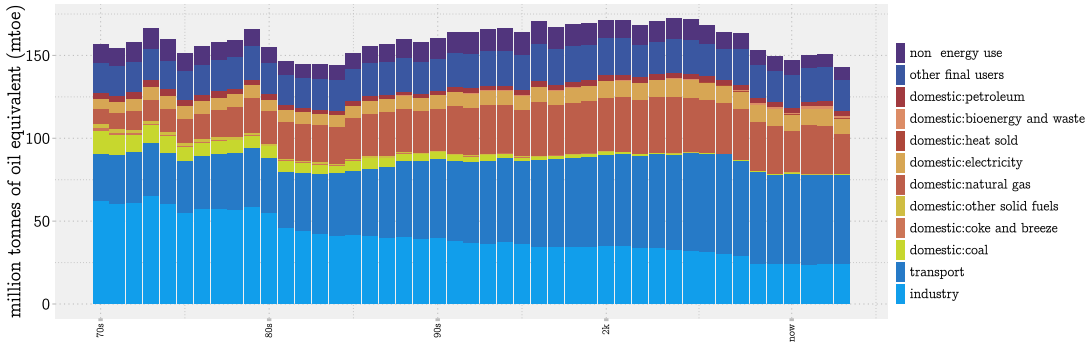
Figure 2.7: UK energy flow chart 2011 in million tonnes of oil equivalent [55]



(a) UK energy demand by sector



(b) UK energy demand by sector, highlighting domestic sector by services



(c) UK energy demand by sector, highlighting domestic sector by fuel

Figure 2.8: UK energy demand by sector, mix and services
Source: [47, 179]

Hence, it can be said that sources of information regarding household appliances should be constantly updated. In this way, not only the energy profiles could be improved but also the link with more specific socio-demographics data sources could be increased. This link provides an opportunity to fully understand the drivers affecting the use of energy in dwellings, and therefore their environmental contribution.

2.1.2 Energy Demand and Environmental Contribution

The UK emitted a total of 564 $MtCO_{2e}$ in 2011, which is 36% below the peak value registered in 1979 and 28% below that of 1990 [58]. The domestic sector contributed 124 $MtCO_{2e}$ of the total emissions (ca. 1650 TWh equivalents) [62]. Two-fifths of these were caused by generation of energy in power stations and the remainder by end-use emissions. Most of this demand was supplied by fossil fuels, as seen in the energy flows drawn in Figure 2.7. These energy flows, which in sum represent 27% of the total energy demand in the UK, have been maintained at similar levels during the last forty years. The estimation of these values is supported, at least, since formal records of energy levels started in the country (see Figure 2.8). The end-use demand is attributable to four key services: 60% space heating, 20% DHW, 17% lighting and appliances, and 3% cooking (see Figures 2.1:k and 2.9). Electricity is expectedly the main end-use supply for domestic appliances, which has maintained a growing rate in the last forty years.

Figures 2.1:g-h and 2.9 highlight the significant changes in installed space heating systems supplied by gas, and in the DHW technologies found in UK dwellings [60, 70]. Incidentally, gas is also the most common source of energy used by UK power stations to generate electricity and for district heating, especially since the 1990s, by the supply shift from coal to natural gas in power stations. This shift helped to reduce carbon emissions significantly (see Figure 2.1:i). Figure 2.10 shows that gas is, directly and indirectly, responsible for some 73% of all space heating, 79% of all DHW, 36% of all lighting and appliance energy, and 55% of all cooking energy [60]. Some houses in rural locations use petroleum or oil and Figure 2.10 shows that they satisfy 8% of total demand for space heating. Electricity represents 22% of the energy used for lights and appliances, 4% for space-heating, 5% for water heating and 15% for cooking. Figure 2.10 exemplifies that other supplies, in Equivalent Primary Energy (EPE), such as oil and solid fuels, including renewable biomass, are also relevant in the national supply share (see Figure 2.1:j).

By 2011, the UK electricity generation mix was 33% coal, 34% gas, 20% nuclear, 3% renewable, 1% oil and, the remaining, other fuels [61]. However, relevant changes have occurred during the last years. As shown in Figures 2.11a and 2.11b, there was a wholesale change in the amount of coal used to generate electricity, replacing it with natural gas in the 1990s. Furthermore, during this time, the inclusion of renewable sources began to rise. In 1990, the first formal record displays a value of 5 TWh of renewable energy for the domestic sector.

Not only gas became the main source of electricity generation, but also replaced oil, wood and coal in the generation of heat into dwellings. These two changes have significantly influenced the estimation of atmospheric pollutants. Despite the estimation of direct emissions have not been trivial, these changes in the fuel mix consequently reduced the estimated amounts of methane and NOx associated to these former sources of energy. Figures 2.11b and 2.12 show such a variation in the fuel mix and its environmental consequences.

Decarbonisation of the Housing Stock

The UK's Climate Change Act aims to reduce its 1990 Greenhouse Gases (GHG) emissions by 80 % by 2050 [183]. To this end, the Committee on Climate Change (CCC) has established a series of incremental targets (or budgets) for the whole energy sector, including the production of 30 % of electricity from renewable sources by 2020, and the reduction of GHG emissions by 50 %

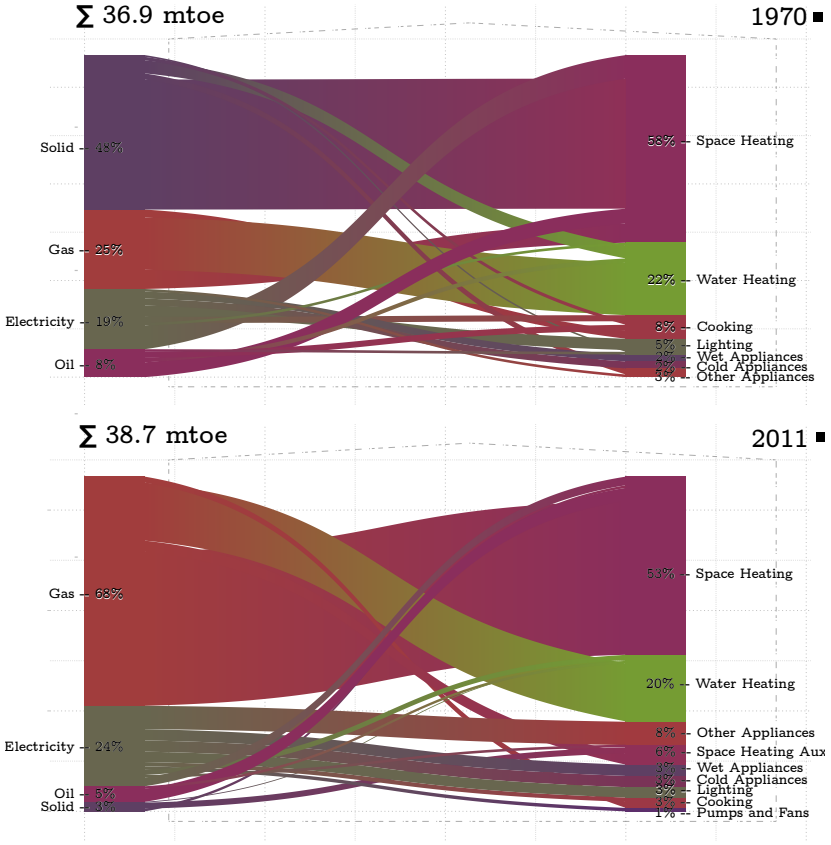


Figure 2.9: Change in domestic final energy demand by fuel supply and demand services per year
[61]

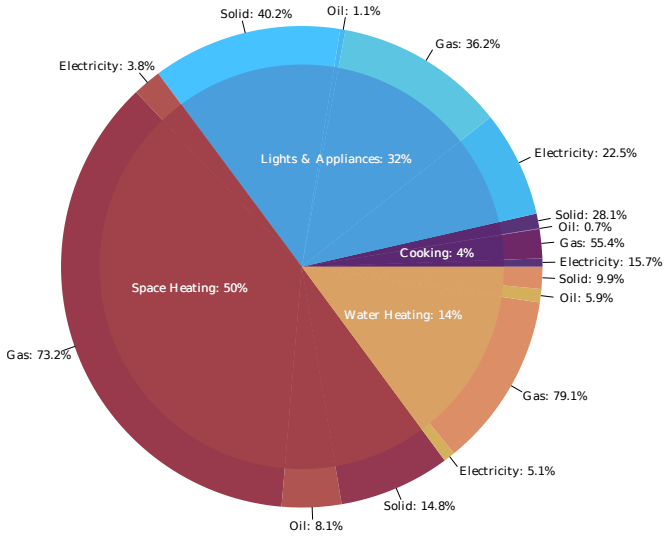
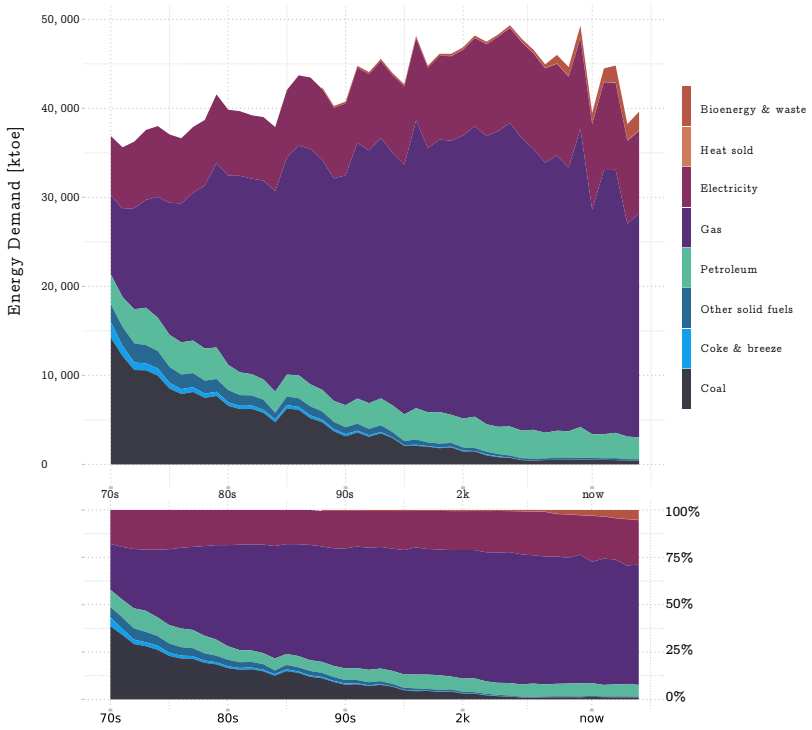
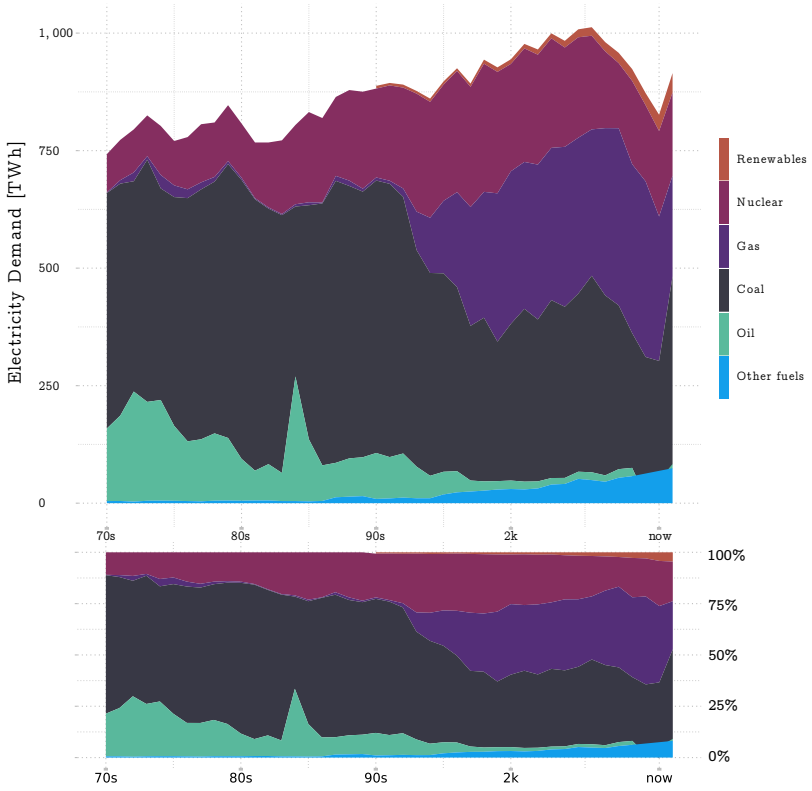


Figure 2.10: Fuel supply by demand services in 2011 EPE: a zoom into Fig 2.7 domestic end-use
[61]; derived from DUKES and CHM



(a) Energy demand by fuel contribution per year



(b) Electricity demand by fuel contribution per year

Figure 2.11: Historic domestic energy demand per year
Source: World Bank–databank.worldbank.org/data/ and [47]

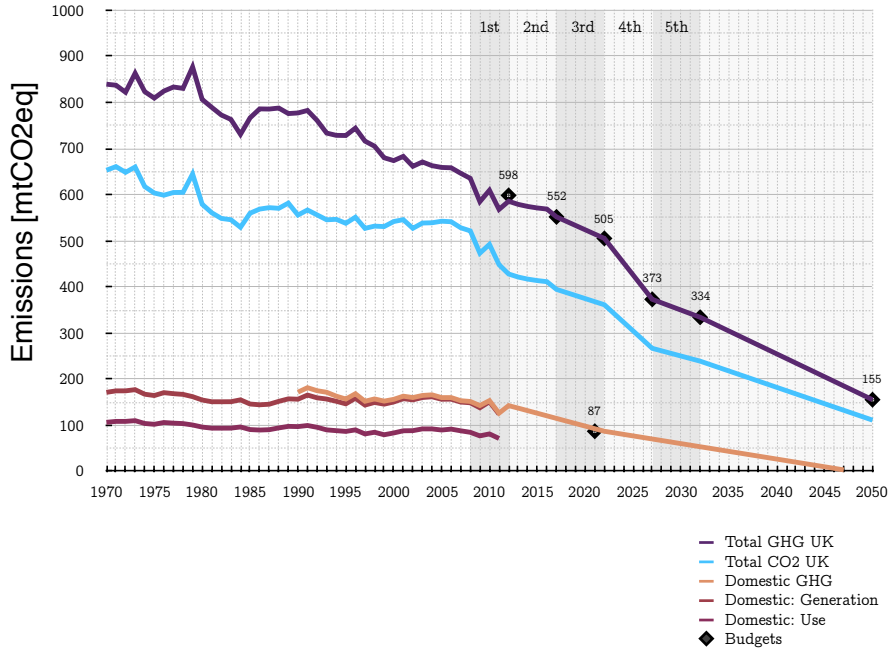


Figure 2.12: Historic carbon emissions in the UK
Source: [183], [38] and World Bank–databank.worldbank.org/data/

Table 2.2: Five-year budgets to revise the decarbonisation process in the UK
Source: [183] and [38]

Budget	**Carbon budget amount**	**% Reduction below base year**
1^{st} Budget (2008–2012)	3,018 $MtCO_{2e}$	23%
2^{nd} Budget (2013–2017)	2,782 $MtCO_{2e}$	29%
3^{rd} Budget (2018–2022)	2,544 $MtCO_{2e}$	35% by 2020
4^{th} Budget (2023–2027)	1,950 $MtCO_{2e}$	52% by 2025
5^{th} Budget (2028–2032)	1,765 $MtCO_{2e}$	57% by 2030

by 2025. These budget aims are summarised in Table 2.2 and highlighted in Figure 2.12.

By 2011, the UK emitted a total of 564 $MtCO_{2e}$, which is 36% below the peak value registered in 1979 and 28% below that of 1990 [58]. Such a reduction was mainly caused by a supply shift from coal to natural gas; by a displacement of industrial activity (primarily to Asia); and by major improvements in the performance of the transport sector [149]. This means that even though the reduction in this period is close to the CCC target, this has largely been achieved in the absence of structural improvements to reduce energy demands. Some of the more significant opportunities for demand reduction are found in the domestic sector, where emissions have been maintained at almost the same level since 1990 [58]. Moreover, the numbers suggest that there has been a significant lack in investment on energy efficient technologies in the domestic sector, despite the increase in policies, programmes and assessments that have been developed in the country since the last forty years, as part of the multiple attempts to tackle the national decarbonisation.

2.2 Policies, Programmes and Assessments

The 1973 oil crisis lead to more stringent energy regulations that continued into the 1980s Housing Act regulations [196, 220, 166]. Figure 2.13 summarises the introduction of policies, programmes assessments, and notable projects in the UK.

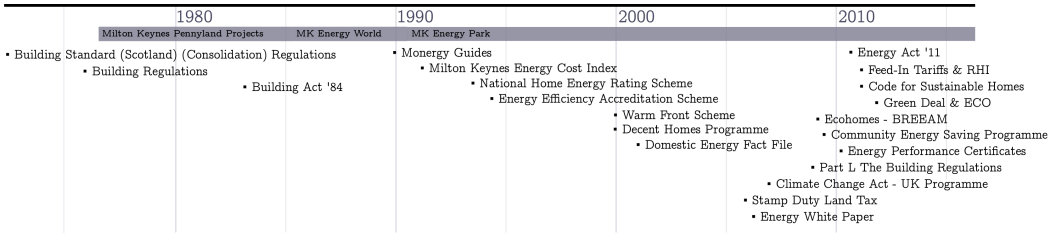


Figure 2.13: Development of policies, programmes and assessments in the UK
Source: adapted from [133]

As aforementioned, the composition of tenures changed during this epoch, due to the government's interest in developing specific-purpose settlements steered by these regulations, which complemented other polices related to safety, environmental and technological issues. These changes in housing tenure and regulation were also correlated with the way policies and energy-related measures impacted the population [19]. A brief summary is presented below to highlight the most relevant outcomes developed in the UK, aiming to improve the housing stock energy performance:

+ ***Building Regulations*** [202, 149] The first directive related to energy-related issues appeared in the Building Regulations 1965, as a consequence of the introduction of constructive technologies. These issues were related with the conservation of energy by measuring (and limiting) the thermal efficiency of materials on the envelope, expressed as the U-value—an indicator of heat transferred per area. However, the application of these standards became effective until the Building Regulations 1976. In this version, in addition to changes on the thermal insulation values, a few concepts related to the study of ventilation were introduced. An achievement in terms of the application of standards was the publication of the Building Act 1984, and the right to make building regulations. During this period, a shift was made towards the concept of energy efficiency in buildings. The powers to make Building Regulations were amended and extended in their scope by two recent Acts of Parliament; the Sustainable and Secure Buildings Act 2004 and the Climate Change and Sustainable Energy Act 2006. The current regulations made under the Building Act 1984 are the Building Regulations 2016.

+ ***Milton Keynes Projects*** [71, 27, 214, 142] Two projects were developed in Milton Keynes starting back in 1976: the Pennyland Project and the Linford Project. The Pennyland Project evaluated 177 houses under conventional properties of social sector; and the Linford Project studied eight high insulated houses. The construction of the houses finished in the early 1980s. These houses were designed, built and monitored as an attempt to evaluate their energy performance; although the results were limited to the evaluation of a trade-off between performance and cost. Still, this project became essential for upcoming studies of energy performance. By 1986, 51 more houses were built to reach an extra efficiency of 30% compared with the regulations. This project—known as the Energy World Demonstration—aimed to the development of labels and procedures in the housing sector.

One of the most substantial outcomes of the Milton Keynes Energy Park was, in 1989, the development of normalised measure of energy efficiency in houses: the Milton Keynes Energy Cost Index (MKECI). In here, the algorithms adopted initially by the project were complemented with Building Research Establishment Domestic Energy Model (BREDEM)-5 (described later in Chapter 3). In sum, the Milton Keynes Projects—see Figure 2.14—studied improvements such as floor and wall insulation, double glazing, condensing boilers, gas-air central heating and heat recovery ventilation systems.

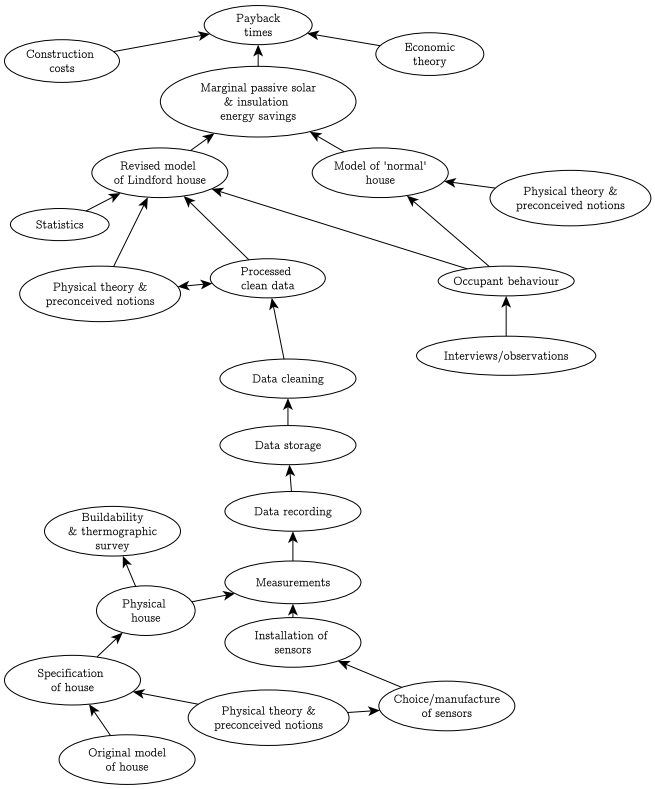


Figure 2.14: Diagrammatic representation of the Milton Keynes projects

+ ***Monergy Guide*** [148, 142, 149] Derived from the improvements defined on the Building Regulations, energy efficient projects and assessment tools, such as BREDEM-7—described later in Chapter 3, the government promoted the idea of energy efficiency in houses for both users and developers. An example of this promotion was the campaign *Get More for your Monergy* in 1986. This campaign defined numerical indicators related to the amount of money assigned to efficiency projects and the number of jobs created, among others. An outcome in the domestic sector was the Monergy Guide, which was as an attempt to cut energy cost by applying better practices. Besides, it is also worth pointing out that 1986 was designated as the *Energy Efficiency Year* in the UK, and motivated the development of further programmes and policies on energy efficiency (see Figure 2.13).

+ ***National Home Energy Rating Scheme Scheme*** [74] Derived from the results of the Milton Keynes Projects, the National Energy Services (then the National Energy Foundation) launched in 1990 the National Home Energy Rating Scheme. Essentially, the aim was to expand the use of the MKECIs to the whole country. By doing so, the National Home

Energy Rating Scheme (NHER) was used for labelling the energy efficiency of buildings for both users and developers. Even though this scheme was developed mainly for commercial buildings, it became the basis of the SAP rank developed in 1995. The scheme evolved to become the Energy Efficiency Accreditation Scheme with emphasis on renewables sources.

+ ***Climate Change Programme (CCP)*** [50, 74] By 2000, the UK Government launched the CCP in response to the commitments to reduce GHG—those agreed on the 1992 United Nations Conference on Environment and Development (UNCED). This organism established that not only the Kyoto targets would be achieved, but also the reduction of carbon dioxide in about 20%. By 2008, the Climate Change Act was published as a law, defining some mechanisms to achieve the agreed reductions. Also, the CCC was created as the body responsible to track and report the status on the reductions.

+ ***Domestic Energy Fact File (DEFF)*** [199, 221] By order to present a systematic report of the way energy is used in the domestic sector, particularly by measuring improvement on efficiency, the DEFF was published in 2003. For some specific topics, detailed year-by-year information was considered since 1970 to the publication year.

+ ***Energy White Paper*** [3] By 2003, the UK Government published the White Paper on Energy—referred as *Our Energy Future: creating a Low Carbon Economy*. This document introduced the concepts of fuel security and carbon intensity into the regulations. The latter impacting directly the standards in the domestic sector. In 2006, more detailed actions towards cutting emissions were developed in the *Energy Review* version. Later on the 2007 version: *Meeting the Energy Challenge*, some new directives were established, especially on new houses, towards the development of zero-carbon buildings, and a set of improvements of appliances. The long term targets of decarbonisation became part of the agenda. These documents, in turn, became the basis of both the Climate Change Act and Energy Act in 2008.

+ ***Stamp Duty Land Tax (SDLT)*** [133] By 2003, SDLT introduced some tax exceptions related to the energy efficiency of buildings; this SDLT was mainly reflected when the buildings were bought or transferred. By 2005, zero carbon homes granted full exemption on the tax. To achieve so, some calculations were standardised, based on SAP-05, such as thermal efficiency, solar hot water heating, generation of renewable energy and energy used for lighting.

+ ***Ecohomes - BREEAM*** [13, 34, 142] A domestic version of the Building Research Establishment Environmental Assessment Methodology (BREEAM) was published in 2008 named *Ecohomes*. It provided guidelines on fabric efficiency, water usage, emissions, recycling issues, lighting and appliances electricity efficiency, etc. In this way, a comprehensive environmental rating scheme was established for homes. The rating was replaced in 2008 by the Code for Sustainable Homes. This scheme, supported by the Department for Communities and Local Government (DCLG), became then the national standard for new houses. Their aims were to improve upon decarbonisation and promote higher standards of sustainable design than those established in the Building Regulations.

+ ***Revision of Part L: The Building Regulations*** [100, 133] Substantial improvements were added in 2006 to the Building Regulations: Part L, Conservation of fuel and power. These were on the definition of insulation parameters, openings, permeability, infiltration, heating efficiency and lighting efficiency. The algorithms were aligned to the SAP—particularly

by creating the reduced version: RdSAP, and further details were included in the derived sections defined as: L1A New dwellings, L1B Existing Dwellings, L2A New Buildings other than Dwellings and L2B Existing Buildings other than dwellings.

+ ***Energy Performance Certificate*** [101, 133] By employing the RdSAP, a method to rank efficiency on houses was developed, including new and existing constructions. This rank allowed the creation of Energy Performance Certificates (EPCs) by 2007, and provided a label ranking from A (very efficient) to G (inefficient). This rank was made to measure both energy demand and estimated carbon emissions. The EPC was added to the Home Information Packs as a useful (and eventually mandatory) set of documents describing the properties of the building to be sold.

+ ***Feed-In Tariffs - Renewables instruments*** [184] By 2008, the Energy Act introduced a mechanism to encourage the investment on low-carbon electricity, by creating a set of special contracts. Depending on the technology used, the agreed tariffs were defined as way to guarantee profits by investing on these sources of energy. By 2011, a similar scheme was launched in order to improve the development of small-scale renewable heat. Despite the Renewable Heat Incentive (RHI) was made for the non-domestic sector, it was expanded to domestic sector later on 2013. Essentially, the scheme created payments for the developers.

+ ***Green Deal & Energy Company Obligation (ECO)*** [86, 184] The Green Deal was established in 2012 as a financial incentive to invest on residential energy efficient technologies. In brief, the Green Deal consisted on recouping the borrowed investment through the energy bills, so that householders might gain confidence in the investment. The expected financial savings should be less than the savings on the energy bill. The Green Deal Programme also developed a framework of advice and guide to enhance the penetration of new technology in the domestic sector. In addition to the Green Deal, the government created a scheme where energy suppliers, via a fund, were able to install energy efficient measures into houses facing difficulties to pay the energy bills—the concept of fuel poverty. These measures were for example: loft or wall insulation, and replacement boilers. Such a scheme, named ECO, was focused on the cumulative reduction of CO_2 emissions in the housing stock. Similarly, in Scotland the scheme was developed as the Action on Carbon and Energy Performance.

+ ***Energy Acts*** [175, 66, 178] Continuing with the attempts to persuade the installation of micro-generation installations in the country, the Climate Change and Sustainable Energy Act was published in 2006. Then in 2008, the Climate Change Act 2008 was introduced as a body to track and control the gaols agreed in the Kyoto Protocol. Also, the Energy Act was released to complement the Energy Paper targets where, in particular for the domestic sector, the main topics included smart metering, feed-in tariffs and the installation of renewable heat technologies. These measures were also echoed by the Green Energy Act 2009 and the Climate Change (Scotland) Act. The update to the Energy act in 2010 established a mandatory social price support and changes on the regulation of the energy market that affected the investments on micro-generation and renewable. The Energy Act 2011 defined the minimum energy efficiency standards a property should have.

It is worth drawing attention that regardless of the slow participation of the country in the diffusion of new technologies investments, some dwelling-oriented efficiency projects have become world-wide relevant in the analysis of housing performance. According to Herring [96], the government in the UK maintained a persistent resistance to improving appliance efficiency. For a long time, the voluntary approach was preferred despite the difficulty in securing support from local manufacturers and the widening gap in efficiency labels; however, this trend was gradually reversed by the end of the 1980s. In addition, not only improvements of appliances were reached, but also better practices were achieved through the promotion of policies, household guides, and financial investments. Since late 1980s—see Figure 2.13, the evolution of energy-related projects has encouraged the development of fit-for-purpose surveys and sources of information to characterise the housing stock in the UK.

2.3 Sources of Information

2.3.1 Scope and Quantity

In the UK, a national census is conducted every ten years, providing a rich picture of the population by household, giving the number of inhabitants, their education, employment and social affairs, and building information including tenure, size and installed heating system[3]. However, the accuracy of the data may deteriorate over time as household composition and characteristics evolve. To address this, the census is updated annually using information from specific topical surveys and regional data, such as the Annual Population Surveys [171]. Census data is also enriched with estimates and direct measures from electricity and gas sales in the domestic sector [58], to increase the resolution of energy-related information. For example, the UK Government's Department for Environment, Food and Rural Affairs (DEFRA)'s Market Transformation Programme (MTP) can model the energy demand of domestic appliances, and make detailed estimations of their usage by calibrating ownership with energy demand [59] although the level of disaggregation is constrained by the available supply subtotals, and cannot describe behaviours affecting the energy use in dwellings.

Limited attempts have been made to identify environmental attitudes and energy-using behaviours. For example, a highly disaggregated study regarding appliances' energy demand was conducted in the 1990s [150], and was extrapolated and calibrated for an estimation of national energy demand but limited to a small sample in the south of the UK. More recently, DEFRA, the Department of Energy and Climate Change (DECC) and the Energy Saving Trust (EST) commissioned a detailed field study of electricity use by 251 households at a one-minute resolution [174, 176, 234]. DECC published these studies as cross-tabulated data [60], which is a source for the Digest of UK Energy Statistics (DUKES). These data calibrate and reconcile HSEMs [31, 142, 110] for a sample of the national housing stock.

Common approaches used to reconcile a representative and statistically valid sample of the housing stock are sampling and clustering. Sampling involves random or systematic selection of units (e.g., households, dwellings, neighbourhoods) amongst the population, whereas clustering involves the identification of groups that share characteristics. Both methods use weights to then match

[3]The collection of national and regional statistics is conducted by ONS, National Records of Scotland, Statistics for Wales, and Northern Ireland Statistics and Research Agency, respectively

Table 2.3: UK regional classification in 2011
Source: Office for National Statistics (ONS)–www.ons.gov.uk/

	Census			LA	NUTS⋆			SOA◦	
	People	**Households**	**Dwellings**	**LA**	**1st**	**2nd**	**3rd**	**Middle**	**Lower**
United Kingdom	62,055,838	27,651,734	27,580,884	406	12	37	140	8,429	41,773
Great Britain	60,267,499	26,903,499	26,832,836	380	11	36	135	8,429	40,883
England and Wales	55,071,113	24,429,618	24,359,880	348	10	32	112	7,194	34,378
England	52,059,931	23,044,097	22,976,066	326†	9	30	100	6,781	32,482
Wales	3,011,182	1,385,521	1,383,814	22	1	2	12	413	1,896
Scotland	5,196,386	2,473,881	2,472,956	32	1	4	23	1,235	6,505
Northern Ireland	1,788,339	748,235	748,048	26	1	1	5	NA	890

† Local Authoritys (LAs) include non-metropolitan districts, metropolitan districts, unitary authorities and London boroughs.

⋆ These values correspond to the Nomenclature of Territorial Units for Statistics (NUTS)-2010 nomenclature, regulated by the European Union.

◦ The output areas define synthetic polygons around an average of 125 households. Lower-Super Output Areas (SOAs) are made of output areas and typically enclose about 1500 households. Middle-SOAs contain an average of about 7200 households.

the population totals. Table 2.3 identifies the characteristics of the three clustering strategies that are most commonly used to describe the UK population, and to account for dwellings and households: Local Authority (LA), Nomenclature of Territorial Units for Statistics (NUTS) and Super Output Area (SOA).

Firstly, the LA classification are helpful in that they represent administrative areas, and thus express relationship with the government and its structures. An LA may have access to specific datasets (public and private), which enhance the calibration of the stock weights [222, 64]. Secondly, NUTS is a geocode standard used for referencing the subdivisions of countries for statistical purposes and is regulated by the European Union. Its different layers align with UK regions; for example, the former Government Offices for the Regions (GOR) are equivalent to NUTS-1 [58]. Finally, the SOA is designed to improve the reporting of small area statistics and are made up of groups of output areas and LAs [222]. They are constructed by clustering households by socio-economic status. These clusters have been used to support the development of national housing surveys because they maintain characteristic attributes (demographic, household composition, housing, socio-economic and employment) and so can be considered as representative subsets of the population.

Furthermore, for the analysis of very specific dwelling properties, reference typologies have been deployed with representative attributes associated with them [123, 76, 39]. Reference typologies include the variability among dwellings and can be linked to the stock typologies, even though their attributes result from extrapolation. Nevertheless, they can also be useful for the calibration and reconciliation of HSEMs with the different sources of information. Besides, one advantage of this approach is that the typologies can be scaled, although the procedure for each subset needs to be correspondingly refined [123, 165, 140]. By contrast, representative sampling provides real examples of the housing stock and can be used to study performance under realistic constraints; for instance, they can support cost-effective analyses based on household energy expenditure. An example of this type of sampling is the English Housing Survey (EHS)—see Box 2.3.1.

Box 2.3.1: EHS

The EHS [47][a] is a valuable source of energy related data on dwelling energy use in England (83.3% of the UK stock). It adopts a clustering method to select a statistically representative sample of 16,150 English dwellings. Data is collected by questionnaire and a visual survey of dwellings' physical properties. The EHS samples are calibrated with a subsample of SOA. The process of sampling and weighting used in the EHS [49] (i.e. selecting, characterising, extrapolating and calibrating) is summarised in twelve stages listed below:

S-1 Address selection probabilities. This is based on the random sample of the Postcode Address File (PAF) (ca. 38 k)

S-2 Address to dwelling relationship. This is the fraction of the PAF sample with the total of PAF addresses

S-3 Dwelling to household relationship. A dwelling is defined as *a self-contained unit of accommodation where the occupants of that accommodation have sole use of all the rooms and facilities*

S-4 Refusals prior to the interview and probability of contact

S-5 Probability of response at the interview survey

S-6 Calibration weighting for the interview survey (population totals). It uses tenure and geographical area

S-7 Refusals prior to the interview and probability of contact

S-8 Probability of cooperation at the dwelling

S-9 Sub-sampling by tenure for the physical survey

S-10 Co-operation at the physical survey

S-11 Calibration weighting to dwelling totals, including an adjustment for new builds

S-12 Calibrating the household weights for the paired sample to those from the full household sample

The EHS publishes its content in specific sections of the survey. In here, the main sections include the difference between primary data and derived data. Following sections include derived data, which use information from samples with physical inspection to fill their fields; sections with derived data also involve the specification of systems and recode specific fields. Evidently, surveyed samples provide more granular information on the properties of both houses and households.

[a] In April 2008 the EHS was created by merging the English House Condition Survey with the Survey of English Housing www.gov.uk/government/collections/english-housing-survey

Figure 2.15 presents a simplification of the process taken by the EHS and other surveys to weight their sampling units [21]. The process employs Census data to calibrate the sample so that those parameters with support of granular information can be matched in the resulting sample dataset. Likewise, the Living in Wales survey (succeeded by the National Survey for Wales), the Northern Ireland House Condition Survey and the Scottish House Condition Survey, all contain equivalent information describing the dwellings of their respective countries, and apply similar sampling and weighting methods.

Other surveys and databases used to characterise the housing stock are contained in Table 2.4.

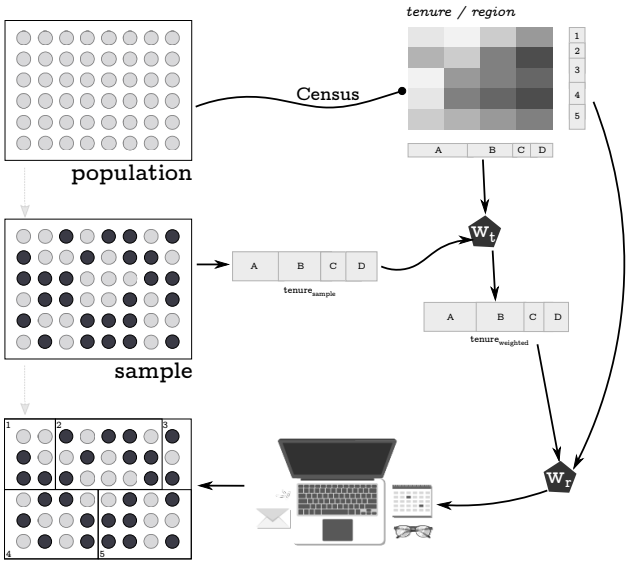


Figure 2.15: Simplified weighting process. W_t (tenure) and W_r (region) include the Census reference to calibrate the partial values from the sample

Sources with smaller samples improve the resolution of the collected information. For example, the Energy Follow Up Survey (EFUS) measures indoor temperature and provides additional socio-demographic information. Furthermore, both the National Energy Efficiency Data-framework (NEED) and the Home Energy Efficiency Database (HEED) analyse and forecast the adoption of retrofit measures across the stock, but reduce their own resolutions. The HEED [87] specifically gives information on retrofitted energy efficiency measures and a wealth of physical dwelling characteristics (property age, property tenure, glazing type, heating system or insulation), but its data is limited to a number of UK regions and so suffers from sampling biases and reductions.

2.3.2 Reductions and Imputations: Common Assumptions in Stock Data

In the context of building performance modelling, it has been argued that as the quantity of required input data increases, so the average quality of this data reduces and the risk of input errors increases [25]. Whilst it may be theoretically possible to acquire data for the 27 million dwellings that comprise the UK housing stock, this is neither viable nor is it necessarily desirable. Conversely, a sampling process may over-simplify the data, if it is not repeated when updates are made. Furthermore, surveyors are constrained by limits on budget, time, or personnel, but computational advances can offset these limitations by increasing the speed of data collection and by improving the number of variables collected.

Sampling is required to form representative datasets of the entire housing stock, with inherent risks of bias and errors needing to be managed. Sampling bias is related to the procedures, criteria, and strategies taken when selecting samples, and is commonly compensated by a weighting process. Buckley et al. [21] describe one such method for weighting nationwide surveys. Sampling errors occur during the process of collecting data and can be random (false occupant responses, survey mistakes, typographic errors) or systematic (instrument error, inaccurate resolution, conversion

Table 2.4: Summary of surveys and datasets

Source	Size	Sampling Units	Building Fabric	Heating Systems	Energy Bills	Meter Readings	Household Composition	Socio-demographics	Household Attitudes	Measured Temperature	Retrofit Measures	Specific Metrics +	Historical Metrics *	Coverage	Support
			Properties												
EHS	∼ 14 k	sample of Postcodes and LSOA	✓	✓	✓		✓	✓						**Eng**	DCLG
EFUS	∼ 2.5 k	extract from EHS physical surveys	✓	✓	✓	✓	✓	✓		✓				**Eng**	BRE + DECC
NSW	∼ 1.2 k	sample of LSOA	✓	✓	✓		✓	✓						**Wal**	UK Data + Wal Gov
NIHCS	∼ 10 k	sample of Postcodes	✓	✓		✓	✓	✓			✓			**NI**	UK Data + NI Gov
SHCS	∼ 18 k	sample of LA	✓	✓	✓		✓	✓	✓	✓				**Sct**	UK Data + Sct Gov
DUKES	∼ 27 M	readings from Postcodes and LA						✓				✓	✓	**UK**	DECC + BIS
Census	∼ 27 M	UK household population	✓	✓			✓	✓						**UK**	ONS
NEED	∼ 25 M	sample from Experian (HEED and EPC)				✓	✓	✓			✓			**Eng Wal** +	DECC
HEED	∼ 15 M	EST	✓	✓							✓			**UK**	EST

Notes: + Include retrofits on insulation, boiler, heaters and windows
* Include metrics from coal, petroleum, gas, electricity, renewables and combined heat and power

errors) [186, 109]. When combined with confidentiality issues, these errors may impose limitations on the use of the datasets. Nevertheless, the process of sampling an entire stock is of such a magnitude that some information must be absent (the data has gaps).

The process of reducing gaps in the collected dataset is tackled either by retaining missing values or by applying imputations: assumptions about values used to fill gaps. Imputations can be arbitrary or systematic, but any replacements can introduce errors. However, these replacements can be calibrated using derived census data, manufacturer specifications, and *in-situ* measurements. But their availability needs to be standardised and centralised based for example on CIBSE guidelines [83] or the Standard Assessment Procedure (SAP)—the UK government's method of assessing the energy performance of dwellings [94]. Additionally, improved values can be normalised with dwelling parameters so that a subset of average typologies is created to fit specific purposes. One of the advantages of this approach is that the typologies can be easily scaled although the procedure for each subset needs to be refined for each case [123, 165, 140].

For example, the SOAs that provide a robust basis for sampling socio-economic clusters in the UK cover around 1500 households per cluster, so that some details are missed or are over-simplified. Here, fabric properties can be inferred statistically for a given SOA, but other properties may differ significantly, even though these details are regulated at LA and ward level where homogeneity can be expected. These differences arise due to locale-specific construction practices. Other properties that are extrapolated (and hence simplified) in SOAs are location, orientation, geometry, and household-specific information. Nevertheless, SOAs provide a sound starting point to study homogeneous groups and to link them with more granular and spatially detailed data. For this reason, the SOAs are used to calibrate some nationwide survey including the EHS. The EHS reduces the sampling errors by making physical inspections of dwellings. Yet the EHS surveys the housing conditions and is not a dedicated energy performance study, which simplifies physical and household properties that feed HSEMs.

In summary, the impacts of information loss from sampling biases have been widely studied and discussed in the literature [103, 67, 132, 109], but little attention has been given to the magnitude of errors attributable to sampling methods. A high level of inaccuracy may be present in datasets generated from censuses and nationwide surveys. Therefore, it is important to track error propagation through each layer of the data, so that specific criteria can be adopted to maintain a workable source of information. An example of this tracking is achieved via *monitoring indicators* that identify the most influential parameters in the composition of a stock of dwellings [165, 140]. *Monitoring indicators* are used to help validate model outcomes. They may be available directly from survey data, or else are derived directly from this data, but without requiring uncertain assumptions. Parsimony is imperative.

The composition of the UK housing stock, as explained in Section 2.1.1, presents a high level of homogeneity in their composition and can be described by typologies well. By studying typologies, the sampling methods can be calibrated, and hence enhance with more valuable information.

2.4 Typologies of the Housing Stock

The study of shapes and sizes enables a valuable preliminary step when comparing the energy performance among the housing stock, despite the processes involved in the characterisation of the stock that imply reduction and loss of information. To this end, two elements have been essential to initially describe the physicality of the housing stock (i.e. the infrastructure) in the UK: type and age of construction. Figures 2.17 and 2.16 briefly exemplify each type, with a view to their most common properties. By identifying some of the components in the façade, as well as some architectonic elements, it is possible to infer the type of dwelling and its performance attributes. Figure 2.18, even though limited to the English stock, shows the composition of the stock in terms of these two typological elements. Dwellings constructed between 1920 and 1979 (i.e. inter-war, post-war and industrial) represent more than 50% of the stock. Most of the dwellings within these epochs are semi-detached. Table 2.5 summarises average values related to the energy performance by these typologies; the support for these numbers will be further described in Chapter 3.

Other infrastructural elements that provide valuable information on the composition of the housing stock are: total floor area (or size), and predominant materials used in the constructions. These properties are highly correlated with typologies and epoch of construction. Their correlation



(a) Bungalow



(b) Detached House



(c) Apartments Block



(d) Terraced houses



(e) Semi-detached Houses



(f) End-of-terrace Houses

Figure 2.16: Building typologies in the British housing stock
Source: © Google, used with permission

is usually found in national assessments and guidelines, such as the CIBSE guidelines [83] and the SAP [94]. Old dwellings are usually larger and less energy efficient than modern ones; this is mainly caused by the employed materials (see Figures 2.1:d and 2.19a). Detached and semi-detached, because of larger spaces in their composition, demand more energy than the national mean; however, detached dwellings are usually under-heated and the energy per floor area is less intense. This is highlighted in Figures 2.19b, 2.20, and 2.21, where the difference in the performance can be noted among the groups.

The location of dwellings is also relevant at providing additional information on the typical configuration of construction practices found in the stock. In particular, the location provides a regional estimation of the annual degree days, which are essentially a representation of external conditions (i.e. air temperature and humidity) in regard to a base temperature [29]. In the UK, degree days are typically used to estimate heating demand during the year with only a few dwellings demanding cooling systems. Figure 2.23 exemplifies the mean temperature with a reference to the typical degree-day values.



(a) pre 1850 - Historic

(b) 1850 to 1899 - Victorian

(c) 1900 to 1918 - Edwardian

(d) 1919 to 1944 - Interwar

(e) 1945 to 1964 - Early Postwar

(f) 1965 to 1980 - Postwar

(g) 1981 to 2002 - Modern

(h) post 2003 - New

Figure 2.17: Building epochs in the British housing stock
Source: [166]

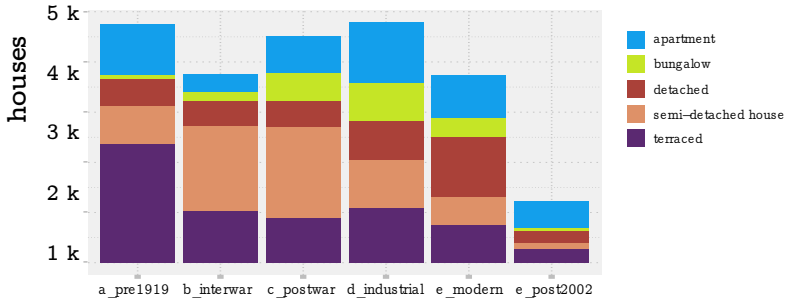
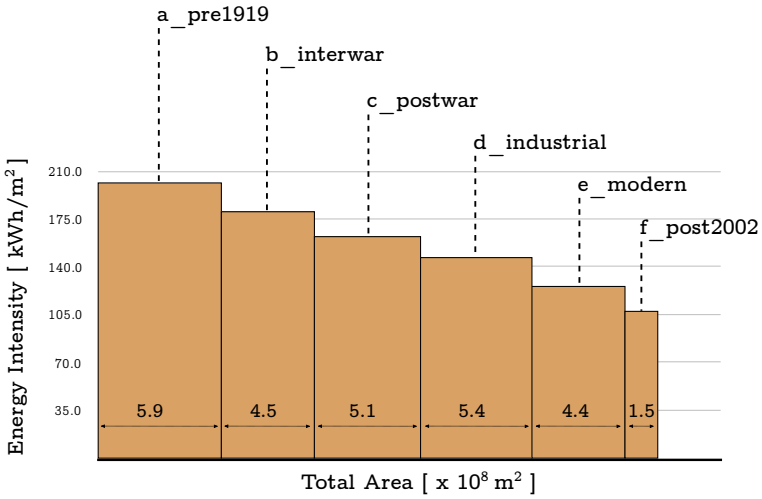
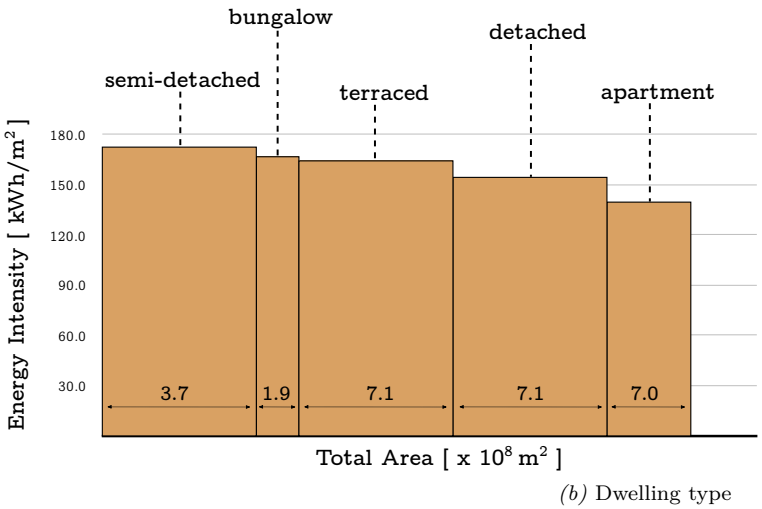


Figure 2.18: Composition of dwellings by type and age of construction
Source: Derived from EHS

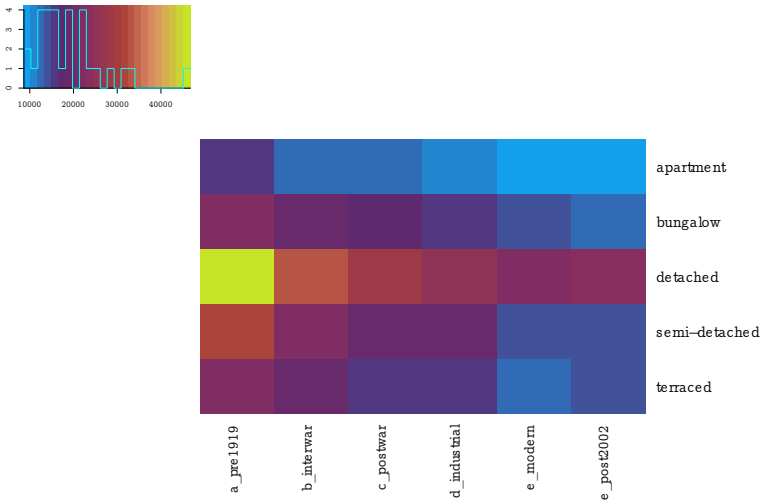


(a) Building epoch

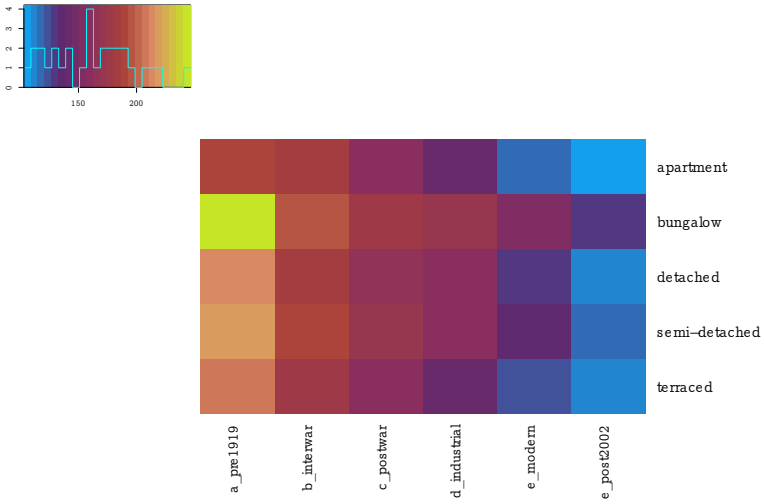


(b) Dwelling type

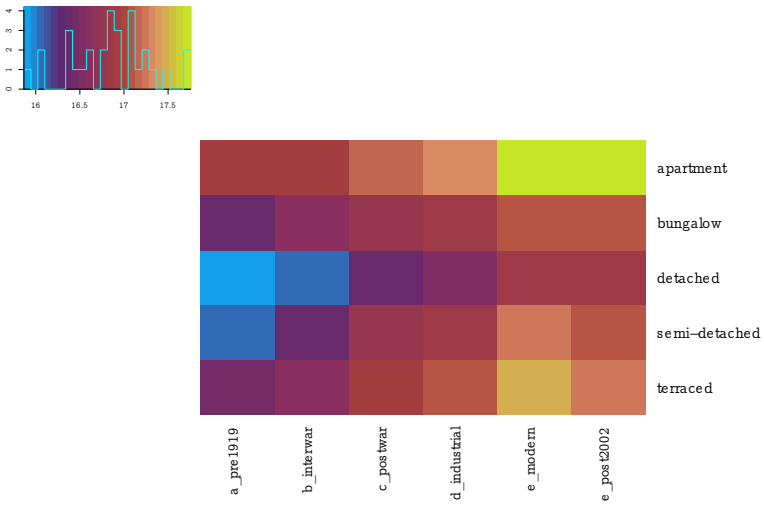
Figure 2.19: Energy intensity by typologies
Source: Derived from EHS and CHM



(a) Energy Demand *kWh*



(b) Energy Intensity *kWh/sqm*



(c) Indoor Temperature *C*

Figure 2.20: Energy performance indicators: type and age
Source: Derived from DUKES and CHM

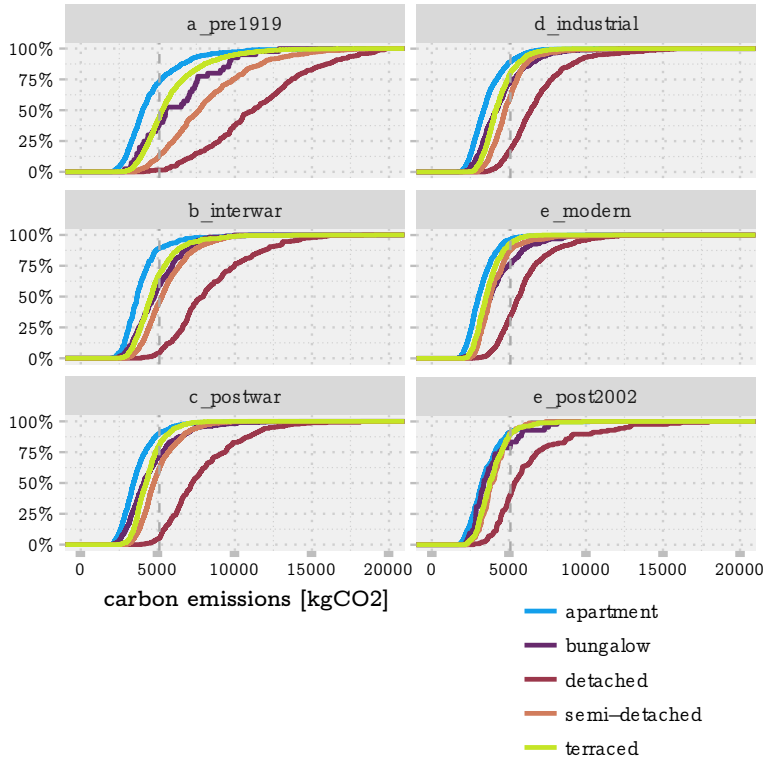


Figure 2.21: Carbon emissions CDF by type and age of construction
Source: Derived from DUKES and CHM

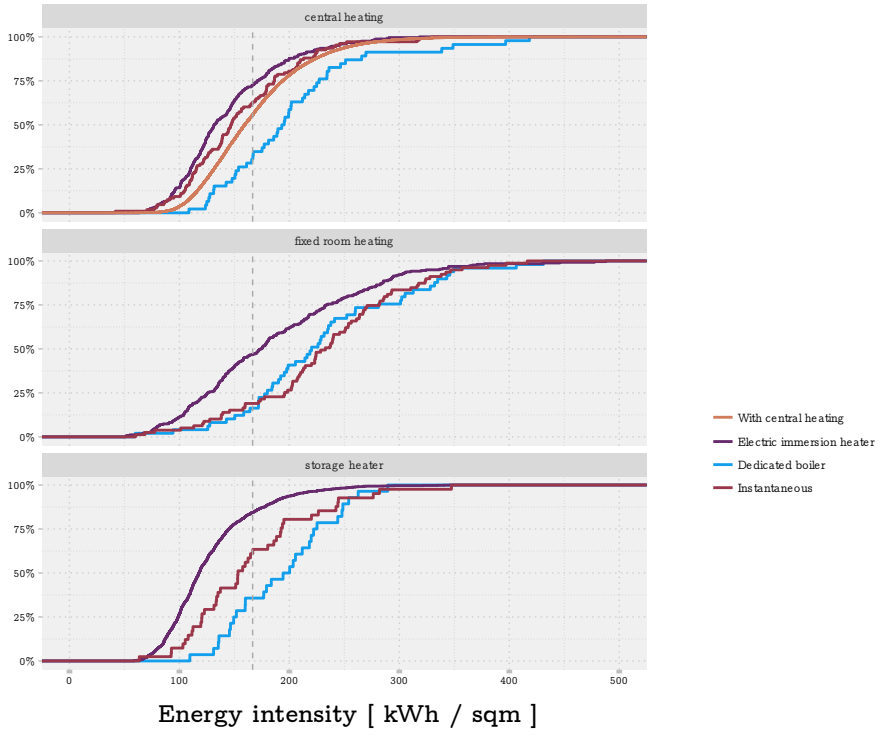


Figure 2.22: Differential energy intensity CDF by heating systems
Source: Derived from DUKES and CHM

Table 2.5: Energy summary estimates by age and type of construction
Source: Derived from DUKES and CHM

	Dwellings (x1000)	%	Indoor Temperature [C]	Energy Demand [kWh]	Carbon Emissions [tonnes CO2]	Energy per Area [kWh/sqm]	Dwelling Heat Loss [W/C]
pre-1919	5,783.1	21.0%	16.4	25,169.2	6.9	206.6	438.0
interwar	4,312.2	15.6%	16.5	21,579.2	5.7	185.9	379.6
postwar	5,519.5	20.0%	16.8	18,381.2	5.1	167.7	296.1
industrial	5,816.4	21.1%	17.0	16,684.2	4.8	153.8	264.7
modern	4,589.4	16.6%	17.3	15,062.2	4.5	130.0	230.3
post-2002	1,560.2	5.7%	17.3	13,269.3	4.2	110.3	205.1
terraced	7,749.3	28.2%	17.3	11,496	3.8	149.1	168.5
bungalow	2,442.4	8.8%	16.8	16,624	4.8	173.4	262.2
semi-detached	7,102.3	26.0%	16.5	28,945	7.8	159.3	529.8
detached	4,633.7	16.6%	16.6	20,756	5.6	176.8	345.8
apartments	5,653.2	20.3%	16.8	18,180	5.0	172.0	289.6
	27,581						

In addition to these (tangible) contextual attributes, the characterisation of the most common heating systems adopted in houses provides useful information to compare the energy demands in the housing stock. Some differences in the energy intensity for dwellings with different heating systems is shown in Figure 2.22. It can be seen that the combination of central heating with electric boiler is usually more efficient, but without central heating, the energy demand is significantly higher. When central heating is installed (ca. 80% of the stock), it is expected that 1) indoor temperatures have a higher fluctuation, and 2) more zones in the dwellings are under-heated.

Moreover, in terms of the structure (or organisational parameters), the tenure is a useful classifier, not only as a common descriptor of the household composition, but also as a link with other sources of information (see Section 2.3.1). Its inclusion is also relevant because it provides, in combination with the physical typologies, a picture of how households utilise energy within different socio-demographic profiles.

To summarise, it is highly important to identify the key components of the housing stock in terms of their composition, and also in terms of their contribution to the overall energy demand and carbon emissions associated. Table 2.6 summarises these key descriptors of the housing stock in the UK, with its most relevant facts described below:

- The UK housing stock is comprised of 27,5 million dwellings (83.3% Eng, 9% Sco, 5% Wal and 2.7% NI)
- Around two-thirds of the stock are owner-occupied
- More than one-third of the stock was built before the 1944
- Around 80% of the stock is comprised by houses; semi-detached houses are the most common type
- Gas central heating is the most common system found in the housing stock with more than three-quarters of the total
- Standard boiler is predominant for hot-water systems

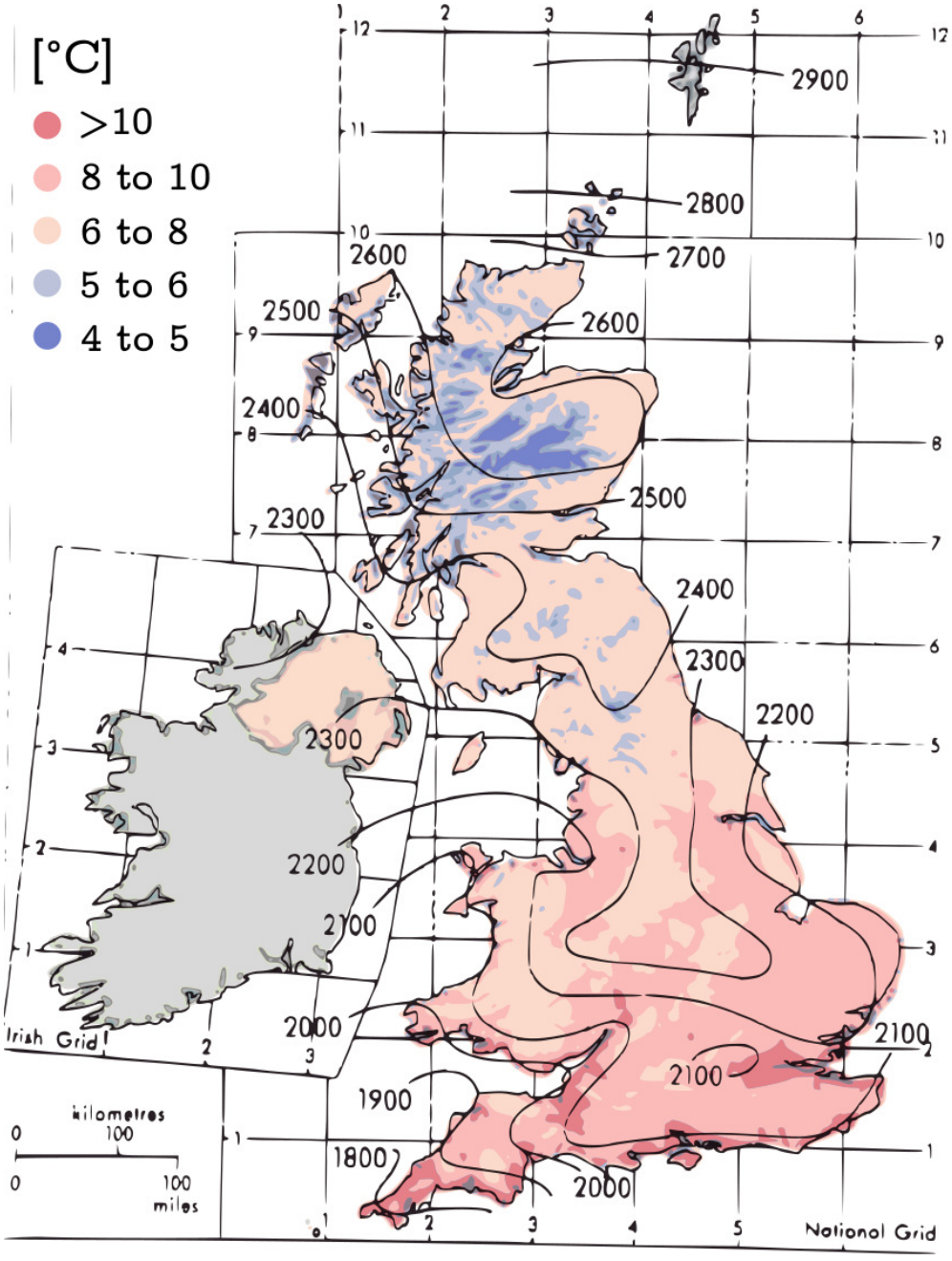


Figure 2.23: Mean annual temperature with degree days map overlay
Source: © MetOffice, used with permission

Table 2.6: Key Statistics of the 2011 UK Housing Stock

	England	Wales	Scotland	Northern Ireland	United Kingdom	UK%
owner occupied	14,908,443	973,530	1,548,513	487,892	17,918,377	65.0%
private rented	4,056,025	188,694	288,019	130,473	4,663,211	16.9%
local authority	1,901,293	88,433	368,166	84,610	2,442,503	8.9%
housing association	2,110,304	133,157	268,258	45,073	2,556,792	9.3%
Pre-1919	4,779,022	417,911	498,525	87,660	5,783,118	21.0%
1919-44	3,791,051	151,336	301,678	68,130	4,312,194	15.6%
1945-64	4,549,261	293,500	542,825	133,960	5,519,546	20.0%
1965-80	4,835,839	256,642	554,357	169,550	5,816,388	21.1%
1981-90	3,747,026	197,337	429,541	215,489	4,589,392	16.6%
Post 1990	1,273,868	67,088	146,030	73,259	1,560,245	5.7%
Detached	5,113,411	383,264	546,683	277,063	6,320,421	22.9%
Semi-detached	7,055,504	429,425	554,100	207,852	8,246,881	29.9%
Terraced (including end-terrace)	5,626,310	384,904	458,423	187,631	6,657,268	24.1%
Purpose-built block of flats or tenement	3,843,072	133,353	850,643	64,502	4,891,569	17.7%
Part of a converted or shared house	981,378	33,010	45,381	7,397	1,067,167	3.9%
Flat In a commercial building	256,459	14,080	13,354	2,491	286,384	1.0%
Caravan or temporary	99,932	5,778	4,373	1,112	111,195	0.4%
No central heating	619,156	31,374	57,286	4,006	711,821	2.6%
Gas central heating	18,106,055	1,055,441	1,835,799	128,656	21,125,952	76.6%
Electric (including storage heaters) central heating	1,904,232	76,672	331,250	26,242	2,338,396	8.5%
Oil central heating	883,230	121,084	140,932	465,107	1,610,353	5.8%
Solid fuel (for example wood, coal) central heating	155,886	26,543	27,316	19,274	229,019	0.8%
Other central heating	372,722	12,342	17,309	4,343	406,716	1.5%
Two or more types of central heating	934,783	60,359	63,065	100,420	1,158,627	4.2%
No Boiler	2,642,248	92,808	429,864	40,354	3,205,275	11.6%
Boiler: Standard/Back	7,168,533	572,087	916,550	606,061	9,263,231	33.6%
Boiler: Combi	4,457,357	473,738	879,493	47,080	5,857,668	21.2%
Boiler: Condensing	2,205,702	26,319	66,703	29,145	2,327,869	8.4%
Boiler: Both	6,502,227	218,861	180,345	25,408	6,926,842	25.1%
less than 50 sqm	2,700,097	69,256	198,382	34,888	3,002,623	10.9%
50 to 69 sqm	5,575,892	241,989	425,105	135,550	6,378,536	23.1%
70 to 89 sqm	6,370,572	411,334	629,785	230,576	7,642,268	27.7%
90 to 109 sqm	3,236,279	250,541	485,984	106,090	4,078,893	14.8%
110 sqm or more	5,093,227	410,694	733,700	240,944	6,478,563	23.5%
solid masonry	6,228,196	434,250	540,904	148,095	7,351,444	26.7%
cavity masonry	14,616,268	854,004	1,490,926	555,990	17,517,187	63.5%
masonry crosswall	273,645	50,682	117,432	20,946	462,705	1.7%
timber frame	638,168	44,066	302,503	17,469	1,002,206	3.6%
other	1,219,789	812	21,192	5,548	1,247,342	4.5%
	83.3%	5.0%	9.0%	2.7%	27,580,884	

- The usable floor area is mainly between 70 and 110 square meters.
- Cavity walls are the most common construction applied to the housing stock with around two-thirds of the total
- The average energy demand is between 13,269 and 25,169 [kWh]
- Correspondingly, the carbon emissions are 4.2 to 6.9 [$tonnesCO_2$]
- The average indoor temperature is 17.2 °C

2.5 Conclusion of *the British Housing Stock*

This Chapter highlighted the relevance of the domestic sector in the contribution to the national energy demand and, consequentially, GHG emissions. More than 80% of the energy demand in the domestic sector is used to satisfy heating requirements; however, whilst a series of improvements have been achieved and projected to reduce heat losses in the stock of dwellings, these have been counteracted by the population growth, and by the adoption of states of comfort that demand higher temperatures and thus more heat. In addition, despite the energy share of household appliances represents one-fifth the whole demand, its contribution is constantly rising due to increased ownership and thus due to more intense use of energy. Therefore, it is important to comprehensively understand the drivers affecting the heating demand in dwellings and the quality of information characterising the stock.

In terms of energy-related GHG emissions, there are three main components that shape such a contribution: grid pureness, fuel supply and distribution, and end-use efficiency. The UK government is making progress to minimise these emissions and this Chapter identified significant attempts to achieve the agreed CCC carbon budgets. First, a significant reduction was achieved by the switch to natural gas to replace coal in power plants in the 1990s, and by changes to the national fuel imports. Second, considerable investments in the installation of renewable technologies, and decentralised energy systems, have been made. Yet gas is still the main fuel used in the stock to supply power plants and to satisfy end-use energy demand. And third, the efficiency of systems and appliances has improved, and energy conservations measures and regulations have been applied; however, the effectiveness of these end-use measures heavily depend upon the decisions taken by householders, which have not been received enough attention to significantly improve the use of energy across the housing stock.

In sum, excluding improvements on the grid pureness, structural changes are lacking in the domestic sector to meet the CCC targets. Noteworthy, it was highlighted that domestic GHG emissions haven been maintained at broadly the same levels in the last forty years and hence they have undermined the effectiveness of decarbonisation strategies. Therefore, the decarbonisation transition still is significantly challenging and there is a long way to be covered.

This Chapter identified that more granular data is needed to characterise the stock covering physicality of dwellings, occupancy and socio-demographic information; and ownership of household goods. For that reason, the strategies to decarbonise the stock need to gain insights in all these components that shape the housing energy demand. Nevertheless, the analysis of data sources shows that it is easy for significant errors to be incorporated into a database. These can occur from a wide number of processes, such as sampling biases, the imputation of missing data,

or the aggregation of information to make it more manageable. One process that is currently unavoidable is the identification of properties shared by a large number of dwellings so that they can be clustered together. Whatever the process, it is clearly important to track the origins of data to avoid error propagation at each stage of aggregation. Finally, it is worth drawing attention that most sources of data are updated periodically, but most housing stocks evolve more or less continuously. Thus, it is essential to continually update databases to maintain their validity and to calibrate them with the natural evolution of the housing stock.

To end, this Chapter summarised the composition of the UK housing stock; provided a review of policies, programmes and assessments developed in the country; and identified the relevance of data to characterise the housing stock. It is concluded that, to improve upon the strategies to decarbonise the housing stock, data needs to be calibrated, validated and available so that models can be benefited from it. However, to achieve this, data can only be validated if it is easily edited and constantly updated.

3 Housing Stock Energy Models

3.1 Modelling Rationale

Models are used to represent reality in a convenient way; especially when the process of gathering information on real phenomena is difficult, or even impossible, to achieve. Models are also conveniently employed because they capture different conditions in manageable, time- and cost-effective ways. However, to achieve this, they require a process of abstraction which implies—per definition—the adoption of assumptions and, therefore, the loss of information[6]. On the one hand, the assumptions improve the functionality and applicability of the model, by providing reliable output variables, in function of the limited resolution of the model; but on the other hand, they can lead to noticeable levels of uncertainty, by overlapping some significant input variables. The processes of representing a large system, without losing the essence of the model in form and function, have historically been the cornerstone—and hence the challenge—in modelling approaches [6, 134]. Figure 3.1 illustrates the abstraction made to represent real phenomena. The Figure shows how a process of abstraction simplifies the representation of a real-world phenomenon, and how such a simplification can also describe the most relevant interrelationships capturing it.

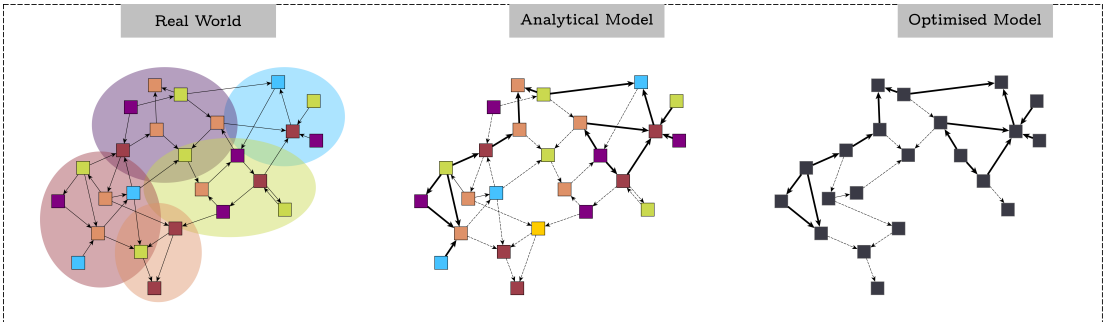
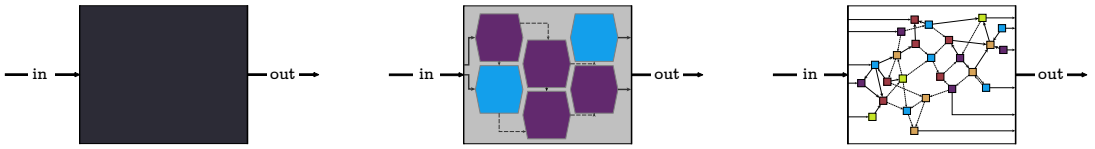


Figure 3.1: Abstraction: how reality is seen through a model
Source: [6]

From the analytical point of view, it is worth drawing attention that such an abstraction can be represented either as a black box (also known as empirical), where the description of internal

interdependencies are not available; or as a white box, where the insights of the model are presented transparently. Further, an existing middle stage can be defined as a grey box. These differences are exemplified in Figure 3.2.

Figure 3.2 also illustrates the typical representation of interdependencies between inputs (i.e. explicit excitations and conditions) and outputs (i.e. states and exits), throughout a system boundary. These interdependencies are defined by levels of both *granularity* and *sophistication* in the models. In fact, as seen in Figure 3.3, it can be claimed that the functionality of the models may be in function of these two domains. Noteworthy, given the fact that information is usually collected by different entities and at different times, both levels are defined by the ability to collect data, and the urgency to provide explanations about the abstracted phenomena.



(a) Black-box model – built upon restrictions or (un-)definitions

(b) Grey-box model – only some modules developed

(c) White-box model – focus on the internal dependencies of the abstraction

Figure 3.2: Usability and reproducibility of multi-layered models

Inevitably, the opted (or available) levels of granularity and sophistication, as well as the adopted ways to represent the empirical model, introduce bias and uncertainties, which, in consequence, impact on the exactitude and the precision of the model. Because of this, although the purpose of the analytical model is to represent the reality simplistically (see Figure 3.1), the rationale applied to these modelling strategies may result in variate levels of under- and over-estimations of parameters.

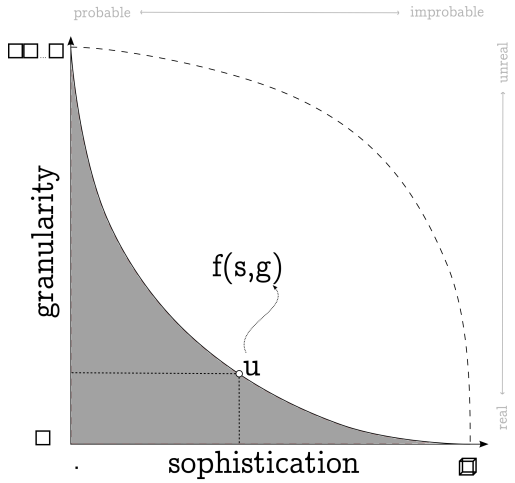


Figure 3.3: Balance between both granularity and sophistication on modelling techniques. The y-axis represents the level of detail which can be captured and parsed to models, whereas the x axis represents the ability to represent real-world data with accuracy, defined by the ability to abstract the phenomenon by maintaining the main properties in that exist in practice. A balance u is defined in function of both.

In this regard, it has been argued that as the quantity of required input data increases, so the average quality of this data reduces and the risk of input errors increases[25]. Figure 3.4 shows that the rate of error increases linearly by the difficulty to gather data, either by the instruments or the people involved; but also an error increases when more elements are applied to the model. An overall error is also shown in Figure 3.4 with a view to representing the combination of both assumptions. This is also pointed out by Clarke [33], who claims that the balance of a functional model affects the sophistication of the adopted algorithms; the author also points out that the accuracy of the model depends on the resources (i.e. computational tools, time and human resources), as well as the chosen abstractions by modellers.

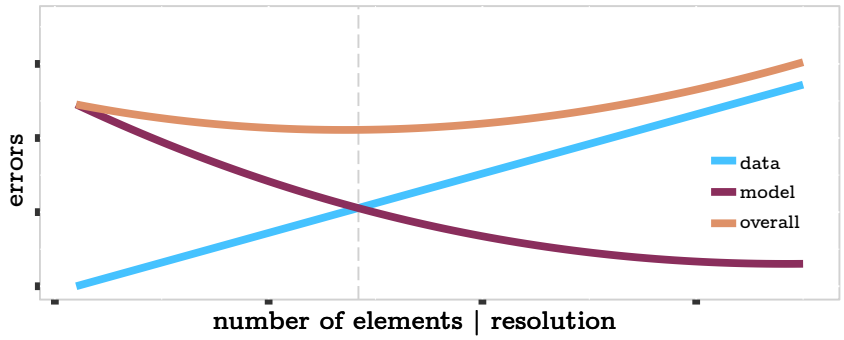


Figure 3.4: Uncertainty expectancy by level of granularity
Source: [25]

A particular condition on modelling approaches appears as a result of the adopted *level of sophistication*. In this regard, when the abstraction to capture the phenomenon requires more independent inputs than the explicit ones, the model is typically considered a *complex* one. Formally, this property emerges either when some properties demanded by the model are unrelated to the common algorithms employed, or when the algorithms are incapable to capture the phenomenon via formalisms [69, 158]. Typically, when a phenomenon is part of a multidisciplinary framework, and hence some specific inputs are required for each discipline, it is more likely that the resulting model is developed on this domain.

It can thus be claimed that modelling strategies are defined by the triad formed by inputs, interdependencies and outputs. These components can steer the scope of these models towards very different approaches. A very common approach among building energy models is the so-called *simulation*, which is defined as the application of a model to represent a phenomenon via numerical results. In this respect, Cacciabue [22] presents a formalisation on both model and simulation terms. Here, the author explains that models are focussed mainly on theoretical aspects and descriptions of the phenomena, and simulations exist around the numerical transformation of the presented arguments. However, in building energy models, it is common practice to use both terms interchangeably, disregarding the semantic differences. Yet, when the models need to be more specific, the distinction is thus relevant to bring about the differences between analytical approaches and simulation capabilities.

Moreover, although the classical definition of modelling and simulation suggests the provision of results when the collection of empirical data is infeasible or impractical, the application of a modelling approach may be rather focused on some other outcomes, such as gaining insights from a particular phenomenon, adapting multiple models, simplifying processes, accounting for

uncertainties and evaluating process of decision-making [7, 120]. It is therefore important to understand the scope of the modelling approaches, a well as the underlying rationales applied to them.

The following section provides a summary of different classes in the most common modelling approaches. This is useful to establish a theoretical background in the study of HSEMs, and to identify the attributes from these approaches to the evaluation of energy flow pathways in the housing stock.

3.1.1 Modelling Domains in Building Energy Models

The rationale applied to these modelling strategies provides the initial approach to develop algorithms, and to demand properties on the inputs. Such a rationale may be described in terms of four domains: a) logic in construction, b) nature of the symbolic model in the collection of data, c) nature of the symbolic model in the interrelationships, and d) unit of observation [134, 12]. It must be considered, nevertheless, that such distinctions are strongly correlated and (some) highly codependent. A description of such domains is provided below:

+ ***By the logic in construction.*** This domain is related the methodological structure to representing and assembling components in the modelling approach from a chosen perspective. This can be from the general to the particular, taking into account the relationships of the whole system, in which a deductive model is adopted. On the contrary, when the parts are used to explain the whole, and the focus is on the constituent elements, an inductive model is suitably adopted. Such *logic in construction* in building energy models is commonly known as top-down (i.e. around the macro) and bottom-up (i.e. around the micro), respectively.

+ ***By the nature of the symbolic model: origin of data.*** This is related to the way data is initially collected. When inputs are taken as a snapshot of the phenomena, whether by simplicity or expertise availability, such variables are defined as static. By contrast, when evolving alterations are considered among the set of inputs, then these variables are considered dynamic. In addition, when the static elements are linked to the body of inputs, the model is defined as analytical, suggesting the use of correlation methods. Conversely, when the model adds such variations within the representation, the model is commonly defined as operational; operational methods usually become the basis of simulation approaches.

  A different classification may be considered by the method employed to obtain the information of the variables. For instance, they can be continuous or discrete, which in turn can be further distinguished between qualitative and nominal. This classification is highly dependent on the collection processes (i.e. scope, measurement instruments, human resources).

+ ***By the nature of the symbolic model: relationships.*** The relationships are defined by the way the model is drawn and the variables are treated in the model. This is close related to the projected expertise by modellers, who may define different abstractions for the same phenomena. In addition, the link between variables may be capable to represent feedback and, consequently, to capture a property of memory on each of the variables. Conversely, the previous state of the variables can be simplified or neglected, but with faster calculations.

Similarly, when the variables are probabilistic independent and every run of the model produces the same outcomes, the system can be considered as deterministic; on the contrary, when the variables operate including a level of randomness and the outcomes are different for each test, then the system can capture a level of stochasticity. Because of this, deterministic variables may be easily controlled rather than their counterparts, which tend to be difficult to track due to their dependence on unpredictable contextual variables, and also due to their unique responses.

+ ***By the unit of observation.*** Subject to the purpose of the model, the *unit of observation* can be fed by intrinsic and extrinsic properties, leading to complex or, at least, more extensive representations. Normally, this unit may be inferred according to the adopted logic in construction, although not explicitly. For example, the unit of observation defines whether to consider the whole stock, specific sectors, individual buildings, households or isolated inhabitants. Therefore, the processes to collect data are different according to the units of observation.

  Besides, once the unit is specified, the concept of *ageing* may be added to the variables. In this way, not only the ability to accurately forecast outcomes is enabled, but also the ability to calibrate the models with more granular units is increased.

The different classes are highly dependent among them, and complement the modelling approaches. In here, the adoption of different units, dependencies and strategies, implies variations on the overall levels of sophistication and granularity of the models (see Figure 3.3). Models employ these rationales to emulate the way buildings behave [32]. Moreover, models cannot be compared in terms of their properties only, but in terms of their usability and appropriateness to the very specific needs that modellers and users can have. Hence, according to the applied rationales in the abstraction processes, models can be used to forecast, to describe, to explain and/or to test the energy performance in buildings.

Domestic buildings, as a subtype of all buildings, satisfy the heating demands in specific ways, especially because they are related to inhabitants' needs, which are unique for each dwelling. As a result, two dwellings with the same physical properties are more likely to have different demands due to inhabitants' behaviours, whereas two non-domestic buildings may have more similar demands if both are controlled by management systems. This is true provided that dwellings satisfy the purpose they were built for, which is to provide shelter from the weather, but also to provide security, comfort, and privacy. In this regard, by making decisions to reach satisfactory indoor conditions, a dwelling may limit flows of daylight, air, noise, moisture, and heat. In consequence, the development of domestic energy models is vast because of the adopted rationales and domains, which can be highly specific about users, buildings and interactions between them.

3.2 Development of Housing Stock Energy Models in the UK

3.2.1 Classification and Use of Energy Models

Historically, domestic energy models have been mainly focussed on the description of heating energy flow pathways. Some of them have been focused on a single part of the building (wall, window, device), others on the whole construction, and others on a group of conjoined buildings. If required, these models have been extrapolated to represent a stock of dwellings. This process of extrapolation is referred to as a Housing Stock Energy Model (HSEM) approach.

As noted earlier, in order to accurately describe the variables involved in the energy flow pathways, at the domestic sector, two wide modelling classes of HSEM have been developed: *top-down* or *bottom-up* [230, 215, 138]. Top-down approaches follow a deductive method, and start at a macroscopic level by considering measures such as appliance efficiency labels, energy price, or weather conditions. Bottom-up approaches follow an inductive method, considering measures such as internal conditions, building properties, or system usage. Each approach can be further categorised by statistical and physical modelling methods. Statistical methods generally apply regression techniques to describe and predict phenomena, whereas physical methods typically utilise simplified analytical models of physical phenomena [16, 215, 129, 138]. Another term used for describing physical methods is: *engineering* [215, 129].

It is common for HSEMs to combine the virtues of these two approaches, by calibrating the predictions of physical models with statistical data that describes a stock; primarily employing bottom-up strategies [128]. Figure 3.5 presents a typical relation between top-down and bottom-up approaches, and their convergence in the evaluation of energy demands. Figure 3.5 also shows that it is common for top-down approaches to be supported by economical studies, whereas for bottom-up ones, they are commonly supported by technical studies.

Figure 3.6 exemplifies the adoption of more detailed classes according to the nature of the symbolic models[1]. The application of a specific approach has depended, hence, on the purpose of the model, on the available tools to collect data, and on the sophistication of algorithms to capture energy flows in dwellings. More specifically, Figures 3.7 and 3.8 illustrate further distinctions adopted by HSEMs to explore particular foci with models. For example, Figure 3.7 illustrates a particular subdivision of these methods by adding a category for individual dwellings methods focused on buildings' detailed properties. And Figure 3.8 presents some of the most popular methods to study energy demands in the stock of dwellings.

In the UK, the first attempts to model energy flows in dwellings were made in the mid-1970s [133, 138], as a consequence to the oil crisis in 1973 [58]. They supported the identification and testing of strategies to reduce building energy demand, as well as the formulation of building regulations to enforce them. These first attempts were constrained by computing power, availability of dwelling information, and the ability to process it [215]. More recently, the adoption of more sophisticated algorithms have improved the accuracy, and usability of HSEMs [18, 109]. This has been facilitated by improved computing power, as well as increased availability and

[1]The International Energy Agency (IEA) has developed a formal update of modelling techniques, which are then adopted on the development of HSEM. An example of such an update is presented on [5], being this a relevant reference to present existing tools developed around the modelling approaches

resolution of representative data.

Figure 3.9 displays the temporal evolution of UK HSEMs, and shows the underlying energy-related policies, guidelines, and national programmes—described in Section 2.2—that motivated their incremental evolution. Most models predict the energy required by systems and appliances to satisfy occupant demands for comfort and services. Incidentally, Figure 3.9 also shows the significant increase in domestic energy models in the 1990s, which corroborates that significance change in the national strategy to improve the domestic sector, contrary to the persistent resistance to investment on energy efficient measures that preceded this period [96]. Advances in computational power significantly contributed to the development of HSEMs, especially due to the ability to incorporate Graphical User Interfaces (GUIs) and visual aids to the core of these modelling tools. A summary of existing HSEMs used to estimate the energy demand of the UK housing stock is described in the following section.

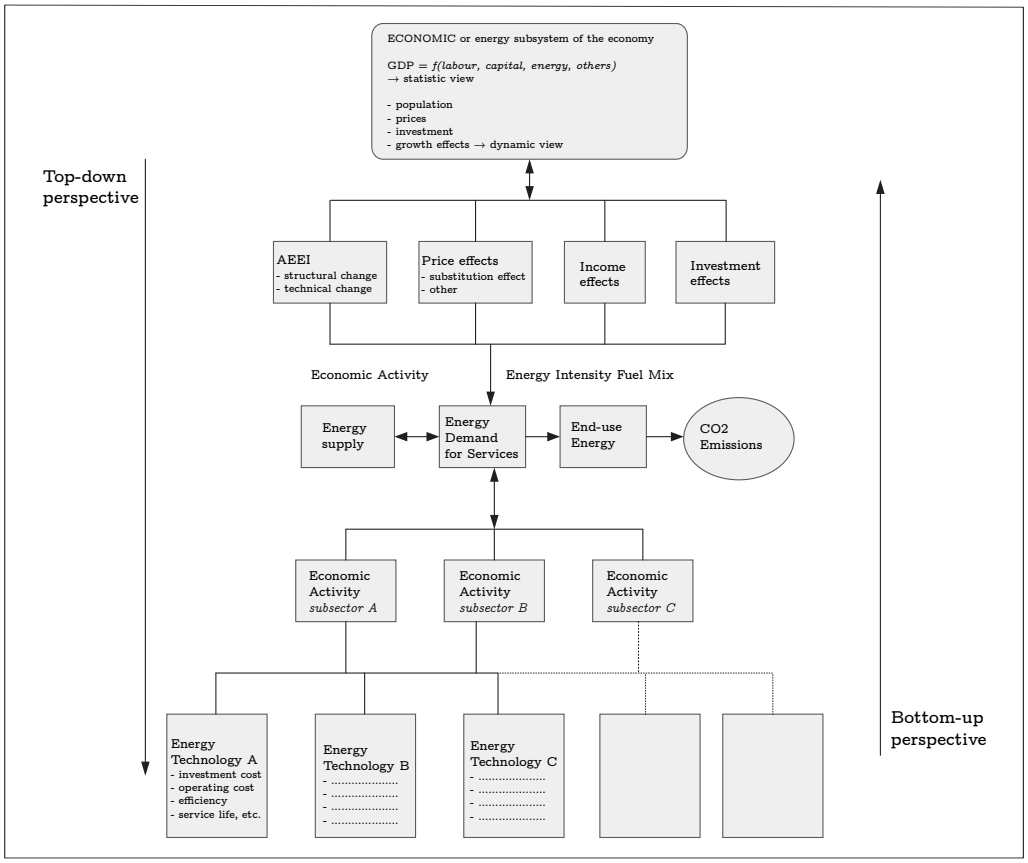


Figure 3.5: Classification of energy performance modelling: top-down and bottom-up
Source: [114]

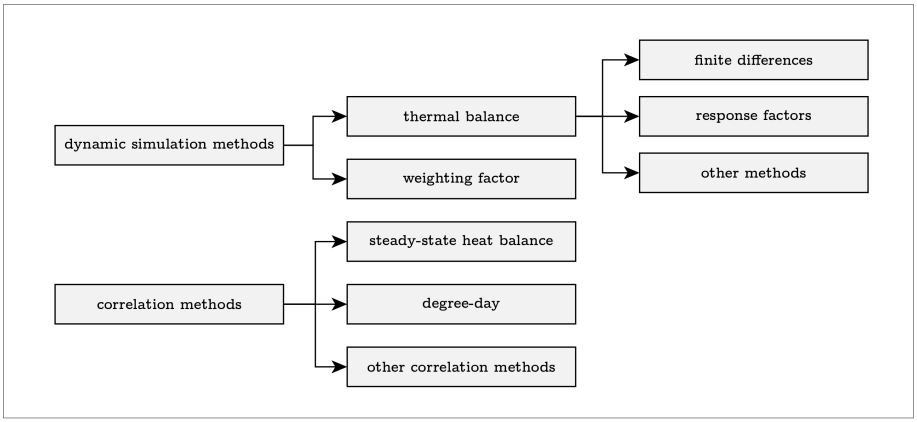


Figure 3.6: Classification of analytic approaches
Source: [113]

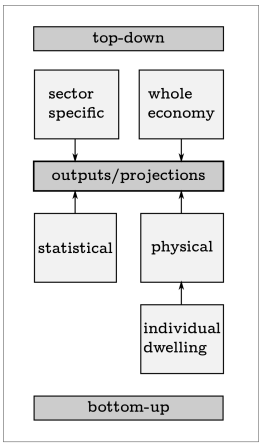


Figure 3.7: Top-down and bottom-up approaches in domestic energy modelling
Source: [138]

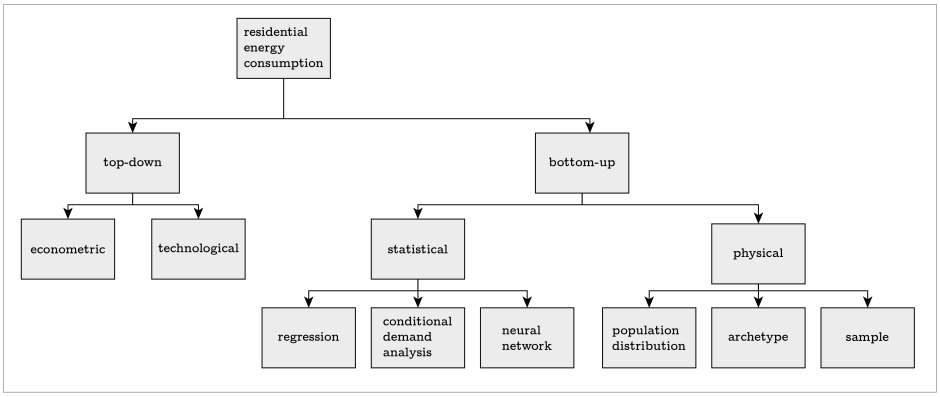


Figure 3.8: Classification of domestic energy modelling approaches
Source: [215]

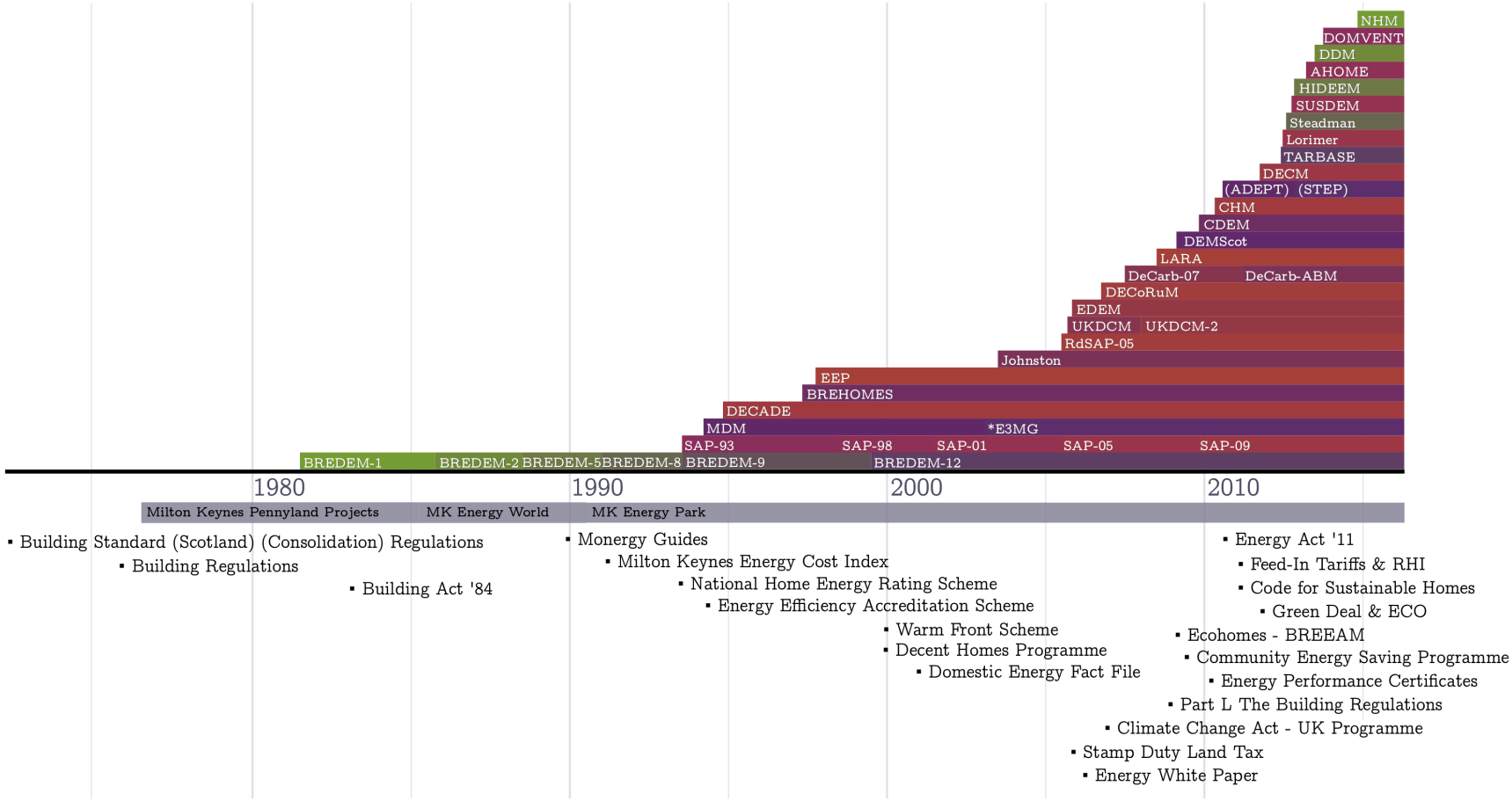


Figure 3.9: Evolution of HSEMs in the UK
adapted from [133]

3.2.2 Overview of Housing Stock Energy Models in the UK

Building Research Establishment Domestic Energy Model (BREDEM) [200, 95, 142] – The first version, BREDEM-1 (1981), was developed as a method for calculating energy use and fuel requirements in dwellings. Only two zones within the model were considered for a temperature balance sub-method; in the second version, BREDEM-2 (1985), a degree-day algorithm was used maintaining simplified considerations of the same model. In BREDEM-5 (1988), also known as *Energy Auditor*, other specific parameters were included in the analysis such as space heating, water heating, cooking, lighting and appliances. Contrary to the previous versions, a multi-zone model was improved here; nonetheless, some models based on the BREDEM still use two zones as a reference. A relevant output of this version was the MKECI—see § 2.2. In the following version, BREDEM-8 (1991), monthly predictions of energy consumption were generated employing a few samples of the EHS. During this stage of development, a robust version of the *Monergy Guide* was generated in order to diffuse energy saving aids to householders. In BREDEM-9 (1992), a degree-day algorithm was systematically applied, and a number of additional features were included such as heating profiles, renewable resources and monthly demand for household appliances. The sample was increased to 50 dwellings selected from the EHS. Finally, the last stable version: BREDEM-12 (1999), was able to make yearly predictions of energy consumption validating their results in 47 dwellings chosen from the EHS. It is worth mentioning that the algorithm was improved by pairing it with the ISO9164[2]; the standard defines a heat balance equation to calculate the annual energy requirements for space heating in dwellings, excluding passive solar design, greenhouses and sun-spaces. Figure 3.10a illustrates the algorithm employed in BREDEM; the Figure shows an example of how to re-conciliate data by combing theoretical with empirical data.

Standard Assessment Procedure (SAP) [57, 133] – This procedure, compiled by DECC and Building Research Establishment (BRE), consisted of a correlation between energy performance of dwellings, and economical-related variables taken from previous versions of the EHS; the outcome of this procedure helped in the development of a regulatory assessment tool, based on the BREDEM algorithm. The application of this tool resulted in the formalisation of an energy performance index (similar to the MKECI), later incorporated into the *Part L: Building Regulations*. By 2005, a Reduced Standard Assessment Procedure (RdSAP) was published in order to project retrofit improvements using estimations of a sample of 55 dwellings. The development of SAP-98 was initially focused on the energy balance of a single zone in 42 dwellings; then in SAP-01, a specific carbon index was added in conjunction with the economical correlation index. In addition, the rating scale was improved for the chosen archetypes. For the SAP-05 version, other parameters were included in the procedure such as lighting calculation, dwelling CO_2 emission rate and environmental impact rating; these parameters were calibrated by testing 177 archetypes. Likewise, a supplementary calculation for the SDLT was added as a result of this updated model—see § 2.2, providing some exemptions to zero carbon homes. Finally, in SAP-09, two relevant factors were improved: environmental impact rating, and monthly space and water heating demand.

Johnston [123] – This predictive BREDEM-based model, considered only two notional dwelling types of dwelling archetypes to project a stock transformation in 2050. Such a transformation estimated a reduction of 82% in carbon emissions. The archetypes to represent the stock were

[2]ISO 9164:1989 Thermal insulation - Calculation of space heating requirements for residential buildings [115]

selected from the EHS and weighted accordingly. Figure 3.10b illustrates the simplified algorithm employed in this model to evaluate scenarios.

Building Research Establishment's Housing Model for Energy Studies (BREHOMES) [198] – This model, developed initially in 1997, was constructed to represent the energy demand in the entire UK housing stock, the model included social driven factors to calibrate the calculations based on the BREDEM. Then in the 2005 version, 1000 dwellings types, extracted from the EHS, were considered for a 2020 projection analysis (later extended to 2050) employing a stock transformation method, and including additional data sources. Finally, the results were compared (and then projected) with DUKES, so that a baseline energy demand might be obtained, to be then analysed by developing scenarios of improved energy demand.

Domestic Energy and Carbon Model (DeCarb) [163, 164] – The first BREDEM-based version, DeCarb (2007), was created to model a stock transformation for 8064 dwelling types clustered by age class defined on the EHS, by validating a set of policy alternatives to achieve 60% reductions. The core used on the model was BREHOMES. Then, by 2011, an agent-based approach was added, aiming to capture controlled and uncontrolled variations over the future and, more importantly, to represent the impact of building occupants through their response within the building and within their community (context). Figures 3.10d and 3.10e illustrate the algorithm employed in DeCarb; the approach shown here is similar to the Johnston's model. The addition in the Agent Based Modelling (ABM) version employs additional parameters to feed the model and to stimulate an agent-based environment.

Energy and Environment Prediction (EEP) [127] – This tool was created to estimate energy use, neighbourhood quality, home hazards and mouldy issues at an urban scale. The information was extracted from surveys and matched into Geographical Information System (GIS) maps. In this way, it was possible to identify potential areas and regions where policies might be effectively applied. The analysis of energy employed the SAP methodology. Figure 3.10f illustrates the hierarchy of sub-models attached to a GIS database.

Domestic Energy Model for Scotland (DEMScot) [197, 24] – This model was created in 2009 to evaluate a set of interventions in the Scottish housing stock such as fabric insulation, upgraded heating systems, and installation of renewable technologies; and to estimate the reduction of GHGs. The DEMScot based its calculations on SAP-05 and selected 3146 archetypes from the Scottish House Condition Survey (SHCS). Prominent attributes of this version were the inclusion of more sophisticated algorithms to represent household appliances (e.g. lighting, cooking devices, devices), to emulate behaviours and to forecast stock evolution (e.g. population growth, demolition rate). The updated version DEMScot-2 included econometric algorithms to emulate price elasticity and rebound effects. These effects include, for example, the increase in internal temperature, due to lower costs for heating; and the more intense use of household appliances, due to the savings from an energy-efficiency improvement.

Community Domestic Energy Model (CDEM) [74] – A group of 47 dwelling types was considered in this stock transformation method, based on BREDEM. In addition to the study of the stock and the projection of scenarios, this model a) implemented a method to limit the stock to regions based on disaggregated location data, and b) developed a two-levelled sensitivity analysis for the characterisation of the stock (i.e. linearity and superposition tests). It is worth noting that the module is focused on the analysis of the most significant parameters that affect

carbon emissions. Figure 3.10h illustrates the algorithm employed in CDEM; the Figure shows the multiple datasets used to create the 47 archetypes, and the calibration of results using Census information.

Cambridge Housing Model (CHM) [110, 142] – This model, developed by Cambridge Architectural Research for DECC in 2010, was generated to estimate household energy consumption. It employed BREDEM at its core and complemented some parameters with the SAP catalogue. The CHM evaluated the EHS, and theoretically the SHCS, to estimate energy use, emissions by fuel type, and a rating scheme related to SAP. In the 2011 version, 14,951 EHS archetypes were used to assess the energy performance of the English housing stock. The CHM adapted 117 variables from the 2936 available on EHS, related to geometry, constructions, materials, and energy related systems within. A relevant outcome of the CHM has been the support of the *Housing Energy Fact File* for DECC—which replaced the *Domestic Energy Fact File*, and the *Energy Consumption in the UK* studies. In here, the results of the CHM are weighted to the UK stock applying a pro-rata extrapolation (i.e. the proportional population Eng-UK-Sco, Eng-UK-Wal and Eng-UK-NI). Figures 3.10i and 3.10j illustrate the algorithm employed in the CHM.

Domestic Energy and Carbon Model (DECM) [31] – A social-related model predictor was deployed in DECM based also on BREDEM and SAP. In this model, household employment status was used to calibrate the energy model projections by estimating accurate occupancy profiles on the 50 archetypes selected from EHS. Given the fact that a statistical based method was used in the calibration, a rebound effect, based on historical data, was added to the predictions from the analytical model. Additionally, a local sensitivity analysis module was added to the core of the Domestic Energy and Carbon Model (DECM), so that the most influential parameters might be identified.

Stochastic Urban Scale - Domestic Energy Model (SUSDEM) [17] – This stock model was developed with the purpose of calibrating micro-level engineering models with macro-level data. The SUSDEM was conceived in three stages: clustering, which allowed the analysis of dwellings with similar characteristics; calibrating by reference values; and calibrating by probabilistic properties in the inputs. Such distinction enabled the possibility to account a diverse range of uncertainties. In addition, it increased the opportunity to test specific inputs, despite the inherent stochasticity on both occupancy and behavioural profiles. Figure 3.10k illustrates the algorithm employed in SUSDEM; the Figure further describes the methods applied to study a small group of dwellings, a medium group and the entire housing stock.

Domestic Dwelling Model (DDM) [157] – This model was developed as visualisation tool to study a stock of dwellings within a neighbourhood. The model used the both SAP and the BREDEM to perform energy calculations. An attribute of this model consisted in the development of a preliminary domestic model ontology via an Unified Modelling Language (UML) catalogue. An application of such an architecture was its straightforward adaptation to a GIS database, which provided inputs, and became a promptly reference to compare the model outcomes. Figure 3.10n illustrates the algorithm employed in DDM; the Figure describes the steps employed to extrapolate changes in the domestic stock energy demand, and to emulate processes of decision-making.

Domestic Ventilation Model (DOMVENT) [126, 124] – This tool was developed to study infiltration and exfiltration in theoretical archetypes. DOMVENT employs the EHS to construct

horizontal and vertical surfaces of the external envelopes and assumes data related to airflow rate, internal pressure and façades properties. Although the model was limited to a very specific area in the abstraction of energy flow in dwellings, it provided basis to generate volumetric models from survey data.

Health Impact of Domestic Energy Efficiency Measures (HIDEEM) [125, 159] – This model was supported by DECC with the purpose of analysing the indoor environmental conditions, and monetising the health impact associated with energy efficiency changes. These conditions included, for instance, temperature, concentrations of particle pollution, second hand tobacco smoke, radon, and risk of mould growth. The HIDEEM was nested with the study of energy efficiency interventions projected for the Great Britain's housing stock.

Agent Home Owner Model of Energy (AHOME) [139] – This model was developed to simulate interactions from the consumer (home-owners), so that a set of external changes might be included in the estimation, such as the effect of carbon tax policies, applied to alter decision-making in consumers; and the diffusion of behavioural changes, related to the investment on energy-efficient technologies. Hence, the AHOME evaluated the interaction between (isolated) individuals and (collective) behaviours, on a highly disaggregated stock model; the model also estimated the reduction of carbon emissions in a long term scenario. Figure 3.10l illustrates the algorithm employed in AHOME; the Figure illustrates the representation of decision-making and the development of different scenarios.

National Housing Model (NHM) [3]– This model, to be published soon, has been supported by DECC and Centre for Sustainable Energy. It aims to test improved policy and technology scenarios. To achieve so, it combines information from EHS and SHCS, and focusses on the properties of buildings and occupants. The NHM is meant to improve on the evaluation of microsimulation parameters and sensitivity analysis.

As mentioned earlier, statistical-biased methods typically rely on correlations and fixed values. Besides, even though they are built as bottom-up models, they are commonly calibrated in alignment with top-down models. These models analyse macro-level data and predict outcomes according to long-term consumption patterns, appliance ownership, weather data, and general energy-related policies.

Domestic Equipment and Carbon Dioxide Emissions (DECADE) [15] – This model, developed in 1994, estimated a transformation of the stock in terms of ownership of goods, power consumption and energy use intensity. The DECADE model evaluated 1000 archetypes taken from the EHS. Essentially, such transformation was estimated by assessing the rate of penetration of new technologies, based on historical data about sales, ownership, technical characteristics and usage profiles. The stock projection was built to improve the process of decision-making, by further analysing both environmental impacts and socio-economical issues; and by emulating the application of policies destined to improve better practices across households. To study the diffusion and likelihood of uptakes, a set of additional topics were added to the DECADE model such as water consumption, chain manufacturing processes and recycling of materials.

[3] www.cse.org.uk

UK Carbon Domestic Model (UKDCM) [13, 98, 3] – In this BREDEM-based model, a 2050 scenario was estimated using 384 multi-zone archetypes from different regions by combining typologies. In the UKDCM-2 version, three main scenarios for the UK housing stock were estimated, by considering fluctuation on climate events, energy price shocks and direct consequences of political changes. Figure 3.10c illustrates the algorithm employed in UKDCM; the Figure exemplifies the combination of datasets to develop the 2050 scenario, the modularity of sub-systems, and the calibration of outputs using carbon emissions.

Domestic Energy, Carbon Counting and Carbon Reduction model (DECoRuM) [79] – This BREDEM-based model integrated specific information of dwellings such as: age of construction, built form, primary uses of appliances, solar potential and location, to predict technology penetration, energy demand and carbon emissions. In brief, the DECoRuM was created as a visualisation tool, able to merge GIS with energy performance information.

Local Area Resource Analysis (LARA) [68] – This project was created to analyse the flows of energy consumption in specific small geographical areas: *UK National Output Area Classification (OAC)*[4]. This model defined some basis to combine socio-economic characteristics of buildings archetypes with highly aggregated information of the places and surrounding environment. Figure 3.10g illustrates the algorithm employed in LARA; the Figure shows that household-related data is fed into each of the households, which are in turn linked to local surveys and census via OACs.

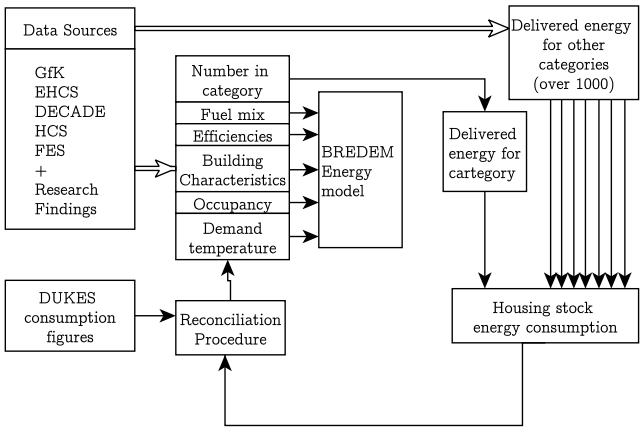
(Energy Systems Research Unit) Domestic Energy Model (EDEM) [34] – An improved analysis of carbon and energy performance of housing (technologies, behaviours and environmental factors) was deployed in this statistical model (2005). The archetypes were chosen and calibrated employing the SHCS. The aim of this model was to implement a model capable to account for archetype analysis, as well as for stock transformation, providing thus a reliable tool for decision-making.

Market Allocation (MARKAL) [3, 138] – An analysis of energy-related policies and estimations was deployed in this MARKAL (2005) predictor. It must indeed be considered that these type of modelling approaches have been extensively developed in the past; however, this model supported by the IEA, was made as a predictor of energy policies in the UK. The model estimated emissions, by testing individual technologies at the bottom-up level. Noteworthy, this tool was originally developed to provide insights for the *Energy White Paper 2003*, being a multi-time period linear optimisation model, able to import different models into a comprehensive one.

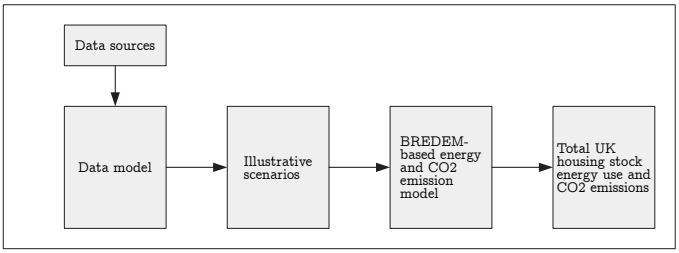
Lorimer's [142] – An adaptable regression of energy use (non-heating systems as lighting and appliances) for English dwellings was deployed in this model by 2012. The model employed only 2 archetypes from the EHS, and it was implemented to match and calibrate values, using more detailed surveys, so that the projection could employ aggregated values. For example, the Lorimer's model correlated electricity meters in small areas with broad census data. By doing so, the information gaps in the datasets were tackled with the use of specific surveys.

Steadman's [131, 87, 211] – The contribution of urban built form on total energy use indicated by gas and electricity consumption was measured in this statistical model (2012) based on the *Home*

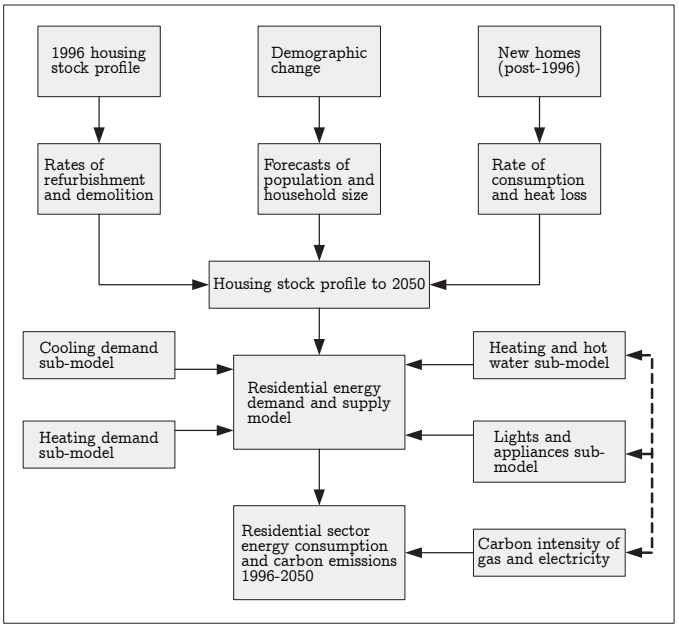
[4]Mainly developed for marketing purposes since 1971, areas of the country with similar characteristics have been defined in terms of six main census dimensions: demographic, household composition, housing, socio-economic, employment and industry sector.



(a) BREDEM - [200]

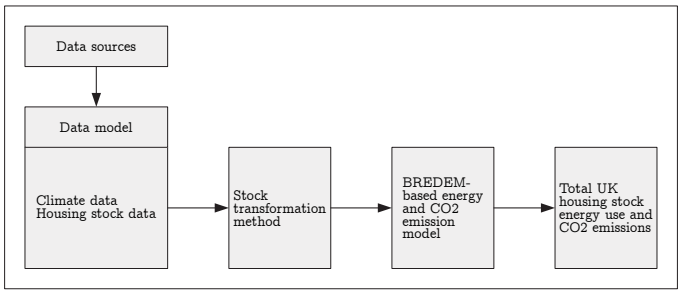


(b) Johnston's model - [122]

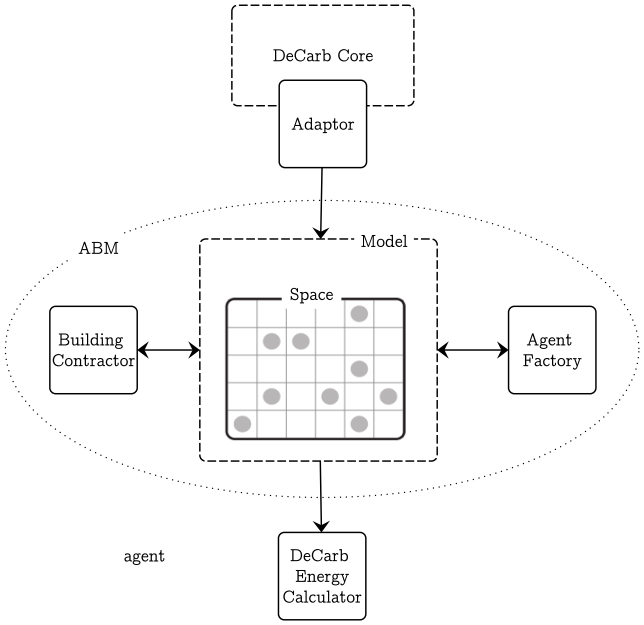


(c) UKDCM - [3]

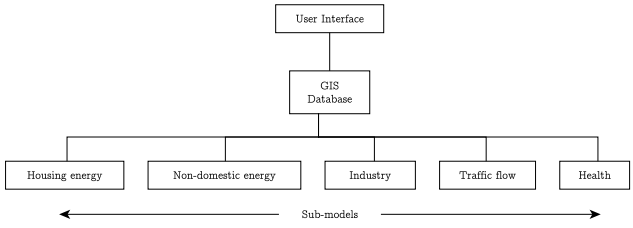
Figure 3.10: Diagrammatic representation of HSEMs



(d) DeCarb - [163]

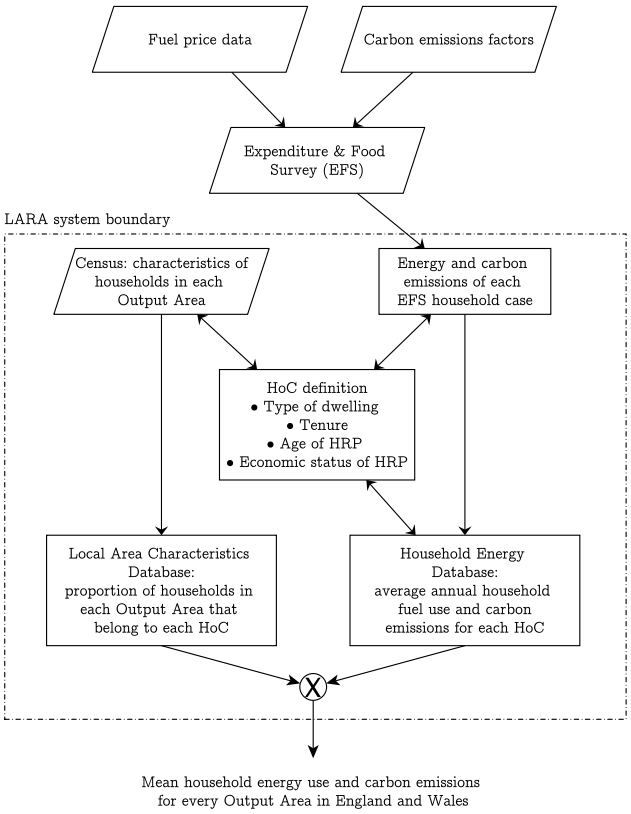


(e) DeCarb-ABM - [164]

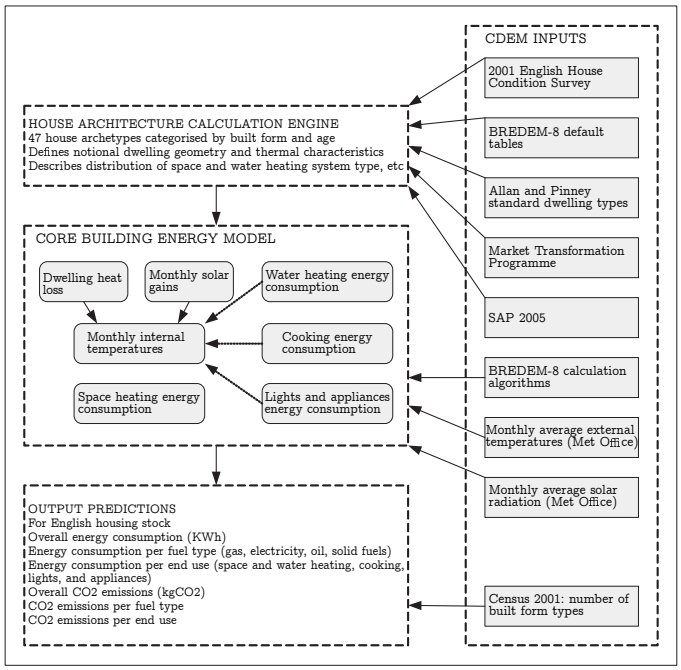


(f) EEP - [127]

Figure 3.10: Diagrammatic representation of HSEMs – continued

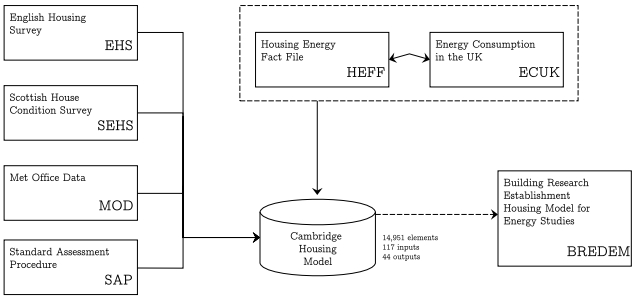


(g) LARA - [68]

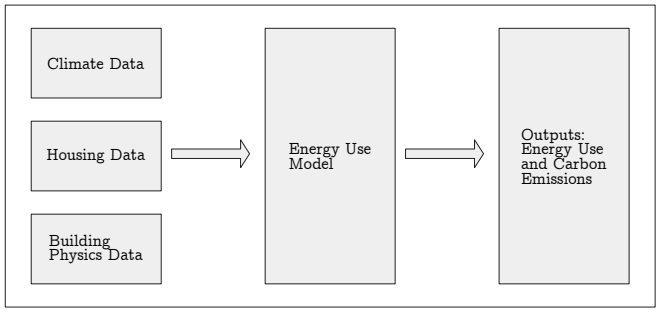


(h) CDEM - [74]

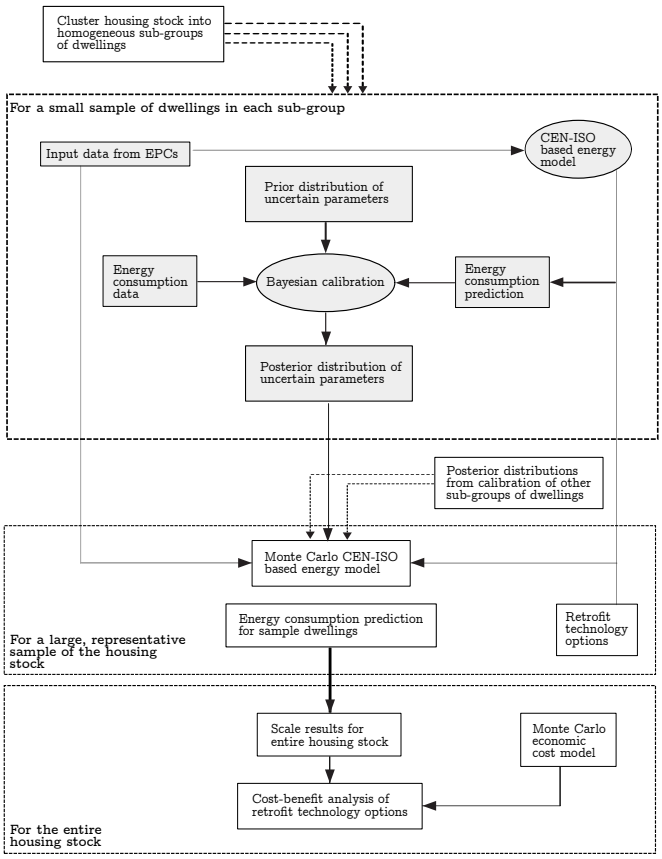
Figure 3.10: Diagrammatic representation of HSEMs – continued



(i) CHM - [110]

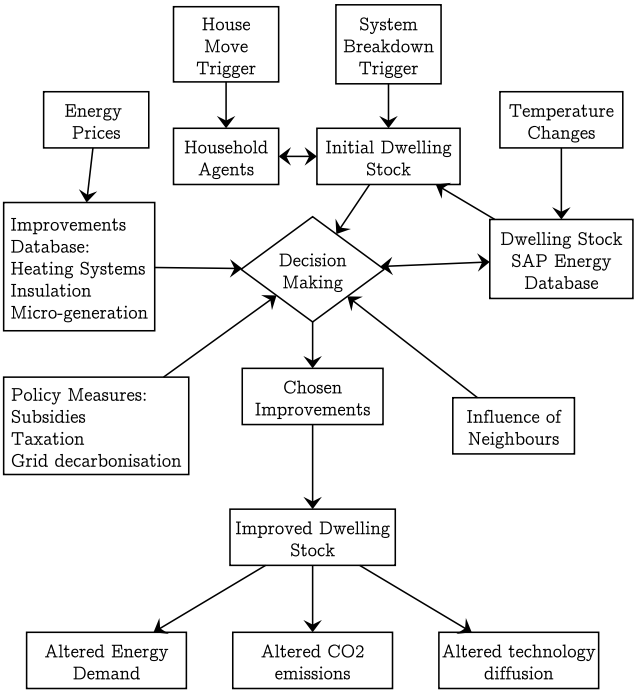


(j) CHM abstract - [110]

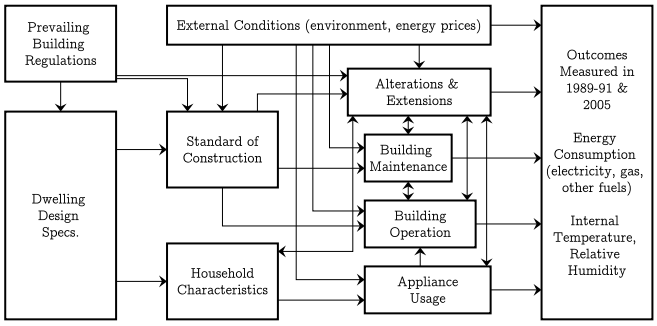


(k) SUSDEM - [17]

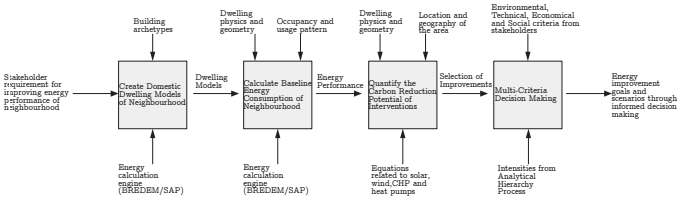
Figure 3.10: Diagrammatic representation of HSEMs – continued



(l) AHOME - [139]

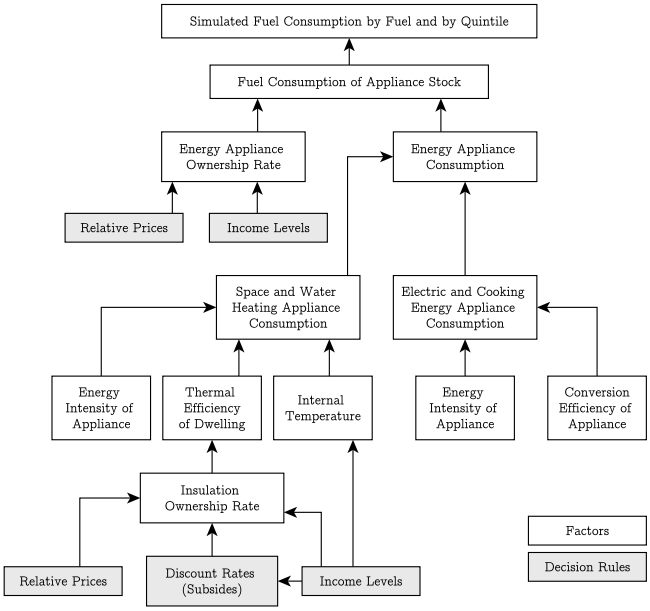


(m) Summerfield's model - [138]



(n) DDM - [157]

Figure 3.10: Diagrammatic representation of HSEMs – continued



(o) MDM - [42]

Figure 3.10: Diagrammatic representation of HSEMs – continued

Energy Efficiency Database (HEED), which represents around 15 million distinct dwellings (see Table 2.4). In sum, this model was built to estimate the impact of energy efficiency retrofits and, therefore, the applicability of energy-related policies, particularly on loft and cavity insulation, glazing and boiler replacement.

Technology Assessment for Radically improving the Built Asset base Domestic Model (TARBASE) [119] – An stock transformation model for carbon-saving improvements and refurbishments was deployed in TARBASE analysing DUKES figures. The model was implemented as an attempt to set the basis of future comprehensive models capable to calibrate top-down models. In here, a relative importance was given into the use of performance indices rather than standard guidelines, which increased the scalability of the model.

Typically, top-down approaches have been commonly associated to economical performance of the stock, mainly due to their ability to include non-energy related variables in an extended way (i.e. socio-demographics values, economical projections, prices of fuels). Sector-specific and whole-economy are two common subtypes used by such approaches [230, 215, 138] (see Figure 3.7). These models typically operate at the stock level applying correlation methods to describe and project macro-level data. A term which is typically used to describe these methods is *econometrics*, emphasising the weight of economical drivers.

Summerfield's [138, 129] – Two sub-models were generated by nesting DUKES figures: the Annual Delivered Energy, Price and Temperature (ADEPT) (2010), which calculated the trajectory of energy delivered to UK households; and the Seasonal Temperature Energy Price (STEP) (2012), which used polynomial multiple regression to relate quarterly delivered energy price index. Both models were intended to be accurate tools for testing the penetration of new technologies on

the market, with improved information on decision-making processes. Figure 3.10m illustrates the algorithm employed in these models; the Figure illustrates the multiple dependencies among the components in the stock, so that different configurations of technologies can be tested, in addition to improved operation profiles.

The Domestic Energy Sub-Model [9, 42] – This model was deployed initially in 1994 and integrated into the Cambridge Multisectoral Dynamic Model of the British Economy (MDM). Then, the 2003 version included a multi-agent dynamic model for the projection of archetypes; this model was part of the Energy Economy Environment Model at the Global Level (E3MG), as an approach for combining top-down and bottom-up models, by considering: atmospheric emissions, electricity supply, combined heat and power, and both domestic and transport energy usage. Figure 3.10o illustrates the algorithm employed in the MDM; the Figure highlights the factors influencing decision-making for each of the sub-systems composing the stock model.

In addition to the described models, Böhringer et al. [16] suggest an approach for defining a market equilibrium as a mixed complementary problem to bridge the gap between the two modelling perspectives (bottom-up and top down), despite the significant improvements to narrowing this with previous models. This approach has been an attempt to predict the evolution of the stock with an increased accuracy of the analysed archetypes, and to map the *decision-making parameters* in a parsimonious model, including policy measures, scenarios testing and micro-parameters.

To end the list, some other models focused on specific sectors or technologies [99], but directly impacting the domestic sector, are presented as following:

- **Energy Demand Model (EDM)**: econometric model to simulate responses of the whole energy system, forecasting aggregate properties and policy interventions at different sectors.
- **Energy Systems Modelling Environment**: a model to estimate space heat demand and use of heat pumps by season and other drivers.
- **Green Deal Household Model**: a predictor of energy changes by the installation of insulation.
- **Data Distribution Management**: despite not exclusively a domestic model, part of this model was made to analyse grid-level pumped storage capacity and yearly correlation between price and capacity in the domestic sector.
- **Transform**: a model to analyse the transmission network demand.

Remarkably, it is worth drawing attention that the differences between models, in terms of the physical and statistical approaches, can be very subtle. Such differences, in practice, are only an initial bias of the models, and they are commonly complemented by their counterpart approaches. For example, to study a stock of dwellings, models can initially consider *groups*, *clusters*, or *typologies* to apply suitable abstractions, which imply an approach bias. Nevertheless, selecting the appropriate model depends on the granularity of the statistical information, which can compensate the process of sampling and so the assumptions in dwelling's geometry, age and context to be later extrapolated [230, 215, 138, 140].

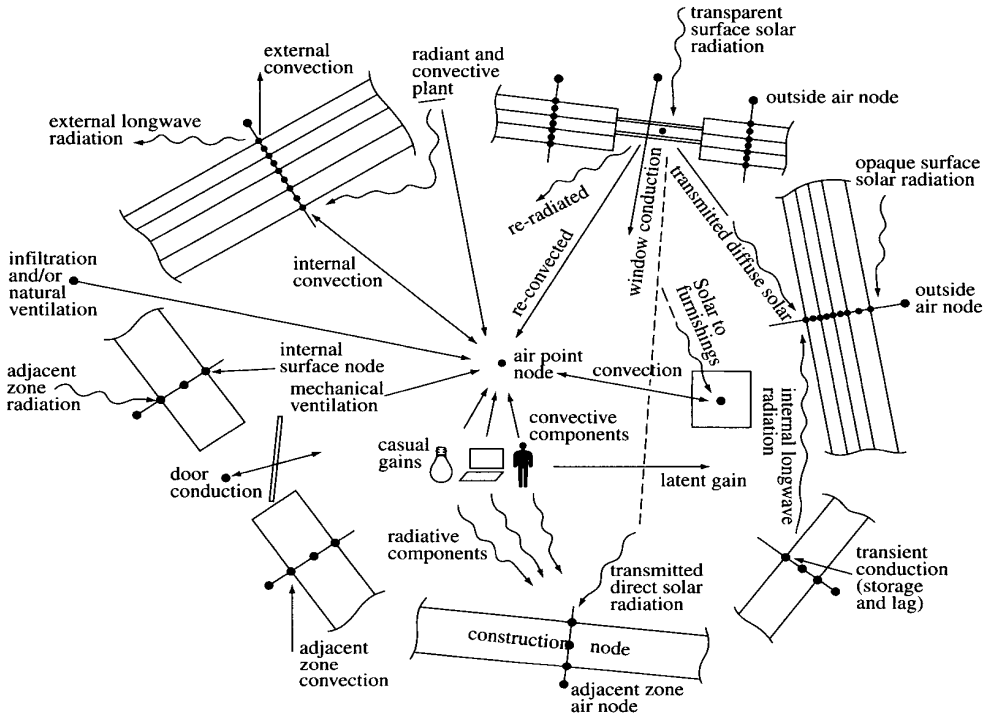


Figure 3.11: Energy flow pathways
[33]

The development of a Housing Stock Energy Model (HSEM) starts with a basic abstraction that captures the energy flow pathways in a single dwelling. This mainly comprises the heat transfer through the envelope (to or from the external environment or conjoined buildings), via conduction and associated surface convective and radiative transfers and by infiltration and exfiltration, as well as the thermal gains from occupants and appliances. A common analogy used to represent these energy flow pathways in a dwellings is via a resistor–capacitor network [113, 230, 5, 215, 138]. In here, currents are equivalent to the heat flows, and each component of the dwelling is connected by nodes. Clarke [33] develops a methodological assessment to represent each of the processes involved in such a abstraction, as exemplified in Figure 3.11.

There are, however, a couple of caveats that emerge when representing these processes in dwellings. Firstly, for a single building, it might be achievable to capture each of these components with relatively high accuracy, whereas for a stock of dwellings, this task may—very possibly—be restricted due to the resources and the time to perform the equivalent (and corresponding) analysis. Secondly, because each process is simultaneously active in practice, the computation demand may significantly increase and, in consequence, undermine the utility of the model too—see § 3.1. Therefore, for reasons of practicality, representations of these pathways (such as heat gains, ventilation and infiltration losses) may be simplified. A key example of a simplified abstraction of energy flow pathways in dwellings is the BREDEM (see Figure 3.12) which, as identified in the summary, forms the basis of many other UK HSEMs. In this way, such an abstraction for a single dwelling can be more conveniently replicated for a given housing stock, yet some abstractions focus on a very specific element of the energy models.

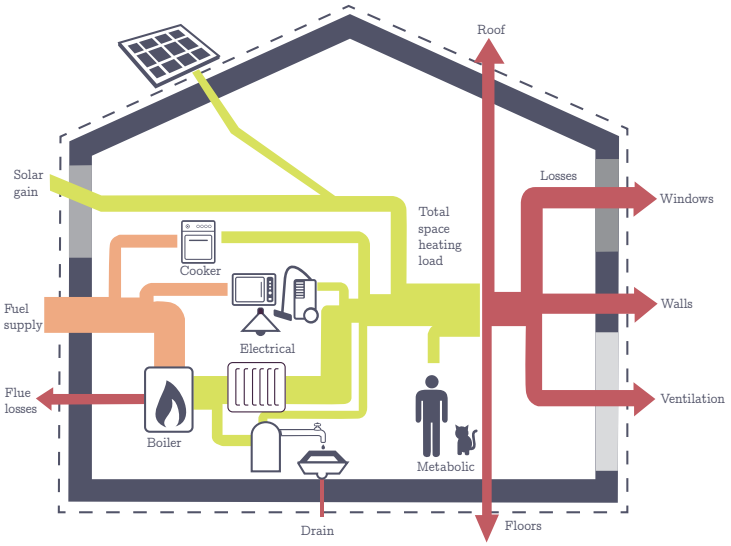


Figure 3.12: Schematic illustration of the energy balance principle used by the BREDEM [200] *Source:*

3.3 The Conformation of Housing Stock Energy Models

There are three important elements that are common to many HSEMs: 1) descriptions of building geometry and fabric; 2) analysis of the inter-relationships between dwelling components and zones; and 3) representation of occupancy and appliance usage within dwellings. The latter element is particularly complex and—as reviewed in § 3.2.2—is neglected by many models.

3.3.1 Building Geometry and Fabric

When modelling a single dwelling, it is possible to describe the attributes of the building with a high level of resolution. However, when modelling a stock of dwellings, only the most relevant properties, such as those for geometry and fabric, are used to describe physical attributes, because they are essential for understanding the energy flows in any building. It is also important to understand their composition (arrangement, layout) and their context (environment, climate).

BREDEM does not require an explicit three-dimensional representation of building geometry. Its energy balance calculation requires that heat gains and losses across each element of the envelope, infiltration, and internal heat gains, are represented as parameters [200]. Thus, the envelope is described by a series of elemental areas and properties supplemented by total floor area, A_T, and height, to determine the internal volume. The CHM [108] improves BREDEM's geometric representation of dwellings by including multiple stories and additional spaces, such as attics and basements. Consequently, additional parameters are derived to describe the physical space, such as living area fractions (allowing heating demand to be allocated more specifically) and linear scalar factors (used to calibrate the solar radiation gains).

The physical space is enclosed by surfaces constructed of materials with different properties. These elemental properties include heat transfer coefficients, solar reflectance values, and thicknesses.

If they are unspecified then default values can be assigned from guidelines (e.g. [95, 54, 83]). In BREDEM, the wall thickness is not a direct input, but it is indirectly included in a general thermal mass parameter, which is used to account for thermal storage of solar gains and gains from intermittently used heating systems. The CHM derives wall thicknesses and material properties from the age of construction and thus provides more detailed building fabric properties than BREDEM.

The resulting heat transfer is a direct function of the (unintentional) infiltration and (intentional) ventilation rates, which are represented by a number of different approaches. BREDEM uses a simple air exchange rate associated with both infiltration and ventilation that is a function of the average regional wind speed. The algorithm for ventilation is improved by the CHM, which includes the number of sheltered sides and a draught factor. More recently, the DOMVENT improved the prediction of infiltration in dwellings by generating façades and including detailed parameters that account for local surroundings [124]. This can in principle be used by HSEMs.

The context of a stock of dwellings defines the general composition of each of the buildings, and the external conditions that interact with them. The DECoRuM correlates the shape of a stock of dwellings with their energy demand. It uses a GIS to estimate energy demand as a function of dwelling size and location and estimates the risk of overheating. However, its ability to model the energy demand of dwellings is constrained by limited data and simplistic algorithms [80]. Similarly, the DDM model [157] links aerial imagery with census data. Both models improve the efficiency with which energy-related measures may be assessed and, by using GIS, can link detailed regional information about dwelling fabric and its context. Steadman's model also includes a built form parameter to explore general retrofits, but this is limited to Middle-SOAs and a very limited stock size [131, 87, 211].

Although dwelling geometry and orientation are both essential for an initial estimate of solar gains, the latter is only considered by BREDEM when estimating the energy contribution of photovoltaic cells, as a correction factor for incident solar irradiation. In contrast, the CHM explicitly considers gains through two theoretical windows to represents daylight-responsive artificial lighting use. Solar loads through walls are not represented by the BREDEM algorithm because of its simplified representation of dwelling geometry and of heat flow.

Therefore, studying building geometry and fabric is thus essential to initially understand the relevant components that influence the energy demand in dwellings. Most of this demand, as mentioned earlier, satisfies heating needs in dwellings (i.e. space, water and cooking). For this reason, it is important to develop an accurate characterisation of the physical factors affecting the heating demand in dwellings.

As a way of demonstration, Figure 3.13 presents a simplistic comparison between two houses with opposite levels of efficiency and measures of energy conservation; in this simplification, where neither metabolic gains nor external gains are considered, the average heat losses are accounted. The Figure provides an example of how influential are the fabric components in the residential heating performance, and provides a reference about the most influential components in dwellings. This explains the relatively prominent developed algorithms to describe heat transfers characterising physical components of the house, which is achieved using resistor–capacitor networks, numeric methods or regression analysis. A review of the rationale applied in HSEMs to capture these energy flow transfers are discussed in the following Section.

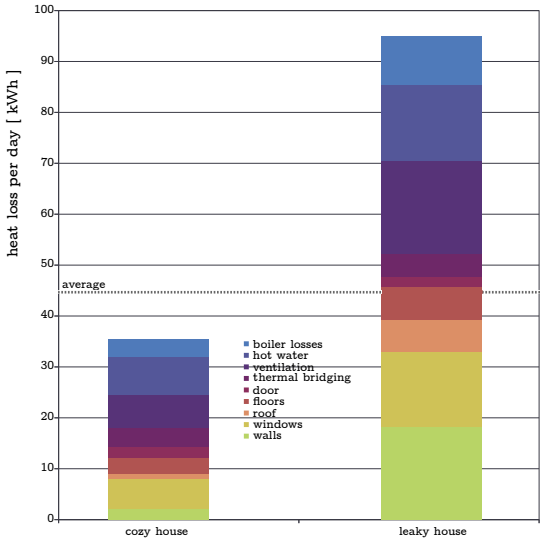


Figure 3.13: Typical heat loss in dwellings
Source: Adapted from [217]

3.3.2 Systems Components and Inter-Relationships

Section 3.3.1 described the conceptual approximations of energy flows used by HSEMs as a function of dwelling geometry and fabric properties, yet considering dwellings as empty boxes, which is a valuable and common initial approach. Below is described the abstract representation of dwelling gains and losses to explore in more detail the performance of a single building that can be extrapolated to a housing stock.

Abstraction of Energy Flow Pathways

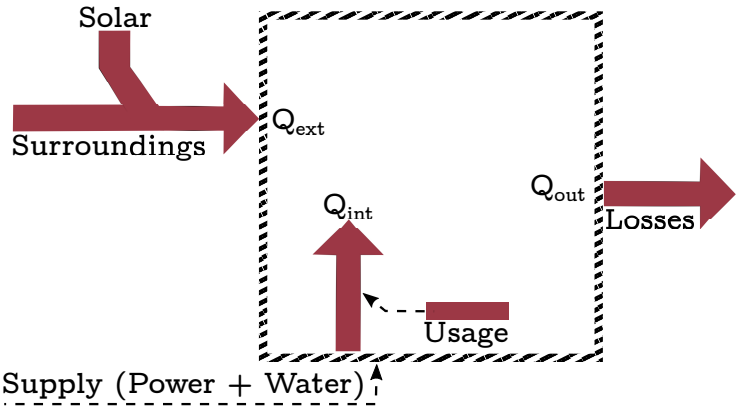
A classic representation of dwelling energy flow pathways is shown in Figure 3.14. Here, a simplified representation of energy flow provides the basis of HSEMs, and a more detailed representation of heat transfer summarises the components involved in the abstraction. It includes the sum of external heat gains $Q_{ext}[W]$ (net solar radiation gains and transmitted heat from the surroundings) and internal heat gains $Q_{int}[W]$ (occupants, appliances and systems), balanced with the energy lost to the environment $Q_{out}[W]$.

$$Q_{ext} + Q_{int} = Q_{out} \tag{3.1}$$

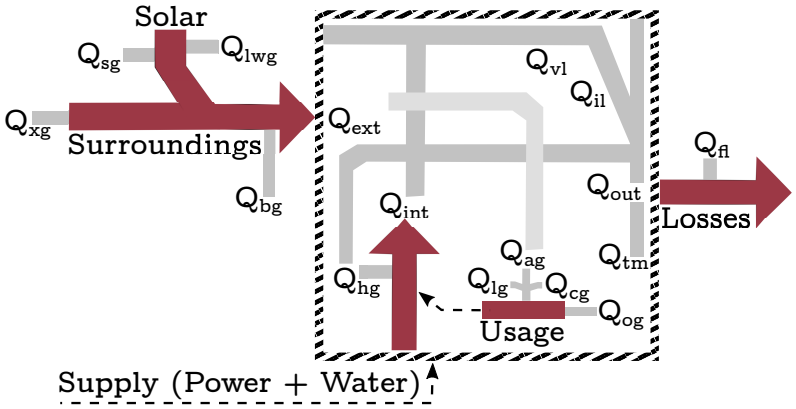
External gains are mostly caused by radiation (direct solar gains Q_{sg} and long-wave radiation Q_{lwg}). Other external gains are attributable to adjacent connected buildings Q_{bg} as transmitted heat from the local environment Q_{xg}.

$$Q_{ext} = (Q_{sg} + Q_{lwg}) + Q_{bg} + Q_{xg} \tag{3.2}$$

Internal gains comprise the indirect effects of lighting Q_{lg}, appliances Q_{ag}, cooking Q_{cg}, occupants



(a) The fabric of empty boxes: location, shape, orientation, ...



(b) System Components and Inter-relationships: context, supplies, usage, occupants, losses, ...

Figure 3.14: Simplified energy flow pathways in dwellings

Q_{og} and, more substantially, the direct effects of heating systems Q_{hg}.

$$Q_{int} = Q_{lg} + Q_{ag} + Q_{cg} + Q_{og} + Q_{hg} \tag{3.3}$$

The right side of Figure 2.9 disaggregates end-use energy demand and in so doing identifies the relative contributions of internal gains to Q_{int}.

The losses, Q_{out}, can be directly transmitted through the fabric, Q_{fl}, lost to the environment via purpose-provided ventilation, Q_{vl}, and infiltration, Q_{il}, or stored by thermal mass, Q_{tm}. In addition, the fabric losses Q_{fl} occur via transmission through those parts of the envelope in contact with the outside air, Q_{ftl}, thermal bridges, Q_{tbl}, and floors in contact with the ground, Q_{gl}.

$$Q_{out} = Q_{fl} + (Q_{vl} + Q_{il}) + Q_{tm} \tag{3.4}$$

$$Q_{fl} = Q_{ftl} + Q_{tbl} + Q_{gl} \tag{3.5}$$

This energy balance is specified by BREDEM [95], ISO13790[5], and related standards (see Appendix A); and is simplified or further elaborated depending on the desired accuracy of a model and the availability of calibration data. Therefore, algorithms vary in each HSEM according to the complexity of the chosen energy flow pathways. A number of notable anomalies and omissions are discussed in Section 3.4.2.

Abstraction of Energy Inputs

External gains (Equation 3.2) are commonly calculated in HSEMs using monthly factors as well as adjustments to represent transmitted gains from solar and long-wave radiation. These typological and contextual properties also relate to the heating systems and, in some cases, households activities (use of appliances, utilisation of spaces in houses that are frugally heated). The consideration of the external environment (Q_{bg} and Q_{xg} in Equation 3.2) and its influence on dwelling energy demand is limited in most HSEMs; for example, BREDEM only considers an over-shading factor. The LARA model includes regional information at SOA resolution but focusses on the estimations of fuel use and associated GHG [68]. Furthermore, the EEP model uses local neighbourhood and home hazard information from the EHS to estimate negative health consequences of poor indoor air quality, such as respiratory diseases [127]. Yet the extent to which it applies Equation 3.2 is poorly defined.

The main contributor to Q_{int} (Equation 3.3) is the heating system (Q_{hg}). Its output is usually estimated as a function of the local climatic conditions and the number of heating degree days using a base temperature (also known as a balance-point temperature). This estimate is then scaled by the variation of the external air temperature over a year. Modern heating systems can improve their performance by adding semi-automatic controls, but their operation is typically seasonal. The CHM includes the SAP catalogue and increases the number of heating systems and their properties to provide a more detailed estimate of the direct heat gains from them. Furthermore, BREDEM estimates an overall monthly water demand that can differentiate between hot and cold water use. The Domestic Hot Water (DHW) demand is estimated using a typical number of heating systems and their corresponding fuel type, and is adjusted by the number of occupants N.

If N is unknown then BREDEM estimates the value as a function of the total floor area, A_T, [93] where

$$N = 1 + 1.76(1 - e^{-0.00035(A_T - 13.9)^2} + \qquad 0.0013(A_T - 13.9); \quad A_T \geq 13.9 \tag{3.6}$$

N is also used to estimate average metabolic heat gains (Q_{og}).

In BREDEM, lighting L and appliances A, and later their corresponding heat gains (Q_{lg} and Q_{ag}), are estimated by assuming a correlation between A_T and N [94]. The power law relationship

[5] BS EN ISO 13790:2008 Energy performance of buildings — Calculation of energy use for space heating and cooling [116]

described by Equation 3.7 has been estimated from electricity meter readings. These lumped gains are then allocated according to the weights of L (0.21) and A (0.79) [15].

$$L + A = 263 \cdot (A_T \cdot N)^{0.47} \tag{3.7}$$

However, this model for L and A cannot replicate the large variations in appliance energy use that are observed in practice [82]. Lorimer [142] has developed a more disaggregated model of lighting and appliance energy demand, but its scope is limited to the HEED and therefore a limited number of UK regions.

In particular, the energy used for cooking tends to be highly simplified in HSEMs, because it is generally only responsible for a minor proportion of dwelling energy demand (see Figure 2.10). In BREDEM, cooking type coefficients are derived to estimate yearly demands aggregated by fuel types. This lack of detail may be problematic, because cooking directly affects indoor air quality and the responses of ventilation systems and householders to it, with corresponding impacts on ventilation losses.

Finally, the losses described by Equation 3.4 are varied in HSEMs. Infiltration (Q_{il}) is represented in BREDEM as a correlation factor derived from material properties. Purpose-provided ventilation (Q_{vl}) is calculated monthly using an adjustment factor that accounts for internal flues and an exposure factor. The CHM ascribes ventilation losses to different ventilation systems (mechanical and forced), thus increasing its variability across the stock, although there are only a limited number of these systems in the model. More detailed descriptions of ventilation are developed in the HIDEEM [159] and the DOMVENT models [125, 126]. Energy transmission through the envelope (Q_{fl}) and thermal bridges (Q_{tm}) are simplified to average values, which result as a function of the materials and the estimated dwelling geometry.

So although Equations 3.1-3.5 are simplified representations of heat flows in dwellings, they do nevertheless represent the principal factors influencing them. Many of the underlying drivers of operational and investment behaviours that directly affect energy flow pathways could and should be linked to them. An HSEM can in principle be used to assess the consequences of householder behaviours that influence decision-making on dwelling energy performance; although reliably predicting the consequences of operational behaviours in the absence of the prediction of the dynamic stimuli influencing them would not be straightforward.

3.3.3 Occupant Behaviour in Housing Stock Energy Models

HSEMs poorly describe the influence of occupants and their behaviour on energy demands [14], likewise local climatic conditions [129]. Statistically biased HSEMs are often more accurate than their physical counterparts, but their predictive capabilities are less flexible [18].

Meanwhile, significant advances have been made in the modelling of occupants' activities and dependent behaviours and their corresponding impacts on thermal and electrical energy demands [187, 227, 117, 151, 194]. But these models do not consistently handle diversity in household characteristics or in housing attributes, both of which can influence activities and behaviours. For example, the preferred living room temperature may be decided individually or following

negotiation with other householders [155, 151]. There is also a relatively poor quantitative understanding of the extent to which investment and day-to-day operational (e.g. use of lights and windows) behaviours can be influenced. For example, decisions to invest in new technologies may be the consequence of peer pressure, social incentives, or taxation programmes to encourage the efficient use of energy in homes.

In the absence of comprehensive stochastic models of occupants' diverse behaviours and the factors influencing them, occupants currently tend to be represented simply as average heat sources (e.g. by BREDEM and CHM), undermining their ability to account for strategies that may drive behaviour change.

In order to represent this behaviour in a HSEM, each occupant (or collective of them) may be represented as a unit with adaptive rules and responses, also known as *agency*. This agency enables heterogeneity in occupants' behaviour to be captured. For example, the AHOME and the improved DeCarb-ABM are able to include householder responses at building and neighbourhood scales [164, 139].

These responses of householders can also vary considerably over time because of socio-demographic changes, such as the ageing of the population. For example, by reconciling the energy demand with national consumption figures, the BREHOMES predicts housing stock changes by also considering social factors. This approach is also at the core of BREDEM [201]. More detail is provided by the DECM [31], the EDEM [34] and the SUSDEM [18], which all include household employment and occupancy profiles to represent collective responses (decision-making), albeit at a regional scale. Steadman's model [211] focuses on the consequence of decision-making and was built to estimate the impact of retrofitted energy efficiency measures and the applicability of related policies, particularly on loft and cavity insulation, and glazing and boiler replacements.

Top-Down and Hybrid Modelling

BREDEM and its derivatives are essentially bottom-up approaches to the energy modelling of the housing stock. But, top-down and most of the statistical approaches have also been developed to support strategic decision-making. Examples of top-down approaches are the ADEPT model and the STEP model [213], and the EDM [99]. The EDM is a sophisticated econometric predictor with the ability to reconcile policies with energy demand. The ADEPT and STEP models can efficiently estimate the consequences of changes to these external factors.

Integrating the virtues of both top-down and bottom-up approaches into a coherent whole may have merit: handling housing and household behaviours in a parsimonious way. An example of this reconciliation is found in the TARBASE [119] which evaluates decarbonisation and refurbishment scenarios using algorithms to represent lighting, appliances, DHW, refrigeration, glazing, insulation and solar technologies. This model also enhances the assumptions applied to Equations 3.2-3.3 by including more parameters for each gain, but using deterministic profiles. Similarly [24], the DEMScot complements the building-physics approach with forecasts related to fuel elasticity and carbon intensities; refurbishment maintenance costs and demolition rates; and direct rebound effects.

HSEMs and their underlying datasets have become progressively more comprehensive in their scope and sophisticated in the methods employed. But there remain some shortcomings when

representing the 27 million dwellings as a function of their geometric and physical properties and from the chosen statistical modelling approaches. These can be overcome by allowing open access to datasets and modelling modules of HSEMs, and by documenting the decisions and assumptions made during their development.

The development of hybrid modelling approaches, identified in Section 3.4, addresses some obvious omissions in the scope of HSEMs to support the formulation of policy measures designed to decarbonise the built environment. The ability to integrate models, and their algorithms, enhances the development of these measures and provides with certainty the choice of an appropriate strategy. The Section also provides the foundations for a systematic evaluation of HSEMs to highlight key areas that require further work.

3.4 Evaluation of the Housing Stock Energy Model

HSEMs are developed to describe the energy flow pathways in a sample of dwellings, and are extrapolated to represent an entire stock. For reasons of practicality, representations of these pathways (such as heat gains, ventilation and infiltration losses) are simplified. However, it is important that these simplifications are appropriate and that the principle of parsimony is respected so that a model is as simple as possible, but no simpler. If it is too simple, its ability to make reliable estimates of performance and inform decarbonisation policies and strategies is undermined.

In previous reviews of HSEMs [215, 129, 172], a set of attributes and gaps have been identified (see Table 3.1). They describe some attributes arising by the modelling rationales and strategies—with their particular features, described in § 3.1.1—to abstract energy flow pathways in dwellings. It must be considered, however, that many of these gaps have since been partially addressed by augmenting existing algorithms and techniques. For example, societal effects have been improved in models that include agency, such as DeCarb and AHOME. They account for human interactions with a dwelling, its systems, and appliances, although these relationships are based on a modest amount of empirical evidence. Furthermore, models such as AHOME, SUSDEM, EDEM and DDM have improved their ability to assess building fabric by adopting GIS to better represent building form, and by including investment scenarios, which improves the assumptions made on form and fabric components.

In addition to these assumptions and considerations of modelling scope and rigour, it is important that an HSEM be modular. Kavgic et al. [129] find that BREDEM is modular but its links are not always transparent. All algorithms should be editable and thoroughly documented, if their scope of applicability is to be well understood.

3.4.1 Evaluation of Functionality

Some concerns about data resolution and modelling abstractions were identified in § 2.3.2 and § 3.3, respectively, highlighting a need to evaluate existing HSEMs to determine their fitness for purpose. To this end, the functionality and accessibility of some twenty-nine UK HSEMs is evaluated. ASHRAE's general framework [5] is used for selecting, or determining the appropriateness of a modelling strategy, using six criteria:

Table 3.1: Positive and negative attributes of housing energy modelling approaches

positive attributes + + +
bottom-up
- ability for typical end-use contribution
- inclusion of new technologies influence
- characterisation of archetypes
engineering
- measurement of clustered energy consumption
- detailed simulation and multiple outcomes
statistical
- projection of budget impact
- determination of socio-economical effects
- inclusion of occupancy group profiles
top-down
- long term forecasting
- analysis of simple input information
- inclusion of macro-economic effects
negative attributes - - -
bottom-up
- reliance on average consumption
- assumptions about typology form and fabric homogeneity
- inability to include granular data-sets
- lack of feedback in archetypes
- onerous computational requirements
engineering
- demand specific software
- poor representation of social and environmental influences on occupant behaviours
- oversimplified calculation procedures and underlying algorithms
- inaccurate descriptions of energy-related decision-making
- require specific methods for analysis of uncertainties and their propagation
statistical
- inflexible models with limited survey size and data gathering methods
- multicollinearity
- limited ability for decision-making impact
top-down
- unlinked methods for the co-simulation of specific modelling algorithms
- inability to represent investments in new technologies
- reliance on historical data to formulate correlations

1) *Accuracy:* Has the model been tested and validated and is it capable of estimating prediction uncertainties?
2) *Sensitivity:* How sensitive is the model to the design options under consideration?
3) *Versatility:* Does the modelling method allow for the analysis of all design options under consideration?
4) *Computability:* How appropriate is the computational time when compared to the resolution of predictions?
5) *Reproducibility:* Is it likely that different users aiming to solve the same problem will make the same input choices and how reproducible are the predictions?
6) *Usability:* How easy is it for a user to make their input choices and to analyse predictions?

Although these criteria are helpful in guiding the review of the principal modules that comprise HSEMs; in particular to determine the ability of an HSEM to successfully model energy flows and the socio-economic drivers that influence them; it is worth noting that they were developed to evaluate energy models of individual buildings rather than of stocks of buildings, and so they do not explicitly address the characterisation of the housing stock that is being modelled. To do this, databases should be evaluated considering their resolution (sample), completeness (scope), coherence (compatibility between datasets), regularity of updates, and the integrity of any underlying databases that have been unified. Nevertheless, the ASHRAE criteria, augmented by an analysis of transparency, are a valuable starting point. They are used here to support the evaluation of both data resolution and modelling sophistication, which is summarised in Table 3.2. In this, transparency or accessibility is evaluated by assigning scores according to whether each of the modules are restricted or absent (black box: 0), open and editable (white box: 1), or somewhere in-between (grey box: $\frac{1}{2}$). The accessibility to the HSEMs' documentation is also considered, as well as the operating system(s) used and software license restrictions. The evaluation also aims to identify the components of HSEMs that could be augmented by the consideration of diverse disciplines that have historically been unrelated to energy models, such as health and socio-demographic studies that explore responses to indoor stimuli.

Some of these models cannot strictly speaking be considered HSEMs, but they have been included because they share common data sources (e.g. DDM and MDM) or computational engines (e.g. RdSAP). The SAP is included in this analysis because some HSEMs directly use elements of it to make predictions (e.g. EDEM, DEMScot, DECoRuM, CHM, DDM and RdSAP), or apply it indirectly via its reduced format RdSAP (e.g. CHM and TARBASE). Furthermore, given the fact that BREDEM is at the core of many HSEMs to perform their energy balance [142, 31], its structure to define the evaluation categories (rows in Table 3.2) is adopted, which correspond to the description given in Sections 3.3.1, 3.3.2 and 3.3.3. Likewise, a differentiation is made among the three categories: fabric, systems and occupancy. The evaluated HSEM are grouped by their modelling approach in their latest version (see Figure 3.9).

By way of demonstration of the evaluation, as set out in Table 3.2, the BREDEM's geometry module is considered. Firstly, its sensitivity, adaptability, usability, and accuracy can be assessed as limited, because it does not explicitly consider the dwelling volume and only retains A_T, a derivative of building geometry. However, this module could in principle be replaced by a more detailed model to improve the accuracy of energy demand predictions. The outputs of this module are inputs to other modules and so its usability is restricted. Hence, four of the criteria are considered to be grey boxes (i.e. $4 \cdot \frac{1}{2}$). Further, this module is highly compatible with other algorithms and data, and so its predictions are reproducible (i.e. $1 \cdot \frac{1}{6}$). In the same

vein, its adaptability depends on the purpose of the model because it can easily be included in isolated algorithms or in stock models (i.e. $1 \cdot \frac{1}{6}$). Overall, the geometry module scores 4 of 6 and is considered to be a white-box module because of the likelihood of future editing and the consequent improvement of its attributes.

Similarly considered are CHM and DECM, which are bottom-up physically biased models, both of which use BREDEM [110]; although DECM does not fully document the algorithms it applies. DECM's focus is on the assessment of uncertainties, which is valuable but limited by its lack of model transparency. Neither model explicitly accounts for dynamic responses to indoor conditions or detailed household representations. The assessment of energy performance and carbon intensity is limited to their sample units, which are SOA and LA, respectively. If these two HSEMs are compared, the CHM ranks higher because it is more transparent, despite the fact that DECM supports scenario projections and uses an index of carbon intensity. However, its limited transparency undermines its functionality.

The DECADE model and DECoRuM are both statistical biased models. DECADE relies heavily on historical tendencies to create a set of scenarios using some fixed conditions [13, 98]. DECoRuM uses BREDEM and SAP to calibrate its estimations, although their application is restricted spatially to a district scale analysis. DECoRuM is essentially a visualisation tool, and so its dependence on physical models provides an easy opportunity to expand its scope [79, 80]. The approaches of both models render them unable to represent geometrical- and contextual-related values. Changes to these values are embedded into their models, which are assessed as partially or fully restricted. Nevertheless, their ability to represent trends and changes in the properties of the elements (i.e. ageing) is highly beneficial and may complement the outcomes from physical-biased approaches. Some other statistically-biased models are ranked lower than DECADE and DECoRuM because of their limited scope or focus, and not because they are less fit-for-purpose.

Table 3.2 systematically rates all HSEMs that are currently available (distributed and/or published) in the UK using a range of core underlying functions to rate and rank them. The rankings are based on the assumption that individual functions (rows in the table) are equally weighted, and so they should be used with caution. Table 3.2 shows that statistically biased and top-down models, which lack fidelity in the handling of physical phenomena, are nevertheless able to consider societal parameters reasonably well. The bottom row of Table 3.2 suggests that CHM can be considered the best performing model. Its relative functionality and transparency may explain why it is used to provide a statistical evaluation of the UK housing stock for the *Energy in the UK* documents [110]. However, CHM and other similar models could be improved considerably by substituting modelling algorithms, for example to support dynamic simulation (i.e. by capturing the time varying behaviour of a system), by substituting the calibration data with more up-to-date or sophisticated alternatives, and by accounting for household attributes that affect decision-making and energy usage in the stock of dwellings.

Table 3.2: Evaluation of models according to the level of transparency as a function of accuracy, computability, reproducibility, sensitivity, adaptability and usability. Models are sorted by year and type of approach Notation: ■: black-box models, ⊠: grey-box models, □: white-box models

	BREDEM-12	SAP-09	BREHOMES	Johnston	RdSAP-05	CDEM	DEMScot-2	CHM	DeCarb-ABM	DECM	SUSDEM	HIDEEM	AHOME	DOMVENT	NHM	DECADE	EEP	*E3MG	EDEM	DECoRuM	UKDCM2	LARA	TARBASE	Lorimer	Steadman	DDM	MDM	(ADEPT)	(STEP)	
Layout Geometry	□	□	□	⊠	□	□	□	□	□	□	□	⊠	⊠	⊠	□	⊠	⊠	■	□	□	□	⊠	⊠	■	□	□	■	⊠	■	**Fabric**
Construction Materials	□	□	□	■	□	□	□	□	□	□	□	⊠	⊠	⊠	□	□	□	■	□	□	□	□	⊠	■	□	□	■	⊠	■	
Ventilation Rates	⊠	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	⊠	□	⊠	□	⊠	⊠	■	■	⊠	⊠	⊠	■	⊠	■	⊠	⊠	■	⊠	■	
Infiltration Parameters	⊠	⊠	⊠	⊠	⊠	□	⊠	□	⊠	⊠	⊠	□	⊠	□	⊠	⊠	■	■	⊠	⊠	⊠	■	⊠	■	⊠	⊠	■	⊠	■	
Context & Orientation	⊠	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	□	⊠	⊠	□	⊠	■	□	■	⊠	⊠	⊠	□	⊠	■	⊠	⊠	■	■	■	
Solar Gains	□	□	□	■	□	□	□	□	□	□	□	⊠	⊠	□	□	■	⊠	■	□	□	□	⊠	⊠	■	⊠	□	■	■	■	**Systems**
Infiltration Gains	⊠	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	⊠	□	⊠	□	⊠	⊠	■	■	⊠	⊠	⊠	■	⊠	■	⊠	⊠	■	⊠	■	
Thermal Mass Gains	□	⊠	□	■	⊠	□	⊠	□	□	□	⊠	⊠	⊠	⊠	□	■	■	■	□	□	□	■	⊠	■	⊠	□	■	⊠	■	
Appliances Gains	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	■	⊠	■	⊠	⊠	□	⊠	⊠	⊠	⊠	□	⊠	□	⊠	⊠	⊠	⊠	⊠	
Lighting Gains	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	■	⊠	■	⊠	⊠	□	⊠	⊠	⊠	⊠	□	⊠	□	⊠	⊠	⊠	⊠	⊠	
Occupants Gains	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	⊠	■	⊠	□	⊠	⊠	■	⊠	■	⊠	⊠	⊠	⊠	■	■	■	⊠	■	⊠	■	
Cooking Gains	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	⊠	⊠	□	⊠	□	⊠	⊠	⊠	⊠	⊠	
Water Flows	⊠	⊠	⊠	■	⊠	⊠	⊠	□	⊠	⊠	■	■	■	■	⊠	⊠	⊠	■	⊠	⊠	⊠	■	■	■	■	⊠	■	■	■	
Light & Noise Quality	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	**Occupancy**
Air Pollutants	■	■	■	■	■	□	■	■	■	■	■	□	■	□	⊠	■	⊠	■	■	■	⊠	⊠	⊠	■	■	■	⊠	■	■	
Draught Risk	■	⊠	■	■	■	■	⊠	⊠	■	■	■	□	■	□	■	■	⊠	■	■	■	■	■	■	■	■	■	⊠	■	■	
Comfort Responses	■	■	■	⊠	■	■	■	■	⊠	■	■	■	⊠	■	■	■	■	⊠	⊠	■	■	■	⊠	⊠	⊠	■	⊠	⊠	⊠	
Peer Pressure	■	■	■	⊠	■	■	■	■	⊠	■	⊠	■	□	■	⊠	□	⊠	□	⊠	⊠	■	□	⊠	⊠	⊠	■	⊠	□	⊠	
Budget Effects	⊠	⊠	⊠	⊠	⊠	⊠	□	⊠	⊠	□	⊠	■	□	■	⊠	□	⊠	□	⊠	⊠	⊠	□	□	□	□	⊠	⊠	□	⊠	
Awareness	■	■	■	⊠	■	■	⊠	■	⊠	■	⊠	■	□	■	⊠	□	⊠	□	⊠	⊠	■	□	⊠	⊠	⊠	■	⊠	⊠	⊠	
			bottom-up *physical* biased													bottom-up *statistical* biased											top-down			
rank	-	-	15	23	17	7	3	1	6	9	13	11	20	4	2	18	14	26	7	5	10	19	22	24	21	12	25	16	27	

3.4.2 Anomalies and Omissions

The inclusion of detailed household attributes is essential for assessing the effects of changes in energy demand when occupants adopt reduction measures, either by investing on new technologies or by modifying patterns of use. Mitton et al. [161] identify two advantages of including this type of household information: the ability to represent ageing elements, and the ability to describe the—typically complex—interactions between them. To date, their inclusion is still problematic and can create gaps in an HSEM. It is easier to insert a replacement module and validate it if it is directly linked to a single open source platform rather than persisting with the current practice of developing and maintaining a plethora of closed-source HSEMs.

Therefore, it is important to consider the representation of occupants' behaviours in response to underlying environmental stimuli or the demand of energy-related services. Behavioural responses can result from the interplay between the building, its systems, and the interactions between them. Currently, the parameters that affect behaviour are blurred in HSEMs (see Table 3.2). The indoor environment of a building may be affected by available fuel sources and geometric shape constraints that are a function of its location. A building can define an occupant's needs, but an occupant may also define the building's environment. For example, larger spaces may demand more energy to keep them warm, but occupants may invest in insulation technology so heat losses are minimised. Conversely, they may alter their utilisation of space and, with time, reduce the volume used and heated during the heating season. A modelling approach that considers householder behaviours as a function of occupancy and environmental stimuli would signify a transition from a simplified energy balance modelling approach to dynamic simulation. An inability to accurately represent these dynamic processes, and the socio-economic processes that influence household decisions, seriously undermines the functionality of HSEMs. Required information could be obtained from socio-economical surveys and used to inform either top-down or bottom-up models.

The representation of domestic appliances in a HSEM requires data related to household ownership, efficiencies or rated powers, and associated usage patterns. To date, there is no HSEM that comprehensively represents these parameters. Current representations of appliance use and its dependency on household characteristics is often reduced to a scalar correlation factor using national stock averages. Any departure from this simplified function (whose theoretical basis is unclear) would negatively impact on its predictive accuracy. For example, Equation 3.7 is incapable of encapsulating factors that affect appliance use. As a result, its associated energy demand is insufficiently granular.

Societal factors could be incorporated from AHOME or Steadman's model, or even substituted with more rigorous alternatives. Improving the modularity of the CHM, and other similar models, would also allow modelling algorithms and calibration data to be substituted with more sophisticated and up-to-date alternatives. For instance, descriptions of dwelling geometry could be improved, perhaps using a similar volumetric method to that used by DOMVENT, yet focussed on ventilation and infiltration rates. Incidentally, several studies have identified that dwelling heating demand is most sensitive to fluctuations in these rates [209, 124]. It has also been suggested [124] that existing algorithms, such as those used by SAP [57] and BREDEM [200], may be overly simplistic and based on limited empirical evidence. This suggests that there is a need for a more rigorous representation of infiltration rates in homes, requiring a combination of rigorous fieldwork and improved modelling. This would then lead to more refined predictions

of health risks in homes.

Some other parameters that are required to model the complex energy interplay within dwellings are neglected in many existing HSEMs. These include lighting level, air quality, and noise. Different ratios of artificial and natural light affect the indoor light quality and activities that depend on it. Similarly, airflow between a dwelling and its external environment is considered via a range of different strategies, and so the pollutant content of that air should be considered. Some activities, such as cooking, will affect indoor air quality and occupants' use of windows and fans to regulate it. Noise is also related to the energy demand of a dwelling. For example, occupants may attempt to moderate external noises by closing windows, or attempt to mitigate internal noise by switching off, or avoiding the use of, fans. It is clear that the consideration of noise and air quality should not be decoupled.

One final factor that is often overlooked is householder health, which a dwelling can both improve and harm. The main purpose of a dwelling is to provide shelter from the weather, but it also provides security, comfort, and privacy. However, these benefits may cause a building to limit flows of daylight, air, noise, moisture, and heat (affecting temperature and thermal comfort). Accordingly, poor indoor environment quality may adversely affect the physical and mental health of occupants. This highlights the relevance and importance of health as an HSEM module.

It is clear that the underlying drivers that affect householder behaviour (i.e. investments and energy using practices) and their interaction with their dwellings are limited in some HSEMs and ignored by most. A consideration of these drivers could lead to real changes in the energy demand of houses, and their omission from HSEMs significantly limits their ability to estimate the varying conditions of housing and households at stock level [139, 107]. Multi-agent simulation techniques would appear to have much to offer in this endeavour [26, 28].

Finally, the ability of HSEMs to predict indoor comfort, and the impacts of associated heating system set-point temperature choices, is limited by the granularity of the underlying energy model. Predicting comfort and the impacts of the allocation of the energy and comfort co-benefits from energy-related renovation investments, meaning that potential energy savings may be offset by improved comfort and health arising form high set-point temperatures, requires dynamic simulation.

3.4.3 Future Functionality

A transparent model clearly describes its underlying algorithms so that its users can comprehend its rationale and either confirm or discard them to adapt the model to meet specific needs. A preliminary open framework rationale is presented in [163], whose processes can only be accomplished under an open-source philosophy. This philosophy requires a significant reinterpretation of the way data is handled, collected, processed, and shared. Hence, emphasis should be placed on the usability of data, so that it can enhance the communication between models and users. For instance, the National Statistics Harmonisation Programme whose rationale is to standardise classifications, definitions, and standards among datasets [170], and the pool of energy-related building data developed by European countries to improve the analysis of regional and local stocks [140]. By improving data collection intervals and the way data is calibrated, the likelihood of its adoption can be increased. Instead of relying on historical data, models can use more

up-to-date and more comprehensive housing and household data. Lorimer [142] finds that the calibration of data can also be used to help implement measures designed to reduce non-heating end-use energy demand as part of national programmes.

Improvements in the transparency of HSEM data and modelling algorithms should be complemented with improvements to modularity. Modularity prolongs the life of a model by allowing it to be improved and delays obsolescence. Such modularity should also facilitate the execution of sensitivity analysis methods that can be used to identify the most important inputs to a model, to highlight areas where further data gathering should be focussed, and where models should be refined to more faithfully represent phenomena addressed by them [98, 109, 43, 124]. These analyses should be complemented by a quantification of the uncertainties in a model's predictions. It should account for both deterministic (building physics) and stochastic (behavioural) phenomena, and the data that underpin their representation. The reporting of errors and omissions enhances the understanding of the accuracy of different modelling strategies and their descriptions of the energy flow pathways at stock level.

3.5 Conclusion of *Housing Stock Energy Models*

Three elements were identified as essential to HSEMs: 1) dwelling geometry and fabric; 2) contained systems and inter-relationships between dwelling components and zones; and 3) occupancy and appliance usage within dwellings. The resolution to capture these elements is defined by the units of observation and the purpose of the models. Still, it has been identified the need of simplifications, so that these elements can be passed to the modelling algorithms. In here, dwellings have been and are the main units used to explore and to develop HSEMs; but the relevance of users and their decisions is fundamental to accurately represent drivers influencing energy flow pathways at stock level. It must be pointed out, however, that the representation of users in HSEMs has been limited due to the inherent complexity involved in the characterisation of their unique behaviours; as a result, the assumptions made to integrate them into HSEMs are significantly greater than the other two components, and hence the propagation of errors attributable to occupants is substantial. Further, the sophistication of HSEMs has, on the one hand, tackled some limitations that constrained previous models; but on the other, they have increased the ability to capture more inputs, which increases the likelihood of data input errors (see § 3.1).

Therefore, HSEMs are benefited not only by increasing their technical capabilities, but also by making accurate descriptions of the assumptions and the methods applied to their rationales. A sound understanding of domestic energy flow pathways and factors influencing them (both housing and household) is essential in formulating policies designed to reduce the energy demand and associated carbon emissions of any housing stock. It is then imperative to identify a parsimonious housing stock energy modelling strategy.

The evaluation of the HSEMs and their functionality shows that very few are completely transparent, which makes it difficult to understand all the methods and algorithms that have been used to make predictions of energy demand and thus to judge their fitness for purpose. This is compounded by the limited scope and simplicity of many HSEMs, which are generally based on simplified steady-state energy balance models (i.e. without changes over time) augmented with statistical models relating to lighting, system and appliance usage extrapolated to the

stock of like typologies. Here, it is important to understand how physical phenomena have been (over-)simplified, or where and on what basis statistical relationships have been incorporated, so that they can be improved in the future, as necessary. Failing to do so undermines the utility of HSEMs and their ability to faithfully predict performance at stock level, or at individual building typologies. An initial approximation of the thermodynamic (im-)balance is useful, but is limited to a partial description of components affecting energy demands in dwellings.

A full understanding of the reasons and factors that affect the energy flow pathways in dwellings is highly linked with the operation that is performed in them, and so HSEM must consider them. A poor representation of these components in HSEMs increases their uncertainty, and so reduces their functionality [74]. Tracing it is important. This is complemented by the findings from Chapter 2, which identified how significant errors can be easily incorporated into a database due to the quality of the data sources; and the need to continually update them in order to maintain their validity.

Future HSEMs should have a modular architecture and be freely available so that they can be edited and updated continuously. This could encourage cross-disciplinary collaboration and increase the rate of improvement to the modelling of phenomena within HSEMs where there is significant prediction uncertainty, such as householder behaviour. Furthermore, whilst some HSEMs specialise in the prediction of a specific aspect of dwelling energy demand, both their utility and longevity could be improved by incorporating them within a larger modular HSEM architecture and by deriving specific metrics for users and their interactions. For instance, approximately two-thirds of homes belong to owner-occupiers and thus energy efficiency improvements may only success when those individuals decide to invest in energy-efficient technologies, which is the major challenge for future HSEMs.

To clarify, it makes no sense to over populate a model when this can be kept simple, or on the contrary, to attempt a proper modelling with only a few variables. It could be tempting to develop a highly sophisticated model at every level of the referred distinctions. However, instead of achieving a reliable and robust model, the result could mean the acquisition of a confounded and restricted one, contradicting the original purpose of a modelling representation. It is important that the simplifications taken to make a model are appropriate, and that the principle of parsimony is respected so that a model is as simple as possible, but no simpler. If it is too simple, its ability to make reliable estimates of performance and inform decarbonisation policies and strategies is undermined.

This Chapter has provided an overview of different housing stock energy modelling strategies, and identified some obvious omissions in the scope of HSEMs to support the formulation of tools designed to decarbonise the built environment. It also provided the foundations for a systematic evaluation of HSEMs to highlight key areas that require further work, summarising conventional and relevant modelling approaches included in them. The most relevant outcomes regarding the evaluation of HSEM are:

I/ the models should be transparent and available for improvement so that their underlying algorithms can be validated;

II/ future HSEMs should have a modular architecture so that each module can be edited continuously; and

III/ databases should continually be updated so they can maintain their validity.

Development and Evaluation Part II

4 Composition and Development of the Simulation Platform: EnHub

Introduction

In the UK, the 28 % reduction in carbon emissions is reasonably close to the aimed Committee on Climate Change (CCC) targets. However, as explained in Section 2.1.2, this has largely been achieved in the absence of structural improvements to reduce energy demands. Some of these opportunities for significant demand reduction are found in the domestic sector. Section 3.2.1 described that efforts have of course been made to improve the performance of the existing housing stock, so they can reduce their energy demand; but these have been mostly counteracted by the construction of new buildings serving a larger population (despite improved construction standards) so that potential carbon emission reductions from these measures are not fully realised in practice. For instance, the carbon emissions attributable to domestic buildings, including end-use energy and power generation, have been maintained at almost the same level since 1990 [58].

For this reason, a full understanding of the energy flow pathways in dwellings, and factors influencing them, has been required to formulate robust policies and strategies [215, 129] to achieve significant reductions in their carbon emission intensity. Therefore, it has been found that special attention needs to be given to both scenario and sensitivity analyses, so that drivers and variables involved in the composition of the stock can be better represented [97, 35]. And also, it has been elaborated that not only the drivers affecting directly the energy demand need to be further explored, but also those exogenous drivers to the energy flow pathways in dwellings [97]. These drivers include, for example, climate change, policy incentives and changes on the households' composition. Most of these drivers have been poorly represented in previous Housing Stock Energy Models (HSEMs).

The functionality of UK HSEMs was further evaluated in Section 3.4.1. The outcome of this evaluation was useful to identify the significant gaps in models that have existed due to their inner granularity and sophistication. Section 3.5 summarised the main priorities to be tackled in future improvements regarding HSEMs.

In addition, and based on the findings related to the stock composition in the UK (see § 2.4), three premises have been defined as essential to shape the development of a robust evaluation platform:

I/ A solid platform of decarbonisation scenarios should be consolidated not only by advancements in computational power, but also by a shift in the way data is processed, shared and published among developers and stakeholders. For example, the characterisation of household appliances need to be constantly updated and, as far as possible, well documented.

II/ Improvements in the energy efficiency of components found in the domestic sector (envelope and appliances) are not enough to fulfil the national carbon budgets for the sector. A better understanding of the processes of decision-making is needed, so that the effectiveness of decarbonisation policies and strategies gains accuracy and confidence.

III/ Better quality performance metrics to characterise the domestic energy sector must be enhanced in order to calibrate and constantly validate models. For example, semi-detached and terraced are both the most common type of dwellings in the UK housing stock; however, it is not clear how influential is the attached wall to the energy performance described in HSEMs.

To accomplish this, a simulation platform is developed with the ability to integrate and enhance, in particular by dynamically simulating the housing stock, the virtues of the existing HSEMs. Dynamic simulation enables to possibility to capture the time varying behaviour of components affecting energy flow pathways in dwellings. As a result, it is possible to represent responses throughout the performance evaluation period. This platform of transparent modularity is defined as the Housing Stock Energy Hub (EnHub). The following sections describe its rationale, composition, modules and functionalities.

4.1 Simulation Platform Architecture

4.1.1 Housing Stock Energy Hub (EnHub) Rationale

The structure of EnHub takes its inspiration from the CHM, which is at the core of the *Energy Consumption in the UK* [60] study. It is identified as the most flexible tool to be adapted and improved, towards the development of a more robust simulation platform. The principle dataset underpinning both EnHub and CHM is derived from the EHS-2011, which comprises 14,951 dwellings, weighted to represent the 21 million houses in England. Employing the EHS as a basis, the results are extrapolated on a pro-rata approximation to match the UK total dwelling population (i.e. applying the proportional rates Eng-Sco, Eng-Wal and Eng-NI; see Table 2.1), as described in the referred study [60]. The principle dataset in the study is also augmented by the Census and the HEED.

The integration and processing of information into the platform, derived from different datasets and catalogues, is implemented using the open-source statistical software `R` [185]. `R` is used due to its computing capability and its ability to develop a programming platform—see Box 4.1.1. By employing the former, EnHub is able to generate a set of volumetric models to be then parsed to the `EnergyPlus` [212] engine where the dynamic simulation of dwellings is performed.

`EnergyPlus` is used due to its ability to process simulations within a highly time resolution and its modularity, that enhances the control on specific properties of components—see Box 4.1.2.

Box 4.1.1: `R` Software

`R` [112, 185] is a free software environment for statistical computing and graphics and is an official part of the GNU ("GNU's Not Unix!") project. It consists of a language plus a run-time *environment* with graphics, a debugger, access to certain system functions, and the ability to run programs stored in script files. The software includes:

1. an effective data handling and storage facility;
2. a suite of operators for calculations on arrays, in particular matrices;
3. a large, coherent, integrated collection of intermediate tools for data analysis;
4. graphical facilities for data analysis and display either on-screen or on hardcopy; and
5. a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

The term environment is intended to characterise it as a fully planned and coherent system, rather than an incremental accretion of very specific and inflexible tools, as is frequently the case with other data analysis software.

For computationally-intensive tasks, `C`, `C++` and `Fortran` code can be linked and called at run time. Advanced users can write `C` code to manipulate `R` objects directly. So, besides being a statistical platform, R is also an environment within which statistical techniques and programming codes are implemented. `R` can be extended via packages which are documented and updated constantly.[a].

[a] Further information may be consulted directly in http://www.r-project.org/

EnHub follows the CHM protocol to transform the EHS datasets into a source of information ready to be employed by an energy modelling tool [108]. This protocol is initially made to mimic the CHM-BREDEM approach; it is also employed because the resulting dataset is linked with the SAP catalogue, where some properties of the archetypes are derived, such as properties on the materials, degree-days and heating system specifications.

The application of such a protocol is thus performed in a modular fashion, following an Object-Oriented Modelling (OOM) approach; essentially a collection of interacting objects, where each object is semantically defined with variables, classes and dependencies [177]. OOM can be represented graphically using UML architecture to provide a clear and unambiguous structure for the model attribution workflow, and to enhance the integrity of the predictions of the algorithms used to describe the energy flow pathways [169]. For example, the use of explicitly defined variables ensures that constants and coefficients are used consistently throughout the platform, and that the processes of encapsulation are achieved.

The UML architecture shows the internal logic of the modules, and can be presented using different types of diagram such as class, use-case, state, sequence, component, or deployment. These diagrams can be easily expanded and complemented. EnHub implements a UML class diagram type, focussed on the characteristics of the modules and their relationships. Similar implementations have been used elsewhere to improve the development of highly detailed libraries of classes to support more robust management of databases and modelling algorithms [164, 157]. This is on the understanding that the UML is, besides explicit, accessible.

It can be claimed that transparency leads to scrutiny. Nevertheless, transparency becomes a virtue in the modelling strategies because each of the components or modules can be comprehended and modified (see § 3.4.1). One of the implications of such a transparent modularity is that the modules are improved, not only when they fail or apply vague assumptions, but also when they require to be aligned with some other modelling approaches and their corresponding scopes. To this end, EnHub uses a central dataset to define typologies of the housing stock, and adapts several modules according to the resolution of energy modelling approaches, so they can be evaluated.

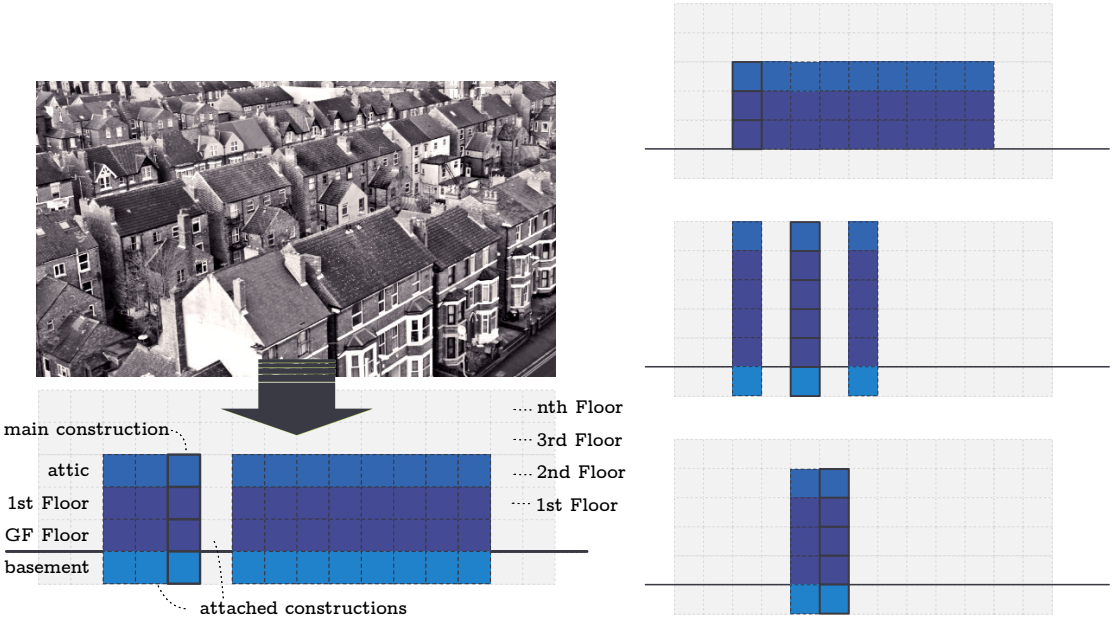
Box 4.1.2: `EnergyPlus` Software

`EnergyPlus` [41, 212] is a modular software designed to perform energy simulation in buildings; it combines the main features of BLAST and DOE-2. The software was created by the U.S. Department of Energy and completed in April 2001. It provides an integrated (simultaneous loads and systems) evaluation for the analysis of realistic system controls, moisture adsorption and desorption in building elements, radiant heating and cooling systems, and inter-zone air flow. `EnergyPlus` has two basic components: 1) heat and mass balance simulation module and 2) building systems simulation module. The first module is based on IBLAST—a research version of BLAST with integrated HVAC systems and building loads simulation. The heat balance module manages the surface and air heat balance modules and acts as an interface between the heat balance and the building systems simulation manager. The surface heat balance module simulates inside and outside surface heat balances; interconnections between heat balances and boundary conditions; and conduction, convection, radiation, and mass transfer (water vapour) effects; the latter is a manager that handles communication between the heat balance engine and various HVAC modules and loops, such as coils, boilers, chillers, pumps, fans, and other equipment/components. Every component is defined by the user for each zone.
In summary, `EnergyPlus` models heating, cooling, lighting, ventilation, other energy flows, and water use; and includes many innovative simulation capabilities such as reduced time-steps, modular systems and plant integrated with heat balance-based zone simulation, multi-zone air flow, thermal comfort, water use, natural ventilation, fenestration calculations including controllable devices, daylighting controls, photovoltaic systems and atmospheric pollution calculations. The outputs may be defined for specific purposes such as: electrical demands in a time period (hourly, monthly, yearly), average conditions (temperature, humidity, pressure), single zone conditions and specific meters [a]. Its most prominent attributes are listed below:

1. modular disaggregation,
2. detailed properties definitions of materials and systems,
3. dynamic profiles with its associated responses,
4. interaction with diverse weather database,
5. specification of results reporting,
6. friendly user interface and capability to be combined with other tools (i.e. drawing interface, GUI, optimisation module), and
7. friendly interconnection with other tools using the code-environment.

`EnergyPlus` is free, open-source software. It is distributed under the terms of a permissive, BSD-style license. The software is compatible with Windows, Mac OS X, and Linux operating systems.

[a] Further information may be consulted directly in http://apps1.eere.energy.gov/buildings/energyplus/



(a) Cuboid description: abstraction of a two floors mid-terraced house with attic and basement. Each square represents a zone for the `EnergyPlus` format. – "Victorian Houses" Licensed under CC-BY2.0. (tinyurl.com/j8ekuvy accessed 15 March 2013)

(b) Examples of cuboid configurations: top - end of terrace house, middle: apartments, bottom: semidetached house with attic and basement

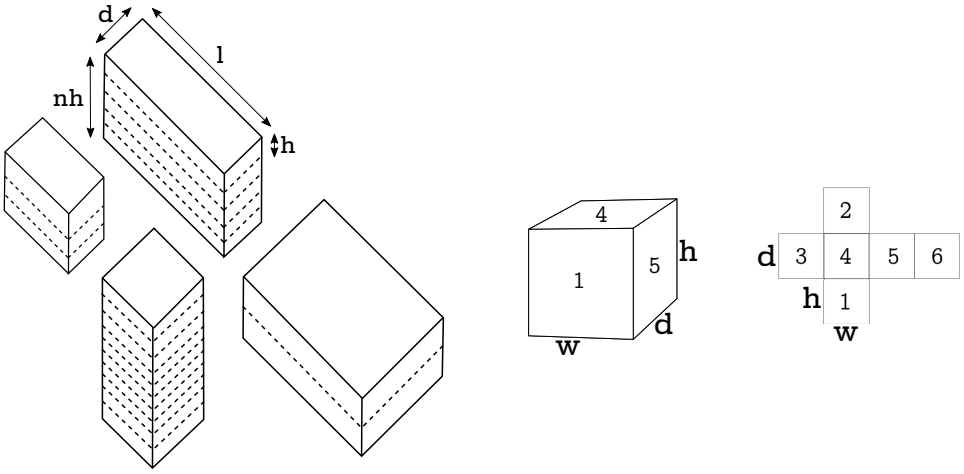
Figure 4.1: Cuboid abstraction: representation of contiguous blocks

4.1.2 Abstraction and Design

As a preliminary step, three facts related to the housing stock in the UK—developed in Section 2.4—are considered developing an initial configuration of the housing stock typologies: 1) heating represents more that three quarters of the energy demand (i.e. for space heating, DHW and cooking); 2) houses constitute 80% of the stock; and 3) gas is the most common fuel to supply heating systems and to produce electricity in power stations. Therefore, an approach able to represent building shapes, and to parse information regarding heating systems (end-use demand and fuel supply) is applied. Besides, it is worth noting that the homogeneity of the UK housing stock is conveniently employed to encapsulate the principal attributes describing the shape of dwellings. Hence, modelling a three-dimensional form of housing typologies in as sets of contiguous cuboids (see Figure 4.1a), enriched with highly disaggregated survey data, is initially opted as a reliable way to encapsulate these facts.

The set of geometrically simplified models are similar to those typologies of buildings from contiguous cuboids developed in the DOMVENT [124] and the Steadman's model [210], as illustrated both in Figure 4.2. An alternative approach would consist in generating each necessary geometry to the model archetype, in which this can be defined either by survey data, or by GIS layouts, yet it would require higher resolution in the collection of data.

In the EnHub platform, semantic attributes are assigned to these cuboids based on survey data (i.e. of envelope properties and household variables). Furthermore, this platform also places emphasis on the influence of household drivers, which have historically been excluded in the



(a) Building block with depth d, length l, number of storeys n and storey height h – Steadman's model

(b) Cuboid dwelling of height H, width W and depth D. Façades 1 and 2, dwelling front and back; façades 3 and 5, sides (exposed if detached or internal if terraced); façades 4 and 6, ceiling and floor – DOMVENT

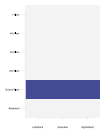
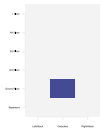
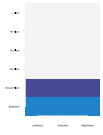
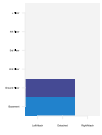
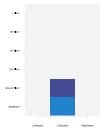
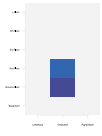
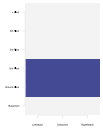
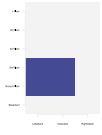
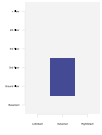
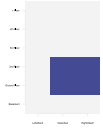
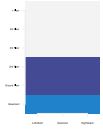
Figure 4.2: Volumetric representation of dwellings
Source: [210, 124]

understanding of the energy flow pathways at the stock level. Including them, albeit in a reduced scale, constitutes a logical next step in the consideration of the most influential motives that affect household decision-making, the stimuli resulting from the interaction between householders and dwellings, as well as the responses to adapt indoor conditions to personal preferences and other individual and group needs. Some of these household drivers are included by deriving specific metrics such as discomfort levels per capita; the proportion of the stock that over- and under-heats—given appropriate descriptions of heating preferences; heat gains per floor area; disaggregated power demand of household appliances; and daily peak heating and power demands.

This geometrically simplified approach enables the platform to account for the number of floors, as well as the presence of attics, basements and attachments (i.e. other dwellings, extensions and conservatories). In this way, the key residential types of the stock can be represented satisfactorily, such as detached houses, semi-detached, end of terrace, mid terraced, apartments and bungalows (see Figure 4.1b). These volumetric properties, complemented with the scalar properties of the buildings, are able to form, in conjunction, a robust configuration of the stock (physical) typologies. The presence of specific heating systems and fuels associated are dependent on the quality of the survey data. They are included as properties of each of the model archetypes and adapted accordingly.

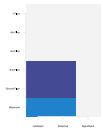
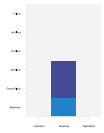
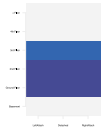
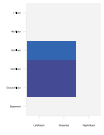
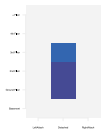
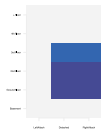
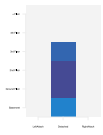
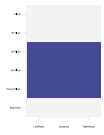
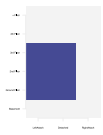
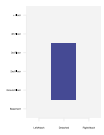
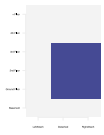
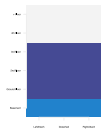
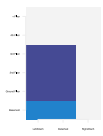
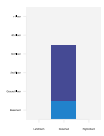
Table 4.1 presents an overview of the main properties describing the physical typologies in the housing stock. Each archetype represents a group of cuboids with similar number of floors (1 to 4), presence of attic (*w*ith or *n*o), presence of basement (*w*ith or *n*o) and attached walls (*l*eft, *r*igth, *b*oth or *n*o). As identified Section 2.4, semi-detached houses are the most common typology in the housing stock; Table 4.1 can represent them in multiple configurations, depending on the number of floors, such as *2FNANBLA*, *2FNANBRA*, *2FNAWBLA*, *2FWANBLA*, *2FWANBRA*.

Table 4.1: Stock typologies represented by cuboids

Archetype	Stock Attributes	Archetype	Stock Attributes
 **1FNANBBA**	1.7M hhlds · 40% built between 1940-1975 · Gas heater owned by $\frac{2}{3}$ · Floor area ∼ 52 m^2 · Household size ∼ 1.34 · 40% Private Tenure · 33% in London	 **1FNANBLA**	1.3M hhlds · 45% built between 1940 and 1970 · Gas heater owned by 70% · Floor area ∼ 58 m^2 · Household size ∼ 1.31 · 43% Private Tenure .
 **1FNANBNA**	1.4M hhlds · 10% built before 1945 · Gas heater owned by 78% · Floor area ∼ 79 m^2 · Household size ∼ 1.39 · 74% Private Tenure .	 **1FNANBRA**	1.2M hhlds · 60% built between 1940 and 1970 · Gas heater owned by 70% · Floor area ∼ 57 m^2 · 60% Public Tenure .
 **1FNAWBBA**	0.2M hhlds · 50% built before 1920 · Electric heater owned by 37% · Floor area ∼ 49 m^2 · Household size ∼ 1.41 · 53% Private Tenure · 50% in London	 **1FNAWBLA**	0.1M hhlds · 31% built before 1900 · Gas heater owned by 62% · Floor area ∼ 59 m^2 · Household size ∼ 1.46 · 67% Private Tenure · 52% in London
 **1FNAWBNA**	0.03M hhlds · 18% built after 1970 · Gas heater owned by 76% · Floor area ∼ 64 m^2 · 52% in South .	 **1FNAWBRA**	0.1M hhlds · 48% built before 1920 · Gas heater owned by 50% · Floor area ∼ 59 m^2 · 61% Private Tenure · 75% in South .
 **1FWANBNA**	0.002M hhlds · 100% built between 1965 and 1974 · Oil heater owned by 100% · Floor area ∼ 152 m^2 · Household size ∼ 1.5 · 100% Private Tenure .	 **2FNANBBA**	3.4M hhlds · 20% built after 1975 · Gas heater owned by 92% · Floor area ∼ 82 m^2 · Household size ∼ 1.5 · 78% Private Tenure · 37% in North
 **2FNANBLA**	3.5M hhlds · 67% built between 1920 and 1974 · Gas heater owned by 92% · Floor area ∼ 92 m^2 · Household size ∼ 1.5 · $\frac{1}{3}$ Public Tenure .	 **2FNANBNA**	2.8M hhlds · 42% built after 1980 · Gas heater owned by 86% · Floor area ∼ 137 m^2 · Household size ∼ 1.5 · 90% Private Tenure .
 **2FNANBRA**	3.3M hhlds · 57% built between 1920 and 1965 · Gas heater owned by 91% · Floor area ∼ 92 m^2 · Household size ∼ 1.49 · 69% Private Tenure .	 **2FNAWBBA**	0.001M hhlds · 100% built after 1980 · CHP heater owned by 100% · Floor area ∼ 100 m^2 · Household size ∼ 1.61 . .

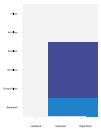
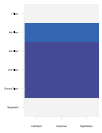
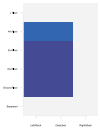
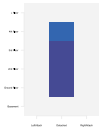
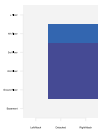
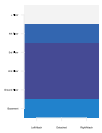
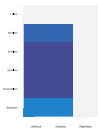
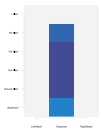
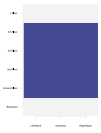
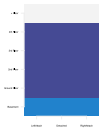
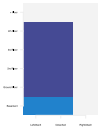
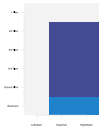
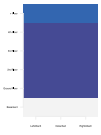
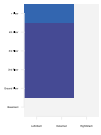
Continued on next page

Table 4.1 – continued from previous page

Archetype	Stock Attributes	Archetype	Stock Attributes
 **2FNAWBLA**	0.004M hhlds · 70% built between 1965 and 1980 · Gas heater owned by 100% · Floor area ∼ 67 m^2 · 70% LA Tenure · 70% LA Tenure .	 **2FNAWBNA**	0.009M hhlds · 50% built between 1945 and 1975 · Gas heater owned by 65% · Floor area ∼ 125 m^2 · Household size ∼ 1.5 · 100% Private Tenure .
 **2FWANBBA**	0.034M hhlds · 57% built before 1940 · Gas heater owned by 90% · Floor area ∼ 99 m^2 · Household size ∼ 1.6 · $\frac{3}{4}$ Private Tenure .	 **2FWANBLA**	0.2M hhlds · 65% built between 1945 and 1980 · Gas heater owned by 88% · Floor area ∼ 119 m^2 · Household size ∼ 1.5 · 90% Private Tenure .
 **2FWANBNA**	0.5M hhlds · 17% built after 1980 · Gas heater owned by 75% · Floor area ∼ 173 m^2 · Household size ∼ 1.5 .	 **2FWANBRA**	0.2M hhlds · 64% built between 1945 and 1975 · Gas heater owned by 86% · Floor area ∼ 106 m^2 · Household size ∼ 1.5 · 93% Private Tenure .
 **2FWAWBNA**	0.002M hhlds · 100% built between 1945 and 1965 · Gas heater owned by 100% · Floor area ∼ 192 m^2 · Household size ∼ 1.51 · 100% Private Owner Tenure · 100% East Midlands	 **3FNANBBA**	0.1M hhlds · 34% built after 1990 · Gas heater owned by 94% · Floor area ∼ 130 m^2 · Household size ∼ 2 · 37% Public Tenure .
 **3FNANBLA**	0.049M hhlds · 44% built after 1990 · Gas heater owned by 100% · Floor area ∼ 131 m^2 · Household size ∼ 1.9 · 82% Private Tenure .	 **3FNANBNA**	0.02M hhlds · 44% built before 1850 · Gas heater owned by 78% · Floor area ∼ 224 m^2 · Household size ∼ 1.5 · 100% Private Tenure .
 **3FNANBRA**	0.031M hhlds · 40% built after 2000 · Gas heater owned by 100% · Floor area ∼ 135 m^2 · Household size ∼ 1.65 · $\frac{1}{3}$ Private Rented .	 **3FNAWBBA**	0.042M hhlds · 76% built before 1920 · Gas heater owned by 100% · Floor area ∼ 118 m^2 · Household size ∼ 1.51 · 42% Private Rented .
 **3FNAWBLA**	0.022M hhlds · 62% built before 1940 · Gas heater owned by 100% · Floor area ∼ 114 m^2 · Household size ∼ 1.19 · 72% Private Owned · 80% in North West	 **3FNAWBNA**	0.028M hhlds · 46% built after 1970 · Gas heater owned by 87% · Floor area ∼ 115 m^2 · Household size ∼ 1.8 · 94% Private Owned .

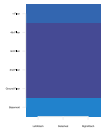
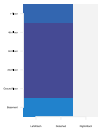
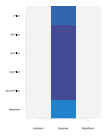
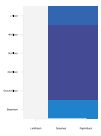
Continued on next page

Table 4.1 – continued from previous page

Archetype	Stock Attributes	Archetype	Stock Attributes
 **3FNAWBRA**	0.009M hhlds · 75% built before 1940 · Gas heater owned by 87% · Floor area ∼ 108 m^2 · Household size ∼ 1.95 · 100% Private Tenure .	 **3FWANBBA**	0.5M hhlds · 69% built before 1920 · Gas heater owned by 94% · Floor area ∼ 123 m^2 · Household size ∼ 1.9 · 28% Private Rented .
 **3FWANBLA**	0.3M hhlds · 73% built before 1945 · Gas heater owned by 96% · Floor area ∼ 151 m^2 · Household size ∼ 1.8 · 96% Private Tenure .	 **3FWANBNA**	0.176M hhlds · 25% built after 1990 · Gas heater owned by 80% · Floor area ∼ 267 m^2 · Household size ∼ 1.8 · 90% Private Tenure .
 **3FWANBRA**	0.3M hhlds · 70% built before 1940 · Gas heater owned by 90% · Floor area ∼ 146 m^2 · Household size ∼ 1.8 · 96% Private Tenure .	 **3FWAWBBA**	0.002M hhlds · 100% built before 1900 · Gas heater owned by 100% · Floor area ∼ 67 m^2 · Household size ∼ 3.5 · 100% Private Tenure .
 **3FWAWBLA**	0.002M hhlds · 100% built between 1945 and 1965 · Gas heater owned by 100% · Floor area ∼ 35 m^2 · Household size ∼ 1 · 100% Private Tenure .	 **3FWAWBNA**	0.009M hhlds · 60% built before 1960 · Gas heater owned by 100% · Floor area ∼ 147 m^2 · Household size ∼ 1.5 · 100% Private Tenure .
 **4FNANBBA**	0.002M hhlds · 100% built after 1990 · Gas heater owned by 100% · Floor area ∼ 163 m^2 · Household size ∼ 1.95 .	 **4FNAWBBA**	0.016M hhlds · 80% built before 1944 · Gas heater owned by 100% · Floor area ∼ 157 m^2 · Household size ∼ 1.61 .
 **4FNAWBLA**	0.005M hhlds · 100% built before 1900 · Gas heater owned by 100% · Floor area ∼ 110 m^2 · Household size ∼ 1.81 .	 **4FNAWBRA**	0.002M hhlds · 100% built before 1900 · Gas heater owned by 100% · Floor area ∼ 365 m^2 · Household size ∼ 2.31 .
 **4FWANBBA**	0.006M hhlds · 70% built before 1945 · Gas heater owned by 100% · Floor area ∼ 152 m^2 · Household size ∼ 2 · 100% Private Rented .	 **4FWANBLA**	0.002M hhlds · 100% built before 1900 · Gas heater owned by 100% · Floor area ∼ 250 m^2 · Household size ∼ 2.65 · 50% LA Tenure .

Continued on next page

Table 4.1 – continued from previous page

Archetype	Stock Attributes	Archetype	Stock Attributes
 **4FWAWBBA**	0.046M hhlds · 100% built before 1940 · Gas heater owned by 100% · Floor area ∼ 122 m^2 · Household size ∼ 1.8 · 51% Private Rented .	 **4FWAWBLA**	0.017M hhlds · 90% built before 1920 · Gas heater owned by 100% · Floor area ∼ 246 m^2 · Household size ∼ 1.61 · 30% Private Rented .
 **4FWAWBNA**	0.005M hhlds · 100% built before 1920 · Gas heater owned by 100% · Floor area ∼ 89 m^2 · Household size ∼ 2.35 . .	 **4FWAWBRA**	0.014M hhlds · 100% built before 1940 · Gas heater owned by 100% · Floor area ∼ 220 m^2 · Household size ∼ 2.2 · 90% Private Owned .

End of **Table 4.1**

Notation: (#)**F**loors-(with)**A**ttic-(with)**B**asement-(side)**A**ttachments

As a preliminary step, the characterisation of physical archetypes provides the basis for developing a volumetric configuration of models to be further evaluated, and is a convenient way to identify the relevance of shapes composing the stock.

Therefore, to make a workable platform, all the modules need to be explicitly integrated onto the platform. This process is referred to as reconciliation, which improves the consideration of the three pillars shaping the housing stock: contextual factors, physical composition and behaviours. The architecture components are based on both the `EnergyPlus Input Data Files (idf)` documentation [40, 212] and the BREDEM structure [95], through its application via the CHM protocol [108]. The OOM approach provides the ability to detach specific sections of the modelling process, to be refined independently. The UML architecture is presented in Figure 4.3

For instance, changes on physical attributes can be compared against average figures of household appliances' usage and the diffusion of energy-related policies; similarly, changes on the household properties can be reflected in the model outputs, depending on the model sophistication. The ability to capture all these variations is used to evaluate the effectiveness of (future) targeted policies, programmes and occupant behaviour, where the analysis of co-benefits may be refined.

It is worth noting that, with regard to the balance between granularity and sophistication, if the scope defined by the simulation platform is related to the whole stock, then a set of limited modelling algorithms might be sufficient at providing general results; however, if the scope is aiming at specific groups, then the algorithms require more sophisticated algorithms. In order to make this possible, each of the modules is linked within the EnHub platform, so that any change in the modules can be reflected along the process, augmenting thus the functionality of the platform.

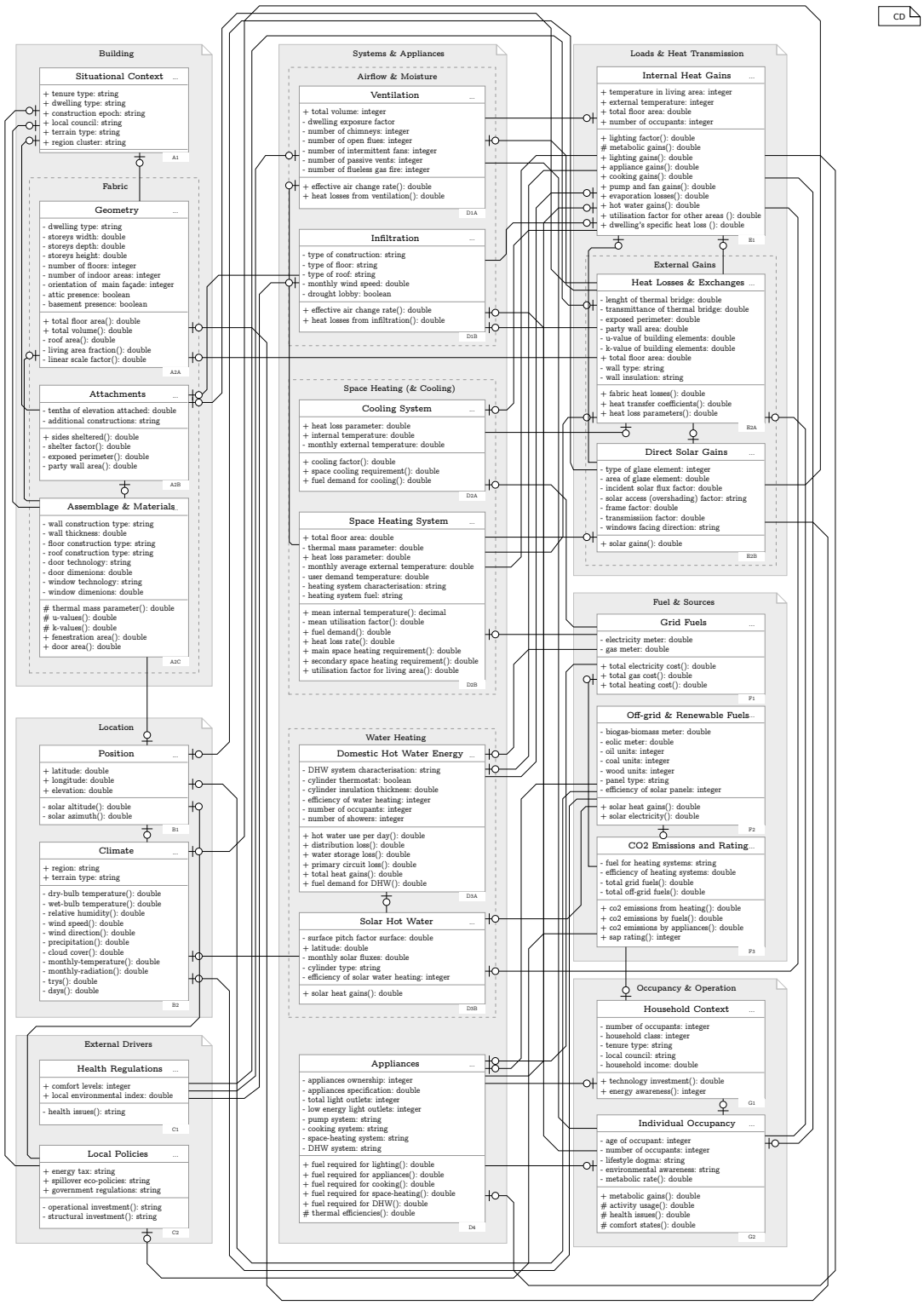


Figure 4.3: EnHub architecture of the modelling platform - see Appendix D for detailed description

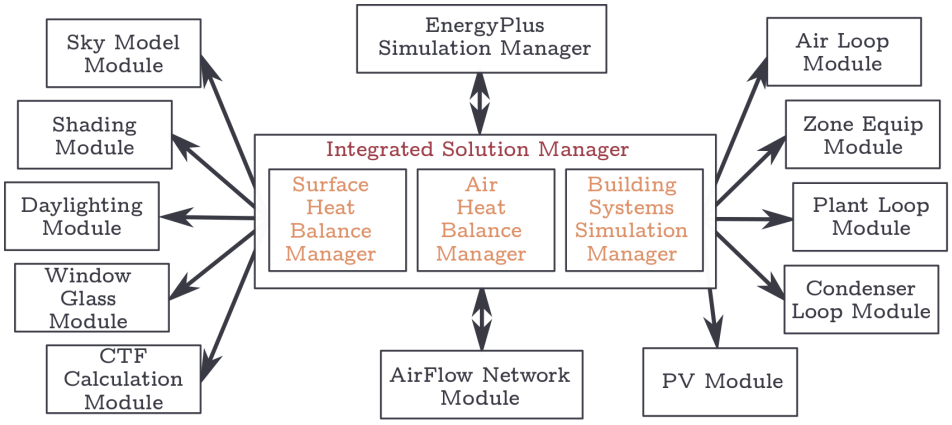


Figure 4.4: `EnergyPlus` program schematic
[137]

4.1.3 `EnergyPlus` Modules

`EnergyPlus` combines multiple sub-modules to calculate (and simulate) the energy flow in buildings. As described in Box 4.1.2, the core of the calculation is based on fundamental heat balance principles [137]. Figure 4.4 outlines the structure of the `EnergyPlus` software. Three components are at the core of the solution manager: surface heat balance, air heat balance, and building systems simulation. Within each of the core components, a number of modules are available for the application of specific tasks, such as sky model, shading, daylighting, window glass, Conduction Transfer Function (CTF), air loop, zone equipment, plant loop, condenser loop, and photovoltaic system. Some available modules are not employed for the calculation of energy flow in residential buildings, namely HVAC loops. The main components are complemented by integrating an airflow network manager, which provides the ability to simulate the performance of an air distribution system, including thermal conduction and air leakage losses.

`EnergyPlus` defines a thermal zone, which does not necessarily corresponds to a geometrical zone. One or more geometrical zones may be contained in a thermal zone (hereinafter called as zone). `EnergyPlus` calculates the energy required to maintain each zone at a specified temperature for a given period. Therefore, the number of zones in a building should be limited by the resolution of the data, as well as by the need to analyse specific indoor performance parameters.

In summary, the configuration of building elements in `EnergyPlus` is described as following: a building contains zones; a zone contains a number of surfaces (walls, floor, roof, ceiling); each surface assigns a material; and each material contains thermal properties. Furthermore, a surface may be split in different types, such as openings (e.g. doors, windows), and surfaces with different materials and uses (e.g. structural walls, partition walls, indoor ceilings, roofs).

In addition to the emulated physical properties, a surface may be defined either as a heat storage surface, or as a heat transfer surface (i.e. a differential of temperature between divided zones). As a result, this affects the way each element is treated by the zone heat balance algorithms. These algorithms significantly improve the number of steady-state methods which have been at the core of previous models, for example BREDEM (see § 3.2.2). The addition of time-related algorithms [216], in addition to volumetric archetypes, improves the study of heat balance in the residential buildings.

The basis of a heat balance, as defined previously in Equation 3.1, involves external gains (solar, other zones), internal gains (appliances, heating systems, occupants), and losses (heat transfer to zones, losses to the environment). `EnergyPlus` simulates the balance dynamically (i.e. in a time frame). Therefore, this balance is expressed as a derivative with respect to time.

$$\dot{Q}_{ext} + \dot{Q}_{int} + \dot{Q}_{air} = 0 \tag{4.1}$$

A zone air balance $\dot{Q}_{air}$ is formed by the differential value between the energy stored $\dot{Q}_{str}$ and the output of air systems $\dot{Q}_{sys}$ (i.e. injection of hot or cold air), which in residential buildings may be highly simplified; here, the heating demand is low as compared to non-residential buildings with large Heating, Ventilation and Air Conditioning (HVAC) system demands, and the cooling demand is mostly negligible across the housing stock.

$$\dot{Q}_{air} = \dot{Q}_{str} - \dot{Q}_{sys} \tag{4.2}$$

The external gains $\dot{Q}_{ext}$ are composed by inter-zone gains $\dot{Q}_{izn}$, convective heat transfer from the zone surfaces $\dot{Q}_{srf}$ and heat transfer due to infiltration of outside air $\dot{Q}_{inf}$.

$$\dot{Q}_{ext} = \dot{Q}_{izn} + \dot{Q}_{srf} + \dot{Q}_{inf} \tag{4.3}$$

where:

$\dot{Q}_{izn} = \sum_{i=1}^{N_{zones}} \dot{m}_i C_p \left(T_{zi} - T_z\right)$

$\dot{Q}_{srf} = \sum_{i=1}^{N_{surfaces}} h_i A_i \left(T_{si} - T_z\right)$

$\dot{Q}_{inf} = \dot{m}_{\text{inf}} C_p \left(T_\infty - T_z\right)$

$m_{i,\text{inf}}$ = mass flow (internal, outdoor)

C_p = zone air specific heat

h_i = heat transfer coefficient

A_i = surface area (internal)

$T_{i,z,\text{inf}}$ = temperature (internal, external zone and outdoor)

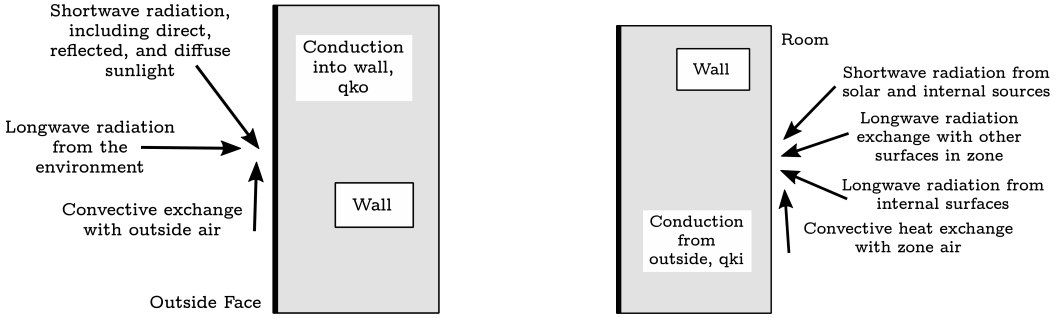
`EnergyPlus` enables multiple methods to calculate parameters used in the balance, such as specific heat C_p and heat transfer coefficient h_i. This is defined by the modeller according to the purpose of the study. As seen, the calculation of $\dot{Q}_{srf}$ is made for each surface $N_{surfaces}$ in the zone. This is one of the major advantages of employing volumetric archetypes to represent dwellings, and to estimate heat transfer.

Because both internal and external gains are dropped into the zone, the resulting zone air balance

(Eq. 4.2) may be expressed as following:

$$\dot{Q}_{air} = \dot{Q}_{int} + \dot{Q}_{ext} \tag{4.4}$$

The method employed to calculate heat transfer from surfaces $\dot{Q}_{srf}$ is selected according to the type of surface, or the zones which divides. Figure 4.5a represents the variables involved in the estimation of heat transferred from the outside. Figure 4.5b represents the ones involved in the internal heat transfer. Adiabatic walls considers similar conditions in both sides so they can be balanced. Yet, the internal gains may affect the corresponding surface. For example, the analysis of surfaces between zones with differing temperatures may represent a controlled zone and an attachment. Here, the attachment will perform a simplified balance, but will affect the common wall as opposed to an adiabatic element or outdoor boundary.



(a) Outside heat *(b)* Inside heat

Figure 4.5: Heat balance control volume diagram
Source: [137]

Most of the estimations regarding conduction is performed by applying CTFs. CTFs define the conduction parameter in terms of temperatures and constant values for each material property. CTFs are solved, in `EnergyPlus`, with state space methods, which express the current heat flux in terms of the current temperature and historic temperature and heat flux values.

For such time-related balances, a finite difference approximation is employed to calculate the derivative functions, defined below, which is solved in `EnergyPlus` via numerical method solvers.

$$\frac{dT}{dt} = (\delta t)^{-1} \left({T_z}^t - {T_z}^{t-\delta t} \right) + \mathrm{O}\left(\delta t\right) \tag{4.5}$$

`EnergyPlus` follows an equivalent consideration of components to estimate moisture balance, which is used to complement the zone air balance [137]. It also follows a similar approach to estimate carbon dioxide and other pollutant transfers. Moreover, a prominent module available in `EnergyPlus` is for the estimation of shading. This enables the analysis of solar radiation, distinguishing between shaded areas and those with direct sunlight. The shading module also affects the estimation of ratios of interior illuminance and glare. Which, in a future stage, may be linked to more sophisticated occupancy and behaviour models.

4.1.4 Enriched Volumetric Generation

To link the EHS with the typology-specific idf modules, a base case idf is developed where variable flags[1] are used to represent variables in the UML classes. The base case idf is a stable file upon which the edition of components are referenced. By using this, each EHS sample unit fills their corresponding property fields and scales the volumetric model, accordingly. For example, Script 4.1 displays a fragment of the idf with specific flag variables to be scaled (c_{var}).

Code 4.1 Example of flags in the typology-specific idf file. Here, three variables are represented: c_CuD (main cuboid depth), c_CuW (main cuboid width) and c_GfH (main cuboid height); they correspond to vertices in the volumetric cuboid, which are scaled according to survey information

```
BuildingSurface:Detailed,
    GF:Roof,                  !- Name
    Ceiling,                  !- Surface Type
    Interior Ceiling,         !- Construction Name
    GF,                       !- Zone Name
    Surface,                  !- Outside Boundary Condition
    1F:Floor,                 !- Outside Boundary Condition Object
    NoSun,                    !- Sun Exposure
    NoWind,                   !- Wind Exposure
    0.0,                      !- View Factor to Ground
    4,                        !- Number of Vertices
    0.000000000000,           !- Vertex 1 X-coordinate {m}
    c_CuD,                    !- Vertex 1 Y-coordinate {m}
    c_GfH,                    !- Vertex 1 Z-coordinate {m}
    0.000000000000,           !- Vertex 2 X-coordinate {m}
    0.000000000000,           !- Vertex 2 Y-coordinate {m}
    c_GfH,                    !- Vertex 2 Z-coordinate {m}
    c_CuW,                    !- Vertex 3 X-coordinate {m}
    0.000000000000,           !- Vertex 3 Y-coordinate {m}
    c_GfH,                    !- Vertex 3 Z-coordinate {m}
    c_CuW,                    !- Vertex 4 X-coordinate {m}
    c_CuD,                    !- Vertex 4 Y-coordinate {m}
    c_GfH;                    !- Vertex 4 Z-coordinate {m}
```

The archetype generation begins with reading in a base case idf and adapting this using variable flags to match the property fields from the EHS; the unused geometric blocks are also removed from the file. Similarly, Script 4.2 reveals how the idf initially considers a combination with all the cuboid blocks (see Figure 4.1a), which are then adapted accordingly. Noteworthy, each thermal zone is used to evaluate air balance, but the list of zones are used to feed parameters related to the wholes building, such as appliances, occupants, fuels and lighting.

Code 4.2 Example of base case in the typology-specific idf file. According to the survey data, the assignment of sub-lists can be adjusted, which allows the study of more specific metrics in the volumetric archetype

```
!-   ===========  ALL OBJECTS IN CLASS: ZONELIST ===========

ZoneList,
    Zone:All,                 !- Name
    GF,                       !- Zone 1 Name
    1F,                       !- Zone 2 Name
    2F,                       !- Zone 3 Name
    3F,                       !- Zone 4 Name
    A,                        !- Zone 5 Name
    B;                        !- Zone 6 Name
```

[1] In programming, a flag is a predefined string or sequence that is used as a reference when a script is read or executed.

```
11
12 ZoneList,
13     Zone:Main,                !- Name
14     GF;                       !- Zone 1 Name
15
16 ZoneList,
17     Zone:Secondary,           !- Name
18     B,                        !- Zone 1 Name
19     1F,                       !- Zone 2 Name
20     2F,                       !- Zone 3 Name
21     3F,                       !- Zone 4 Name
22     A;                        !- Zone 5 Name
```

Therefore, employing a base case idf is convenient due to the precision in the coordinates that `EnergyPlus` demands. Further, the idfs can be combined in any order, but it is convenient to follow the sequential structure of the `EnergyPlus` format [40, 212]. The construction of models involves the nine following sections:

- *A. Simulation Parameters :* Here, general information is fed into the model such as the *building name*, which is also the global identifier used in the analysis. The *timestep* is adjusted according to the resolution of the dataset and expected outcomes. The *timestep* of one hour is chosen based on the availability of data to support internal gains. For example, Section 2.1.1 presented the average profile in energy-use of domestic appliances, which can be adapted to such a timestep resolution.

  In addition, if there is an external file or function defined at a different resolution (e.g. weather file, solar incidence data, activity profiles), a process of inter- or extra-polation is performed, so that the timesteps correspond. This implies, consequently, that some information can be missed in the process, if the resolution of the datasets is different; but on the other hand, it can add with flexibility to the process. For example, by linking information from different datasets where the resolution is limited or scarce.

- *B. Location and Period :* This module assigns the *location* to link calculations of sun position with the climate context. This value is used to directly assign weather files onto the models. To this end, a library of standard EnergyPlus Weather (EPW) files is referenced (and identified in the idf) according to the region where the model archetype is allocated. These files are synthesised Test Reference Years (TRYs), where each month represents an average month over a period of around 20 years. Each month may come from a different year, and so they are combined using a cubic spline method [223]. TRYs are appropriate for the prediction of average annual energy usage and average incidences of thermal comfort. The EPW files include hourly or sub-hourly information regarding time zone, dry-bulb temperature, dew-point temperature, relative humidity, barometric pressure, direct normal radiation, diffuse horizontal radiation, total and opaque sky cover, wind direction, wind speed, as well as some auxiliary data such as rain, snow, elevation, peak heating and cooling design conditions, holidays, daylight savings period, typical and extreme periods [137].

  The simulation period is defined for an entire calendar year, which could be potentially reduced, by simulating specific days or by selecting seasons, as is the case with design days used for the characterisation of HVAC. However, the difference in the extrapolation of this mode and the one-year period at such a timestep resolution is not significant. In addition, by simulating for an entire year, metrics related to indoor conditions can be applied directly (i.e. energy peak demands, over- and under-heated areas).

Table 4.2: Climate data summary for the EHS regions
Source: SAP Table U1-3

(a) Mean external temperature (°C) at sea level

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
London	4.5	5.0	6.9	9.4	12.2	15.3	17.3	17.2	14.6	11.0	7.4	4.5
South East	4.3	4.7	6.5	9.1	12.1	14.8	16.9	16.9	14.3	10.8	7.3	4.5
South West	6.7	6.8	7.9	9.7	12.4	15.1	16.9	17.1	15.2	12.4	9.5	6.9
East Midlands	3.9	4.4	6.2	8.6	11.5	14.3	16.0	15.9	13.5	10.2	6.7	3.8
West Midlands	5.2	5.6	7.1	9.3	12.3	14.9	16.6	16.5	14.4	11.2	8.1	5.2
North West	3.1	3.7	5.1	7.4	10.1	12.8	14.5	14.3	12.1	9.0	5.8	2.9
North East	3.1	3.6	5.0	7.2	9.8	12.7	14.7	14.6	12.4	9.1	5.7	2.9
Yorkshire and the Humber	4.0	4.4	5.9	8.3	11.1	14.0	15.8	15.6	13.4	10.0	6.7	3.8
East of England	4.6	5.1	6.9	9.5	12.5	15.5	17.7	17.6	15.0	11.4	7.6	4.6
	4.4	4.8	6.4	8.7	11.6	14.4	16.3	16.2	13.9	10.6	7.2	4.3

(b) Wind speed (m/s) for calculation of infiltration rate

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
London	6.3	6.1	5.3	5.0	4.7	4.4	4.5	5.1	5.8	6.0	5.9	20
South East	4.4	4.3	3.7	3.7	3.4	3.4	3.2	3.4	3.6	3.8	4.1	31
South West	6.0	6.2	5.6	5.5	4.9	5.1	5.0	5.7	6.5	6.4	6.9	36
East Midlands	4.0	4.0	3.5	3.4	3.2	3.1	3.0	3.1	3.4	3.5	3.7	26
West Midlands	5.5	5.2	4.6	4.5	4.2	4.1	4.2	4.3	4.9	5.0	5.2	27
North West	5.7	5.5	4.7	4.5	4.1	3.7	3.6	4.2	4.7	5.2	5.1	22
North East	6.2	5.9	5.1	4.9	4.6	4.2	4.3	4.9	5.4	5.9	5.7	25
Yorkshire and the Humber	5.2	5.0	4.4	4.2	3.9	3.9	3.7	4.0	4.3	4.6	4.9	25
East of England	4.5	4.4	4.1	4.0	3.6	3.7	3.7	3.8	4.0	4.0	4.3	31
	5.3	5.2	4.6	4.4	4.1	4.0	3.9	4.3	4.7	4.9	5.1	27

(c) Mean global solar irradiance (W/m2) on a horizontal plane

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
London	20	46	92	147	200	193	184	149	103	57	27	15
South East	31	57	100	164	200	223	209	176	130	75	40	25
South West	36	63	113	178	217	239	214	187	139	80	44	28
East Midlands	26	53	96	155	192	207	198	163	118	70	33	21
West Midlands	27	55	101	161	204	222	198	162	123	69	33	21
North West	22	49	94	153	193	197	189	153	112	60	29	19
North East	25	52	97	154	198	199	190	159	115	64	31	20
Yorkshire and the Humber	25	52	94	146	185	193	180	151	111	63	29	20
East of England	31	58	102	170	208	225	210	177	130	76	40	25
	27	54	99	159	200	211	197	164	120	68	34	22

Similarly to the CHM, the regions are defined by those available in the EHS, which correspond to the NUTS-1 layer (see Table 2.3). Figure 4.6 shows the available EHS regions, as well as the associated EPW files used in the simulation. Table 4.2 presents different climatic values with respect to the EHS regions, such as external temperatures, wind speeds and solar radiations. Figure 2.23 also illustrates the slight differences in such values. Nevertheless, the inclusion of climate data might be further improved by adopting the number of regions defined in both Chartered Institution Building Services Engineers (CIBSE) and SAP catalogues, that specify differences for about 18 and 21 regions, respectively.

- *C. Schedules associated :* The process of constructing idfs requires the definition of *schedules*



Figure 4.6: EHS regions with EPW reference

to emulate dynamic actions of devices and occupants. These include time-related events, such as occupant presence, or the use of appliances, window openings, heating systems, and water fixtures. In addition, it has been identified that, historically, previous HSEMs have neglected or over-simplified the levels of occupancy and patterns of use in their representation of activity in dwellings (see § 3.4.2).

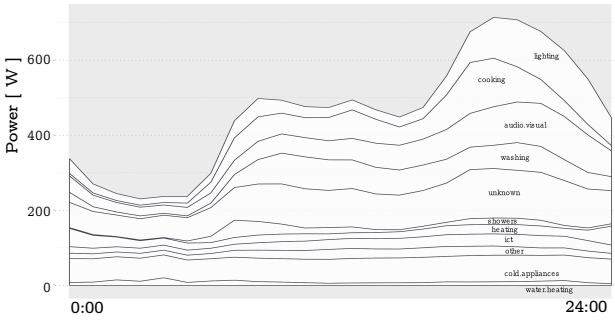
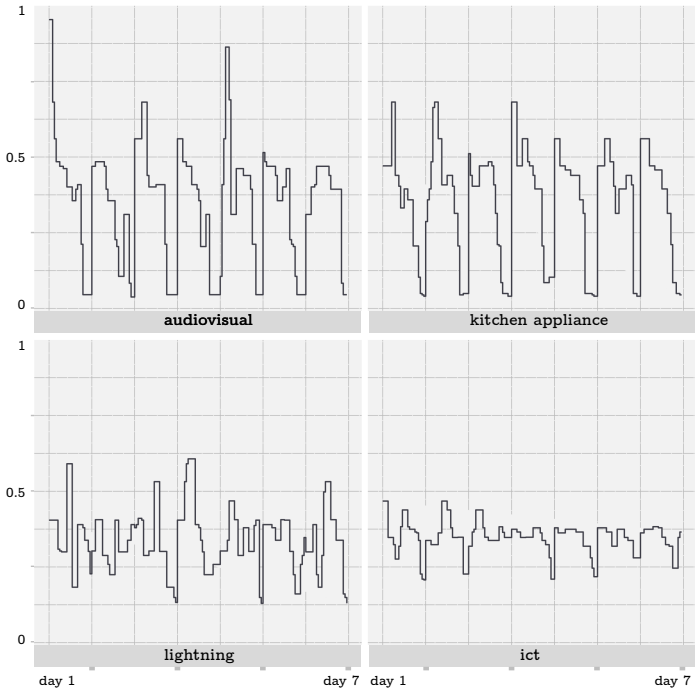
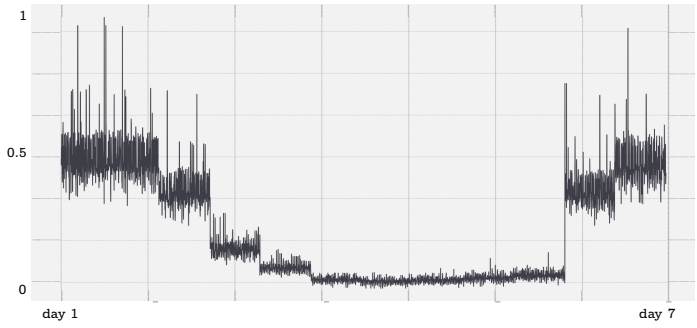


Figure 4.7: Average 24-hour electricity use profile for owner-occupied homes
Source: [60]

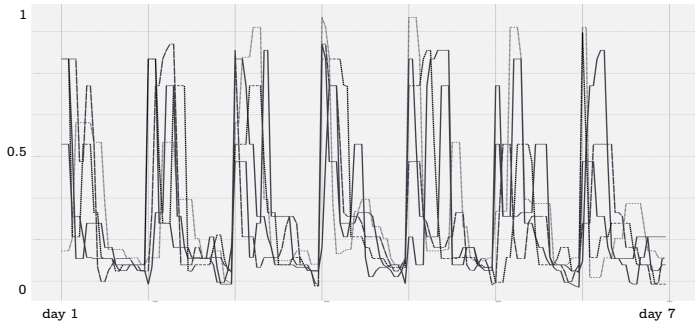
In the idf profiles, this is improved in two stages. First, by considering the typical profile shown in Figure 4.7, the corresponding profile for each system appliance is replicated for an entire year at an hourly timestep resolution. Here, it can be noted that the most stable profiles are those associated with cold and wet appliances. When associated to their corresponding power demands, the yearly energy demand matches the equivalent outcome from a constant value; however, it includes a more explicit representation of the daily contribution, not only to the energy demand but also to associated heat gains. Second, each household-specific schedule is improved by employing a probabilistic function that is dependent on the properties for each archetype. The procedure to form this function is



(a) Four profiles



(b) Seasonal profiles



(c) Probabilistic profiles

Figure 4.8: Example of associated EnergyPlus Input Data Files profiles

described in Pseudo-code 4.3.

Pseudo-code 4.3 Probabilistic profile generation

```
1. extract profile from national average daily profiles
  M_n -> extract(data: average usage); n elements

2. extract representative household parameters
  Hh_p -> extract(data: household); p population

3. assign a probabilistic function to each profile
based on average daily usage
  d_n -> P(M_n)

4. create random profile based on p-functions for a given period
  loop:
    for i:=1 to 1016 do
      for j:=1 to n do
        for k:=1 to 365 do
```

In this way, the probabilistic profiles augment the ability to study detailed responses in the simulated energy flow pathways; and, more importantly, include sophisticated models of occupancy capturing levels of stochasticity on each profile. As a result, the patterns in here are shaped by the properties of each archetype such as household composition or dwelling type. This means that even though the profiles follow similar patterns, and are based on national average figures, they are unique for each dwelling archetype. The resolution of the schedule is independent of the model, but they are conveniently paired with the timestep at an hourly resolution [40]. Currently, schedules are constant for an entire year, but they can be seasonal.

Figure 4.8a illustrates four different profiles associated to time-related components. These profiles are made for an entire year and can be shaped to seasonal periods. For instance, Figure 4.8b illustrates the profile for a heating system, affected by the seasonal period. Figure 4.8c shows different profiles associated to the same component (i.e. a probabilistic pattern describing the activity or the use of a household appliance), which enables the possibility to analyse, albeit limited by the algorithms, specific behaviours on a given group of archetypes. It also enables the possibility to employ more sophisticated algorithms that represent disaggregated household appliances' usage.

Pseudo-code 4.4 Example of a material catalogue in the typology-specific idf file. The thermal properties (L,k,d,c) for each material (m) are adjusted according to survey data and the SAP catalogue

```
Material,
     b_Cavity,                !- Name
     Rough,                   !- Roughness
     b_L,                     !- Thickness {m}
     b_k,                     !- Conductivity {W/m-K}
     b_d,                     !- Density {kg/m3}
     b_c;                     !- Specific Heat {J/kg-K}

Material,
     c_Heavy concrete block,  !- Name
     VeryRough,               !- Roughness
     c_L,                     !- Thickness {m}
     c_k,                     !- Conductivity {W/m-K}
     c_d,                     !- Density {kg/m3}
     c_c;                     !- Specific Heat {J/kg-K}
```

Table 4.3: Building properties for England and Wales
Source: SAP Tables S3 and S6

(a) Wall Thickness (mm)

	Wall Construction Type / Epoch	**before 1900**	**1900 - 1929**	**1930 - 1949**	**1950 - 1966**	**1967 - 1975**	**1976 - 1982**	**1983 - 1990**	**1991 - 1995**	**1996 - 2002**	**2003 - 2006**	**2007 +**
1	**Stone: granite or whin (as built)**	500	500	500	500	450	420	420	420	450	450	450
2	**Stone: sandstone (as built)**	500	500	500	500	450	420	420	420	450	450	450
3	**Solid brick (as built)**	220	220	220	220	240	250	270	270	300	300	300
4	**Stone/solid brick (external insulation)**	290	290	290	290	310	320	340	340	370	370	370
5	**Stone/solid brick (internal insulation)**	290	290	290	290	310	320	340	340	370	370	370
6	**Cob (as built)**	540	540	540	540	540	540	560	560	590	590	590
7	**Cob (external insulation)**	590	590	590	590	590	590	610	610	640	640	640
8	**Cob (internal insulation)**	590	590	590	590	590	590	610	610	640	640	640
9	**Cavity (as built)**	250	250	250	250	250	260	270	270	300	300	300
10	**Filled cavity / Cavity with insulation**	250	250	250	250	250	260	270	270	300	300	300
11	**Timber frame**	150	150	150	250	270	270	270	270	300	300	300
12	**System build (as built)**	250	250	250	250	250	300	300	300	300	300	300
13	**System build (external insulation)**	320	320	320	320	320	370	370	370	370	370	370
14	**System build (internal insulation)**	320	320	320	320	320	370	370	370	370	370	370
15	**Metal Frame**	100	100	100	200	220	220	220	220	250	250	250

(b) U-value (W/m^2K)

	Wall Construction Type / Epoch	**before 1900**	**1900 - 1929**	**1930 - 1949**	**1950 - 1966**	**1967 - 1975**	**1976 - 1982**	**1983 - 1990**	**1991 - 1995**	**1996 - 2002**	**2003 - 2006**	**2007 +**
1	**Stone: granite or whin (as built)**	2.40	2.40	2.40	2.40	1.70	1.00	0.60	0.60	0.45	0.35	0.30
2	**Stone: sandstone (as built)**	2.10	2.10	2.10	2.10	1.70	1.00	0.60	0.60	0.45	0.35	0.30
3	**Solid brick (as built)**	2.10	2.10	2.10	2.10	1.70	1.00	0.60	0.60	0.45	0.35	0.30
4	**Stone/solid brick (external insulation)**	0.60	0.60	0.60	0.60	0.55	0.45	0.35	0.35	0.30	0.25	0.21
5	**Stone/solid brick (internal insulation)**	0.60	0.60	0.60	0.60	0.55	0.45	0.35	0.35	0.30	0.25	0.21
6	**Cob (as built)**	0.80	0.80	0.80	0.80	0.80	0.80	0.60	0.60	0.45	0.35	0.30
7	**Cob (external insulation)**	0.40	0.40	0.40	0.40	0.40	0.40	0.35	0.35	0.30	0.25	0.21
8	**Cob (internal insulation)**	0.40	0.40	0.40	0.40	0.40	0.40	0.35	0.35	0.30	0.25	0.21
9	**Cavity (as built)**	2.10	1.60	1.60	1.60	1.60	1.00	0.60	0.60	0.45	0.35	0.30
10	**Filled cavity / Cavity with insulation**	0.65	0.65	0.65	0.65	0.65	0.65	0.35	0.35	0.45	0.35	0.30
11	**Timber frame**	2.50	1.90	1.90	1.00	0.80	0.45	0.40	0.40	0.40	0.35	0.30
12	**System build (as built)**	2.00	2.00	2.00	2.00	1.70	1.00	0.60	0.60	0.45	0.35	0.30
13	**System build (external insulation)**	0.60	0.60	0.60	0.60	0.55	0.45	0.35	0.35	0.30	0.25	0.21
14	**System build (internal insulation)**	0.60	0.60	0.60	0.60	0.55	0.45	0.35	0.35	0.30	0.25	0.21
15	**Metal Frame**	2.20	2.20	2.20	0.86	0.86	0.53	0.53	0.53	0.45	0.35	0.30

- *D. Surface and Constructions :* The model is populated according to the characteristics of the envelope: walls, floors, roofs and windows. The properties of *constructions* and *materials* are taken from the CHM documentation. Also, the openings are adjusted and enriched with their physical properties. Table 4.3 shows an example of the catalogue employed, which is embedded into the model for the `EnergyPlus` (see Code 4.4).

  In `EnergyPlus`, each construction is defined as a layered configuration of materials, in which their thermophysical properties are specified for each of them. By combining these layers, a construction can represent any typical configuration of constructions found in the stock, such as cavity wall, solid wall, timber frame and steel frame. This enables a relatively

simple substitution of constructions attached to a given surface, and opens the possibility to further evaluate highly specific changes in the properties of the materials.

- *E. Thermal Zones and Geometry :* Once the absent geometries are removed from the cuboid, this is scaled and adapted according to the relation between opening areas and walls, as well as façade dimensions and internal heights. The zones are used to link existing elements and calibrate energy flows in the whole cuboid. This applies the BREDEM basis but now based on an explicit volumetric representation. In addition, each zone includes the relation between openings (i.e. windows, doors, accesses) and total floor area, which enhances the ability to study solar gains, as a function of the location of the dwelling archetypes, as opposed to the assumptions described in Section 3.3.1.

  As aforementioned, a base case idf is used to adapt each volumetric archetype (see Code 4.2). A different alternative would be to build up a geometry applying coordinates starting in a $(0,0,0)$ origin and specifying each coordinate for all the elements in the building (i.e. walls, floors, roofs, windows, doors), which, incidentally, represents the primitive process performed in the `EnergyPlus` configuration. Its application would, in theory, ease the development of the geometry, but the precision required the software is highly sensitive (i.e. 10^{-6}). For example, each zone and component require to match their counterparts in detail; however, although the volumetric construction may follow a logic sequence, its application requires a process of validation which may affect the functionality of the process and increase the computational time considerably. Moreover, for the cases with contiguous cuboids, each surface must contain its values following the idf standard. This approach, in practice, can be enhanced by defining an idf base case—assuming a limited number of shapes, which is validated and tested independently. By doing so, each shape can then be scaled and the shared properties of the elements can be maintained.

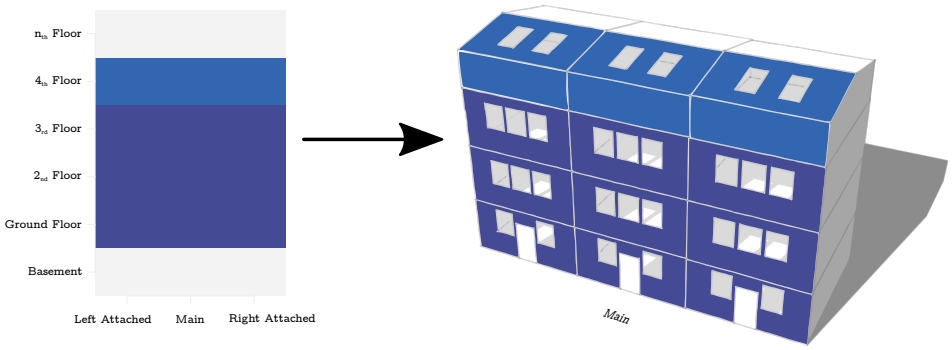


Figure 4.9: Example of a volumetric cuboid for simulation. In `EnergyPlus`, the attachments are not simulated with the same rigour as the main zone, yet they provide additional parameters for the consideration of boundaries

An example of the explicit volumetric representation is shown in Figure 4.9. At present, the resolution of the geometry is defined by the EHS granularity. The EHS does not consider the internal layout and the characterisation of detailed components (i.e. windows, doors, walls) is broadly considered. In the idf formation, this is improved by scaling the proportion of opening areas regarding the total floor area for each block. In addition, each zone is able to link information regarding appliances and occupancy, and may split the simulation results into each of them.

- *F. Internal Gains :* In terms of the internal gains, a catalogue of appliances (i.e. cooking devices, lighting, household appliances) and heating systems is included in the base case idf. Their specifications (i.e. power, efficiency and heat gains), and corresponding fuels, are shaped according to their EHS unit. Similarly, household water services are included in the `idfs`, although they are reduced to total values for a yearly demand. The `idf` structure also requires the definition of infiltration and ventilation parameters at this stage, for which the values are derived from BREDEM and the CHM.

  Finally, occupants are included with their associated (and theoretically) metabolic gains, and usage patterns. For this, a normalised value for the occupancy is initially estimated employing a distinction in terms of number of adults and children in the household, which is an application of the *OECD equivalence scale*, used in economics to represent income[2]; here, the equivalised Household Size (eHS) is defined by counting unity for one (adult) member of the household (H_a), α is an adjustment value for each subsequent adult, whereas β for every child in the household (H_c). This relation results in $eHS = 1 + \alpha \cdot (H_a - 1) + \beta \cdot H_c$. This is helpful to make a distinction between households with equal size but with different composition, in which a dissimilar energy intensity is expected.

  According to the levels of education and tenure type, provided by the EHS, a specific intensity is assigned to the model archetype. However, this module is only used for the study of occupancy profiles and One-at-a-time (OAT) sensitivity analyses. The standard module only differentiates in terms of the adults and specifies their theoretical metabolic properties. Here the occupancy is split in three different profiles according to the properties of their members with each profile (related to inner socio-demographic circumstances), following a similar approach to that of the DEMScot model. Here, occupants can be directly linked with specific intensity values, if provided, such as energy demand per person, or floor area per capita.

- *G. HVAC :* A set of sub-modules are added to the model according to the heating system, namely central heating with gas, gas room heater, electric room storage, or other systems that use oil or wood. The modules are simplistic enough to make a functional system; yet they lack in the internal network that exists on some dwellings.

  Also, despite the heating devices are simplified to central units, these may be replicated or scaled for each of the zones. In here, a fixed set-point is defined at 20°C for each of the archetypes; this value represents the average thermostat setting in the UK [53]. This value, however, is also compensated by the associated schedule in the heating systems, which may add active levels to the system, disregarding the set-point values. The module assumes both main and secondary systems, and the values are adjusted as specified in the EHS.

  A summary of typical heating system configurations is provided in Table 4.4. In BREDEM, these configurations used to define the heating efficiency, to estimate fuel demand, carbon contribution, electricity demand for pumps and fans, as well as indoor temperature adjustments. `EnergyPlus` requires higher resolution when simulating HVAC systems; however, for the domestic buildings represented in the idfs, the complexity of these systems can be simplified. Still, it must be considered that a typical configuration of heating systems in dwellings comprises a network where a centrally-heated system is linked to DHW devices, which significantly affects the internal heat gains across the dwellings.

[2] www.ons.gov.uk - People, population and community → Personal and household finances → Income and wealth → Family Spending → Equivalised income

Table 4.4: Fifteen high-level heating system types, used in the modelling process
Source: CHM [108]

.	Main heating system
1	Gas standard
2	Gas - combi
3	Gas back boiler
4	Oil standard
5	Solid boiler (house coal/anthracite)
6	Electric boiler
7	Electric storage
8	Electric room heater
9	Warm air - gas fired
10	Warm air - electric
11	Community heating without CHP
12	Community heating with CHP
13	Biomass boiler
14	Ground source heat pump
15	Air source heat pump

- *H. Water System :* Specifically, DHW values are not fully detailed in the EHS. For this, an assumption on the proportion of hot and cold water is defined in the BREDEM. These values are then passed to the properties of the water demand, and to the associated heat gains of the DHW systems. At present, the model does not integrate central heating with the assigned DHW system. The logical next step, in a future version of the platform, would be to represent such a combination of heating systems which, in practice, represent the most common configuration found in UK dwellings. In this stage, the model only characterise the boiler units so that the heat gains are defined in the simulation. A virtue of the architecture modularity is that multiple DHW devices may be defined in this module, which enables the possibility to integrate more sophisticated algorithms or databases that contain such information.

- *I. Outputs :* The outputs are specified to obtain comfort conditions, carbon emissions and heat gains. Their resolution is defined by the timestep and the simulated period. The associated carbon emissions are based on the DEFRA's GHG conversion guidelines [52]; however, average values are used instead of regional-specific ones. Indicators of thermal comfort are also computed in `EnergyPlus`; at present, these are based on the ASHRAE-55 standard [40], but they could in the future be readily augmented or substituted with more appropriate indicators of instantaneous comfort or indeed of over- or under-heating risk. They provide a general overview of the conditions in the defined zones and their indoor conditions.

Such a methodological composition of an idf, supported by the OOM approach, is thus useful to generate enriched volumetric archetypes based on survey data, which in this case is the EHS-2011; and semantically attributed to represent constructional, occupancy, appliance and system characteristics. The main components (that form the typology-specific idfs) are summarised in Table 4.5.

Table 4.5: `EnergyPlus` typology-specific components with UML reference - see Appendix D for detailed description

UML blocks

idf sections	A1	A2a	A2b	A2c	B1	B2	C1	C2	D1a	D2b	D2a	D2b	D3a	D3b	D4	E1	E2a	E2b	F1	F2	F3	G1	G2
A	x	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
B	x	.	.	.	.	.	.	x	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C	.	.	.	.	.	.	x	x	.	.	.	.	.	.	.	.	.	.	.	.	.	x	x
D	x	.	x	x	x	.	.	x	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
E	x	x	x	x	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
F	x	.	.	.	x	x	.	.	x	.	.	.	.	.	x	x	x	x	.	.	.	.	.
G	.	.	.	.	.	x	x	x	x	x	x	x	.	.	.	.	.	.	.	.	.	x	x
H	x	.	.	.	.	.	.	x	.	.	.	.	x	x	x	.	.	.	.	.	.	x	x
I	x	.	.	.	.	.	x	x	.	.	.	.	.	.	.	x	x	.	x	x	x	.	.

A id: Simulation Parameters
- *timestep* resolution
- *building name* as global identifier
- *surface convection* algorithm
- *heat balance* algorithm

B id: Location and Period
- *location* information
- linked weather file
- simulation *period*
- seasonal attributes

C id: Schedules associated
- linked *schedules*
- simulation *timestep*

D id: Surface and Constructions
- *materials* catalogue
- *constructions* catalogue
- specific openings properties
- specific values for walls, floors, roofs

E id: Thermal Zones and Geometry
- opening areas and walls
- façade dimensions
- internal heights
- zone lists

F id: Internal Gains
- occupants properties
- equipment properties

G id: HVAC
- main heating system type
- additional heating system
- heat zone assignment

H id: DHW System
- dwelling water demand
- heat gains from DHW

I id: Outputs
- energy demand by end-use
- fuel demand
- heat gains
- comfort estimates

As a result, the EnHub extends its capabilities to the study of specific changes on the properties of their components defined in the idf details. For example, an evaluation might be focussed on the thermal properties of a construction material; on the solar heat gains in a specific zone; on the usage intensity of an appliance; on the thermal heat gains of occupants; or on any value which can be defined in the idf.

In other words, besides the composition of archetypes for stock evaluation, this approach enables the application of further parametrisation in each of the modules of the model archetypes (in similar studies, this approach is also known as *what-if analysis*). This is one of the major advantages of employing the statistical tool `R`, which facilities the application of more sophisticated data analyses. In present day practice, two methods are mainly considered for parametric studies in the domestic energy performance domain: *optimisation* and *Sensitivity Analysis (SA)* [134, 204, 88]. These methods, essentially, apply a process of controlled variability in a set of chosen variables of the archetypes. Optimisation is commonly used for early stage buildings design (e.g new constructions) that require specific cost-effective measures. By contrast, SA is commonly used to explore the relative importance of a component, either in real (empirical) performance, or in its abstraction (modelling), and to track uncertainties arising in its representation.

In this case, the EnHub platform enables a SA module to validate the algorithms and to support the development of scenarios in the study of the housing stock. It is worth noting that, although this is not explicitly a platform to study retrofit, the methods to explore uncertainties may be similar. The application of either approach is determined by the purpose of the model. The EnHub employs the SA modules to improve the study of energy demand and carbon emissions in the stock of dwellings.

4.2 Additional Functionalities of the Platform

4.2.1 Verification, Validation and Calibration of the EnHub Platform

As stated in Chapter 3, *simulation* is defined as the application of a model to represent a phenomenon via numerical results. It also was highlighted that a number of assumption are taken to abstract such a phenomenon. EnHub employs external algorithms such as BREDEM and `EnergyPlus` modules. As a result, the attributes contained in these tools are propagated, affecting (positive or negatively) the functionality of the whole platform.

To establish a level of confidence in the functionality of these external tools, and hence the EnHub one, three activities are considered: verification, validation and calibration [143, 139]. *Verification* identifies the ability to develop an abstraction (i.e. developing a numerical model, building a software algorithm, processing a survey), and to accurately reflect the specifications of the chosen algorithm. *Validation* determines how adequate is the abstraction of the real phenomena. For this, a comparison between of the simulated results and empirical values is typically performed. Not only to determine the gain confidence in the estimation of numerical results, but also to investigate the appropriateness of the algorithm to arrive to similar outcomes. *Calibration* is similarly performed to determine the likelihood to represent a real-world phenomenon with numerical results. However, because a number of assumptions are taken during the abstraction process, a bias is expected in the model. Therefore, a calibration helps to reduce the variation caused by this process.

It is also worth noting that these activities are enhanced in the algorithms are transparency and editable (see § 3.4). This helps to directly identify the assumptions and sources applied in the model development, and also to perform iterative methods to investigate the response of the algorithms, namely Sensitivity Analysis (SA) and Uncertainty Analysis (UA) [23].

EnHub performs a verification of the platform by employing the `EnergyPlus` modules and by comparing against the CHM. Both the validation and the calibration is limited to the resolution of the datasets that provide values to represent a volumetric archetype, albeit in a reduced configuration. A future stage of the EnHub platform would consider the calibration with more appropriate survey data, so it can each component of the UML architecture.

4.2.2 Sensitivity Analysis and Uncertainties

Sensitivity Analysis (SA) techniques are primarily used to quantify input uncertainties [141, 44, 147, 45] and to measure the variation of model outputs relative to their inputs [190]. SAs are important for tracking errors and assumptions in model inputs and for determining input influence. When applied to housing stocks, they can be used to test the effectiveness of strategic changes to them, and to improve the development of stock models [74, 23, 130].

There are several methods applied in building simulation to measure sensitivity and error propagation [191, 218, 209], with the following four being the most used in building simulation:

+ ***Screening Methods*** These methods select a group of variables, based in levels, without reducing the output variance. By applying small variations, screening methods identify the

most important inputs contributing to the output uncertainty. The OAT approach has been extended to screening methods [193, 92], which compare the uncertainties in the inputs to those in the outputs. The Morris method (also called the *elementary effects* method) is the technique most often used in this domain. Here, each input variable takes a discrete number of values that are chosen within limits defined by their statistical properties. The method provides a measure of the mean value, μ, which estimates the overall effect of the input on the output, and a measure of the standard deviation, σ, that estimates second- and higher-order effects of the input variable [162, 193].

+ ***Uncertainty Analysis*** A comparison between design parameters and the model output is evaluated in this method [147]. Essentially, the analysis provides a reference of confidence intervals to study the output variables of the model. In sum, the uncertainty analysis aims to provide three aspects: magnitude of the uncertainty, ranking of influential parameters, and inherent properties of these parameters.

+ ***Calibration*** This method performs a simple, but methodological comparison between empirical and modelled data [209]. By applying this, the magnitude of the error can be used to provide certainty in the results, and to steer the set of improvements to develop a more accurate model. The calibration open the possibility to test the accuracy of the adopted algorithms, and to improve their performance by determining the error magnitude. At some extent, this provides the basis for further validation of both model and survey data.

+ ***Local and Global Sensitivity*** These methods measure the effect of parameters on the model output [105, 218]. The local sensitivity approach (also called *differential sensitivity analysis*), evaluates partial derivation of the output in relation to inputs. Global sensitivity analyses, by contrast, evaluate all the variables simultaneously, and so are able to determine the range or extent to which the model can be affected by variation on the input variables. The most widely used technique for the analysis is the regression method. Both approaches assume a linear correlation in the model variables, which may affect the quality and usability of the outcomes. As a result, the outcome from these methods may provide a general (and fast) overview of the model response in terms of the controlled variation of inputs; however, they can also misrepresent some of the influential parameters, especially those correlated with a significant degree of dependency.

Each approach conveys specific applications and can be complemented among each other. Tian [218] claims that regression methods are the most widely used method for SA in building energy analysis, particularly when the inputs are independent. However, this requirement can be problematic when two or more model inputs are correlated. Therefore, it is common to combine both global and local SAs to determine the level of significance of the inputs.

A common local SAs method is the OAT approach, where each input parameter is varied at a defined interval while the others are held constant [84, 74]. Firth et al. [74] summarise the four steps required for the local SA:

1. each input j defines its variation resolution k_j as convenient;
2. according to the variation resolution, a change is applied for an input Δk_j;
3. the model is run with the changed input so new outputs i are derived; and

4. a (typically normalised) cross-comparison is performed with each set of inputs and outputs generated to generate sensitivity coefficients.

The evaluation of these steps may be calculated using a finite-difference approximation based on central differences [74, 109]. The sensitivity coefficients are generated according to the number of inputs m and the number of outputs n.

$$\frac{\partial y_i}{\partial k_j} \approx \frac{y_i(k_j + \Delta k_j) - y_i(k_j - \Delta k_j)}{2\Delta kj}, \qquad i = 1,\ldots,n \;\; and \;\; j = 1,\ldots,m \tag{4.6}$$

Likewise, the application of a SA may be applied for a time-related evaluation. Here, the effect of the output is represented in Equation 4.7, where a nominal values is represented by X_i, in a given interval ΔX_i and the output y variables is accounted for a given period t.

$$S_i = X_i \cdot \frac{\Delta y_k(t)}{\Delta X_i} \tag{4.7}$$

The application of these steps provides a promptly overview of the influential parameters in the model. It helps to identify how the input variables affect in the uncertainty of the model output. Figure 4.10 illustrates a classic approach to study both UA and SA. UA is useful to describe the range of possible outcomes given a set of inputs. If the nature of the inputs convey a certain uncertainty, this is thus propagated in the outcomes. This approach, nevertheless, cannot provide a rank or a description of the most influential parameters. Such a specificity is given by applying SA techniques. It summary, at some extent, it can be said that UAs provide a measure of error whereas SAs provide a measure of the input effects. Nonetheless, the application of these methods is restricted to the quality of the data, and the processes involved in their collection. Besides, in agreement with Saltelli et al. [192], the usability of SA consists in scrutinising the purpose of the analysis, and not necessarily in measuring the capability of the method employed. Therefore, a SA should preferably be accompanied by expertise and complementary methods of validation.

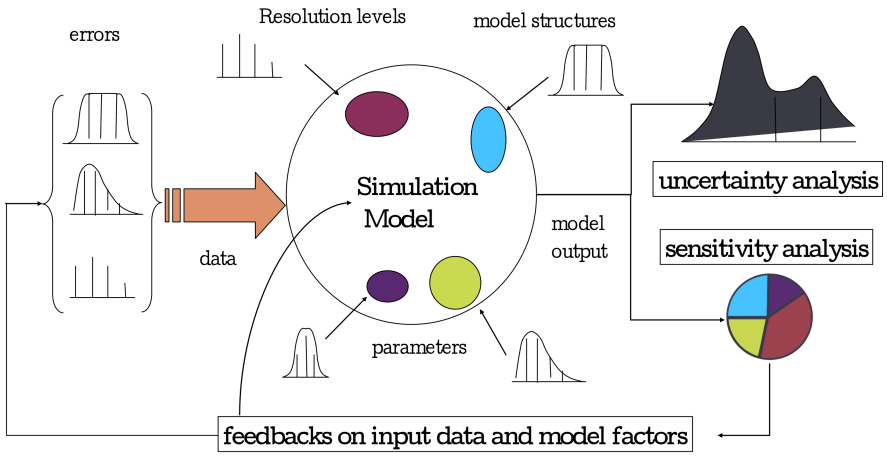
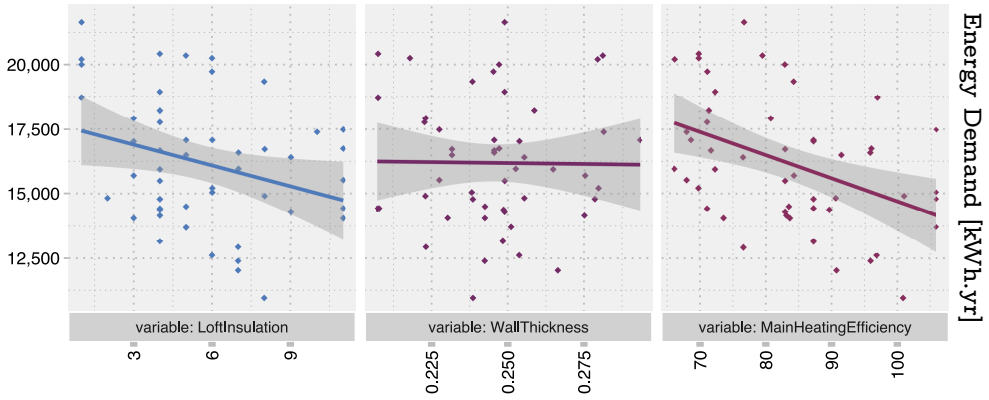
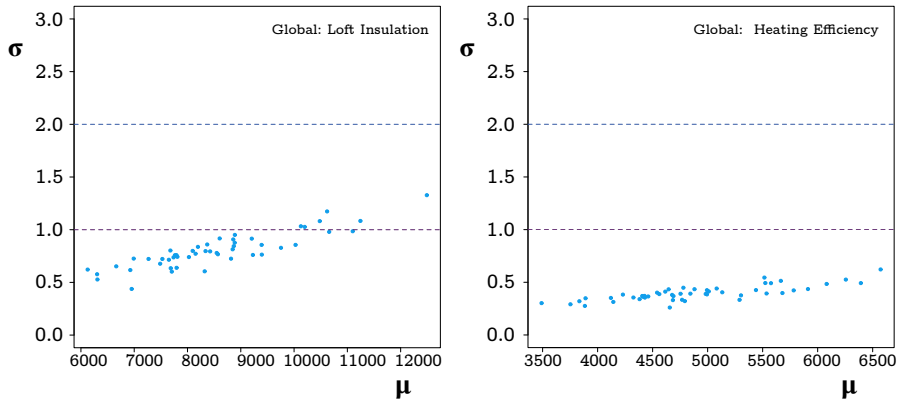


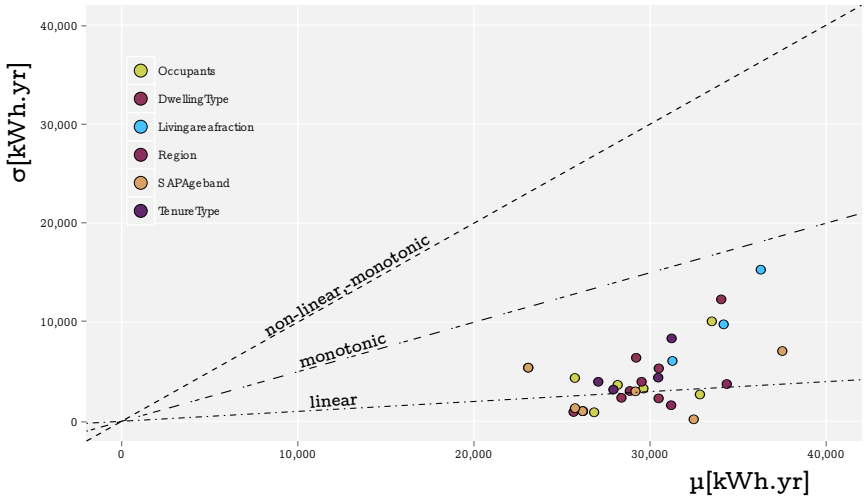
Figure 4.10: An engineer's vision of UA, SA
Source: Public domain image (au: A. Saltelli)



(a) Global SA - insulation, thickness, efficiency. The Confidence Interval (CI) is determined by applying *locally weighted scatterplot smoothing.*

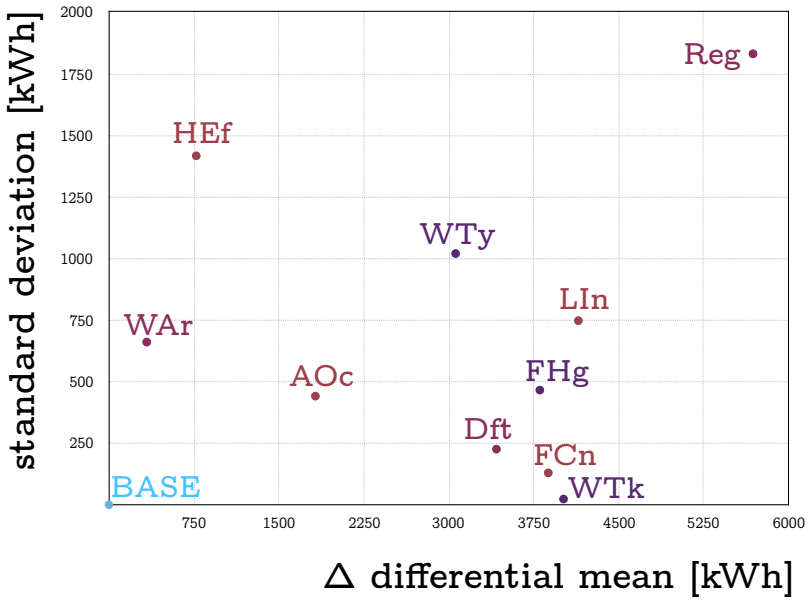


(b) Local SA - insulation, efficiency. The variation below the lower horizontal mark is not significant for the output, whereas the values above these mark may indicate a significant influence.

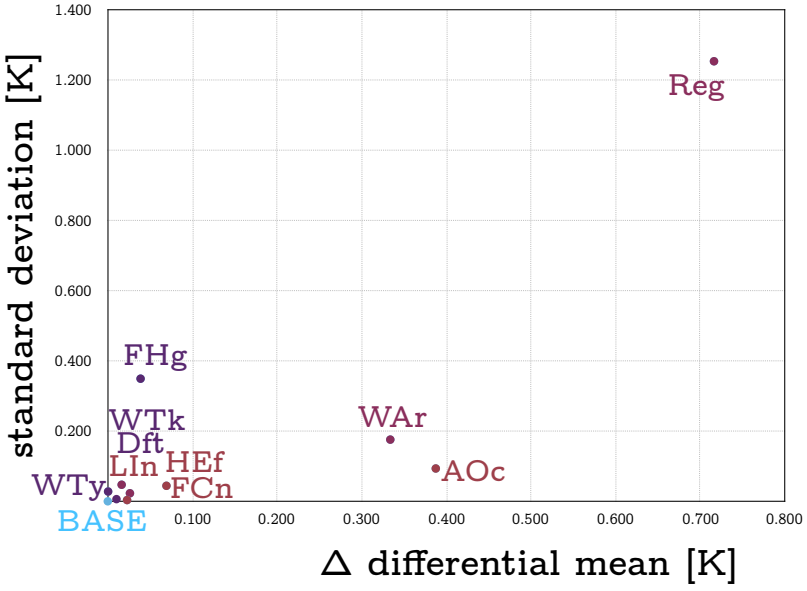


(c) Global SA - categorical and numerical variables. Values below the first diagonal are not significant, below the second diagonal may be significant and so require further exploration, below the third diagonal are highly significant to the output.

Figure 4.11: Examples of the SA module: local and global



(a) Screening Method - energy demand reference



(b) Screening Method - indoor temperature reference

Figure 4.12: Examples of the SA module: screening method
Notation: AOc: AdultOccupants; Reg: *Region*; LIn: *Loft Insulation*; FCn: *GF Construction*; FHg: *GF Storey Height*; WAr: *Windows Area 1F*; HEf: *Main Heating Efficiency*; WTy: *Glazing Technology*; Dft: *Draught Infiltration*; WTk: *Wall Thickness.*

EnHub Sensitivity Analysis module

To the enriched cuboids (i.e. those with detailed information on geometry, systems and householders), global and OAT analyses are performed to identify the most influential parameters in the dataset, many of which have been identified in previous studies [147, 74, 109, 130]. In EnHub, they are directly included in the simulation platform, by enabling a SA module which can employ survey data and idf parameters to perform the analysis. For example, Figure 4.11 illustrates the potential to conduct SAs by controlling changes to the properties of elements contained in a catalogue of properties used to generate idfs, and by modifying the reduced data set. Similarly, EnHub is able to evaluate changes to properties that are related to the external environment, such as weather files and solar incidence tables. Figure 4.11a shows that changes in loft (sometimes known as an attic) insulation and heating efficiency significantly affect the energy performance more than only increasing wall thickness (indicated by their gradients). Figures 4.12a and 4.12b compare two output variables by applying a screening method, which makes it possible to combine variables of different classes. Each input variable defines a dimension for the analysis in which a trajectory of perturbation is defined. The trajectory is chosen by advancing within the same dimension or to one other. This allows comparisons between the discrete values of each input variable with those of the output. There is widespread agreement that this method is versatile and useful [193, 92], but the computing time increases significantly and so its application to large datasets is limited. EnHub accomplishes this analysis by selecting variables based on expertise, and on their use in national programmes and policies.

The results of the SA module are consistent with those provided in previous studies [74, 109], where a rank of parameters establishes that the differential of temperatures (internal and external) is the most sensitive component. This is followed by the efficiency of heating systems (space and DHW), and infiltration rates. U-values are also prominent factors, especially for walls; as well as the number of occupants, although these value is limited to a limited range, provided by its statistical significance, and disregarding the intensity of associated activities.

4.2.3 High Performance Computing and Cloud Computing

It must be considered that a common restriction for parametric analyses is the computation cost. By way of illustration, OAT analysis might be applied to test 40 variables chosen according to a set of influential parameters found in other HSEMs (102 input parameters are considered in the previously referred study by Hughes et al. [109]; and 27 in Firth et al. [74]); from which then each variable can be assigned with about 100 unique values (i.e. 40×100). The extent to which the number of values are assigned, depends on the properties (or variability) of the input variable. Moreover, this theoretical analysis can be applied for ten different archetypes or scenarios, which results in a large number of cases, and so the computational cost increases (i.e. $40 \times 100 \times 10$). This cost—also identified in § 3.4 as one of the gaps in previous HSEMs—can be minimised by detaching these processes from the platform and executing them in a different environment, according to the hardware capabilities

It is worth noting that both `R` and `EnergyPlus` can be paired with a GUI or run in Command Language Interpreter (CLI) mode. By using the latter, both can be controlled from within a shell[3]. The purpose of using a shell is that 1) the platform can be detached from the

[3]A shell is a program which provides an interface between the user and the operating system. Such an interface

Operating System (OS) and 2) a set of low-level scripts can systematically control the process. In addition, the whole process can be performed in a stand-alone mode, so that each step may be conducted independently. This opens the possibility of parallelisation, and so the University of Nottingham High Performance Computing Facility (UoN-HPC) can be called upon to accelerate the computation tasks (see Box 4.2.1).

Box 4.2.1: HPC Facility

A High Performance Computing (HPC) facility can be defined as a group of individual computers that share resources (i.e. memory and processors) in order to perform significantly large calculations, in a way that typical desktop computer or workstation would be limited or overloaded. Such a group is referred to as a cluster of nodes, where each node contains an array of cores (or processing units).
The general arrangement of the HPCs is such that the individual nodes can communicate each other, which alows the development of robust serial or and parallel applications. [136] define the high-performance computers as machines with a good balance among the following elements:

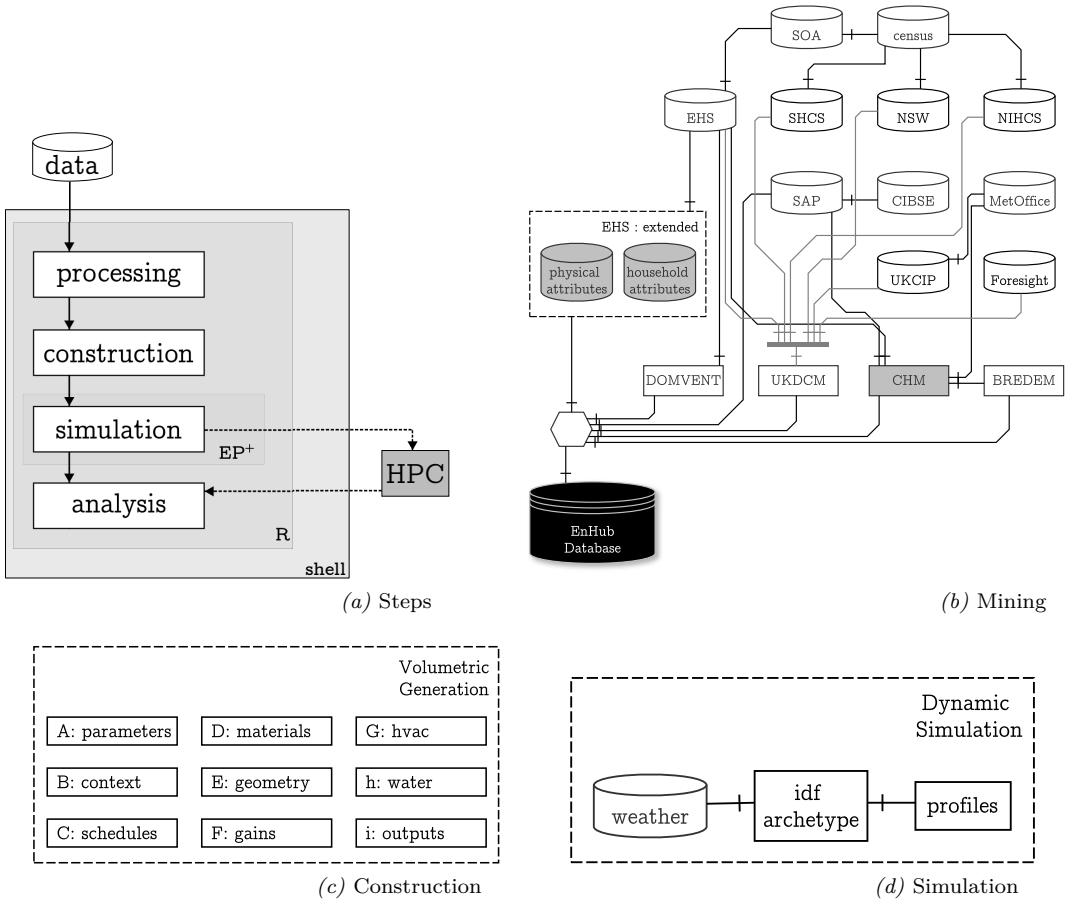
- Multi-staged (pipelined) functional units
- Multiple central processing units
- Multiple cores
- Fast central registers
- Very large, fast memories
- Very fast communication among functional units
- Vector, video, or array processors
- Software that integrates the above effectively

Therefore, an HPC facility is used to scale multiple processes running in series or parallel, for solving complex computational problems in external environments. It is worth noting that, although not equivalent, the HPC is close related to cloud computing systems, which eases the collaborative work.

By doing so, even if the process is not executed in parallel, detaching this from the core of the platform, improves the computation time by using an optimised hardware. Besides, in order to successfully run the process in the HPC, the scripts are required to be bug-free, which, in a way, emulates a process of validation in the whole simulation workflow.

In sum, the ability to detach the process and use an optimised hardware, not only improves the performance of the set of SA, but also the whole simulation performance. For instance, the generation of archetypes in EnHub is made in less than two seconds; this includes the volumetric construction; the assignment of profiles for each time-related component; and the assignment of a weather file. Each simulation requires approximately twenty-six seconds to complete. As a result, if each EHS instance is simulated, the whole simulation would be performed in 116 hours (i.e. $(2+26)[s] \cdot 14951 \cdot (1[h]/3600[s]) = 116[h]$). However, by optimising the hardware performance, the simulation time can significantly be reduced. For example, a modern computer (i.e. post 2013) may contain at least three usable cores which can be employed for parallel computation, and thus the time can be reduced to 38 hours. By using the UoN-HPC, the processing time is reduced by

may be called via a CLI or via scripting files.



(a) Steps *(b)* Mining

(c) Construction *(d)* Simulation

Figure 4.13: Simulation Flowchart: coupling `EnergyPlus` and `R`

about two orders of magnitude. The UoN-HPC makes available more than 100 nodes[4], yet these are available on-demand and their availability may be restricted. Still, the housing stock energy simulation time may be reduced to a processing time below two hours.

4.3 Housing Stock Energy Simulation

The main steps to perform energy simulation are summarised in Figure 4.13. Here, the mining of data is also used to select whether the analysis is focused on the housing stock, or on the sensitivity of variables employing a single archetype. These are the steps used to perform the housing stock energy simulation using each EHS case. Its application provides three main outcomes:

i/ calibration of the modelling parameters against national figures;

ii/ identification of redundant parameters arising from the resolution of the modelling resolution;

iii/ provision of computational flexibility by improving software and hardware communication.

[4]A node is usually formed by a group of processors and a block of memory [78].

By 2011, the *UK's Housing Energy Fact File* [179] reported a final energy use of 1597 TWh. The domestic sector contributed with 452 TWh. Further, because of the pro-rata calculations made in such study—see § 3.2.2, the estimated energy demand for the UK may be split as follows: 383.5 TWh Eng, 40.6 Twh Sct, 22.8 Twh Wal and 12.0 Twh NI. This value corresponds to the one measured and reported by DUKES. Consequently, the estimated carbon emissions associated to this domestic energy demand were 124 $MtCO_{2e}$. In here, the building physics calculations were performed employing the CHM; however, CHM is thought to over-estimate gas demand in about 20% and under-estimates electricity demand in about 5% [181]. As reported in Section 3.3, the abstraction in the CHM is unable to represent drivers of decision-making that affect the energy performance in the archetypes, and the level of accuracy is thus undermined.

The results of the EnHub simulation for each instance in the EHS, and therefore the English housing stock, sum an accumulated energy demand of 424 TWh. Table 4.6 summarises the energy demand in terms of typologies and provides some metrics related to the volumetric representation of dwellings. The obtained total energy demand is 10.7% greater than the adjusted DUKES value. Section 4.3.1 compares these values with those obtained from previous studies. Here must be considered that some of these figures employ survey data from different years, and so the comparison needs to be applied with caution. Nevertheless, this value is consistent with the sub-totals for each instance. Figure 4.14 illustrates the contribution to the energy demand by typologies, where it can be seen that older constructions demand more energy than the rest. In addition, the Figure shows that detached houses are the major contributor to the energy demand intensity among the housing stock.

Table 4.6: Energy performance summary by age and type of construction, using dynamic simulation

Typology	Dwellings (x1000)	%	Indoor Temperature [C]	% indoor comfort	Mean Energy Demand [kWh]	Carbon Emissions [tonnes CO2]	Energy per Area [kWh/sqm]	Total Energy Demand [TWh]
pre-1919	4,738,847	20.8%	18.3	32.3%	23,496.8	5,067.7	255.4	111.3
interwar	3,761,547	16.5%	17.6	31.1%	21,227.8	4,632.0	230.7	79.8
postwar	4,501,900	19.8%	17.1	31.6%	17,783.1	4,373.2	193.3	80.1
industrial	3,298,143	14.5%	16.9	32.0%	16,666.1	4,372.0	181.2	55.0
modern	5,223,255	23.0%	17.0	30.5%	15,599.5	4,261.7	169.6	81.5
post-2002	1,230,448	5.4%	16.9	31.4%	13,786.0	4,293.4	149.8	17.0
apartments	4,627,572	20.3%	16.3	31.1%	11,319	4,059	120.9	52.4
bungalow	1,995,531	8.8%	15.0	31.1%	15,848	4,126	172.4	31.6
detached	3,785,814	16.6%	18.2	35.6%	26,543	5,492	293.9	100.5
semi-detached	5,917,249	26.0%	18.3	31.3%	19,596	4,523	211.7	116.0
terraced	6,427,974	28.2%	18.7	29.0%	17,157	4,299	186.5	110.3
	22,754,140				18,093			424.7

For the housing stock evaluation, the EnHub platform evaluates every instance in the EHS, derives indicators related to the volumetric composition of archetypes, and provides summaries by typologies. As previously identified, the computational cost increases not only by increasing the resolution of volumetric archetypes to describe the stock, but also by augmenting the resolution of their emulated parameters affecting the energy flow in them. An example of the relatively granular profiles assigned to each of the archetypes is presented in Figure 4.16. In here, a noticeable fluctuation in the profiles is revealed. This is important for linking specific responses in the performance of energy flow in dwellings, especially when the functions to define such profiles are

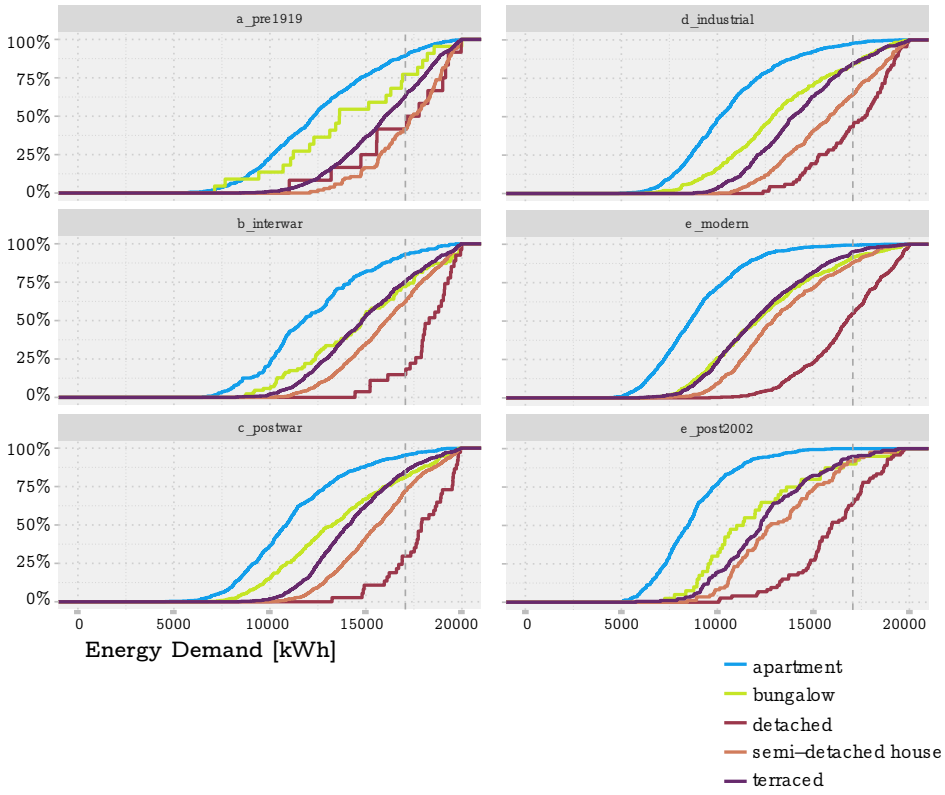


Figure 4.14: Distribution of energy demand by typologies

linked with more sophisticated modules. A summarised description of the modelling process for one archetype is described in Appendix C.

Therefore, the outcome of the housing stock evaluation suggests that some of the simulated archetypes are highly correlated, and the difference in their final estimated energy demand is attributable to the scalar impact of their properties (physical and estimated appliances). Therefore, there is significant potential to identify redundant parameters describing the stock, so the number of archetypes can be reduced. This is made not only to improve the simulation time, but also to increase the potential of further parametrisation and specific analysis of dwellings within the same modelling approach.

4.3.1 Similar Energy Performance Studies

As described in Section 3.2.2, the BREDEM applies steady-state methods to estimate the heat balance in buildings. Additionally, although the consideration of household appliances and occupants follows a similar logic in such a balance, their characterisation is based on empirical information, which is then adjusted via a power law relationship (see Eq. 3.7). Therefore, a between metered data and calculated data is expected. This variation is used to calibrate (see § 4.2.1), and also to identify errors in the model.

An initial comparison between CHM and DUKES, using 2009 EHS and DUKES data, reveals an over-estimation (or error) of 10.5% [109]. Most of the variation is explained by the algorithms

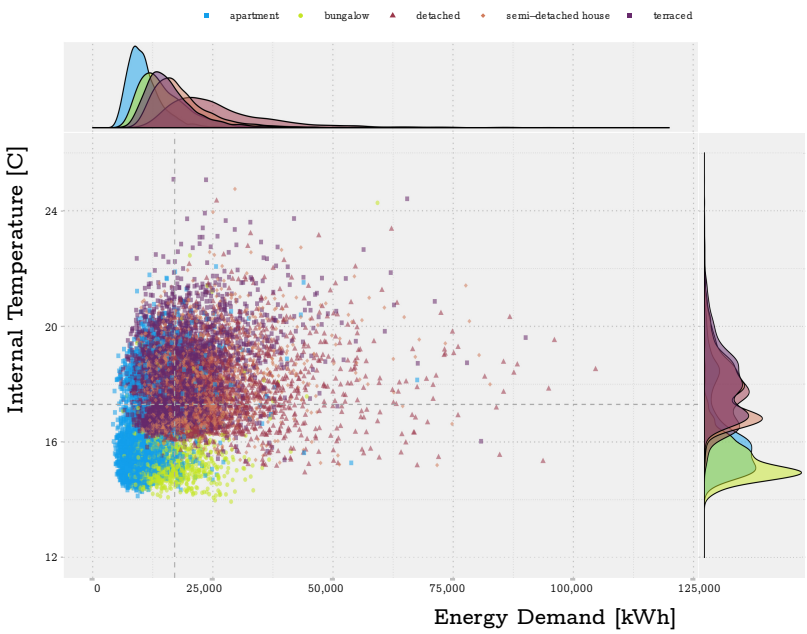


Figure 4.15: Relation between energy demand and indoor temperature by typologies

to represent indoor temperatures and heating devices. Other contributors to this variation are clearly identified in the employed survey data: the EHS. Further, a comparison between the CHM and the NEED readings, with 2010 data, reveals differences (by typologies) between 1 and 23% [182]. Nevertheless, in addition to the error attributed to the simplifications, the study describes other four causes of variation between results:

1. NEED meter readings are weather-corrected by the Energy Companies; which undermines the ability to quantify the magnitude of errors by using raw data.
2. The resolution adopted to characterise weather data is dissimilar between CHM and NEED; which introduces a level of uncertainties caused by the methods to inter- and extra-polate values.
3. The granularity of data regarding occupant behaviour is scarce in survey data and absent and unverifiable in datasets accompanying meter readings. This introduces a range of uncertainty in the average values representing occupancy in dwellings.
4. NEED employs additional datasets describing fuel costs, to better predict use of energy in terms of disposable income, which is limited to income bands in the CHM. Although this can be compensated in the aggregated values, the study of subsets may be inaccurate in the modelling approach.

To complement the outcomes provided in the *UK's Housing Energy Fact File* [179], where the contribution to housing stock energy demand is broken down by typologies, other studies that have employed similar sources of information to elucidate the expected energy performance (i.e. EHS, Census, DUKES, NEED), although limited to subset of the stock or with the addition of modelling algorithms.

Taylor et al. [216] evaluate a sample of houses located in the North East of England using

`EnergyPlus`. The study evaluates three dwelling types of dwellings (terraced, semi-detached and detached) with a number of variations in their properties (size, geometric detail, glazing ratio, schedules) to investigate the level of significance associated to the various configurations. The results indicate that, by using simplified versions of the archetypes, the annual heating energy demand is underestimated in up to 27%. The selection (and deployment) of dwelling archetypes, based on the EHS, would represent only one of these configurations. In consequence, it is important to consider that an error may be caused and propagated, due to the resolution of the dataset (see § 3.4).

Further improvements were undertaken for the study of a similar stock [91]. Here, the evaluation compares the performance with the CHM. The variation between the models is estimated in about 34%. Still, the consideration of behaviours is undermined by the resolution of the datasets, and the strategies to emulate activities in the represented performance.

The construction of archetypes has thus been undermined by the simplification of behaviours. He et al. [90] develop an application for the evaluation of a stock of dwellings with enhanced occupancy profiles. The stock is evaluated dynamically using `EnergyPlus`. Here, the Richardson's model [187] is embedded to emulate stochastic occupancy profiles. The study is not conclusive about the extent to which different profiles affect the annual energy demand; however, a reduction of 20% is estimated for the application of these profiles. As a result, the study offers an example of how a module, which describes occupancy, may be replaced with sophisticated algorithm for better representing housing stock energy demand. Likewise, further improvements may be applied to handle diversity in household characteristics and in housing attributes (see § 3.3.3).

4.3.2 Energy Performance Indicators

HSEMs provide with specific outcomes depending on the purpose of the studies. In similar fashion, EnHub defines a number of outcomes related to the performance of dwellings, but linking indoor conditions provided by the volumetric representation of the stock. A summary of these outcomes are presented below:

- **Energy demand**: this value provides the sum of system-specific totals for each archetype. They can also be obtained by using the fuel-specific totals. The energy demand is obtained at a time-step resolution for one calendar year. `EnergyPlus` interpolates the differential values existing within the timesteps. Table 4.6 provides the average energy demand in dwellings of 18,093 kWh. Figure 4.14 illustrates differences in the energy demand by typologies; detached houses demand 46% more than the average, whilst apartments 37% less.
- **Power demand peaks**: this indicator provides the maximum value demanded for each system. This value is valuable for the study of energy generation and distribution, especially when analysing specific subsets of the population.
- **Heat demand**: This indicator provides a measure of the energy flow in dwellings. The value can be used to describe some long-term profiles and areas, where under- and over-heat may require special intervention. As illustrated in Figure 4.16, the heating demand mainly occurs during winter. Even though this represents a less than 6 months per year, its contribution of the annual energy demand is roughly four-fifths. The skewed average illustrated in Figure 4.15 shows low temperatures; however, it needs to be considered that

these values include unoccupied areas in dwellings, besides the absence of internal layouts which may affect these values.

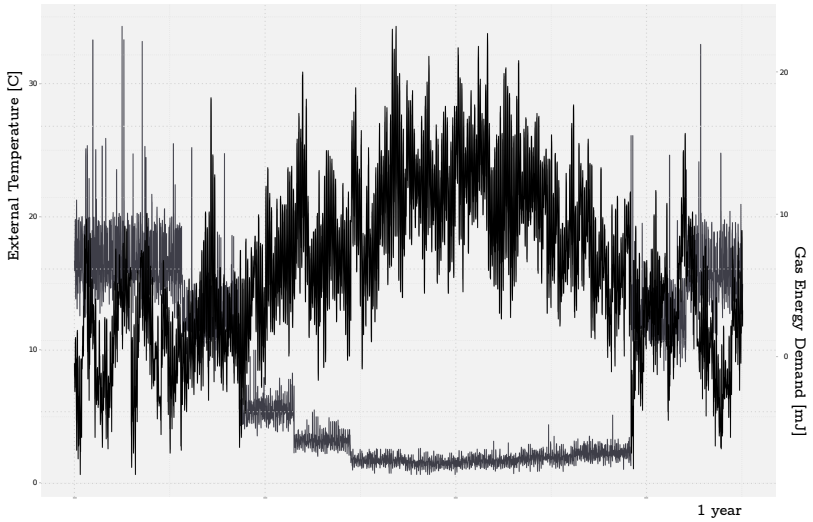


Figure 4.16: Relation between gas energy demand and external temperature for one archetype

- **Comfort values**: despite its limited algorithms employed, as described in § 4.1.4, this indicator provides a valuable overview of the estimated performance based on the assumptions regarding acceptable indoor conditions. Most dwellings in the UK housing stock suffer from indoor conditions below the levels stipulated (and suggested) in Building Regulations; however, there are many factors that are still absent in the analysis of comfort conditions, for example: a) the indoor temperature is not dependent on the context of the building (i.e. urban, suburbs, countryside or other settlement); b) the thermal inertia which may affect the comfort sensation on occupants is neglected; and c) the internal layout and hence the number of unused areas within dwellings is not considered. The consideration of these factors substantially affects the perception of comfort and its evaluation along the housing stock. Table 4.6 provides an (average) indicator of comfort conditions, applying a scale of indoor performance to determine whether the studied room is in comfort, cooler or warmer. The results indicate that, annually, the average comfort conditions are reached in only about a 30% for dwellings.
- **Carbon intensity**: although this might be an extension for the energy demand value, this indicator enables the possibility to explore the influence of adopted conversion factors (which do not necessarily correspond to the catalogue of GHGs provided by the UK government, used in HSEMs). Table 4.6 provides an average value of 4059 tCO2.

4.4 Conclusion of *Composition and Development of the Simulation Platform*

The EnHub platform provides with both greater fidelity and utility compared with existing approaches that are based on simplified energy balance calculations. Furthermore, this platform can be readily augmented to predict the uppermost potential of decarbonisation interventions,

as well as indoor improvements related to comfort and health, in relation to the relevance of decision-making drivers influencing such improvements. For example, this platform is able to apply the BREDEM algorithms; and it is also able to replace models where the assumptions lack or are very general. Similarly, the ownership and usage of appliances can be represented by combining information related to households and dwelling types, which can provide more consistent and useful information onto the generated model archetypes, so that policies and measures can be tested accurately.

In addition, fluctuations on the estimated responses of their occupants can be further specified in function of socio-demographics. This provides the possibility to measure the level of response of household specific changes and household-related constraints for their implementation. These derived variables are also in function of a set of indicators which provide more useful information such as discomfort levels per capita, heat gains per floor area, estimated household appliances power, maximum temperatures, and carbon intensity values. In this regard, it is worth highlighting that the EnHub modularity allows the comparison of carbon intensities associated to each of the fuels considered. This enhances the ability to study the implications of changes in the fuel mix at both end-use and power generation levels. By following a similar approach, the EnHub has the potential to link forecast models. These models may represent changes in the composition of households, changes in the cost of fuels, cost of new technologies, and demolition rates. As a result, the platform is able to represent (and calibrate) the evolution of the stock and, more importantly, some of the direct rebound effects that appear in practice.

To support this flexibility, EnHub is conceived as a modular open-source platform. An open-source simulation platform not only improves the sophistication of modelling approaches by making each component editable and verifiable, but also boosts inter- and trans-disciplinary collaboration, by providing a standardised hub to combine data from different perspectives. The housing stock decarbonisation in the UK is a process that requires constant update and validation. EnHub provides the ability to support such a long-term strategy for its successful accomplishment.

To sum, this platform integrates the virtues of existing housing stock energy models, in particular their representation of the composition and attribution of the stock—albeit in reduced form—with mechanisms to quantify uncertainties and to explore the potential impacts of policies and strategies to decarbonise the UK housing stock. Additionally, because of the modularity of their components, a set of replacements are applied to each of the algorithms. And finally, supported by the application of the statistical tool (`R`), the platform is able to perform parametric applications, for which two of them are considered: sensitivity analysis and evaluation of scenarios. The former is considered as a support to validate the model in multiple stages of the modelling platform; the latter will be described in Chapter 5, yet with additional elements to improve the functionality in the platform workflow.

5 Workflow for Efficient Housing Stock Representation

Introduction

Chapter 4 described the process employed to define the basis of the simulation platform, and the composition of volumetric archetypes used to represent the housing stock. This Chapter described the composition of the EnHub platform, which improves virtues of the CHM. Such improvements, besides the more granular and transparent architecture of the former, can be summarised in three aspects:

I/ EnHub utilises the dynamic simulation program `EnergyPlus` for energy performance predictions, while CHM uses BREDEM, a simplified energy balance model.

II/ EnHub employs a process of statistical data reduction to reduce the number of typologies to simulate, while CHM evaluates every instance of the EHS datasets, with corresponding redundancy.

III/ EnHub's architecture is readily extensible to also model households' responses to socio-economic drivers influencing investments and changes to behavioural practice that impact on net building energy demand.

Therefore, on the one hand, by employing the statistical tool `R`, in combination with the dynamic simulation engine `EnergyPlus`—both controlled conveniently in CLI—the ability to increase the level of detail in the development of archetypes is accomplished; this includes the evaluation of indoor specific metrics and, more importantly, the emulation of behaviour profiles which can be linked to more specific household-related algorithms. However, on the other hand, the resulting resolution increases the computational cost significantly (simulation time and amount of memory required). For this, the EnHub platform increases its capabilities by allowing the employment of an HPC facility; spreading computational cost. Nevertheless, to efficiently evaluate the effectiveness of housing stock decarbonisation scenarios, it is desirable to reduce the number of archetypes representing the stock. To this end, a process of data mining is required. This is made to reveal redundancy in the dataset, so it can be reduced; and to validate the processing of information. The following sections describe this process, and the methods applied to reduce the dataset.

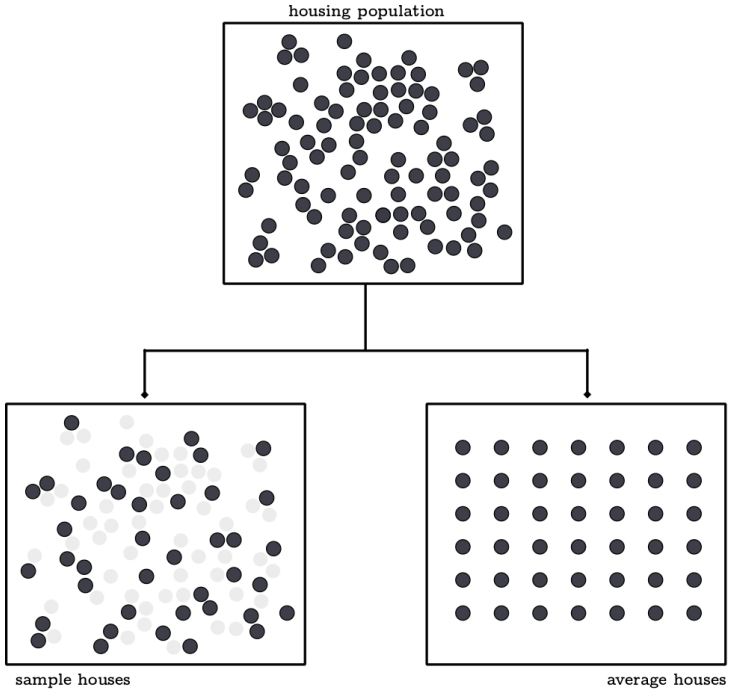


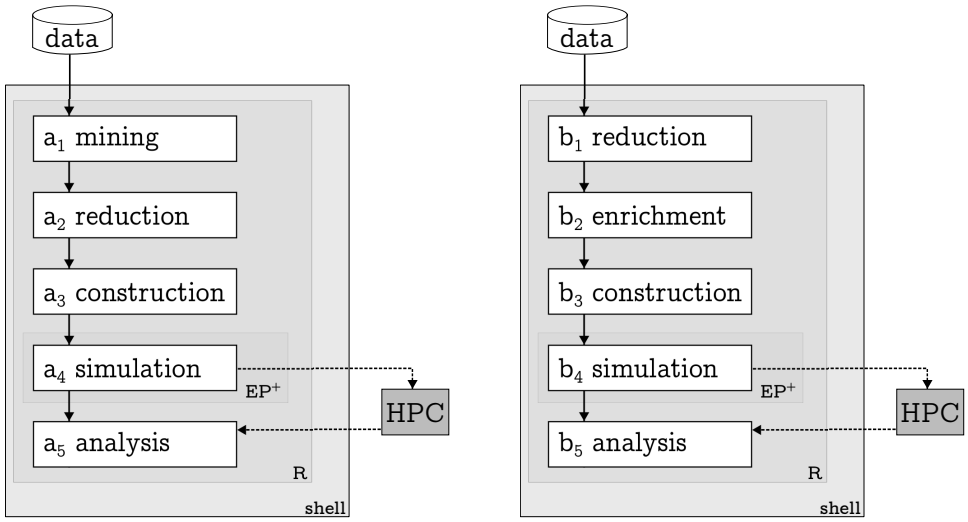
Figure 5.1: Conceptual strategies to reduce the housing stock population

5.1 Methods: Statistical Analysis and Engineering Models

To provide a level of confidence to the EHS dataset, a data mining process is performed. This is typically defined as the process of discovering and presenting associations, dependencies and grouping by applying analytical techniques [72]. Data mining involves the application of diverse methods to predict and/or classify the typically complex databases [89]. Predictive methods usually apply regressions although a highly developed sub-type includes hierarchical models (e.g. trees, additive models, neural networks) [215]. Correspondingly, classification methods usually apply cluster and ordination analyses, and are mainly used to study large databases. Both predictive and classification methods employ the emerging relations of proximity[1] of variables, and can be adapted and complemented between them.

In addition, for housing stock studies, the chosen units of observation may be directly derived from real cases (representative sample) or deployed as average archetypes (synthetic sample) [111, 164, 11, 140]. To capture a sample of the population, both processes imply randomisation, which can be as sophisticated as required; for instance, common methods employed in this domain are Markov-chain Monte Carlo (MCMC) sampling and Latin Hypercube Sampling (LHS) [4, 233]. Representative samples employ extrapolation and weight calibration to make robust cases, which is usually achieved by applying classification techniques. Average archetypes, by contrast, apply predictive methods to define associations, and to derive fit-for-purpose units of the stock. Both techniques, illustrated in Figure 5.1, provide specific attributes, even though they are weighted to represent the population. For example, average archetypes can be easily adapted (and scaled) to different functional subsets of the population; by contrast, sample archetypes may

[1]The term *proximity* is a general term used in statistics that covers the numerous methods for the estimation of distance, dissimilarity coefficients and metric inequalities.



(a) Sample Houses *(b)* Average Houses

Figure 5.2: The EnHub structure for the analysis of scenarios

be easily calibrated with empirical data by maintaining the properties of the original dataset. In practice, using both techniques is possible and its combination helps to adopt multiple sources of information.

To reduce the dataset, the process described in § 4.3 is augmented with an additional step. However, such an addition can be applied to define improved samples or average archetypes. For instance, Figure 5.2a initially ranks the dataset, and then applies an improved sampling method to reduce the number of archetypes. Conversely, Figure 5.2b initially applies reduction, based on predictive methods, and then calibrates the dataset by performing association and hierarchical algorithms. The following steps $a_1 mining$ and $a_2 reduction$ describe the process adopted by Figure 5.2a, steps $b_1 reduction$ and $b_2 enrichment$ describe the process followed by Figure 5.2b. Steps $a - b_3$, $a - b_4$ and $a - b_5$ have been described in Section 4.1.4, and only a summary is provided in the following sections.

5.1.1 Step a_1 Mining

The first step when considering a dataset with inputs and outputs is to test its linearity. This is useful for revealing similarity in the inputs, statistical independence of the output variables, and—for the case of multiple variables—clusters of data. To this end, two parameters are essential to define the level of linearity among variables: proximity and correlation coefficient. These parameters are defined by assuming a linear model describing the relation of data, expressed by Eq. 5.1. Here, the dependent variable y (also known as response or output), is a function of one or more independent variables X_i (also known as predictors or inputs). The independent variables are affected by their corresponding correlation coefficients β that indicate their level of significance. The initial adjustment (or scale) of the model is defined by β_0. An error, for each predictor, is denoted by ϵ. Incidentally, it is worth noting that predictor values usually represent

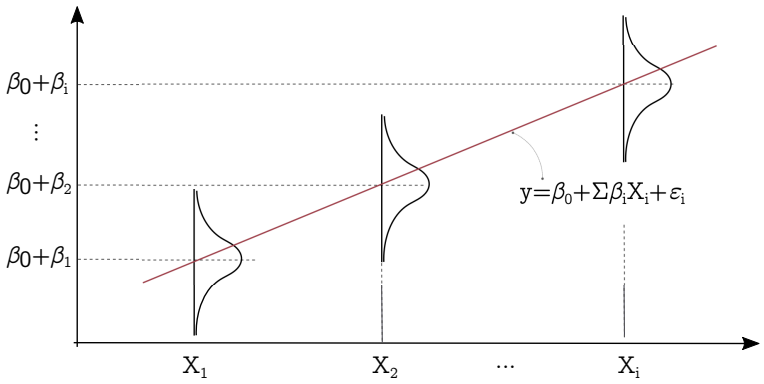


Figure 5.3: Linear model relationships

normal distributions. Figure 5.3 shows a general representation of Eq. 5.1.

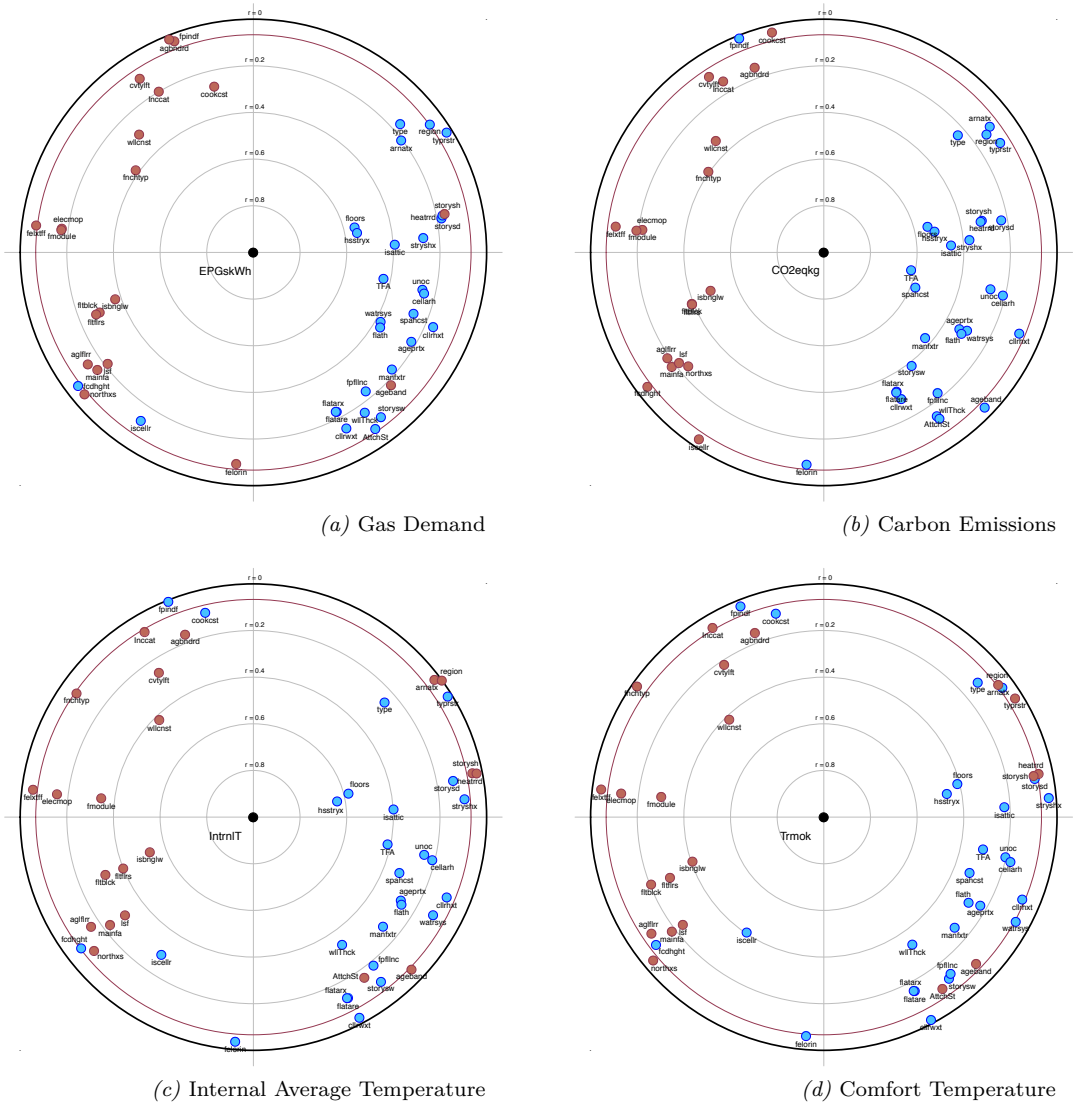
$$y = \beta_0 + \Sigma\beta_i X_i + \epsilon_i \tag{5.1}$$

Depending on the structure of the dataset (i.e. types of variables, number of cases, weights, and survey design), β can be appropriately enhanced and scaled. The proximity values are obtained from their simultaneous association (see Appendix B).

Previous studies of nationwide surveys have investigated the usability of the linear association [8, 87, 152], and there appears to be agreement in the value of employing robust methods to account for the influence of categories, and more importantly, to identify redundant variables in terms of their level of significance [215, 181]. Robust methods apply criteria to reduce the impact of multicollinearity, partial effects, and unclear distributions [106]. Nevertheless, it must be considered that the application of statistical regressions in large surveys, such as the EHS, is susceptible to misrepresent these partial effects and, hence, to introduce errors. Their application should be employed as preliminary guideline, and should be complemented with expertise, as well as with alternative methods such as hierarchical models and cluster analysis.

In this case, a linearity test is performed via Focused Principal Component Analysis (FPCA) and Generalised Linear Models (GLMs), with a view to identifying potential overlaps among the variables. FPCA is used because it provides a virtual rank of the variables, which can be used to reveal collinearities and clusters of variables [73]. This method extends the classic approach of Principal Component Analysis (PCA) that applies scores to the variance of the data and, by extracting the proximity of those most highly ranked, identifies redundant variables. This method is also useful in testing for correlations between variables, the creation of clusters, and the measurement of variability [118]. In brief, FPCA extends the PCA approach by providing an overview of the main predictors and similar proximities among the variables. Further, because non-linear relations and mixed effects are expected in the EHS, FPCA is useful in ranking variables including categorical variables, yet due to the dimension of the EHS, the outcomes are merely indicative and should be complemented with alternative methods.

Figure 5.4 illustrates a significant association of input variables to four corresponding foci: gas demand, carbon emission, internal average temperature, and comfort temperature range. Here, the shorter the distance between a variable and the focus, the stronger their correlation [73]. The



(a) Gas Demand

(b) Carbon Emissions

(c) Internal Average Temperature

(d) Comfort Temperature

Figure 5.4: FPCA visualisation
Because the estimated energy demand in the models is the sum of fuel energy demands, (a) considers the main fuel used in the stock. To identify other variables affecting the overall energy demand, (b) considers the associated carbon emissions, which is directly dependent of this. For (c) and (d), the internal temperature captures the response of the volumetric archetypes to the influence of variables.

colour of the variables indicates whether their correlation with the focus is positive or negative, where red is positive and blue is negative. Also, variables in the same quadrant are positively correlated, whereas those in opposite quadrants are negatively correlated. For example, Figure 5.4c describes the degree of correlation of variables with the internal average temperature *IntrnlT*. Here, both the total floor area (*tfa*) and wall type (*wllcnst*) are at a similar distance from the focus and are located in opposite quadrants. Therefore, they are negatively correlated to each other, and have a similar magnitude of correlation with their focus. However, *tfa* is blue and so is negatively correlated with the focus, whereas *wllcnst* is red and is positively correlated with the focus. These relationships can be explained by the characteristics of different housing typologies, where older dwellings, constructed pre-1920, incorporate less energy conserving materials in their envelope, and are generally bigger [166]. Finally, the surrounding coloured radial line indicates the limit of acceptable associations although the high number of variables considered in this case makes this wide.

To tackle the problem of multicollinearity, indicated by the association in Figure 5.4, GLMs are applied to eliminate redundant variables. GLMs extend ordinary linear models by including a distribution in the expected response and by assuming that more than one variable is dependent [153]. In this way, as with FPCA, GLMs rank the variance in the variables in addition of the level of significance[2]. The response variable may be associated to its probability distribution $E(y)$. This is represented in GLMs by including a link function f between the distribution of the response and the predictors $f(E(y))$. Therefore, a GLM is expressed by Eq. 5.2.

$$f(E(y)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \tag{5.2}$$

The GLM is applied to the full EHS dataset, even though the resolution of some variables is limited, such as those indicating the presence of a physical component (i.e. yes/no); and even though some variables represent similar values, such as specific dimensions (e.g. storey and cellar depths, façade height and heights of each floor). The dataset employs the dynamic simulation engine to evaluate their energy performance.

In consequence, the linearity test is complemented by performing a process of backwards elimination to remove redundant variables. This process consists of ranking the variance of the variables, and systematically removing those having a minor influence. The level of significance of the whole model becomes the criterion used to determine when the model sufficiently represents the stock. Because the GLM splits categorical values in its classes to independently evaluate each of them, some uncommon sub-categories may be eliminated in the process, but then be re-included in the outcome regarding a main variable. For instance, this is the case when considering the main heating system, where 95% are central gas and electric systems, and so the less common systems (e.g. oil, wood, and coal) are initially removed from the GLM outcome.

By studying the four results of the linearity test presented in Figure 5.4, and the results in Table 5.1, it is possible to obtain different ranks according to the chosen outputs. Here, the analysis reveals that fabric components, particularly the geometric parameters, are essential for the model, so that a relatively explicit representation of them is appropriate.

[2]The level of significance (or $p-value$) is a trigger at which the null hypothesis is rejected. Although not mandatory, the trigger at 0.05 is commonly employed.

Table 5.1: GLM rating on BREDEM

	Electricity Demand	Gas Demand	Internal Temperature
TFA			−0.007***
			(0.001)
northaxis	−2.143**		−0.001
	(1.015)		(0.0003)
arnatxother urban centre	508.858	−1,226.125	−0.152
	(441.301)	(976.276)	(0.138)
arnatxrural	665.680	−4,597.320***	−0.827***
	(603.577)	(1,376.714)	(0.188)
arnatxrural residential	1,384.238**	458.260	0.704***
	(624.434)	(1,401.394)	(0.197)
arnatxsuburban residential	759.530*	−672.271	−0.123
	(426.704)	(960.061)	(0.137)
arnatxvillage centre	1,813.287***	−1,721.748	−0.315*
	(574.279)	(1,289.697)	(0.183)
fmoduleHouse (single unit)		−247.084	−0.232
		(1,065.942)	(0.247)
fmodulePurpose built flats (multiple units)		1,625.671*	0.742**
		(930.768)	(0.334)
felorienNorth		3,274.620***	0.454***
		(799.526)	(0.113)
felorienNorth-east		3,141.772***	0.175 (0.133)
		(955.304)	
felorienNorth-west		3,283.262***	0.353**
		(990.398)	(0.138)
felorienSouth		1,627.948*	0.238*
		(877.513)	(0.122)
felorienSouth-east		2,547.355***	0.287**
		(985.400)	(0.140)
felorienSouth-west		2,130.764**	0.439***
		(912.852)	(0.130)
felorienUnknown		19,412.910***	−2.370***
		(7,019.450)	(0.538)
felorienWest		1,669.975*	0.326**
		(915.189)	(0.130)
regionEast Midlands	96.216	−59.336	−0.674***
	(533.030)	(1,174.576)	(0.162)
regionLondon	604.959	176.792	−1.014***
	(431.783)	(952.167)	(0.138)
regionNorth East	958.309	−829.361	−1.572***
	(587.260)	(1,277.758)	(0.180)
regionNorth West	968.633**	1,310.224	−1.791***
	(424.720)	(944.440)	(0.132)
regionSouth East	279.611	−1,187.339	−0.593***
	(419.133)	(927.938)	(0.132)
regionSouth West	1,141.305**	−1,653.824*	−0.150
	(446.948)	(983.404)	(0.140)
regionWest Midlands	189.217	780.760	−1.351***
	(513.133)	(1,145.227)	(0.161)
regionYorkshire and the Humber	293.968	1,950.620*	−1.534***
	(464.006)	(1,029.346)	(0.141)

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
Inccatlow	−2,272.356**		−0.948***
	(883.929)		(0.265)
Inccatmed	−1,296.777**		−0.568***
	(564.885)		(0.179)
ageband1900-1918		8,482.547***	0.409*
		(2,262.874)	(0.231)
ageband1919-1944		−11,876.630	0.239 (0.236)
		(8,643.790)	
ageband1945-1964		−430.885	0.054 (0.236)
		(845.892)	
ageband1965-1974		−21,779.950***	−0.668**
		(4,980.188)	(0.275)
ageband1975-1980		−2,959.396**	−1.012***
		(1,240.613)	(0.309)
ageband1981-1990		−2,817.250***	−1.237***
		(1,035.777)	(0.294)
ageband1991-1995		−1,940.470	−1.342***
		(1,541.450)	(0.333)
ageband1996-2002		−4,849.826***	−1.686***
		(1,605.351)	(0.346)
agebandPost 2002		−3,972.823***	−1.200***
		(995.753)	(0.292)
agebandPre 1850		−1,230.242*	0.062 (0.104)
		(743.533)	
typeFLAT - Non-residential plus flat			−0.235
			(0.443)
typeFLAT - Purpose built			−0.602*
			(0.320)
typeHOUSE - Detached			−0.704***
			(0.250)
typeHOUSE - End terrace			−0.502*
			(0.256)
typeHOUSE - Mid terrace			−0.670***
			(0.258)
typeHOUSE - Semi-detached			−0.505**
			(0.248)
typeHOUSE - Temporary			−1.482**
			(0.717)
isattic		2,646.479***	0.330***
		(525.979)	(0.080)
iscellar			1.418***
			(0.115)
floors		6,312.159***	−0.436***
		(971.071)	(0.163)
AttachStateLeft		2,572.780***	−0.262***
		(579.114)	(0.095)
AttachStateNone		−2,276.800	−1.400***
		(1,963.311)	(0.292)
AttachStateRight		2,672.694***	−0.515***
		(640.036)	(0.099)
heaterredGas			0.253*
			(0.152)

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
heaterredOil			0.318*
			(0.175)
heaterredOther			0.198 (0.228)
heaterredSolid			0.733***
			(0.162)
agebandredPostInd			0.482*
			(0.276)
agebandredPreInd			−0.463**
			(0.226)
flatarea	−598.199**	−1,552.052***	−0.135*
	(249.916)	(564.090)	(0.080)
typewindouble-glazed- UPVC		−1,842.613*	
		(1,000.875)	
typewindouble-glazed- wood		−1,633.188	
		(1,213.154)	
typewinmixed types		−813.180	
		(1,208.819)	
typewinsingle-glazed- metal		−5,474.360***	
		(1,637.496)	
typewinsingle-glazed- UPVC		−3,061.403	
		(3,826.228)	
typewinsingle-glazed- wood casement		−903.250	
		(1,094.423)	
typewinsingle-glazed- wood sash		−1,342.717	
		(1,100.831)	
felcavlfQuestion Not Applicable	5,684.806**	−3,015.674	
	(2,509.387)	(5,396.284)	
felcavlfYes	−95.957	1,245.668**	
	(250.376)	(542.601)	
typerstrflat			0.223 (0.197)
typerstrmansard			0.246 (0.276)
typerstrmixed types			0.162 (0.251)
typerstrpitched			−0.047
			(0.168)
felextffQuestion Not Applicable		−8,095.438**	0.480 (0.519)
		(3,940.779)	
felextffYes		2,739.233	0.522***
		(2,169.354)	(0.167)
felextlfQuestion Not Applicable	−5,705.763**	3,111.386	
	(2,510.862)	(5,404.663)	
felextlfYes	−1,232.903**	3,559.246**	
	(555.795)	(1,695.976)	
felextbfQuestion Not Applicable		−2,164.365*	
		(1,228.670)	
felextbfYes		−5,080.463**	
		(2,066.522)	
cavityfront	1,927.267**	−1,171.927	
	(953.862)	(762.740)	
cavityleft	−1,654.055*		−0.199
	(854.971)		(0.137)
cavityright	−1,904.343**		
	(871.570)		

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
wallconstruction		−314.297***	−0.037***
		(55.034)	(0.008)
constructiondate1900-1918	−133.618	−10,180.700***	
	(389.421)	(2,356.010)	
constructiondate1919-1944	−629.732*	11,454.980	
	(328.798)	(8,660.342)	
constructiondate1945-1964	73.834		
	(321.574)		
constructiondate1965-1974	100.749	20,850.630***	
	(343.880)	(5,003.358)	
constructiondate1975-1980	−360.280		
	(494.311)		
constructiondate1981-1990	−1,244.199***		
	(415.128)		
constructiondate1991-1995	−1,441.712**		
	(660.545)		
constructiondate1996-2002	−1,099.831		
	(706.980)		
constructiondatePost 2002	−699.838*		
	(398.058)		
constructiondatePre 1850	652.560**		
	(322.442)		
wallThickness	3,158.325		1.632**
	(2,137.725)		(0.679)
gasmopn/a = No gas		−2,034.873***	
		(745.060)	
gasmopPre payment		−2,675.785***	
		(689.754)	
gasmopStandard credit		−1,629.660***	
		(611.569)	
fpflgbNot in FP - basic income definition	581.527*	2,222.139***	
	(334.254)	(751.629)	
elecmopPre payment			−0.056
			(0.079)
elecmopStandard credit			−0.151**
			(0.069)
fpindf	−2,135.674*		1.340***
	(1,289.204)		(0.376)
fpflgfNot in FP - full income definition	−768.360**	−1,698.766**	
	(356.503)	(759.975)	
fplihcgap	1.549**		
	(0.617)		
fplihcqdtHigh income low costs		−2,222.905***	
		(576.592)	
fplihcqdtLow income high costs		46.263	
		(578.766)	
fplihcqdtLow income low costs		−1,821.565**	
		(751.431)	
agehrpx	127.945***	−26.602	
	(44.833)	(19.553)	
ageoldx	−120.718***		
	(44.934)		

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
unocUnder occupying			0.104*
			(0.060)
agepartx			0.003 (0.002)
finchtypCommunal/CHP	−8,148.802***		−0.715***
	(1,757.142)		(0.241)
finchtypElectric ceiling/underfloor	2,747.625*		0.430 (0.342)
	(1,415.476)		
finchtypQuestion Not Applicable	1,976.628		−0.304
	(1,981.790)		(0.394)
finchtypRoom heaters	4,747.898***		0.148 (0.183)
	(1,247.743)		
finchtypStorage heaters	2,087.242**		0.287*
	(1,012.540)		(0.164)
finchtypWarm air	−146.093		0.430*
	(993.967)		(0.260)
findistyRadiators	−2,159.239**		
	(974.605)		
findistyUnderfloor	6,250.357***		
	(1,639.902)		
hhincx		−0.414***	
		(0.100)	
hyeargrx		0.161***	
		(0.049)	
finnofir		826.143**	
		(328.304)	
finmhfueCommunal - from boiler	−4,143.596**	901.710	
	(1,963.889)	(4,178.862)	
finmhfueElectricity - 10hr tariff	−4,048.115	5,551.307	
	(2,520.113)	(5,280.428)	
finmhfueElectricity - 24hr tariff	−844.776	8,629.020	
	(3,029.931)	(6,210.958)	
finmhfueElectricity - 7hr tariff	−1,850.275	6,891.481	
	(2,206.837)	(4,542.809)	
finmhfueElectricity - standard	−2,155.658	6,901.943	
	(2,229.460)	(4,615.493)	
finmhfueGas - Bulk/LPG	−1,377.994	8,704.779*	
	(2,304.633)	(5,002.130)	
finmhfueGas - Mains	−1,569.778	11,998.340**	
	(2,500.844)	(5,246.026)	
finmhfueOil	−6,372.868***	−2,000.661	
	(2,302.414)	(4,332.658)	
finmhfueQuestion Not Applicable	−5,810.017**	11,896.630**	
	(2,828.069)	(5,792.274)	
finmhfueSolid fuel - anthracite	−1,579.105	−4,690.303	
	(3,423.123)	(7,223.806)	
finmhfueSolid fuel - coal	−4,261.434	17,046.870**	
	(3,324.294)	(7,301.591)	
finmhfueSolid fuel - smokeless fuel	−5,820.910*	−673.376	
	(3,195.725)	(6,400.846)	
finmhfueSolid fuel - wood	−2,461.404	11,823.600**	
	(2,340.024)	(4,930.865)	

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
finchbcd	−0.337**	0.525 (0.344)	
	(0.166)		
finwhcylQuestion Not Applicable		−2,867.514	
		(1,892.430)	
finwhcylUnknown		479.019	
		(415.612)	
finwhmms12.5mm	−1,383.850*	−2,191.281	
	(745.333)	(1,652.363)	
finwhmms25 mm	−141.379	−1,302.315**	
	(247.521)	(545.759)	
finwhmms38 mm	−2,126.907***	1,565.841	
	(659.002)	(1,428.598)	
finwhmms50 mm	−2,219.884**	542.889	
	(891.607)	(1,781.393)	
heat4xfixed room heating	−3,641.989***		
	(1,147.232)		
heat4xstorage heater	−1,318.605		
	(1,076.429)		
mainfuelbulk LPG	1,945.978	−17,664.080**	
	(3,598.372)	(7,855.263)	
mainfuelcommunity heating from boilers	11,088.280***	−1,658.345	
	(3,564.820)	(7,420.785)	
mainfuelcommunity heating from CHP/waste heat	6,962.044*	−4,040.558	
	(3,633.674)	(7,844.030)	
mainfuelelectricity (10 hr. off peak)	6,386.972*	−7,916.335	
	(3,274.319)	(7,105.703)	
mainfuelelectricity (7 hr. off peak)	6,685.988**	−10,389.010	
	(3,211.061)	(6,887.299)	
mainfuelelectricity (7 hr. on peak)	4,714.173	−1,898.283	
	(4,626.694)	(9,227.019)	
mainfuelelectricity (standard tariff)	7,450.393**	−11,035.800	
	(3,158.378)	(6,885.515)	
mainfuelgas (mains)	1,653.429	−6,828.595	
	(2,781.317)	(6,036.646)	
mainfuelheating oil	5,545.196	−9,192.091	
	(3,462.841)	(7,521.782)	
mainfuelhouse coal	3,559.605	−23,362.920***	
	(3,749.324)	(8,232.796)	
mainfuelsmokeless fuel	4,299.161	−8,032.143	
	(4,278.794)	(9,274.553)	
mainfuelwood	2,558.473	−17,550.650***	
	(3,068.993)	(6,657.916)	
watersysElectric immersion heater	1,875.594***	−1,504.898	0.315**
	(444.175)	(951.205)	(0.135)
watersysInstantaneous	2,587.915***	−750.204	0.551***
	(528.159)	(1,106.023)	(0.160)
watersysWith central heating	809.960	696.070	0.191 (0.146)
	(504.763)	(949.491)	
cellard	−245.193**		
	(105.304)		
cellarh	8,290.003**		2.424*
	(3,889.835)		(1.275)

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
housestoreyx		−2,084.330**	0.235*
		(860.011)	(0.122)
cellarw		396.782	
		(283.693)	
isbungalow			−1.979***
			(0.163)
storeysw	277.440***	571.876***	0.113***
	(49.060)	(133.378)	(0.019)
storeysd		343.387***	0.040**
		(109.681)	(0.016)
storeysh	−1,445.879**	6,945.203	2.005***
	(648.238)	(4,331.311)	(0.669)
cellarhxtra	−4,052.833*		−1.214*
	(2,248.383)		(0.699)
storeysdxtra	97.679*		
	(58.020)		
facadeheight	273.043***		0.149***
	(51.018)		(0.037)
cellarwxtra		2,532.359***	0.490***
		(730.670)	(0.107)
storeyshxtra		−8,748.733**	−2.617***
		(3,931.356)	(0.611)
mainfa		−9.688**	−0.004***
		(4.637)	(0.002)
mainfaxtra			0.012***
			(0.004)
aglfloorarea			0.0001**
			(0.0001)
lsf		812.543***	0.117**
		(266.981)	(0.055)
flatblocks	−823.425***	−362.117**	−0.500***
	(123.492)	(174.361)	(0.095)
flatfloors		764.245	0.287***
		(579.051)	(0.095)
flatw		−773.434*	
		(424.761)	
flatd		−807.233*	
		(424.288)	
flath	−939.991	6,037.718***	−0.826***
	(691.929)	(1,524.396)	(0.231)
flatareax	605.104**	1,640.602***	0.131 (0.082)
	(255.388)	(584.958)	
fpbasinc	0.065 (0.046)		
fpfullinc	−0.092*	−0.082	−0.00001
	(0.051)	(0.056)	(0.00000)
allincx		0.167**	
		(0.068)	
fuelexpn	−2.211***	−2.563*	
	(0.674)	(1.310)	
wathcost	4.012***		
	(1.153)		

Continued on next page

Table 5.1 – *Continued from previous page*

	Electricity Demand	Gas Demand	Internal Temperature
spahcost	2.931***	3.599**	0.0003***
	(0.721)	(1.420)	(0.0001)
cookcost	14.192 (9.501)		0.008***
			(0.002)
litecost		12.320***	
		(2.821)	
eHS	5,262.710**	−994.118*	
	(2,624.621)	(587.506)	
occupancyEUSILC	−4,829.164*		
	(2,602.369)		
attachleft	−391.851*		
	(236.908)		
eHIindpoverty		866.333*	
		(479.746)	
Constant	−1,444.060	−20,724.350*	17.058***
	(6,124.506)	(10,611.130)	(1.676)

Note: *p<0.1; **p<0.05; ***p<0.01

As mentioned earlier, archetypes are represented as sets of contiguous cuboids that account for the number of floors and the presence of attics, basements and attachments such as extensions and conservatories, which allows the possibility to represent the key residential types of the stock, now based on an explicit volumetric representation.

5.1.2 Step a_2 Reduction (or Sampling)

By generating volumetric archetypes, model complexity (in terms of data inputs) and computational cost increase. It is then useful to explore strategies to reduce the archetypes. The first approach consists of considering geometrical configurations of the relatively homogeneous housing stock. This approach allows us to capture attached elements in the typologies. For instance, the cuboid configuration, developed in EnHub, is different for mid-terraced and end of terrace houses. Table 5.2 gives eight reduced variables, C_{1-8}. Variables $C_1 - C_4$ (given above the dashed line in Table 5.2) are used for this stage, and so there are 64 possible archetypes. Nevertheless, despite the ability to represent physical typologies using this combinatorial approach, the reduction leads to a misrepresentation of significant properties of the stock, such as the predominance of gas heating systems, period of construction, tenure type, and region, which are correlated with fabric properties, such as construction type, materials and available fuel. By including these properties, the relevant variables identified in both the FPCA and the GLM can be preserved, so that the stock can then be better characterised.

The location is used to identify an appropriate weather file used by the `EnergyPlus` engine, and because the resolution provided by the dataset is limited (i.e. NUTS-2 level), it is explicitly defined for each archetype. Furthermore, the location is also used to estimate the number of heating degree days, and is related to the typical construction types of dwellings.

Table 5.2: Reduction variables. The initial reduction considers the combination of the first group of variables (i.e. $C(4,1)_{floors} \cdot (C(2,1)_{basement} \cdot C(2,1)_{attic} \cdot C(4,1)_{attachments} = 64$). In addition to this preliminary combination, other variables are identified in Figure 5.4 regarding physical properties and heating energy demand. However, these variables are also correlated with typological elements such as region, built age, heating type and tenure. The selection of the group of variables below the dashed line results also from previous studies that use these parameters to identify unique typologies. It is worth mentioning that some combinations are absent in survey data such as pre 1920 construction with electric heating system, and modern construction with other fuels. These combinations are removed.

Category	**Value**
no. floors	[GF, 1F, 2F, 3F]
basement	[present, absent]
attic	[present, absent]
attachments	[left, right, both, no]
regions	[north, central, south]
heater type	[gas, electric, other]
tenure	[private, social]
built age	[pre 1920, pre 1975, modern]

These variables, summarised in the bottom of Table 5.2, are also employed to reduce the dataset by applying LHS. It is worth noting that the properties are considered in a reduced form, otherwise the combination of each category, as specified in the EHS, would increase the size. Therefore, LHS is applied varying the sample size, the number of variables, and the iteration size. This sampling strategy is convenient because it can handle both numerical and categorical data. The algorithm employed in the generation of a LHS selects n and compares between the random sample x and the original dataset X. The method consists of a search algorithm based on heuristic rules [160], where the number of variables k is the same for both X, x and then η, which represents the true-matrix of those cases existing in both X and x (see Eq. 5.3 below).

$$\eta = \begin{bmatrix} n_{11} & n_{12} & \dots & n_{1k} \\ n_{21} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ n_{n1} & n_{n2} & \dots & n_{nk} \end{bmatrix} \tag{5.3}$$

The objective function defined in Eq. 5.4 minimises the difference between the random sample η and the ideal random sample η_i (i.e. all the cases that exist in X).

$$O = \left\| \eta - \eta_i \right\| \tag{5.4}$$

The results of the LHS indicate that by using the variables $C_1 - C_8$ (see Table 5.2), the objective function stabilises with samples over 600 elements and converges with over 1000. Both reference variables and the weighting values are used to optimise the objective matrix for each EHS sampling

unit. This reduction process shows that when only a few variables are considered, the resulting dataset is highly affected by the sample size, whereas the dataset with more variables stabilises much faster. During the sampling process, described in [208], some properties are removed from the dataset, such as specific heating devices and uncommon materials.

As a result of this LHS, 1016 unique units are obtained that are able to represent the main composition of the housing stock; a 15-fold reduction of the EHS dataset, whilst adequately representing its structure. The combination of variables in Table 5.2 results in 3456 unique typologies; however, as mentioned, some of these unique elements are absent in the survey data and very possibly in real world composition. In addition, other elements represent only a few archetypes as compared with more common typologies. Therefore, the sample is more likely to discard these elements, but maintaining the overal composition of the EHS. Figure 5.5 outlines the performance of samples using a different number of pivotal variables and an equal number of iterations. The selection of the main pivotal variables is, preferably, with the most heterogeneous from both Table 5.2 and EHS, such as the attachments variable.

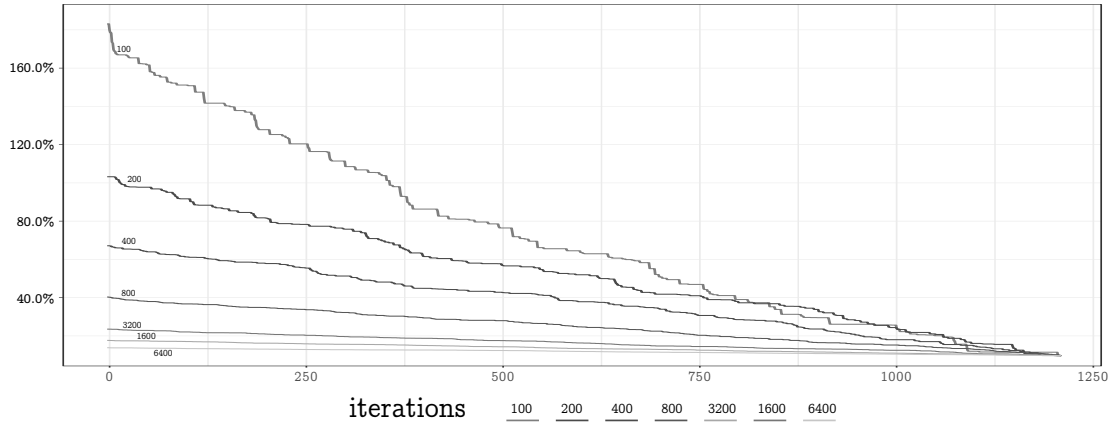
An alternative method of creating a reduced stock is to associate median parameter values to the 64 units formed from 7 of the 8 categorical variables ($C_1 - C_6$ and C_8) given in Table 5.2 (C_7 is excluded here because it is too broad), but this risks an inadequate description of the structure of the stock for a modest computational saving.

Figures 5.6a-5.6c suggest that the loss of information through statistical reduction is not substantial in the representation of the structure of the stock. For instance, Figure 5.6a shows that the statistical properties of the variables remain the same despite the noticeable reduction of elements in the sample. Similarly, the impact in the characterisation of categorical variables is not significantly affected as seen in Figure 5.6c. Similar results are found for all the input variables forming the dataset.

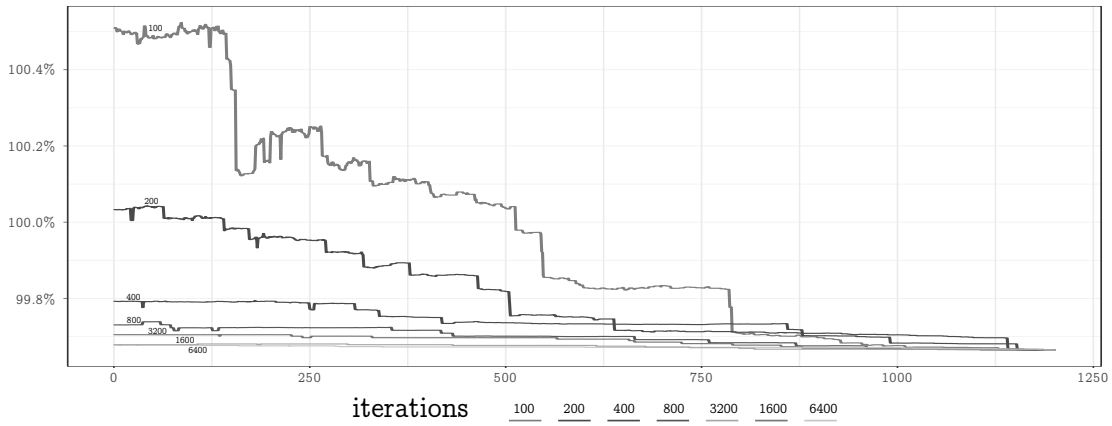
The resulting variability in the dataset is supported by performing a Kruskal-Wallis Test (KWt) across all the variables. This test, as an alternative to the ANalysis Of VAriances (ANOVA), identifies whether the variables in the reduced dataset maintain the original information or not. The KWt compares the correlation of its variables, so these can be numerical or categorical; it is also used when the examined groups are of unequal size. The KWt is defined by Equation 5.5 where c represents the number of samples (in this case comparing the two datasets), n_i is the number of observations in each sample, $N = \sum n_i$ sums all the observations combined, Ri is the sum of the ranks in the ith sample, and $T = (t-1)t(t+1)$ is the mean of the tied ranks.

$$H = \frac{\frac{12}{N(N+1)} \cdot \sum_{i=1}^{c} \frac{R_i^2}{n_i} - 3(N+1)}{1 - \sum \frac{T}{(N^3-N)}} \tag{5.5}$$

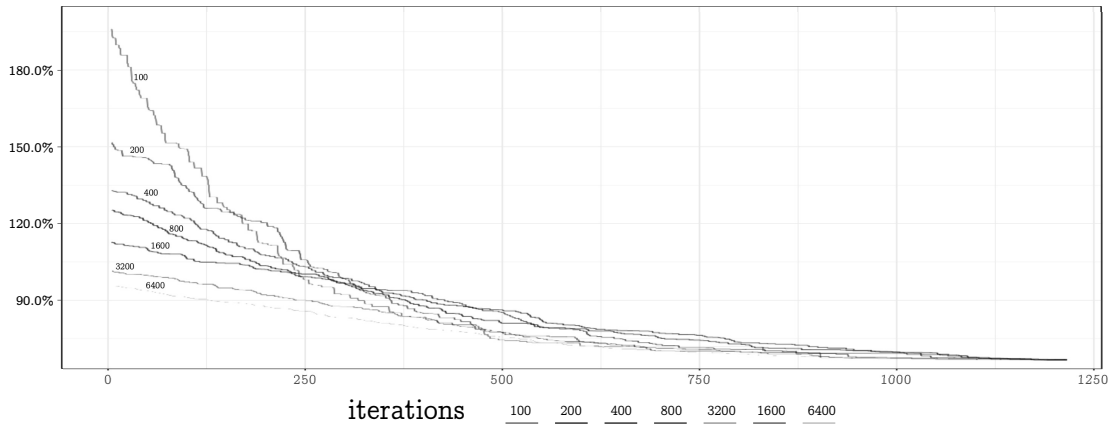
H is a measure of the relative proximity in the correlation of both groups. Large values of H indicate that the datasets are significantly independent. Because the KWt assumes a distribution, then H can be used to measure the level of significance defined by the distribution under the null hypothesis; this hypothesis establishes that the difference between the medians in both groups is null (or quasi-null). Therefore, a value below the level of significance suggests that the groups, in the given category, are different. The purpose of employing an enhanced sampling method is to



(a) 3 pivotal variables - from Table 5.2



(b) 7 pivotal variables - from Table 5.2



(c) 40 pivotal variables - from EHS

Figure 5.5: Conditioned LHS iterations. Equal number of iteration, different pivotal variables.

Table 5.3: KWt summary

variable	name	p.value	signif
felextrf	External insulation (right side)	0.027541	signif
.attach_right	Attachment (right side)	0.029337	signif
ageoldx	Age of oldest person	0.031374	signif
unoc	Under occupancy	0.038761	signif
pyngx	Age of youngest person	0.043826	signif
.main_fa_xtra	Main Extra Construction	0.047755	signif
finnofir	Total number of open fireplaces	0.051575	not
.attachR	Attachment (right side)	0.056553	not
DHWSystem	DWH system	0.112757	not
mainfuel	Main Fuel	0.120774	not
finchtyp	Main Heating System	0.129305	not
finmhfue	Main Fuel	0.153498	not
⋮			
.insulation_front	Insulation (front wall)	0.909020	not
FstFStoreyHeight	1F height	0.944843	not
GFStoreyHeight	GF height	0.965938	not

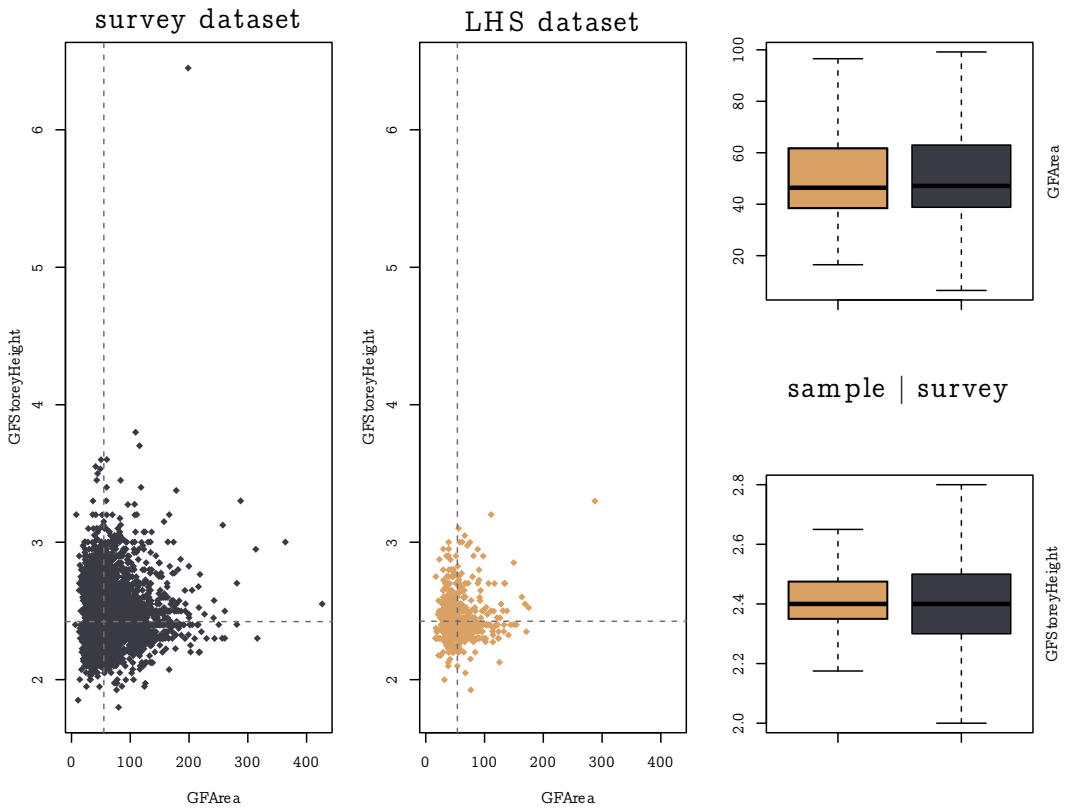
represent the same information with a smaller dataset.

Table 5.3 shows that six of the input variables have a p-value < 0.05, showing that they are significant and so should be included in the reduced dataset; however, none of them directly affect the composition of the volumetric cuboids. Because the EnHub platform uses the reduced data set, a significant variation is expected in the level of detail of non-primary components, such as secondary heating systems and boiler parameters. To determine whether EnHub's predictions of energy performance are sensitive to these parameters, a OAT sensitivity analysis is applied (described in Section 4.2.2). It finds that the dwelling typology significantly impacts energy performance; but this variability, expressed by H, is more sensitive to the contiguous cuboids whose geometry is configured and scaled using the reduced dataset.

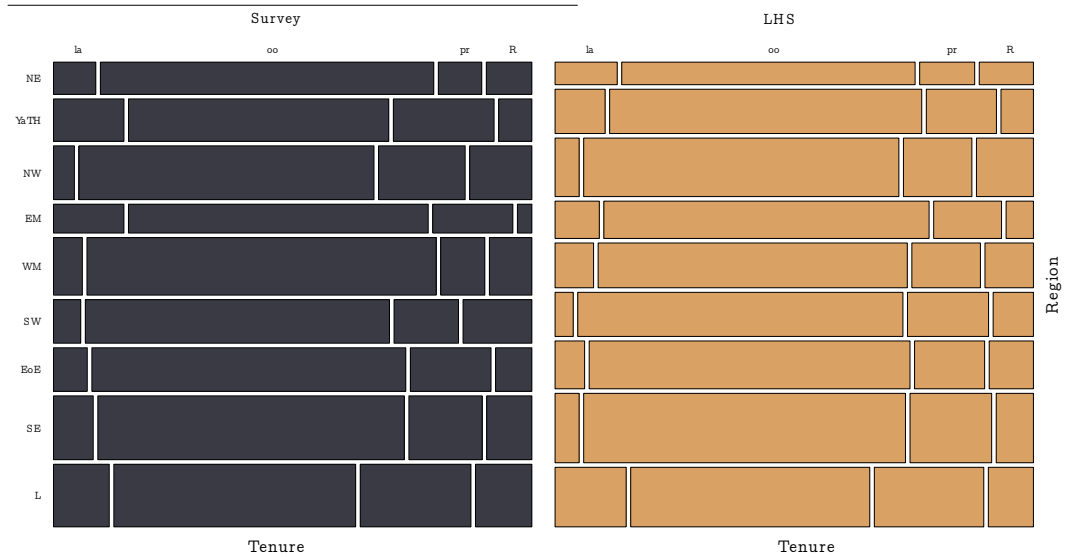
Solar energy technologies and socio-demographic variables are identified by the OAT analysis as significantly relevant to the models describing energy performance, and so they are directly included into the model archetypes.

The original units of the dataset are directly derived from the EHS, which in turn uses the PAF and the Lower-SOA to define their weights, and are used to represent the whole English housing stock. These units are selected randomly and merge information from near PAFs, and the extrapolation is delimited by census subtotals. The inherent bias in this method affects solar-related parameters, and the definition of representative households. Thus, improving the quality of information about these variables would undoubtedly improve their significance in a future analysis.

Once the dataset has been reduced, the units are re-weighted, and their tenure and regional information are re-calibrated following a method employed by many other nationwide surveys [21]. This method compensates for the assumptions made in the sampling process which, as in the EHS, introduces a loss and bias of information [109]. The impact of this reduction affects the performance of the simulation algorithms, because their inputs may not exist or may be over-generalised. Hence, there is virtue in developing a modular approach to improve components

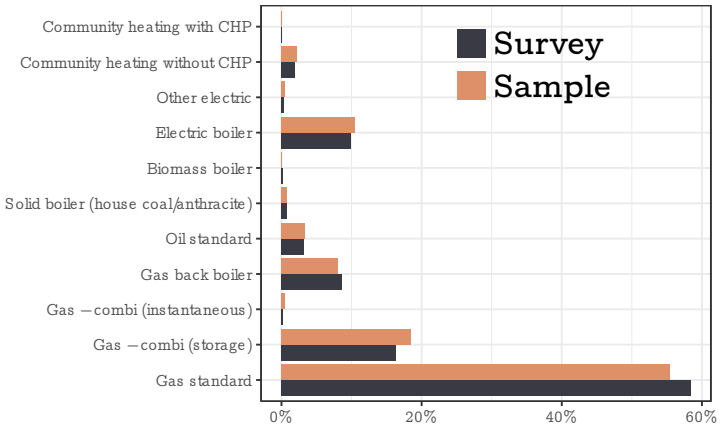


(a) Ground floor dimensions

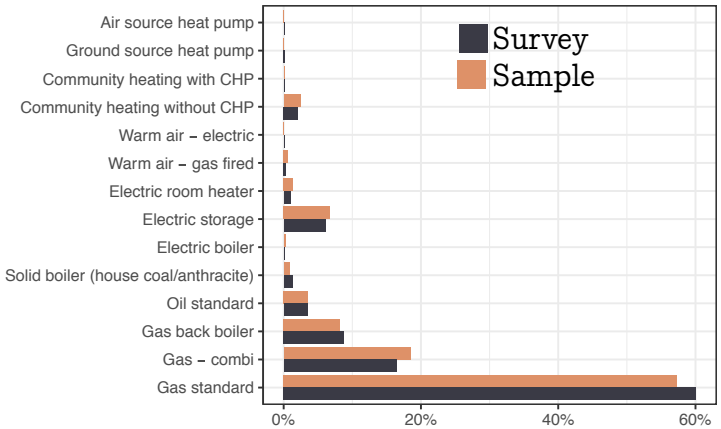


(b) Location and region composition

Figure 5.6: Comparison between the EHS and the reduced dataset



(c) DHW systems



(d) Space heating systems

Figure 5.6: Comparison between the EHS and the reduced dataset – continued

independently.

An alternative approach, described in Steps 5b(5.1.3) and 5c(5.1.4), starts by reducing the number of elements (see Figure 5.2). Then each archetype is enriched with the main attributes for each case found in the EHS data.

5.1.3 Step b_1 Statistical Reduction

The criterion to perform a statistical reduction is based on the homogeneity of the housing stock typology in the UK. Here, top rows in Table 5.2 are used to summarise the variables used in the reduction process. The initial reduction considers the combination of the first group of variables (i.e. $C(4,1)_{flr} \cdot (C(2,1)_{bsmt} \cdot C(2,1)_{attc}) \cdot C(4,1)_{atch} = 64$). These variables are selected due to their promptly availability in the datasets, whereas more detailed information regarding floor layouts is usually (and expectedly) absent.

As described in § 6.3.4, it is possible to obtain a representative dataset with samples between 600 and 1200 elements. Thus, the applied reduction includes the second group of variables that define the representative factors shaping the housing stock (bottom rows in Table 5.2). These

factors are identified by applying regression techniques. Likewise, it must be noted that not all the combinations described in Table 5.2 (i.e. $64 \cdot C(3,1)^3$) are found in the stock. As a result, by employing this method, the reduction of dwellings comprises a total of 865 units.

The reduced dataset of 865 archetypes has the potential to represent the EHS; however, at this stage the reduced dataset does not include additional information from the survey. These are archetypes that represent the combination of Table 5.2 variables only, without any additional information embedded into them. Therefore, it requires a process of enrichment and calibration, to capture the properties of the housing stock. The process is based on predictive methods, and depends on the chosen typologies to infer their corresponding information.

5.1.4 Step b_2 Enrichment and Calibration

This step calibrates the reduced dataset by assigning the most common value to each variable. To this end, the following sequence is applied: (a) use of partition trees techniques, and (b) re-weighting of the resulting archetypes to represent the entire housing stock. The partition trees replicate the distribution of each variable (i.e. categorical, nominal or numerical data) taking as initial reference the variables presented in Table 5.2. Partition trees are generated by dividing the dataset into a number of groups, each division being chosen in function of the chosen variable of proximity. A relevant attribute of this approach is that it does not require the assumption of a distribution in the evaluated variables [89].

In this step, each variable is iteratively derived from the predominant cases found in the EHS, so that the 865 units are able to represent the EHS sample (14,951), and therefore the stock total (21 million dwellings). This approach is capable to maintaining the overall conformation of the housing stock. It must be considered, nevertheless, that the EHS is already losing some information by using pseudo-synthetic data derived from the SOAs [109], but it is able to define homogeneous groups of the population. The sequence for this step is exemplified in the Pseudo-code 5.1[3].

Pseudo-code 5.1 Enrichment sequence

```
1. identify type of variables for each input
  t -> function=type(v)

2. group inputs by variables
  C -> function=group(v)

3. adjust prunning method for each group
  C_par -> function=prune(C)

4. iterate employing its corresponding partition tree
  loop:
    for i:=1 in {cat,num,log,str} do
      for j:=1 to i_n do
        function=partition(i,j)

5. subset original dataset by typologies
  sub -> function=typologies(D)

6. match reduced dataset with subset
  M -> function=match(red,sub)
```

[3]The process of pruning the trees consists of applying a segmentation to the partition trees and measuring the resulting variation. Pruning techniques optimise the resolution of the partition trees [89].

```
7. calibrate resulting match with original dataset
  function=calibrate(M, D)
```

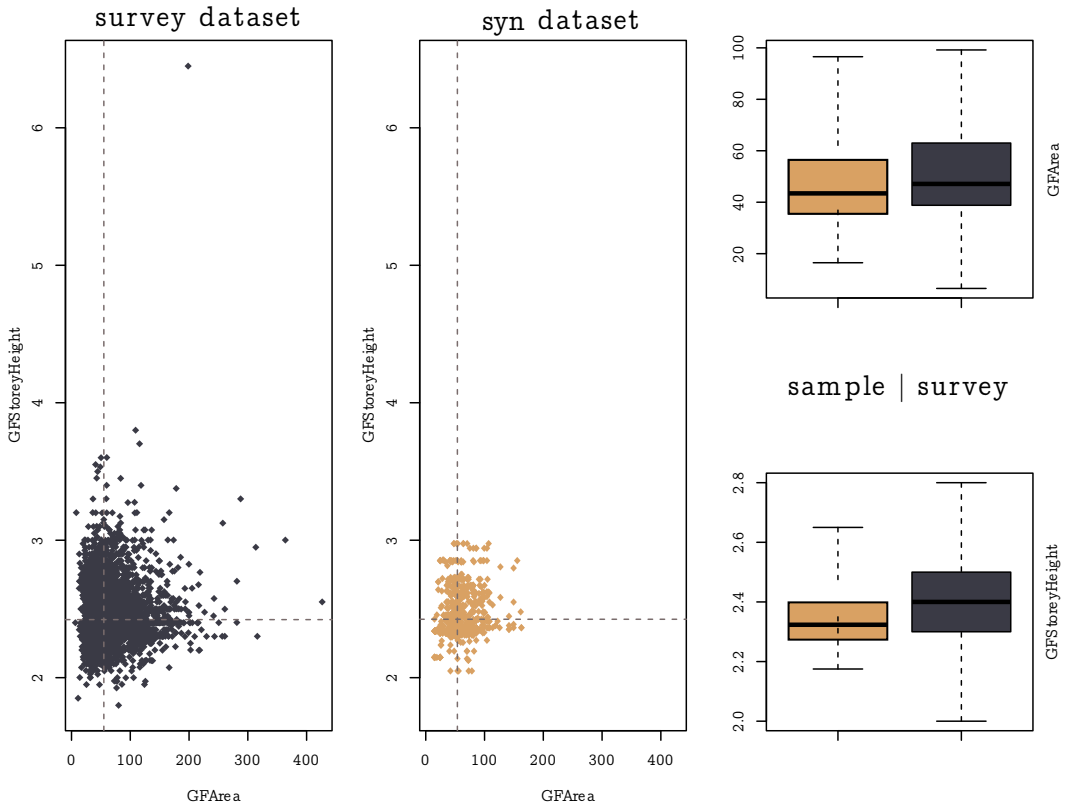
Figures 5.7a-5.7d illustrate a comparison between the EHS and the reduced dataset. It can be seen that the resulting categories maintain the balance after the reduction process. The Figures show how the reduction of variables may slightly alter the variability of the original dataset; however, by applying the enrichment sequence, the dataset can still replicate the most common cases of the original dataset. The results of the partition tree are presented below:

Output Tree

```
   Conditional inference tree with 22 terminal nodes

Response:  aagpd1011
Inputs:  .isattic, .iscellar, .floors, .AttachState, .region_red, .ageband_red, .heater_red, Inc.cat
Number of observations:  14386

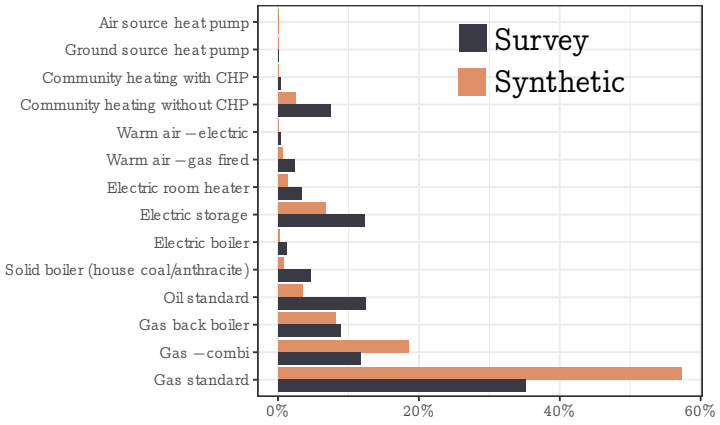
1) .floors <= 1; criterion = 1, statistic = 297.637
  2) .region_red == {Central, South}; criterion = 1, statistic = 87.769
    3) .AttachState == {None}; criterion = 1, statistic = 46.758
      4) Inc.cat == {high}; criterion = 0.999, statistic = 17.925
        5)*  weights = 36
      4) Inc.cat == {low, med}
        6) .heater_red == {Gas, Oil}; criterion = 0.969, statistic = 15.388
          7)*  weights = 583
        6) .heater_red == {Electricity, Other, Solid}
          8)*  weights = 104
    3) .AttachState == {Both, Left, Right}
      9) .ageband_red == {Hrtg}; criterion = 1, statistic = 37.421
        10)*  weights = 274
      9) .ageband_red == {PostInd, PreInd}
        11) .heater_red == {Electricity, Gas, Oil, Solid}; criterion = 1, statistic = 26.487
          12) Inc.cat == {high}; criterion = 0.998, statistic = 16.7
            13)*  weights = 103
          12) Inc.cat == {low, med}
            14)*  weights = 2121
        11) .heater_red == {Other}
          15)*  weights = 254
  2) .region_red == {North}
    16) .AttachState == {None}; criterion = 1, statistic = 35.641
      17)*  weights = 229
    16) .AttachState == {Both, Left, Right}
      18) .heater_red == {Electricity}; criterion = 1, statistic = 33.196
        19)*  weights = 202
      18) .heater_red == {Gas, Oil, Other, Solid}
        20) .ageband_red == {Hrtg}; criterion = 0.999, statistic = 17.893
          21)*  weights = 40
        20) .ageband_red == {PostInd, PreInd}
          22) Inc.cat == {med, high}; criterion = 0.979, statistic = 11.905
            23)*  weights = 726
          22) Inc.cat == {low}
            24)*  weights = 31
1) .floors > 1
  25) .AttachState == {None}; criterion = 1, statistic = 167.857
    26) .region_red == {Central, South}; criterion = 0.982, statistic = 12.17
      27) .ageband_red == {PreInd}; criterion = 0.964, statistic = 10.766
        28) Inc.cat == {low, med}; criterion = 0.999, statistic = 19.24
          29)*  weights = 316
        28) Inc.cat == {high}
          30)*  weights = 110
      27) .ageband_red == {Hrtg, PostInd}
        31)*  weights = 958
    26) .region_red == {North}
      32)*  weights = 473
  25) .AttachState == {Both, Left, Right}
    33) Inc.cat == {low, high}; criterion = 1, statistic = 132.258
      34)*  weights = 817
    33) Inc.cat == {med}
      35) .region_red == {Central, South}; criterion = 1, statistic = 77.037
        36) .isattic <= 0; criterion = 1, statistic = 25.507
          37)*  weights = 3839
        36) .isattic > 0
          38)*  weights = 413
      35) .region_red == {North}
        39) .ageband_red == {Hrtg}; criterion = 1, statistic = 35.259
          40)*  weights = 333
        39) .ageband_red == {PostInd, PreInd}
          41) .isattic <= 0; criterion = 1, statistic = 28.908
            42)*  weights = 2197
          41) .isattic > 0
            43)*  weights = 227
```



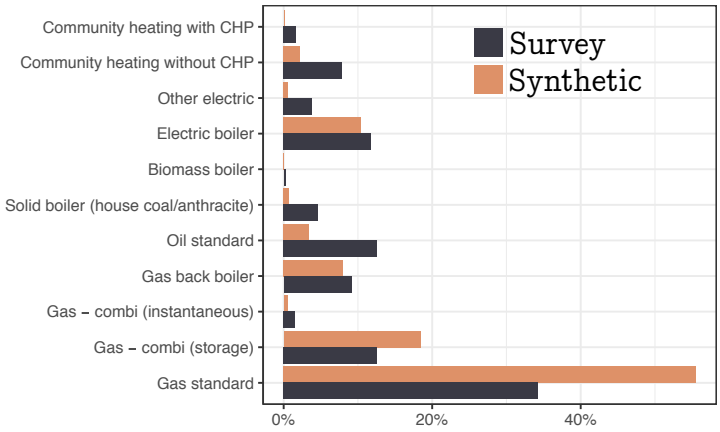
(a) Ground floor dimensions

(b) Location and region composition

Figure 5.7: Comparison between the EHS and the synthetic dataset



(c) DHW systems



(d) Space heating systems

Figure 5.7: Comparison between the EHS and the synthetic dataset – continued

In similar fashion to the step described in § 6.3.4, in order to measure the significance of the reductions, a KWt is applied to each of the variables. The results of this test show that 10 variables present a p-value below 0.05, which usually is a good indicator of significance for the comparison of independent tests. The KWt results show that 10 variables present a p-value below 0.05, which usually is a good indicator of significance. One of the affected variables is the dwelling typology (p=0.38), but we also consider this as a premise in the conformation of the volumetric archetypes, by generating an all-encompassing cuboid. Other variables that are significantly different in the reduction are those related to the solar technologies (p=0.29), and properties of the secondary heating systems (p=0.13). This is expected because we are affecting the level of detail to account for secondary and additional systems by retrieving the main components in the reduced dataset.

5.1.5 Step $a - b_3$ Construction of Archetypes

Once the reduced dataset is defined, the construction of volumetric archetypes follows the same algorithm described in § 4.1.4, including their corresponding weather files and activity profiles. Each of the archetypes is linked with the extended version of the EHS; by doing so, the socio-demographic variables can be used to shape household-specific profiles and to emulate

decision-making, in a later stage. The socio-demographic variables include information such as income of the occupants, age and composition, education and fuel expenditure.

The archetypes use an identifier representative of the parameters $C_1 - C_7$ which is used to combine scenarios of the same archetype or subsets of the dataset. In addition, the reduced dataset is used to re-calibrate the assignment of geographical locations using the Census information (see the weighting process illustrated in Figure 2.15).

In respect of the construction of archetypes, the main difference between both reduction approaches is that, an average dwelling is formed using statistical parameters (i.e. median, mean, standard deviation), and so the archetypes are scaled accordingly from the 865 average dwellings; whereas the 1016 sample archetypes are selected from the original survey. The methods can be combined and used complementary to each other, yet the basis is based in the example provided in Figure 5.1. Incidentally, it was described that the EHS follows a process of sampling and physical inspection [47]. This means that only a fragment of the survey may be composed of empirical information. In sum, it can be said that *average dwellings* are more likely to be theoretical, whereas *representative cases* are more likely to be real. This implies that real-istic dwellings may be more easily linked with socio-demographic information, because they can match specific values from (real) households and (real) dwellings.

The process described in Figure 5.2a is chosen as the more versatile method in which the support for changes in the dataset can be replicated with ease. This method is also opted because the sample units may be directly linked with geographical information or complemented with census information.

5.1.6 Step $a - b_4$ Dynamic Simulation

In this step, the reduced dataset is simulated for an entire calendar year at an hourly resolution. This relatively coarse grained temporal resolution is reasonable, given that we are not explicitly modelling ventilation nor accounting for occupants' responses at this stage; we are simply utilising infiltration and ventilation schedules.

As described in § 4.2.3, each simulation takes around 30 seconds to complete using 1 core in a high-end machine. However, given that our stock is represented by around 1000 archetypes, given by the chosen reduction method, to which we also wish to define and test decarbonisation scenarios—so that we perform multiple rounds of simulation—the workflow is coupled with the UoN-HPC, enabling the simulation of 64x models in parallel, and hence reducing the simulation time considerably (i.e. 1000·30s = 8.3hr; ∴ 64x = 8m.)

With this in mind, it is worth noting that, although the use of an optimised hardware might increase the performance in more than 64x, it may be impractical if the period for analysis is flexible or wide. In consequence, the generation of results at such a pace may not be necessary. Therefore, a reliable request of HPC resources might be limited to 16x or 32x cores. In short, the request of resources needs to be planned so that the high performance is also reflected in the whole study.

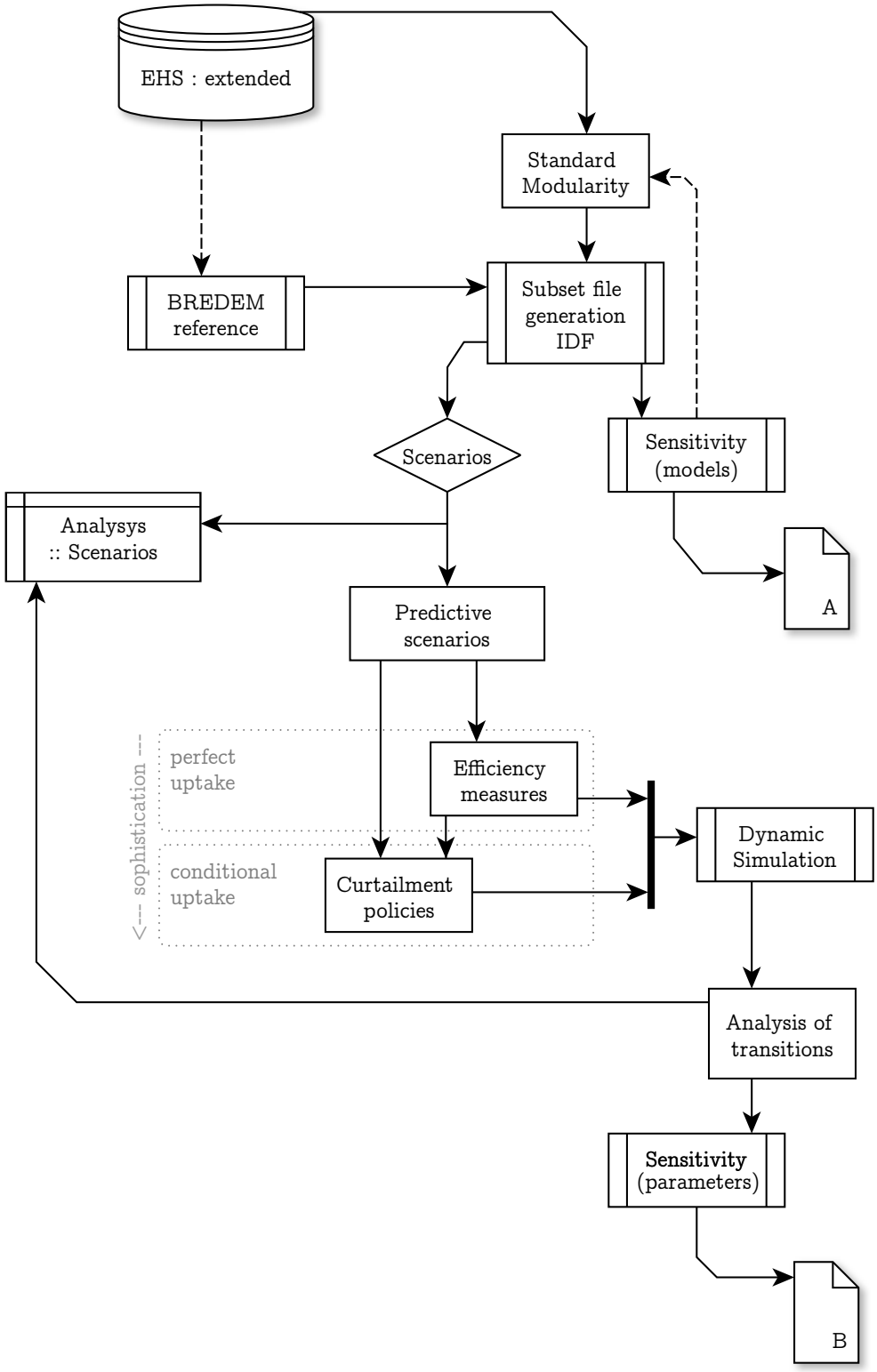


Figure 5.8: Simulation and analysis flowchart: development of scenarios. The processes of verification, validation and calibration are applied within the sensitivity modules

5.1.7 Step $a - b_5$ Iterative Analysis

Each batch of simulations, encompassing the reduced dataset, is conveniently stored employing a typical database structure (i.e. multiple datasets linked by common identifiers); this process is represented by the flowchart described in Figure 5.8. Both inputs and the resulting scenarios are central for the validation of EnHub. For instance, the first stage (A) focuses on the data resolution, whereas the second (B) on the modelling appropriateness and the fluctuation in energy demand as a result of the study of different scenarios. The Figure also shows that the validation is supported by SAs. For reasons of efficiency (and practicality), this analysis is also carried out utilising the UoN-HPC although requesting only a single enhanced core of the facility, following a similar argument as described earlier.

5.2 Scalability of the Platform

An attribute of this platform, derived by its modularity and its ability to detach processes, consists of a relatively high level of scalability. The concept of scalability may be defined as a property of a platform, tool or application, to perform reasonably well, when the inputs change in size (usually an increase), and also when they change in their properties or context. The scalability of a simulation platform typically deals with the required amount of input parameters that can (or should) affect the performance of the model. Zhao et al. [233] evaluate sampling methods in terms of their effectiveness to replicate estimated energy demands. Although the study is applied to evaluate the performance of commercial buildings, it provides a valuable example to evaluate the scalability of a model in terms of the model outcomes. The level of scalability provided by the EnHub platform can be summarised by the following attributes:

I/ EnHub is able to perform a highly disaggregated SA on the algorithms describing a component of the energy flow in dwellings; as well as the analysis on controlled changes to the properties of the archetypes.

II/ EnHub opens the possibility to rapidly evaluate different datasets, either by generating a synthetic dataset, or by extracting and calibrating a sample to test different parameters.

III/ EnHub can evaluate changes on the reduced dataset to test the effectiveness of interventions (or changes) at stock level.

IV/ EnHub can replace occupancy and behaviour algorithms with more sophisticated approaches, despite the potentially significant increase in computational cost.

With regard to the dynamic simulation module, EnHub has the potential to mimic more sophisticated applications. For example, the parametric analysis might be complemented with the advances developed in some specialised and fit-for-purpose softwares such as `jEplus`[4], and the optimisation algorithms developed in there. Likewise, a very specific analysis of indoor responses might be coupled with applications such as `BCVTB`[5], that improve the active feedback required

[4] `jEPlus` is an open-source tool originally developed for managing complex parametric simulations using `EnergyPlus` [232]

[5] `Building Controls Virtual Test Bed` is a software environment, based on Ptolemy II, for co-simulation [226]. For example, it can integrate `EnergyPlus`, `Matlab` and specific applications

to provide a more accurate representation of indoor performance. Although, at present, such application is limited by a single dwelling or by a limited stock of houses.

Furthermore, as discussed in Chapter 3, a logical next step would consist in scaling the functionality of EnHub. Here, the consideration of multi-agent simulation techniques appear to have much to offer in this endeavour [26, 28]. This functionality has been accomplished directly between the dynamic simulation engine `EnergyPlus`, and improved communication protocols. For example, by using a dedicated interface (e.g. a fit-for-purpose GUI), by developing a master programming platform (e.g. EnHub), or by adapting a more specific—but sophisticated—technique such as Functional Mock-up Interface (FMI), which have received special attention in recent years [104]. Its consideration enables the inclusion of more detailed and specific household information. By way of demonstration, the Nottingham Multi-Agent Stochastic Simulation (No-MASS) extension [28] that, using such a FMI protocol, is able to include socio-demographic characteristics into the volumetric models and perform highly disaggregated dynamic simulation under an agent-based environment. The approach complements the modelling of occupants' activities and dependent behaviours, and their corresponding impacts on thermal and electrical energy demands [187, 227, 117, 151, 194].

In sum, the potential to scale the EnHub platform is considerable, and its open-modularity supports the straightforward edition of each of its component parts.

5.3 Conclusion of *Workflow for Efficient Housing Stock Representation*

This Chapter has described two broad methods to reduce the EHS, by identifying redundant variables and by applying methods to generate synthetic datasets. Both methods have the potential to be employed in the evaluation of scenarios and analyses at stock level. Such a development is supported by the inclusion of SA and statistical models, used to validate each of the components, as well as the assumptions used in the evaluation.

Concerning the reduction methods, the sampling one (in this case via LHS), is convenient because the archetypes can be quickly updated and the process can be conveniently compared with the original survey. Average dwellings have, nevertheless, the potential to directly combine surveys from different years, and so the derived data becomes more accurate. To employ different years in survey data, and to incorporate them into models, it is required a further process of validation regarding the underlying methods to deploy the survey. Yet, at present, this is not considered in EnHub.

In sum, this Chapter has described that, even though the sampling size is significantly reduced, the application of such a systematic approach in the EnHub platform opens the possibility to evaluate scenarios based on dynamic simulations (i.e. including indoor-specific metrics, activity profiles, shape-related performances). Likewise, the platform enables the possibility to scale the dataset and, consequently, to study subsets of the population, particularly where more disaggregated information can be included. One such example would be to support the evaluation of detailed housing stock decarbonisation scenarios for a single town, city or urban region; perhaps supplemented with additional surveys relating to locale-specific housing typologies.

Application Part III

6 Analysis of Strategies to Decarbonise the British Housing Stock

Introduction

Chapter 2 described the contribution of the domestic sector to national carbon intensity. Here, heating was highlighted as the main contributor to energy demand in dwellings; the Chapter also identified a set of common assumptions in stock data, which are important to consider while defining alternatives to reduce the stock energy demand. In Chapter 3, it was identified that, despite the number of enhancements in the description of energy flow pathways in HSEMs, its representation is not enough to fully evaluate the reduction in energy demand and associated carbon emissions. Decision-making drivers are important to describe the way energy is used in dwellings, and to study the interventions and changes associated to household conditions, but they have been poorly represented in HSEMs. In consequence, the strategies to improve energy performance of the housing stock has been undermined, making some of them unsuccessful.

In addition, because of the limited resolution of data and the sophistication of modelling algorithms to represent energy flow pathways in HSEMs, a significant level of redundancy is found in the modelling approaches. Therefore, Chapters 4 and 5 presented the basis of EnHub, an open-source platform that is able to improve HSEMs, by linking, validating and replacing specific modelling modules; and by mining the datasets employed. In this way, a process of calibration of data and models is developed to support the development of an evaluation platform of the housing stock in the UK. This platform is scalable in two directions. Firstly, it can perform a highly disaggregated SA on each of its components, which enables a process of calibration and potential collaboration of different perspectives and disciplines. Secondly, the platform can improve on controlled parametrisation to evaluate different scenarios of decarbonisation, which provides information to better design policies and assessments for the housing stock.

The following sections describe the evaluation of scenarios to measure the potential of energy-related uptakes in the housing stock. The evaluation distinguishes between perfect uptakes and conditioned (or assumed) penetrations. Furthermore, the evaluation includes a proposition of social simulation complement to predict penetration in response to specific policies and strategies; this complement is designed to simulate the potential of energy-related uptakes.

Table 6.1: Opted scenarios for perfect and conditional uptakes

id	*Scenario*	*Status*
A: perfect uptake		
A01	**Solid wall insulation**	Non-insulated: 47%; semi-upgraded: 21%; efficient 32%
A02	**Loft insulation**	Non-insulated: 45%; semi-upgraded: 35%; efficient 20%
A03	**Double glazing**	Non-efficient windows: 45%; semi-upgraded: 35%; efficient 20%
A04	**Cylinder insulation**	Non-insulated:15%; below 50mm: 50%; efficient:35%
A05	**Draught proofting**	Below permitted: 65%; efficient: 35%
A06	**Low energy lights**	Non-efficient lighting: 35%; semi-upgraded: 50%; efficient 15%
A07	**Efficient household appliances**	Non-efficient appliances: 50%; semi-upgraded: 45%; efficient 5%
A08	**User Behaviour**	Behaviour stimuli in the use of appliances, lighting and metabolic gains.
B: conditional uptake		
B01	**Solid wall insulation**	This represents: 60% of the A01 upgrade.
B02	**Loft insulation**	This represents: 45% of the A02 upgrade.
B03	**Double glazing**	This represents: 56% of the A03 upgrade.
B04	**Cylinder insulation**	This represents: 46% of the A04 upgrade.
B05	**Draught proofting**	This represents: 33% of the A05 upgrade.
B06	**Low energy lights**	This represents: 40% of the A06 upgrade.
B07	**Efficient household appliances**	This represents: 40% of the A07 upgrade.
B08	**User Behaviour**	Test different assumptions on occupancy and minimum EUIs

NOTES: Assumptions based on both SAP catalogue and The Building Regulations. **01** emulates U-values below 0.3 (W/m^2K); **02** emulates a thickness of 270mm and U-value below 0.2 (W/m^2K); **03** emulates a U-value of at least 3.1 (W/m^2K) as established in The Building Regulations; **04** emulates a cylinder insulation over 55mm; **05** emulates a maximum infiltration level of 0.25 ach; **06** emulates low-energy lighting at 100%; **07** emulates low-energy for average household appliances.

6.1 Definition of Scenarios

The workflow summarised in Figure 5.2 is used to test two CO_2 reduction perspectives that are adapted from national policies [202, 75, 60, 63]:

a *perfect uptake*: where the properties in dwellings (i.e. fabric components and appliances) are upgraded to the most efficient plausible values, regardless of their current values or the associated cost and disruption; and

b *conditional uptake*: where the upgrade is limited to a fraction of the stock that is represented by a given typology, representing real life constraints such as cost, household income, disruption and the initial values of the parameters of interest (e.g. previous improvements).

The scenarios are summarised in Table 6.1, which describes the current status of the (evaluated) dwellings, as well as the variations that are then parsed to the model; with each scenario considered in isolation. A description of the scenarios and the status of the housing stock is provided below:

1 Solid wall insulation: This scenario emulates the adoption of a maximum level of insulation for each wall material. The EHS reports that more than 60% of dwellings have, on average, wall thicknesses between 220 and 260 mm. Cavity and solid walls are the main constructions found in the stock. Typically, a cavity wall is built with two layers of brick, but some internal layers are made of concrete. A solid wall is mainly built with brick or stone. Due to the

thermal efficiency provided by cavity walls, even without proper insulation, there is a clear correlation between wall construction type and built epoch. The evolution in construction types in the housing stock was highlighted in § 2.1.1. Other types of construction can also be insulated (i.e. timber-frame, pre-fabricated concrete or steel-frame). Here, thermal properties of these materials are directly affected. Usually, a wall thickness below 26 cm indicates whether a wall has cavity or not. The EHS, however, publishes a number of dwellings with cavity below this level. This can be attributable to some inaccuracy in the survey, to properly collect this information. The SAP catalogue (see Table 4.3) indicates that, the U-value for solid and cavity walls corresponds to 2.1 W/m^2K in old constructions and up to 0.3 W/m^2K in modern ones. For semi-exposed walls, for example in flats and terraced dwellings, the U-values decreases to 1.14 W/m^2K and 0.27 W/m^2K. Insulated cavity walls decrease this value to a maximum of 0.6 W/m^2K [57].

2 Loft insulation: This scenario emulates the adoption of a maximum level of insulation for each loft construction. The EHS reports that around 50% of houses have a loft insulation level between 100 and 200 mm. The Building Regulations establish that an improved level of insulation consists of thickness above 270 mm and U-value below 0.2 W/m^2K. At present, the loft U-value, as published in the SAP catalogue, is found between 0.25 and 2.3 W/m^2K. Depending on the use given to the loft, it can be insulated in the loft floor or in the roof. A loft which is not designated to be used as a room, provides more flexibility to install multiple layers of insulation, besides the feasibility to make the installation.

3 Double glazing: This scenario emulates the upgrade of glazing technologies in dwellings. The EHS reports that around 20% of dwellings include an improved window technology (i.e. double or triple glazing). The typical U-value for single-layer windows is between 4 and 4.8 W/m^2K. A double glazing window reduces this value to a range between 3.1 and 1.9 W/m^2K. Therefore, 3.1 is the minimum U-value established in Building Regulations. The resulting U-value, and, consequently, the efficiency of windows, are determined by the thermal properties of the material of the window (typically glass), the gas contained in the gap (typically air), as well as additional components in the window such as curtains, which may decrease it in up to Δ 1 W/m^2K, and a given frame configuration, which increases the value in a typical factor of 0.7. For example, average values for frames and sashes (excluding the window) are: 2 W/m^2K wood, 2.4 W/m^2K plastic and 4 W/m^2K aluminium [83, 56].

4 Cylinder insulation: This scenario emulates the improved insulation on the boiler cylinders. Usually, this improvement is accompanied with the insulation of the pipes connected to the cylinder. For dwellings with a boiler cylinder, as reported in the EHS, 74% use factory spray foam whereas the rest use a loose jacket technology. Around one-quarter of the insulated cylinders have an insulation level of 25 mm. The Building Regulations establish that a cylinder should contain more than 50 mm of insulation.

5 Draught proofing: This scenario emulates draught proofing values. More than 90% of dwellings are estimated to have an infiltration value of 0.85 ach. The Building Regulations establish a maximum infiltration level of 0.25 ach; although, there is discrepancy whether a reduction should be homogeneous for all dwelling's components (i.e. infiltration in windows, doors and structures). In addition, it is important to improve the infiltration values but maintain appropriate levels of ventilation. It is assumed that all dwellings in the stock are ventilated naturally. Consequently, besides the insulation, draught proofing should be accompanied by further analysis of the layout and its relation with the context environment.

6 Low energy lights: This scenario emulates the installation of low energy lighting bulbs in dwellings. The Building Regulations establish that at least 90% of dwellings' lighting should be efficient technology. By 2011, the EHS reported that 25% of dwellings had no evidence of low-energy lighting; whilst only 10% of dwellings fulfilled the requirement established in the regulations. Compact Fluorescent Lamps (CFLs) and Light Emitting Diodes (LEDs) are the two most common types of efficient lighting technologies. The former is, however, more efficient and more versatile to replace traditional light bulbs. Their efficiencies, as compared with traditional ones, go from 75% to 99%.

7 Improved electrical appliances: This scenario emulates improvements in the overall efficiency of household appliances in dwellings. Incidentally, appliances' efficiency is not specified in the EHS. Therefore, the minimum value provided by their energy-efficiency labels cannot be accurately populated. The energy demand associated to appliances is calculated by applying a power-law model, derived from electricity meter readings (see § 3.3.2). This estimation is complemented with the characterisation of appliances provided in the *Energy Consumption in the UK* [60] study. For this scenario, an improved efficiency is assigned to each of the appliances included in the corresponding catalogue (see § 4.1.4).

An additional scenario (# 8) is tested to understand the possible implications of different occupancy profiles and specific Energy Use Intensities (EUIs). This is made possible by assigning multiple profiles based on their corresponding probabilistic distribution, so that the values are constrained by the conditions of each archetype; these are based on average profiles defined in the *Energy Consumption in the UK* study [60]. Here, multiple profiles are defined, based on their existing probabilistic distributions, so that the values relate to each archetype, and to each element associated to an activity (see Section 4.1.4).

Heating System Configurations and Other Measures The resolution of the EHS limits the ability to evaluate heating system configuration. This is the case, in particular, of the most common configuration found in the housing stock, that comprises gas central heating and conventional boiler (see § 2.4). Therefore, the study of such a system is not included in the platform at this stage of evaluation. Its consideration requires a process of calibration supported with locale-specific data. It must be considered, nevertheless, that the contribution of heat gains to the indoor conditions is significant, and so the consideration of a disaggregated heating system configuration is in the immediate agenda of the EnHub platform. Other configurations and measures which are not considered for the scenarios involve: the layout arrangement, which is not detailed in the EHS; and a summary of outdoor elements for each dwelling, which affect heating, lighting and wind.

As described in the Flowchart 5.8, each scenario is developed using `R`. In here, a unique dataset is formed with the properties to be replicated in the construction of volumetric archetypes. This enables the possibility to detach the processes, so they can be run on more suitable hardwares and more convenient times (see Chapters 4 and 5). As a result, each scenario is developed and evaluated within a defined environment, which opens the possibility to replicate the evaluation at a different time and in a different hardware. The created volumetric archetypes (comprised of `idf` and activity profiles) may or may not be stored. They are initially stored to quickly compare the simulation performance in different environments, but the modified properties for each archetype

are embedded in the generated dataset.

The evaluation consists of the evaluation of each scenario (A1-A7 and B1-B7 in Table 6.1), plus the activity-related scenario that is run 12 times. As a result, the whole process is made in 220 hr uninterrupted. By using the HPC, and by including the queue time in the facility, the whole process is completed in 3.4 hr (see Eq. 6.1).

$$
\begin{aligned}
&1016 \cdot 30s = 8.4hr = T \\
&T \times S_{A1-7} \cdot S_{B1-7} \cdot (12 \cdot S_8) = 220hr = P \\
&\therefore P/64x = 3.4hr \quad (6.1)
\end{aligned}
$$

6.2 Evaluation of Scenarios

6.2.1 Results and Findings

Table 6.2 and Figure 6.1 summarise the outcomes from the perfect and conditional uptake scenarios whose measures are given in Table 6.1. They show that when perfect uptake is assumed, A02 and A01 have the greatest impacts, which improve fabric performance. Other important scenarios are A05, A08, and A06, partially counteracted by increased heating loads. When household constrains are considered in respect of scenarios SA02, the potential emissions reduction is reduced by around 40%. The constraints considered in measures B01-B07 are related to the disposable income of a household or to the increased costs associated with more technically challenging cases, such as housing in conservation areas, solid masonry wall in old buildings, or the insulation of loft spaces that are used for storage purposes with limited accessibility.

Table 6.2: Scenarios results: the potential of interventions

(a) Perfect Uptake

	gov-ref	BASE	A01	A02	A03	A04	A05	A06	A07	A08	A01+02	A01+03	ALL
Energy Demand [mtoe]	377	428	365	322	423	423	390	399	415	394	319	361	310
Internal Temperature $[K]$	290.5	289.1	0.26	0.12	-0.07	0.02	0.14	0.14	0.00	0.32	0.28	0.19	0.26
Emissions $[\frac{tCO_2}{hh}]$	4.6	5.7	-18%	-31%	-2%	-1%	-9%	-7%	-3%	-8%	-32%	-20%	-35%
Energy per area $[\frac{kWh}{sqm}]$	207	261	222	196	258	258	237	243	253	240	194	219	188
Energy per capita $[\frac{TWh}{p}]$	10.0	11.0	9.4	8.3	10.9	10.9	8.4	8.4	11.0	10.1	8.2	9.3	8.0
Uptakers [%]	-	-	68	79	45	65	65	85	95	-	79	68	100

(b) Conditional Uptake

	gov-ref	BASE	B01	B02	B03	B04	B05	B06	B07	B08	B01+02	B01+03	BLL
Energy Demand [mtoe]	377	428	390	365	425	425	407	412	421	462	368	391	363
Internal Temperature $[K]$	290.5	289.1	0.16	0.07	-0.04	0.01	0.08	0.08	0.00	0.18	0.15	0.11	0.14
Emissions $[\frac{tCO_2}{hh}]$	4.6	5.7	-11%	-19%	-1%	-1%	-5%	-4%	-2%	8%	-17%	-11%	-19%
Energy per area $[\frac{kWh}{sqm}]$	207	261	237	222	259	259	248	251	256	249	224	238	221
Energy per capita $[\frac{TWh}{p}]$	10.0	11.0	9.4	8.3	10.9	10.9	9.6	9.6	11.0	10.5	9.5	10.1	9.3
Uptakers [%]	-	-	36	32	24	29	52	84	94	-	36	36	100

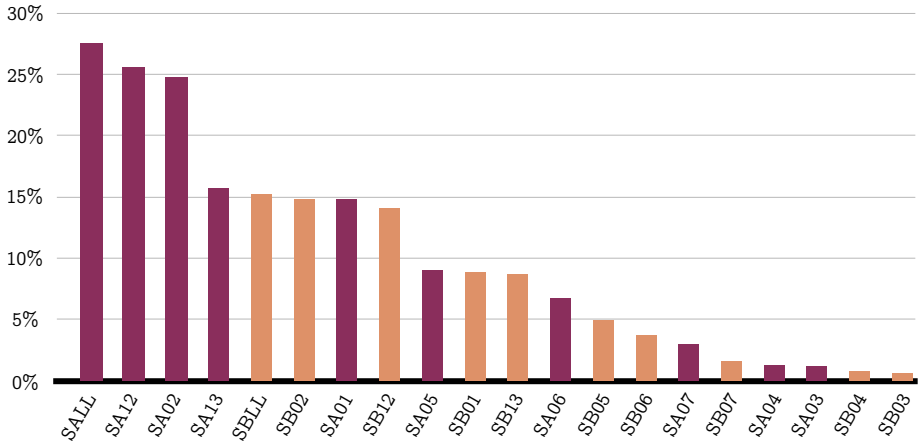
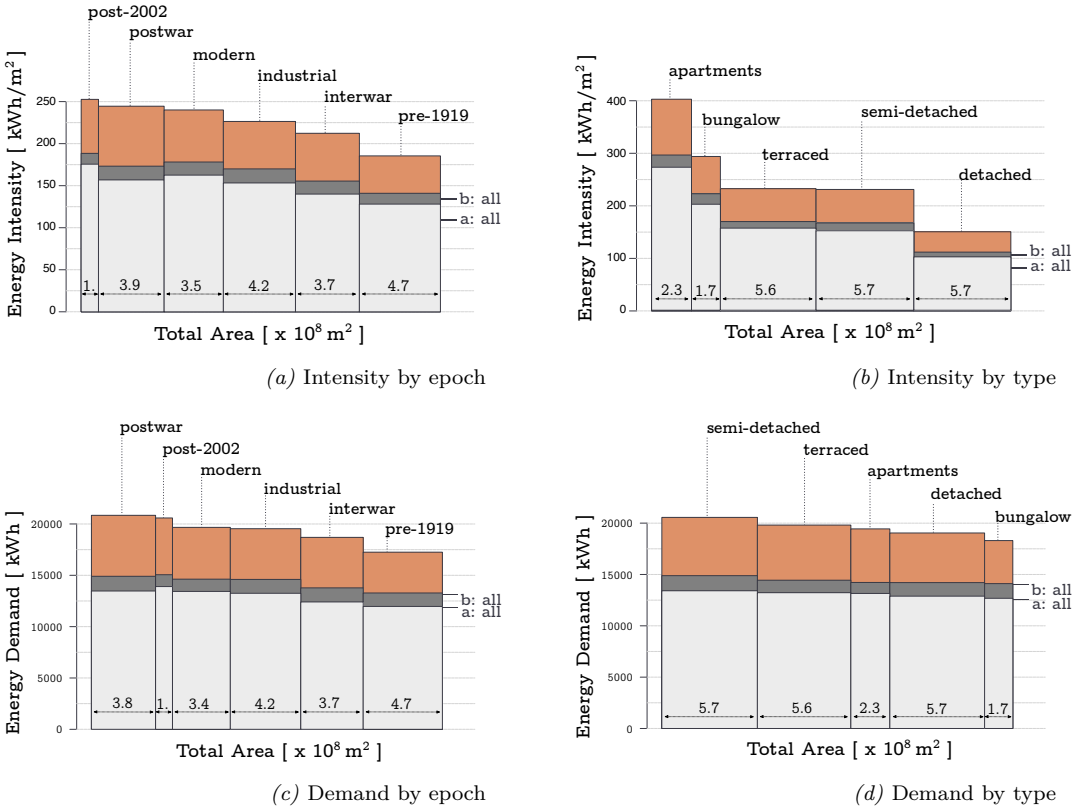


Figure 6.1: Summarised energy demand reduction
S01: wall insulation, **S02**: roof insulation, **S03**: glazing, **S04**: boiler insulation, **S05**: draught proofing, **S06**: lighting, **S07**: appliances, **S08**: occupancy



(a) Intensity by epoch

(b) Intensity by type

(c) Demand by epoch

(d) Demand by type

Figure 6.2: Reduction of energy by typologies (stock indicator)

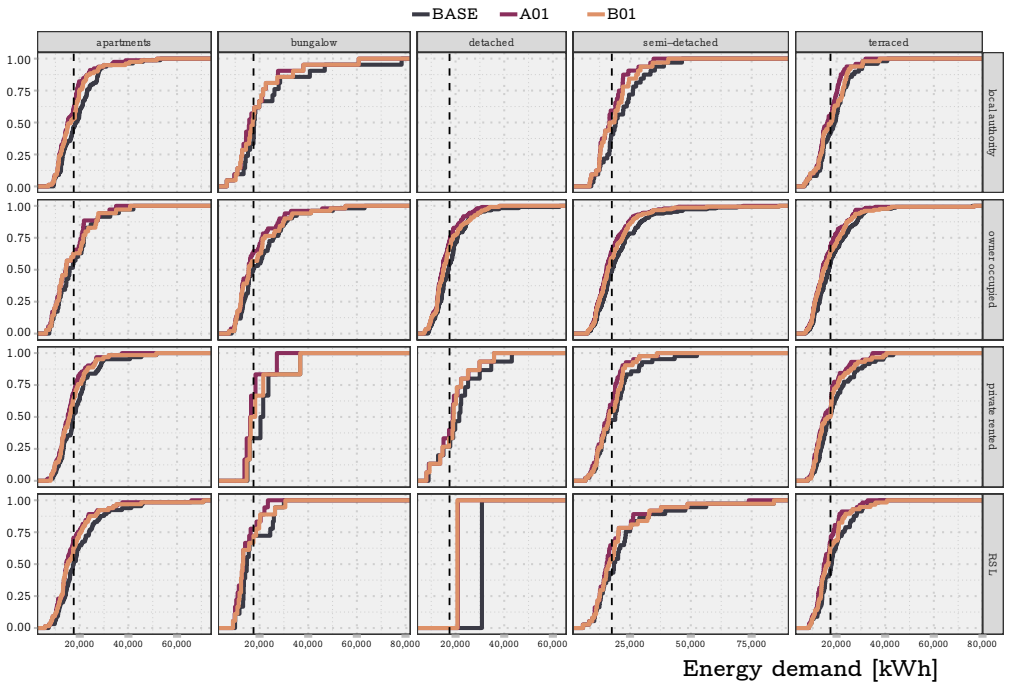


Figure 6.3: Typological comparison in energy demand by tenure

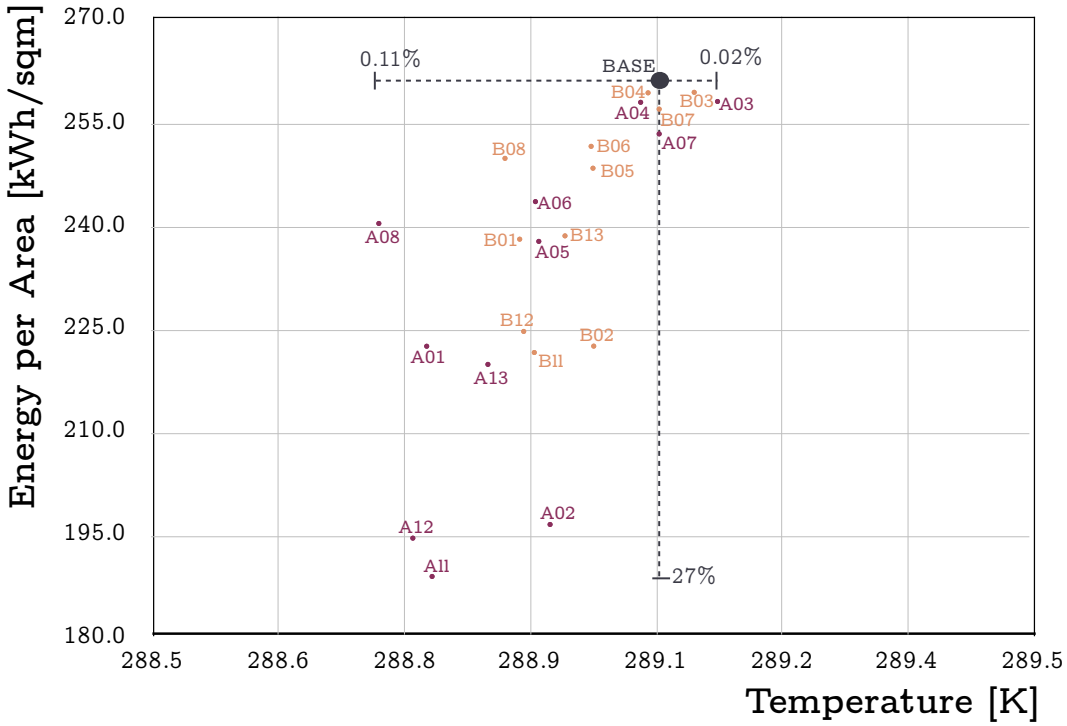


Figure 6.4: Variation in energy intensity and indoor temperature

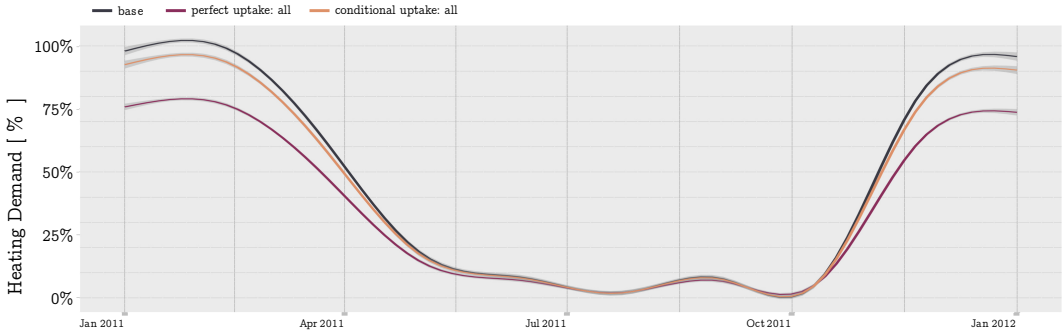


Figure 6.5: Estimated heat energy intensity

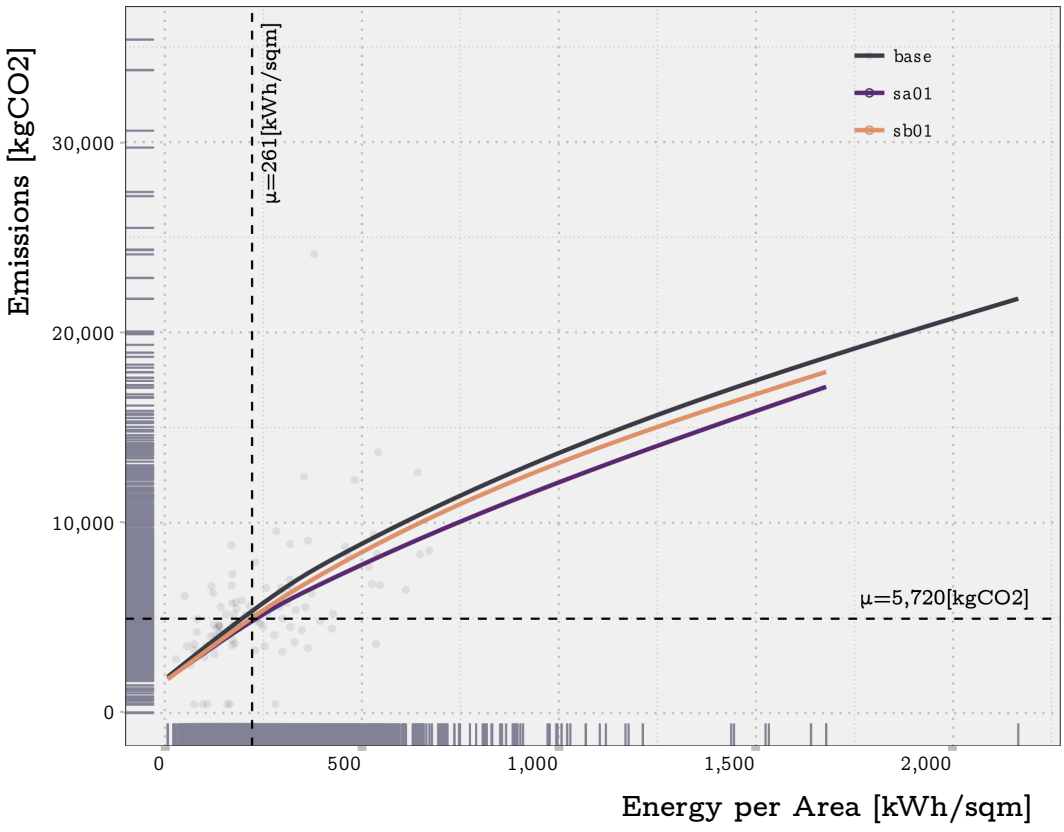


Figure 6.6: Estimated carbon emissions

The combined scenario (A01+A02) has a relatively high impact on energy demand reduction. Table 6.2 shows that their combination is roughly equivalent to all the scenarios together (All). However, it must be noted that some benefits of the individual scenarios are overlapped and consequently the outcome is not the arithmetic difference between scenarios. Likewise, the combination of A01+A03 increases the potential to reduce energy demand, but is limited by the overlap between them.

Figure 6.2 outlines the potential in energy demand reduction by typologies, but considering total usable area for each of them. Although there may be some bias caused by the EHS, and then by the reduced dataset, it is worth noting that the typical (or base) energy demand in dwellings may represent a more intense use in smaller areas.

Improvements to the fabric may be more likely to be viable for old and large dwellings because of the relatively poor performance of their fabric elements, and their relatively large surface areas. However, sometimes the spaces and zones within these type of buildings have been re-adapted and cannot be modified. For example, it is difficult to install energy efficiency measures in a loft that has been designated for storage. Figure 6.3 shows that owner-occupied dwellings are highly affected when household constraints are considered when assessing uptake. It also shows that flats are more amenable to the adoption of uptakes. On one hand, this can be attributed to their small floor areas when compared to other dwelling types, but on the other hand, this could be caused by the limited information provided by the EHS which means that energy flow pathways cannot be modelled properly in flats. Scenarios A01 and B01 follow a similar trend, where the potential for improving solid walls may be hindered by a dwelling's physical properties.

Figure 6.4 suggests that renovation measures may have a significant impact on user thermal comfort satisfaction, as the mean indoor temperature is elevated (relative to the base conditions) by around +0.8 [K]. This is due to indoor air temperatures that exceed the heating system set-point when it is on (due to excess solar gains), or are higher when the heating system is off (due to better conservation of energy). Any co-benefits to carbon emissions and comfort levels that arise from renovation measures are ignored here. However, EnHub can support the examination of these socio-technical impacts in several ways; for example, by improving usage schedules to reduce their intensity. In terms of the overall potential to reduce stock energy demand, Figures 6.5 and 6.6 exemplify the reduction in the heating and carbon intensity, which is expectedly more noticeable in the cold season, and is directly correlated with the energy bills among the population.

Scenarios A06-07 and B06-07 show that lighting and appliances are low-cost upgrades with a quick return on investment; however, they are not very effective. Increasing the efficiency of appliances is clearly a quick and easy way to reduce energy demand. Quantifying the transition rate to new appliances is possible using algorithms that account for appliance ownership and use. It is also possible to derive indoor performance metrics and advantageous to do so; for example, there may be some improvements (e.g. glazing) where the return of investment is long, but the change to the indoor environment is clear. Consequently, other metrics of satisfaction and health can benefit from the modelling of the uptake of new technologies.

6.2.2 Decarbonisation Strategies and Implications

The study of scenarios is useful because it enables the quick and convenient exploration of limits to policies or strategies. The testing of scenarios that consider the conditional uptake of measures gives an understanding of uptake constraints, although they do not shed light on probable outcomes. This would require a thorough understanding of the underlying drivers affecting the uptake of energy-related improvements by households, and the associated trade-offs between energy demand, costs, comfort, and convenience. EnHub will be continually improved to facilitate more granular analyses and to make it a centralised, open platform capable of evaluating and prioritising decarbonisation policies and strategies.

There is a need for evidence that supports the development of models that can predict investment behaviours and changes to behavioural practice, and to embed this information into platforms, such as EnHub. Then, a detailed analysis of predictions would enable the identification of types of houses, and/or the spatial concentrations of housing, that may support targeted renovation programmes.

6.3 Further Evaluation: Household Decision-Making

The study of scenarios provides a forecast, ofttimes optimistic, of how convenient is the transition towards a low-carbon pathway for the housing stock. Nevertheless, such a transition requires the stimuli of a number of decision-making triggers to predict the penetration of energy-related measures. Moreover, not only those elements emerging from each unique dwelling, but also those comprising the context of dwellings in a neighbourhood with their interactions. Such a representation enhances the ability to evaluate the aforementioned implications and strategies to improve the housing stock. In particular, by gaining insights on the trade-offs between energy demand, costs, comfort, and convenience, that affect households uniquely and boost or hinder the transition of the stock.

To incorporate investment behaviours and changes to behavioural practice, further consideration of household decision-making is required, even if this carries with specific parameters for each dwelling. To this end, EnHub performs an exploratory prototype to evaluate such uniqueness (or agency) in the stock, so that the successful adoption of policies and interventions can be better represented.

6.3.1 Triggers of Decision-Making

As it has been stated in previous Chapters, the domestic sector in the UK has contributed roughly one-fourth of national carbon emissions in recent years [58]. Hence, in the national strategy to reduce emissions by 80% according to 1990 values [183], as established by the UKs Climate Change Act, the domestic sector plays a major role and must reduce its carbon emissions at generation, distribution and usage of energy. Particularly, the sector is expected to reduce its emissions in at least 40% by 2030 and to provide 13% of heating energy demand with zero-carbon sources [37].

More than three-quarters of the energy in the domestic sector correspond to the heating demand [58], which is predominantly supplied by gas, either as a direct supply of space heating, hot water, and cooking; or for the generation of electricity in power plants [60] to satisfy the energy demand from the housing stock. However, such a demand, as well as their interventions to minimise it, are susceptible to changes over time. These changes include trends in specific technologies, shifts in policies and taxes related to dwellings, population growth, rebound effects, and some other unpredictable changes, which in sum, can diminish the accomplishment of the long-term energy reduction targets and their corresponding carbon emissions [149]. These phenomena have been roughly integrated by few HSEMs.

Moreover, the domestic sector plays a key role in achieving the energy reduction targets and involves two main types of interventions: technical and operational [144, 168]. Technical interventions improve the efficiency of systems and appliances (e.g. efficient lightings, standby savers), affect the fuel mix by opting for renewable technologies (e.g. solar panels, biogas plants, wind turbines), and alter the distribution of fuel supplies (e.g. district heating, decentralised power plants); on the other hand, operational interventions include the adoption of better practices (e.g. turning off lights when not in use, lowering the set point), and the promotion of improved conservation of energy (e.g. installing smart controllers, shifting to efficient technologies). These interventions can be highly correlated and can complement each other.

Unfortunately, each of these interventions may convey specific hurdles and barriers that diminish or completely interrupt the reductions in the energy demand. For example, the *Energy Efficiency and Value Project* [189] explains the relevance of household attributes that struggle to invest in long-term interventions, due to their high degree of uncertainty and clear benefits. These attributes significantly hinder the strategies to decarbonise the housing stock, and sometimes, play against them.

Additionally, some empirical studies have demonstrated that technical interventions do not necessarily reduce the energy demand in the long term [132], because some crucial properties in dwellings are not affected significantly; this includes the fact that only households with higher incomes are more prone to install energy-saving measures [1]. Thus, the niche for successful interventions is limited and the involved processes of decision-making become problematic to evaluate and stimulate in practice.

Yet a limited number of these processes of decision-making and interventions in the stock of houses have been developed in the past. Sousa et al. [208] in their review, find that actions taken in the UK to study (and forecast) the energy demand in the domestic sector have been limited, including programmes, assessments and HSEMs. To improve upon the processes of decision-making and the accurate description of interventions in HSEMs, it is then fundamental to be more specific on the components that are used to shape household's characteristics such as income, average ages, composition, and tenure. Most of these characteristics have been included in econometric studies [19, 20, 219], but not well linked to physical models. Despite the obvious impact of policies in households and householders [81].

Furthermore, it has been found that describing the energy flow pathways in dwellings is, despite essential, not enough to accomplish substantial changes in the domestic stock. Drivers affecting the households' decision-making, such as socio-demographics, comfort and health, become the trigger in the successful adoption of policies and measures to minimise the use of energy in

dwellings. To do so, the housing stock energy models are bound to couple these dynamic triggers involved in the operation and use of the dwelling spaces.

A valuable approach found in the literature to represent these triggers consists of using *agents* to represent individuals or households [164, 139]. These agents are created according to a set of rules defined by the context which, in turn, may provide uniqueness in each of the agents. The agents are capable, not only to fulfil the conditions required to trigger the adoption of measures, but also (and very commonly) to add a level of randomness in the responses. By doing so, the outcomes derived of this agent-based approach allow dealing with the effects of uncertainty and stochasticity, which is usually found in real world phenomena.

6.3.2 Agency in Housing Stock Energy Models

One of the soundest alternatives to enhance HSEMs with household and individual attributes is found in ABM approaches. ABMs can operate at varying levels of aggregation. Therefore, they can conceive households and individuals as agents, so that they can interact within a given environment [77]. Besides, depending on the resolution of these agents, they can predict individuals' responses based on their personal properties; or they can predict the diffusion of knowledge or choices by comparing with other agents [205, 30].

ABMs in HSEMs have been used to improve local interventions by stimulating unique responses and by including feedback. As a result, this has increased the ability to draw specific ways for developing policies [164, 231]. The responses at micro-level can shape the pattern of diffusion at macro-level [205]. Hence, the parameters describing these agents may be stimulated by specific strategies, including the way different networks can be adopted to represent real-world connectedness [154].

The study of networks and their topologies has been scarce in HSEMs, but their usability has recently received attention as a method to link geographic distribution of households and householders, and as an opportunity to explore the dynamic responses of agents [225, 135]. Its consideration facilitates a way to link richer information from population databases into the ABMs. A clear example of this approach is made in the DeCarb-ABM model [164], in which the connectedness of neighbours is used to stimulate the level of interactions. Such a connectedness incorporates an ability to stimulate a level of interaction, which can be the representation of active communication (e.g. word of mouth) or passive communication (e.g. imitation).

Agents can represent the uniqueness found in households and householders. But, despite their ability to present before our eyes emerging phenomena, they can add a substantial level of uncertainty (and consequential complexity) into the models. Yet this can be used to better understand the energy-related interventions in the stock of houses and to provide robust tools to test decarbonisation policies in the UK.

To set up the ABM environment, it is important to define the purpose of the model and the sophistication applied to the model. Generally, ABMs have no set state to achieve [167], they can be used to test responses and emerging dependencies in the defined environment. The resulting performance of agents, at their corresponding different levels of aggregation in a ABM environment, is influenced by five stimuli domains: *social, educational, regulatory, financial and technological* [188, 224]. These stimuli evolve and shape the metabolism of the environment, its

evolution and its limits. To this end, four questions must be answered to systematically develop an agent-based study:

I/ How do agents act on these stimuli?

II/ How do agents respond and what factors influence these responses (e.g. age, disposable income, education, housing condition)? How are co-benefits apportioned?

III/ How are the responses modelled?

IV/ How can these responses be modelled dynamically (i.e. iterative scheduling of policies; temporal variations in responses, when transitions occur)?

Attributes and Rules for Agents

The stimuli shape the processes of decision-making for households (and individuals). Their representation, nevertheless, requires the definition of quantifiable attributes and the assumption of a certain level of independence among them. These stimuli uniquely affect each household, depending on the understanding of the phenomenon, which is usually biased by cultural and personal issues. In order to test the processes of decision-making at the stock level, some attributes (or stimuli) are inferred for each household/dwelling archetype, and represented as state machines. These attributes are:

- a level of *payback uptake*, which is based on the potential investment with regard to the current household income;
- a level of *comfort potential*, which is based on the resulting outcome in the indoor conditions by modifying the building fabric;
- a level of *network influence*, which is derived from the number of neighbours, their distance, and their current situation in terms of improvements so that they become influencers to their peers;
- a level of *social attitude*, which is based on the household composition (size, age of head of household, scholarly) and is also translated onto a level of pro-activeness to adopt measures; and
- a level of *typology-aptitude*, which is inferred from the typology, region and tenure, as delimiters for improvements ability.
- a level of *household-uniqueness* which represents a level of stochasticity in the responses of reach household, and provides more specific triggers to represent the processes of decision-making.

A state machine defines the current state of the agent, perhaps split by attributes, and includes a decision path to remain in the current state or move to a different state. A primitive diagram of such a process is represented in Figure 6.7.

An extended approach considers states and rules. Their combination define a different state. Figure 6.8 illustrates the interaction with other agents in a given environment. Because the

agent encapsulates both states and rules, it increases its flexibility and autonomy [167]. The Figure suggests the composition of a mesh in which each node behaves uniquely, as a result of the combination of attributes, as well as the definition of rules.

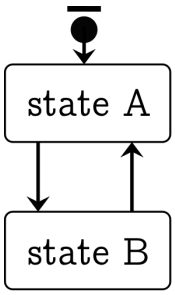


Figure 6.7: Agent state machine diagram

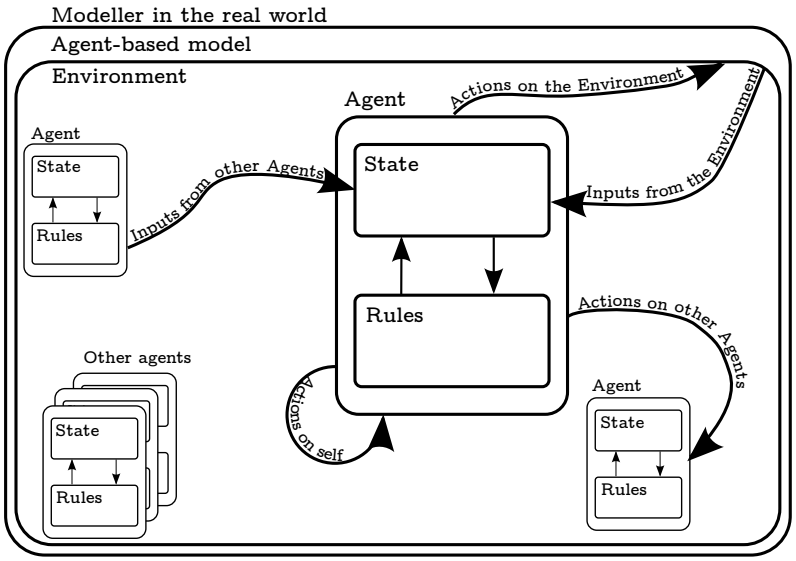


Figure 6.8: Structure of an agent-based model meta-diagram *Source:* [167]

Capturing (and processing) changes

Each of these attributes is then affected by a specific level of relevance or weights. Realistically, these weights are unique for each household, but in practice, a common weight is assigned for a specific stock of houses. In order to improve upon these weights, previous research has identified different strategies that shape the way each driver is weighted in the final outcome. Sopha et al. [205] develop a framework to allocate strategies describing household attitudes in the population; such an allocation is based on empirical data.

Here, these strategies can be steered by imitation, where the external attributes and some personal attributes have a more significant relevance than the rest; by peer-pressure, where the influence of external attributes have the greatest weight; by ability, where both payback time and potential improvements define the relevance of the attributes; or by pro-activeness, where the expectation of long-term investments and general improvements are essential, rather than

the immediate economic gains of the process. Figure 6.9 exemplifies the relative weight adopted by each strategy—based on socio-demographic profiles [205]; the abscissa exposes the relevance of each driver. For example, by considering peer pressure as adopted strategy by a household (i.e. top strategy in Figure 6.9), then the most relevant attribute becomes the network, followed by context, payback and policy; and finally with minor relevance are both fuel poverty and potential comfort.

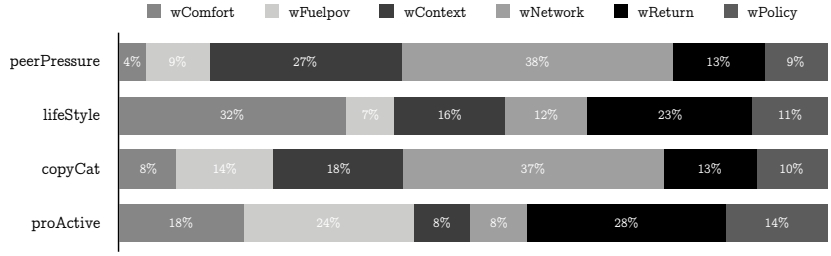


Figure 6.9: Weight of attributes for the four strategies present in the stock of dwellings. The extent to which each strategy represents a sector of the populations is based on socio-demographic information contained in the survey data (i.e. disposable income, education, age of inhabitants, fuel poverty)
Source: Adapted from [205]

From the—perhaps idealistic—regulation and planning point of view, these strategies define the way policies and measures are received. And, in consequence, they provide a measurable framework for the UK government to invest, promote, regulate or forbid energy-related measures in the housing stock.

In sum, the study of the penetration of new policies and measures involves an understanding on how the adoption is spread among the population, which is consequential of the adopted strategies and the inner responses that each individual agent adopts in the process. These responses, despite their uniqueness, can be represented under three main components: context, individual attributes and operational value. These elements have been further explored in the application of the theory of planned behaviour (TPB) approach [2, 195]; however, they tend to stimulate beliefs which requires more granular data and, therefore, they are limited when used at stock level. Nevertheless, the way in which the assignment of weights is applied to each element, can be replicated via an *ad-hoc* model, in which the parameters are adjusted for the particular case of a given dataset. Currently, not a single model exists with all the capabilities of agency included [120]. From the evaluation of HSEMs (§ 3.4), only two models include agency to represent decision-making in their forecasts: DeCarb-ABM and AHOME. The former studied different responses from occupants increase or decrease the energy use intensity in future scenarios. The latter studied the effect of carbon tax policies, applied to alter decision-making in consumers; and the diffusion of behavioural changes, related to the investment on energy-efficient technologies. Both scenarios applied to split the level of response among households.

Essentially, the TPB model can be expressed as the sum of weighted parameters (see Equation 6.2), which comprises the intention or attitude (att), perceived behavioural control (pbc) and personal norms (pn). The aforementioned six attributes are linked to these weighted parameters according to their nature. Each resulting parameter is affected by a level of importance or relevance (imp), defined by the aforementioned strategies; and scaled to a value which is employed to minimise of

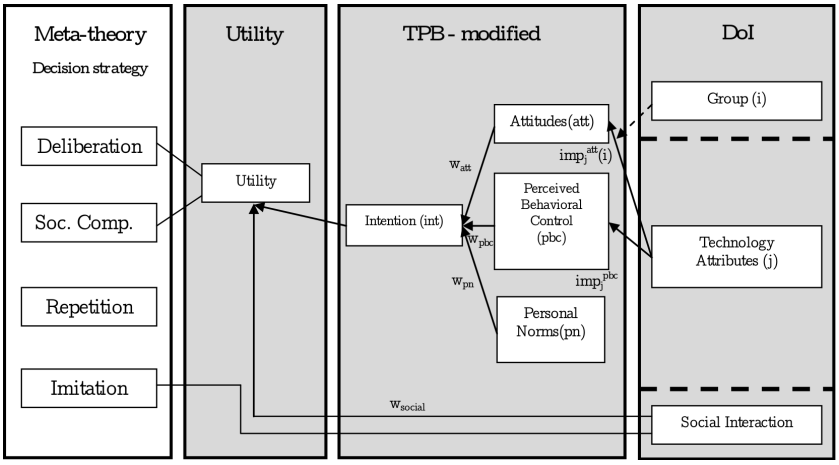


Figure 6.10: Adoption-decision model
[206]

maximise the parameter (i.e. from 0 to 1). These relationships are outlined in Equations 6.3, 6.4 and 6.5, correspondingly.

$$tpb = \omega_{att} \cdot att + \omega_{pbc} \cdot pbc + \omega_{pn} \cdot pn \tag{6.2}$$

$$att = \sum imp^{att}(i) \cdot value \tag{6.3}$$

$$pbc = \sum imp^{pbc}(i) \cdot value \tag{6.4}$$

$$pn = \sum imp^{pn}(i) \cdot value \tag{6.5}$$

The TPB model can be enhanced by including direct social influence, as a way to link the relevance of being contained within a social environment. Sopha et al. [205], in a similar application, explain that the combination of these weights can be used to measure the level of intention (to uptake), and an emulated social feedback. To clarify, Figure 6.10 outlines the composition of the TPB model for the ABM approach, expressed in Equation 6.6.

$$tpb = (\omega_{att} \cdot att + \omega_{pbc} \cdot pbc + \omega_{pn} \cdot pn)(1 - \omega_{soc}(k)) + \%_{peers} \cdot \omega_{soc}(k) \tag{6.6}$$

The scalar factors (*att*, *pbc* and *pn*) contained in the equations are normalised for the whole model (see Equation 6.7). Here, ρ represents the treated variable (s) and the reference (i).

$$value = \sum_{s}^{r} \left| \frac{\rho_s - \hat{\rho_r}}{\hat{\rho_r}} \right| \tag{6.7}$$

Where the social weight is derived from each element in the stock of dwellings and passed to the model. This weight is expressed in Equation 6.8.

$$\omega_{soc}(k) = \sum_{e=0}^{k} e \cdot \delta_e \tag{6.8}$$

The social interactions are captured as nodes in the network. Each node can be assigned with a specific value for all the elements. Equation 6.9 shows an example of this network approach that is then used as to evaluate the adjacency among the elements. These are defined by the resolution of the survey data and the available information regarding their location, for which more in § 6.3.3.

$$\mathsf{imp} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{6.9}$$

Thus, the TPB approach, describe above, can be adapted to include the relevance of context, individual attributes and operational values, and can be employed to represent households that interact with one another.

6.3.3 Network Environment

Context

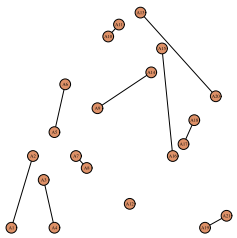
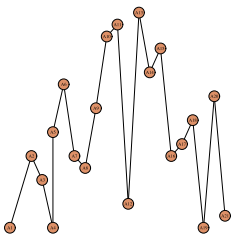
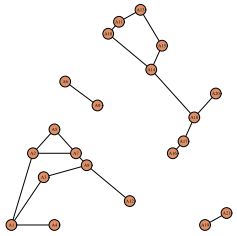
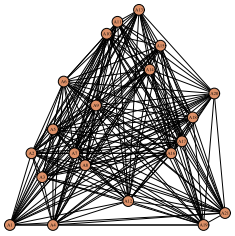
Previous literature has studied diverse topologies adopted while representing a stock of buildings (and particularly houses). Natarajan et al. [164] include the network topology to represent neighbours which influence the network composition. For example, by including this, the final diffusion of an uptake can be steered in terms of the influence of both distance and level of influence of neighbours [46], or lack of them [10]. Additionally, each individual may adopt a specific weight according to the strategy and the network topology [205]. McCullen et al. [154] explore different attributes and their influence of neighbours in their decisions.

In consequence, the network topology or configuration is essential in the way diffusions are studied

in the stock of houses. This topology is defined by the level of randomness of the connections to create the (small-world) network [225, 146]; and depends on the assumptions taken to describe the stock in the modelling approach. For instance, four topologies used to characterise the spatial network are represented in Table 6.3. The different connections among dwellings increase thus the complexity for evaluation.

In practice, each configuration corresponds to the ability to react to stimulus and appropriately change or remain in a given state. For instance, a type-(d) unit, from Table 6.3, may more susceptible to the influence of other units, yet the level of influence may be lesser than a relationship between (a)-type elements.

Table 6.3: Network topologies: different configurations based on their connectedness

Type: A Pairs and Isolated Elements	**Type: B** Sequential Single Connection
$A_1 - -A_2, A_3 - -A_4, \ldots, A_{n-1} - A_n$	$A_1 - A_2, A_2 - A_3, \ldots, A_n - A_1$
	
Type: C Distance Relationship	**Type: D** All Connected
$A_1 - A_4, A_1 - A_{10}, \ldots, A_n - A_r$	$A_1 - A_2, A_1 - A_2, \ldots, A_1 - A_n, \ldots, A_n - A_{n-1}$
	

Here, a distance based configuration is adopted because of two aspects: 1) it allows to study the aforementioned relevance of close communities to spread an improvement; and 2) household/dwelling archetypes establish a link not only by geographic position but also by other socio-demographic elements, which can be augmented with more specific attributes registered in the sources of information.

Individual Attributes

Dwellings can be categorised by the share of physical properties, tenure type, and location, despite their aforementioned uniqueness which exists in practice. Additionally, the household occupying

them add a unique use for each dwelling. This quasi-random behaviour of their occupants can be determined as collective (e.g. peer pressure influencing the installation of solar panels), circumstantial (e.g. local incentives or sales in efficient devices), biological (e.g. occupants needs according to age and health conditions), or cultural (e.g. habits and patterns). In terms of the aforementioned balance, housing models might define their level of granularity in terms of individuals, households, communities or regions, and the adopted models might include predefined responses, evolving elements, collective effects, policies adoption, technological diffusion.

Natarajan et al. [164] argue that dwellings can be considered a passive object and hence the interaction with other dwellings is defined by the assumed neighbourhood configuration. Therefore, the resulting interactions can be highly relevant or quasi-null. Furthermore, households affect such a passiveness and, combined with the other dwellings, demand specific requirements of health, comfort, leisure and activities for each of them. These requirements may be interpreted through patterns, depending on the attributes of their occupants, but also as complex interactions within the occupants and their environment.

A first step to capture this uniqueness or individual attributes consist in providing weights to the household attributes. These values are income, head of household age, household size, head of household scholarly and tenure type. As a result, these values can be used to define a function of behavioural factors (mimicking the attitudinal models) [30].

Operational Value

Based on the performance of the archetypes, described in the analysis of scenarios (§ 6.2), a projection of energy-based improvements (or decrements) for each dwelling may be inferred. This is supported by the preliminary study made in EnHub. In consequence, an operational value is given to each element depending on the differential improvement as compared with the initial state of the platform.

The operational value includes the potential payback of each uptake, and is complemented with changes in the indoor condition, which is one of the virtues of the dynamic simulation developed in EnHub. Moreover, although the indoor conditions are limited by data granularity and model sophistication, these are used to steer the emulated outcomes, provided by the dynamic simulation of volumetric archetypes.

In sum, the operational value represents an indicator (of possibility) of the benefits derived from all applied interventions to the housing stock. This value becomes a referential guide, when used by agents, to stimulate the mechanisms of decision-making and interactions.

6.3.4 Agent-based Uptakes

As aforementioned, ABMs are used due to their ability to represent stimuli and feedback on the elements of a population at a given period. However, even though they improve such responses, the need to represent the triggers of decision-making via a set of rules. These rules define the initial environment of the stock and provide the method that each agent applies to react during the process.

Setup and Calibration

EnHub defines a simulation period between 2011 and 2050, to compare with the CCC budgets (Figure 2.12). During this period, the platform run simultaneously for yearly simulations (i.e. all the evaluated households are treated independently, and only after the evaluation for one year, the resulting values from each household are used to feed the dataset) and test different strategies for specific uptakes.

First, eight different population strategy profiles are tested; this is the assumption of different profiles (of adoption) of the population. In this way, it is possible to capture a population profile with an additional bias in the reactions based on peer pressure; or a population with a balance profile among the stock. Second, different weights are tested for each of the strategies presented in Figure 6.9. This aims to capture the sensitivity of the platform to capture the strategies and attributes that each agent adopts. These attributes can be used to represent the influence of policies and measures which can be measured only by implementation. This is a step into the strategies to decarbonise either as an active or a passive member of the stock (i.e. household or regulator).

For example, the main strategy among the population may be based on imitation; however, if imitation is not representative in the internal weights to emulate its relevance, then the overall strategy may be undermined. This constraint, in fact, may be a realistic hurdle in the diffusion of strategies in the housing stock. Furthermore, the platform adopts a distance-based network topology. It is worth noting that despite the limited resolution of the units, and the obvious over assumption on the spatial configuration, this topology ([c] in Table 6.3) provides a useful environment to test the interaction of agents.

This platform provides a tool to check and verify the performance by running dynamic simulation to re-calibrate the uptakes in the stock. Yet is constrained by the resolution of the archetypes and the origin of the data sources considered in EnHub. By applying this ABM prototype, the long-term decarbonisation targets (and possible detours) can be represented in the platform.

Adoption Process

The main virtue of including a network configurations is that it enables the ability to represent the level of diffusion in the uptakes of energy-related measures. This results in the creation of (community) centres that become stronger influencers during time.

Additionally, not only the positive diffusion is captured in the process, but also the possible isolation and hence consequential restricted adoption among the agents. Depending on the strategy adopted, this adopts a very specific relevance in the process of decision-making for the agents. The adoption process is used to measure the potential of uptakes compared to the carbon budgets (see Figure 2.12), despite the evolution of households is not detailed, at present.

Energy-related measures and potential uptakes

First and foremost, the ABM prototype is based on the physical-based modelling platform EnHub. In other words, the results provided in § 6.2 define the potential to reduce energy demand

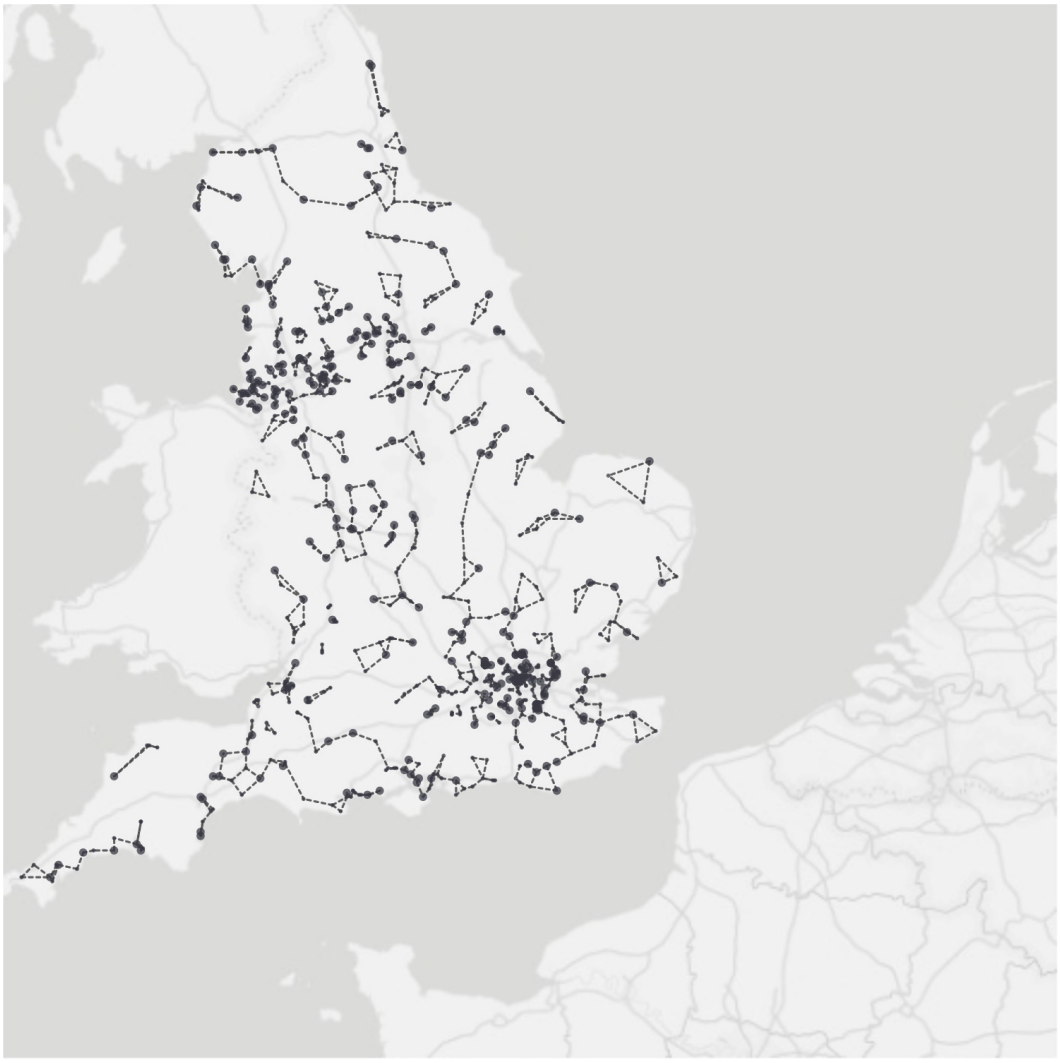


Figure 6.11: Map of households represented as agents

and associated carbon emissions. This potential adds information to each household unit, now represented as a static agent, and is used to complement the information related to the uniqueness of the units. Besides, although the platform was developed for a single year, it considered different scenarios of interventions. These interventions are included in the agent configuration as potential outcomes for each of them. The scenarios consider fabric improvements and appliances upgrade; changes in supply and distribution are not considered.

The energy-related reference provides information to improve payback and indoor improvements for each of the agents and each scenario (i.e. by including the potential of energy demand reduction and improved indoor heat gain). This reference is augmented by including other attributes that may increase the chance of uptake despite the energy-related potential. In this way, the exogenous influencers can be included in the platform.

In the long-term, it can be seen that the number of adopters increases at a faster rate until the stabilisation state is reached. However, this is also defined by the adopted level of decision-making

of each of the agents. We test different levels for the outcome of Equation 6.6. This level might represent the incentives provided by the UK government to invest in technological upgrades in the stock of houses.

Figure 6.11 represents agents in an environment limited by the reduced EHS. Given the fact that these household-agents are obtained following the sample reduction method (see Sections 5.1.1 and), the location is assigned in line with GIS information, that reveals geographical centroids of SOAs. Therefore, there is potential to scale the resolution of the agents with their corresponding location. It can be seen that each dwelling archetype/agent can be influenced by the local set of rules or by the influence of external elements (i.e. other agents and the context). Depending on the local evaluation, the response of each agent changes the status of the agent for the next stage. The retreat of technologies is not represented in the long-term nor the addition of new agents. The latter is based on the assumption that changes in the population (buildings, occupants) need to satisfy a zero-carbon demand.

In brief, a platform that increases the triggers of decision-making is outlined in this ABM prototype. The platform also accounts for the creation of communities of influence among the stock of dwellings. A set of rules are defined to capture the relevance of different attributes and strategies that can be shaped in the long term. And finally, an agent modelling approach is linked, by incorporating the potential improvements in each of the household/dwelling units, with a platform of dynamic simulation that can be updated and increased in its resolution.

6.3.5 Agent-based Outcomes, Discussion and Further Directions

Outcomes

Figure 6.12 provides an overview of the set of interventions evaluated within an ABM environment. Here, the uptake of interventions is slowed after a few iterations, which can be explained by a very narrow potential in the expected improvements. It can also be slowed by the isolation of some agents, which, by contrast, become the factor for further improvements after such period of stabilisation. The Figure shows that the uptake rate still remains positive, which can be explained by the level of feedback influence and rules representing agent interactions.

This approach is useful, albeit limited by the number of assumption in the formation of agents and rules, in that it can compare emulated attitudes of a sector of the population, and it can provide information to the extent to which a policy may find a niche for implementation. By comparing different strategies and emulated interventions (see Figure 6.12), it is possible to pay attention to the changes in the adoption of a specific subset to identify the parameters that influence such a behaviour.

Figures 6.13a and 6.13b outline the relevance of triggers regarding each of the scenarios, which may easily motivate to the adoption of a given uptake. Figure 6.13b shows that the scenario A02 gradually becomes the most successful intervention; however, this is also influenced because of the value assigned for each household once they opt for one of the interventions. Such a level may help in the motivation to improve, despite being initially an unlikely intervention. Some of the bumps illustrated in the Figures are attributable to the adoption of improvements of the nearest agents, and also to the emulated randomness affecting them. Nevertheless, all the Figures reach

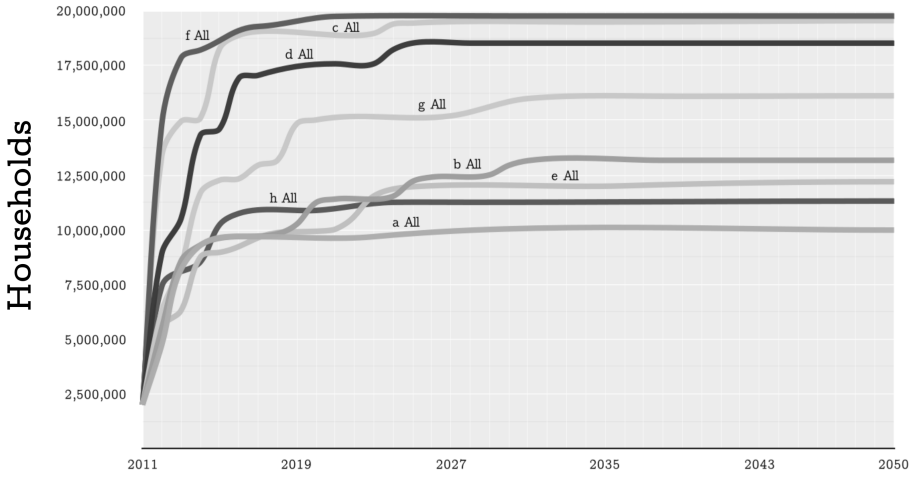


Figure 6.12: Differential adoption by strategies, in the emulated ABM interventions

to a state of balance where the triggers remain on the same levels.

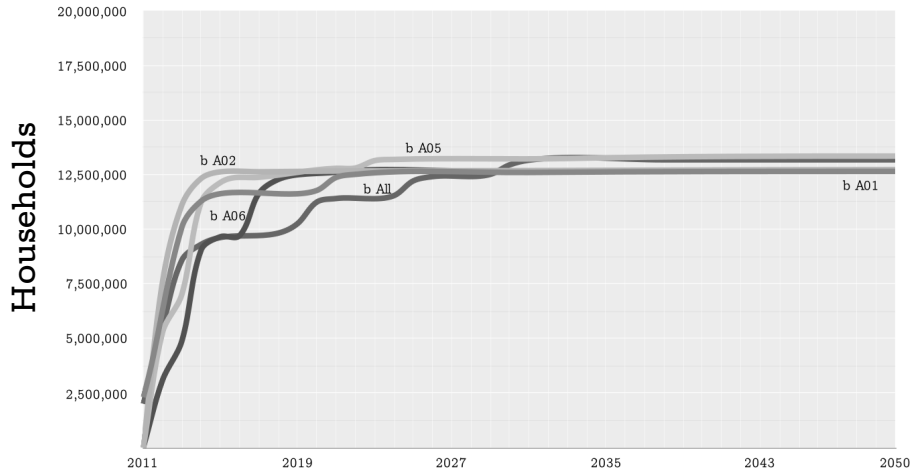
Figure 6.14 illustrates the diffusion of adoptions by considering a network environment. The map represents the allocation of agents in the synthetic sample of the English stock. The status (of the adoption) represents any improvement during the evaluation. The differences between Figured and can be subtle, but still some adoptions are reached.

Obviously, although it may be tempting to adjust the rules to increase the adoption of interventions, it is important to identify processes of stabilisation in the long-term run. In this way, it is then possible to study a subset of the population or a specific energy-related measure. Figure 6.14 summarises the successful adoption, yet limited by the aforementioned assumptions. The maps also reveal how those units that are closely located, are more likely to develop communities of intervention (and diffusion).

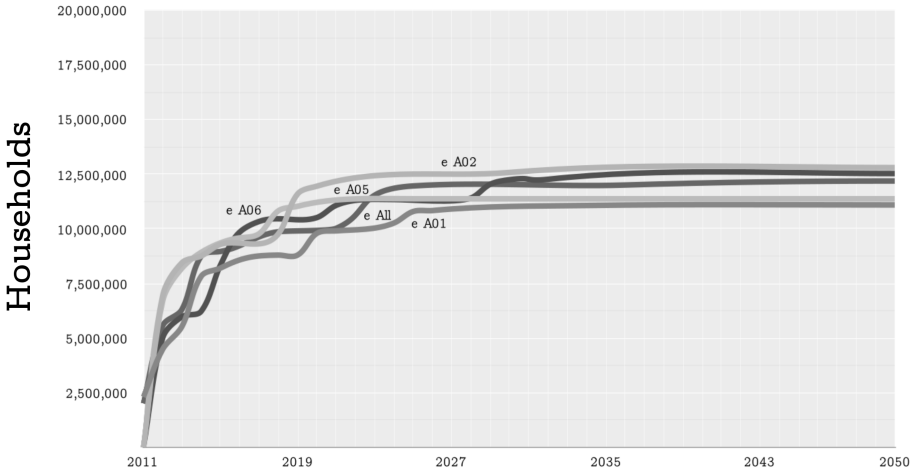
Restricted Uptakes

The UK housing stock is meant to reduce its emissions by at least 40% by 2030 and to continue with the reduction rate for the next years. It has been highlighted that the current level of emissions has been maintained at broadly the same level since the last twenty years. So any significant reduction is not likely to be straightforward.

The stimuli competing against the decarbonisation of the stock are population growth, rebound effects, fuel security and high consumer lifestyles. Population growth involves an increment of the population but with decreasing household size, which in turn represent the demand of more dwellings with base levels of energy demand. Rebound effects result in the increment of appliances ownership despite improved indoor comfort and improved efficiency in these appliances. Fuel security drives changes on the availability of fuels and hence their resulting demand in the market with the cost and taxes involved in the process. Also, high consumer lifestyles represent inadequate levels of comfort and well-being that can increase the heating energy demand in the



(a) strategy B (55%, 12%, 22%, 11%)



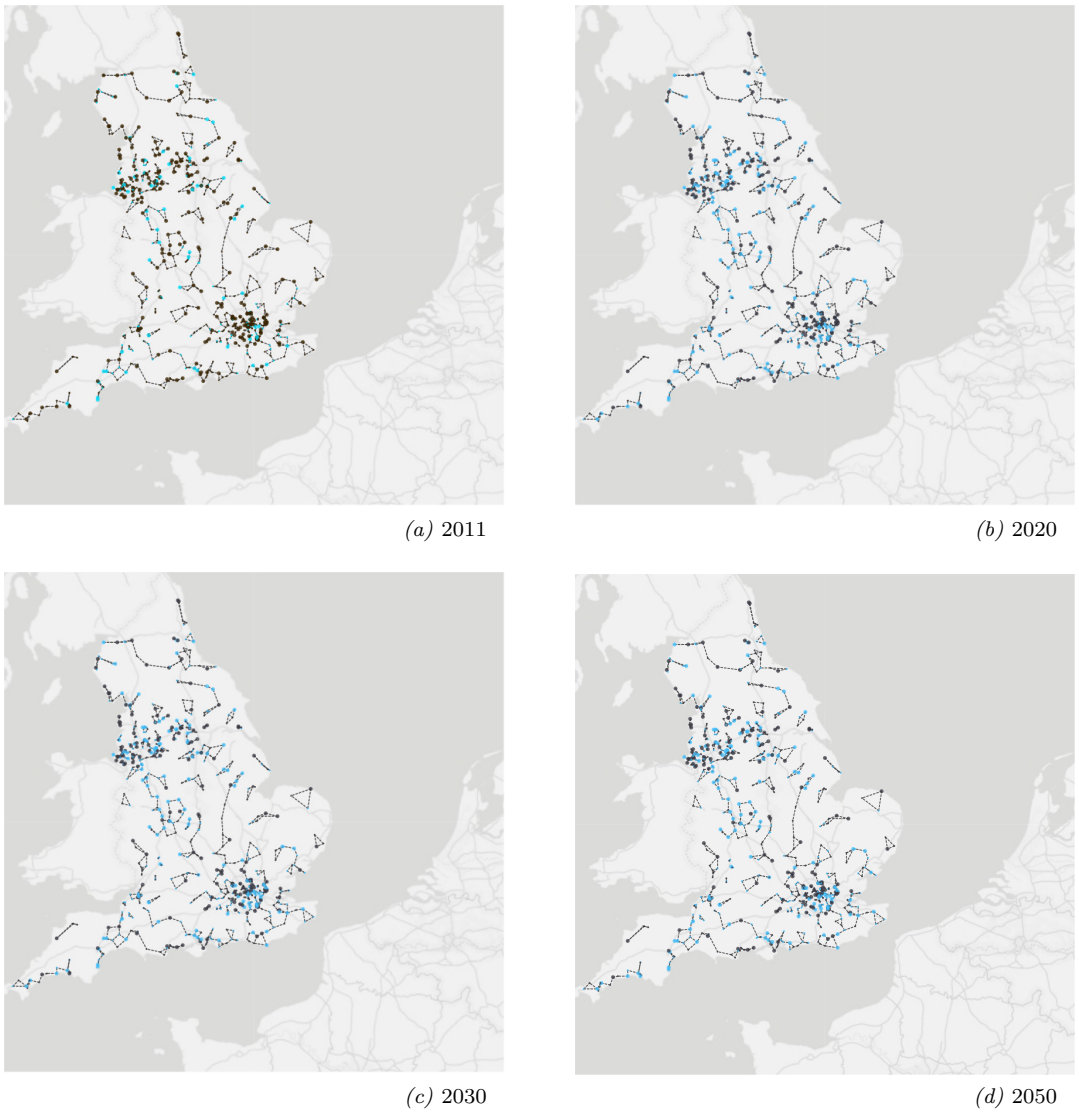
(b) strategy E (10%, 10%, 70%, 10%)

Figure 6.13: Differential adoption by weights, in the emulated ABM interventions - different proportion of the population with attributed strategies

stock.

At present, it must be considered that there are some restrictions in the platform that are not currently detailed. These restriction require a process of validation and calibration to provide with more accurate and up-to-date information. The gaps are summarised as following:

- changes in the weights of the strategies in function of time
- housing (de-)growth in a fixed or variate rate
- changes in the weights of contextual attributes and their link to capture physical composition



(a) 2011 *(b)* 2020

(c) 2030 *(d)* 2050

Figure 6.14: Uptakes in the ABM emulated interventions

attributes (so to reflect government taxes, policies and housing investments)

- changes in the weights of contextual regulations and their link to impact behaviours (so to reflect campaigns, advertisement and attitudes)

- increments on the level of randomness for each household in the weights and triggers of decision-making

Such omissions undermine the evaluation; nevertheless, by identifying significant parameters in the study, it is possible to identify where efforts should be placed for further characterisation of the stock.

Further Directions

This platform provides a visualisation tool for testing strategies and the potential uptake of energy-related interventions. The platform can represent feedback, peer pressure and strategic decisions that in turn affect the decision-making processes and result in the reduction of energy demand at the stock level. As a proof of concept, the platform can be detached from the agency process to be further evaluated in a dynamic energy simulation. The units are maintained as in the previous stage of EnHub and they can be updated according to the available information.

If the improvements on the generation of energy are summed, as well as some other changes in the distribution, such as district heating and decentralised power plants, the decarbonisation targets become more achievable. However, some of these changes depend, to a great extent, on the attitude of the government not only to boost, but also to maintain, the decarbonisation rates on the domestic sector.

6.4 Conclusion of *Analysis of Strategies to Decarbonise the British Housing Stock*

This Chapter has presented an application of the EnHub, an open-access modular platform for the dynamic simulation of the UK housing stock, with a view to efficiently evaluating the effectiveness of national decarbonisation strategies. Such an evaluation has explored the possible implications of interventions, by comparing the potential of technically based uptakes with some restricted scenarios, emulating typical real life constraints. The latter approach considers household-related information, and includes triggers for the estimation of energy performance in dwellings. As a result, the Chapter identified that contextual constraints can substantially hinder potential uptakes, the extent to which is currently poorly understood. In this way, and data permitting, the EnHub platform can enable the interplay between energy, comfort and fuel poverty to be systematically explored, as well as the effectiveness of strategies to bring about changes in investment and day-to-day behavioural practices (e.g. more controlled use of appliances, lower indoor temperatures) and the technologies destined to support these changes. This is essential in the formulation of robust decarbonisation policies and strategies.

It was identified that insulation is, technically, the most successful intervention to improve the energy performance in the stock. Well insulated dwellings reduce heat losses and, in consequence, reduce the heating demand in them. Interestingly, when household constraints are included to the scenarios emulating the ability to insulate dwellings, the potential energy demand reductions is curtailed to approximately 40%. Some of these constraints are made in function of common facts found in practice. For example, lofts are used for storage, lofts are inaccessible, walls are modified for a particular purpose, or any other scenario that makes the intervention problematic. Similarly, some restriction can be found if the household falls within the limits of the so-called *fuel poverty*—despite their accuracy and veracity–which limits the uptake by householders, even in the technical improvements are convenient.

In the analysis, the least successful scenarios involve glazing improvement and cylinder insulation. To some extent, this can be attributable to the quality of the data characterising dwellings, which do not specify dimensions for windows in detail, nor consider additional details besides the

cylinder. Besides, it was also recognised that some of these interventions slightly reduce energy demand, but considerably improve indoor conditions. Therefore, the first part of this Chapter highlighted the relevance of incorporating more specific metrics and, above all, decision-making parameters to improve the evaluation.

The second part of the Chapter improved the analysis of uptake adoptions in the housing stock, by developing an ABM prototype to represent drivers of decision-making, that directly influence the adoption rate, with the support of dynamic simulation. For this approach, despite the positive change in the use of energy in dwellings, the most obvious findings to emerge from this study are that (a) the effectiveness of uptakes to reduce energy demand in dwellings is determined by the process of household decision-making; and (b) by defining different triggers affecting the uptake rate in across the stock, it is possible to emulate the influence of exogenous attributes, which are not clearly accounted for in bottom-up engineering models. Furthermore, these attributes can eventually be enhanced with more specific models of household behaviour, lifestyle and economics, that in turn capture the uniqueness of households among the stock.

Although the findings should be interpreted with caution, the ABM approach has several strengths. For example, the provision of metrics to evaluate such a transition, increases the functionality of the platform and boost inter- and trans-disciplinary collaboration, by increasing the ability to test the performance of dwellings more specifically; further, the representation of processes of decision-making is essential to support the design, application and validation of policies, programmes and assessments, to improve the energy demand in the housing stock and, consequently, reduce the associated GHGs. This is enhanced by the open and modular platform that supports promptly edition and calibration of the algorithms quantifying the energy performance of the UK housing stock. Besides, there is agreement in that a simulation platform provides more value by evaluating the performance of buildings, by testing changes in their components, and by evaluating benefits impacting householders; rather than predicting changes in the stock without describing the process of transition.

The application of the ABM prototype reveals that a number of parameters (exogenous to those typically represented by energy models) can be included onto the core of the evaluation, to evaluate triggers of decision-making at macro-level. Although there is considerable scope for improvement, this platform represents a promising start in the representation of the evolution of the housing stock, and the evaluation of strategies to decarbonise it.

Closure Part IV

7 Summary and General Conclusions

7.1 Summary and Research Questions

Without a doubt, there is an imperative need to act towards the reduction of carbon emissions in the atmosphere. Much of these emissions, in the UK context, are associated with energy use where, in turn, a quarter belongs to end-use in the domestic sector. Of course, with this in mind, the last years have seen important advances through the development of strategies to quantify, and thus reduce, such emissions. One of the most noticeable advances, though perhaps obvious, has been the development of energy models. Here, the primary focus has been on heating energy demand to regulate indoor conditions for thermal comfort.

To represent this, early stock models evolved from models of individual dwellings based on a simplified energy balance. Indecently, this balance was mainly focused on heat transfer, and the ability to predict fuel energy demand to satisfy the heating demand. These models evolved to better represent the stock (data) and the underlying phenomena (models). In particular, stock models were able to incorporate some *aggregated* strategies to support policy analysis. The ability to predict *what-if* situations enabled the link between models, policy tools and programmes to forecast energy demand. Moreover, because the development of models was benefited by advancements in computational power, these models were able to process more information, and to do more in less time; however, the sophistication of algorithms to represent energy balance in dwellings has evolved at a different pace. For example, the ability to apply dynamic models to represent dynamic balance is not well supported, which introduces bias by considering simplified components; stock models lack in transparent modularity to validate their underlying algorithms and assumptions, which hinders their validation and functionality; policy responses cannot be tracked without updated databases, which relies on the accuracy of the corresponding snapshot of the population; the potential to efficiently use hardware acceleration is not well studied in the development of stock models; and finally, the granularity to parse survey data into models is not necessarily appropriate for the purposes of policy support and long-term strategies. To effectively characterise and evaluate housing stock performance as a whole, it is necessary to improve in all these disruptions, and to identify the most representative elements characterising both dwellings and their uses.

Research Questions

*Q1: Which **infrastructure parameters** are the most important for characterising dwelling typologies, and for modelling energy demand?* Properties of the envelope (i.e. materials, insulation, geometry), as well as attachment conditions, were identified as essential when capturing the energy flow in dwellings, either by convection or infiltration. This was initially addressed in Chapter 2, by summarising the components that together shape the UK housing stock, with respect to its typologies, and its contributions to the national energy demand. The Chapter also presented the sources of information that have been used to characterise all the existing dwellings.

*Q2: How appropriate is the resolution of **domestic stock surveys** and datasets for the description of the UK housing stock and national carbon intensity?* Not only the resolution is important to characterise the stock, but also the methods employed to deploy the corresponding sources of information. This was also covered in Chapter 2, and later further elaborated in Chapter 3. Here, the scope and quantity were analysed to determine the properties of the sources of information. In addition, Chapter 6 employed the granularity of the surveys to evaluate an agent-based prototype; however, it could be improved significantly with more appropriate resolution of these sources.

*Q3: In what ways can the datasets be **optimised to analyse strategies** to reduce energy demand, and how well is the loss of information (or propagation of uncertainties), in derived analyses, tracked?* Because there is redundancy found in the datasets, which are adapted to provide parameters to stock energy models, a process of optimisation can be considered. Chapter 5 described the process of data mining to improve the datasets. This resulted in a reduction of the dataset, which allowed the ability to perform a more efficient study of the stock, as well as to identify significant gaps in the datasets. Particularly, two methods were addressed here, and compared in terms of their versatility, when applied to analyse strategies of decarbonisation in the housing stock. The Chapter highlighted the processes applied to the two most common methods for survey data reduction (and optimisation), such as sampling and average archetypes, and the use of statistical methods to validate and calibrate such approaches; it also described the process to perform the simulation and to analyse the results systematically. Previously, Chapter 4 presented the development of an open-source platform able to link, validate and replace specific modelling modules, to be used as support of long-term strategies in the decarbonisation of the UK housing stock. The Chapter also described the metrics used to validate and use HSEMs, as well as the ability to perform parametric analyses, so that more specific studies can be performed, such as Sensitivity Analysis (SA). It was also highlighted that the integration between the sources of data and the modelling algorithms is susceptible of gaps and, hence, of improvements. This suggested further optimisation of the datasets, and the ability to evaluate both technical and operational phenomena in the housing stock. A relevant attribute of EnHub was described here, with regard to the use of optimised hardware to perform parametric analyses. Chapter 4 concluded by identifying the most notable attributes of the domestic sector, and by highlighting that, to improve the use of sources of information, these need to be calibrated, validated and available so that models can be benefited from them.

*Q4: Which are the main elements for the validation of **tools to estimate energy performance** and associated carbon emissions, and how can these be improved?* The potential of a centralised platform to provide, use and update information, was found as a cornerstone towards the validation, and hence robustness, of an evolution tool. Obviously, there is an imperative need to increase the resolution of survey data, in particular to provide information regarding the

physical variables and contextual factors, which was addressed in Chapters 3 and 4. Here were reviewed the modelling strategies that have been used to evaluate the UK housing stock. First, a theoretical revision was made regarding models' rationale, with a view to identifying the principal attributes derived form it. Then, the Chapter reviewed those models developed in the UK since the 80s, when the strategies to improve (thermal) performance in dwellings became part of the national agenda. The Chapter also identified and corroborated the most notable gaps and virtues in such modelling tools. This was useful to present a framework for improvements and, more importantly, for the development of a comprehensive evaluation platform.

Q5: How would it be possible to represent policy outcomes, as well as energy conservation measures, through different abstractions regarding ***curtailment behaviours*** *(better practices) and* ***investment behaviours*** *(technological improvements)?* A sub-routine which captures variation in behaviours associated was addressed in Chapter 6. This open the possibility to include more sophisticated algorithm to represent household conditions, which can also reflect the processes of decision-making. In sum, this Chapter presented a series of strategies to emulate behaviours, and account for constraints—most of them based on household data—that influence the likelihood of effective energy-related measures.

Q6: How should the effectiveness of policies and ***strategies****, addressing these behaviours, be validated?* In addition to more up-to-date data, the relevance of providing more elements that affect performance was highlighted in Chapter 6. The second part of the Chapter described an ABM prototype to evaluate probable impacts of decarbonisation policy measures, through the modelling of different agent personae, weighted attributes influencing their behaviours, and interconnections between them. The validation of strategies is a process that requires constant revision and edition. Such a process is facilitated by the Housing Stock Energy Hub (EnHub).

7.1.1 Results and Findings

Chapter 3 identified that affecting the physical components and the technology of heating systems is not enough for a significant reduction in the energy demand, nor are changes in drivers affecting the use of energy, which are related to individual and collective behaviours, as well as their associated stimuli. The rationale employed in the development of EnHub made possible to represent both elements via a systematic evaluation. For example, by incorporating more valuable information representing the stock, such as the consideration of attachments via contiguous volumetric cuboids, the consideration of activity profiles for dynamic simulation, and the consideration of constraints affecting the processes of decision-making.

The scenarios compare the technical potential with a more restricted (and stratified) one. Here, the outcome of the evaluation states that carbon emissions are highly sensitive to housing insulation levels (between 14% and 31%). However, the stratified approach reveals that this improvement is not necessarily the most likely scenario to be applied at a national scale, due to its cost, particularly in relation to hard-to-treat properties. This is the case, for example, of old houses that may have potential of being intervened but the implementation is restricted. Conversely, reducing infiltration and draughts appears to be a promising strategy, despite a relatively limited reduction to emissions (9%). Initial investigations suggest that occupants' behaviours may also have a significant influence on reducing future emissions. A variation of $\pm$ 8% was identified in the energy demand due to occupancy-related activities. In this respect, although the overall

objective was to evaluate the decrease in energy demand, it is important to consider that usage patterns may result in an increment in the energy demand. Nevertheless, this is relevant to provide certainty in the development of housing stock energy models.

In sum, it was found that, inevitably, the perfect uptake is convenient for the trade-off between energy demand and cost, because is based on the technical potential. However, it is consequently affected when some conditions (on both households and dwellings) are included. This suggests that the more detailed conditions, particularly on the households' characteristics, the higher constrained adoption of improvements and, in consequence, the worst case scenario; however, on the other hand, these constraints aim to further develop a tangible and amenable platform for testing energy-related measures and decarbonisation strategies in the housing stock. This results in a more robust tool for scenario analysis. An attempt to represent this was achieved in the application of the ABM evaluation, where the ability of householder to stimulate the adoption of improvements was captured, despite the modest sophistication of the model.

To end, the constraints incorporated into the model affect the adoption rate in about 40%. Simultaneously, the consideration of other households related one each other can increase the adoption rate. The evolution of the EnHub platform moves from the study of technical potential to constrained uptakes, and then from the possible scenarios to the probable scenarios based on different level of aggregating of data. Therefore, the characterisation of households and householders, in addition to significant improvements to HSEMs, and their algorithms, is required to improve the evaluation of the housing stock. This further development introduces considerable complexity to the platform. Hence, the relevance of an open platform where it is possible to track the responses at different levels. These responses can also be calibrated with external models, usually top-down approaches, so the agents can be stimulated in more detail. Yet the transition from 124 $\mathrm{mtCO_{2eq}}$ to minimal levels by 2050 is still precipitous, as stated by Oreszczyn et al.[173], but the inclusion of detailed and external stimulus allow us to test, for example, the rebound effects which have been the most noticeable hurdles towards the successful decarbonisation of the housing stock in the UK, and the likely impact of specific policies (e.g. in terms of the reduction in emissions arising from specific taxation and subsidy packages; from technological innovation; from educational campaigns).

7.1.2 Description of the Possible and Probable Scenarios

This Thesis has elucidated that a number of factors are misrepresented while analysing energy flow performance in dwellings, which consequently propagates uncertainties in the extrapolation to a stock of dwellings. Some of these factors are limited due to the granularity of the sources of information, and some due to the sophistication of the modelling approaches. However, it has been seen that a balance between granularity and sophistication provides a reasonable understanding of the implications that exist in practice. Therefore, improvements are needed but not all of them are fully achievable. Table 7.1 summarises crucial improvements that would benefit HSEMs and distinguishes between possible and probable improvements.

Although none of the improvements is straightforward, possible ones are mainly benefited by advancements in computational power and big data analysis. However, as mentioned in § 3.5, not only by augmenting computational capabilities are the models more functional, but also by altering the way data is shared and used. Chapters 4 and 5 improved upon this, yet with significant

Table 7.1: Improvements for Housing Stock Energy Models

(possible)
· scenarios with more granular data (layout and geometry data)
· scenarios with occupant's data (profand socio-demographics)
· changes in household population (household sincome band, employment)
· household appliances' specification (ownership, power demand, efficiency labels)
· disaggregated household fuel expenditure (seasonal readings, smart meters)
(probable)
· maintenance and retrofit works costs (repairs and alterations)
· occupants feedback (comfort levels, adaptive preferences)
· occupants ageing (changes in the household composition, changes in their activities)
· characterisation of internal layout (internal walls, furniture)
· characterisation of network heating system (pipe network, chimneys and wall air-vents)

data gaps. Possible improvements are also facilitated by changes in the regulation of data, which might boost the enhancement of highly disaggregated data at stock level. Alternatively, probable improvements require significant improvements in the modelling approaches, but also on the regulation. In this way, the evolution of the stock could be re-captured and validated in future years.

7.2 Discussions and Reflections

Regardless of necessary thoughts about quality of life and discarding a few exceptions, it can be said that every one of us has a place to dwell. And despite the similarities that can be shared amongst dwellings in the stock, each place is unique in its kind and generates specific demands and specific transformations of energy into the environment, either as pollutants or as direct heat. These demands are caused by household's needs in relation to health, wellness, activities and use of technology. Because of this, the development of HSEMs focusses on the ability to fully understand the energy flow pathways in dwellings; to capture factors influencing these flows; and to explain the emerging connections from such needs.

Particularly, the explanation of connections increases the reporting of errors. Nevertheless, this results convenient to improve both modelling approaches and data sources, which, as a result, increases the level of trustworthiness in HSEMs to test strategies and changes across the stock, and highlights the areas where datasets should capture more information.

HSEMs are fed by multiple datasets. In England, the main dataset is the EHS; in Scotland the SHCS, in Wales the *Living in Wales* and in Northern Ireland the NIHCS. All these datasets characterise households and dwellings following similar methods for designing, surveying, weighting, and publishing information. However, they struggle when selecting an optimal sample of the population. Because the basis to capture heterogeneous subsets of the population might be convenient enough for representing some household attributes, but not all dwelling components that would benefit HSEMs. Besides, they were mainly developed to capture information in terms of socio-demographics, and not exclusively to evaluate energy performance. To date, these datasets have evolved and provide with more information to model energy flow in dwellings, but there still are significant gaps for this endeavour. For example, there is a considerable gap in the

information regarding daily energy use, indoor layout, glazing areas and air leakage.

This Thesis, particularly in Chapter 3, presented three main (and complementary) groups essential for representing drivers influencing the energy performance in dwellings: physicality, appliances and behaviours. The Chapter also identified that, despite the minor contribution to the energy demand by household appliances, the activities associated to their use are highly influential to determine the energy performance in dwellings. Such activities are shaped by household context and individual behaviours.

When analysing improvements for the housing stock, changes on dwellings' fabric are the typical—and perhaps obvious—parameters to be tested. However, some of these changes are constrained by real condition of dwellings and the likelihood (of householders) to invest on such modifications. Householders are influenced by cost-benefits in fuel bills, comfort and health. They decide whether to invest on short- or long-term improvements; and whether to improve on technology or on better practices. It is therefore essential to augment the representation of householder's conditions to overcome omissions and over-assumptions in HSEMs, and, as a result, to increase certainty in their outcomes and functionalities. In brief, even the most notorious improvements can be unsuccessful if there is lack in the ability to evaluate the processes of decision-making, and the adoption of such changes in a more realistic way.

Interestingly, not only the decision-making towards the (positive) adoption of measures is essential to improve HSEMs, but also the (negative) adoption of technology is susceptible of further representation. The latter is related to the rebound effect that occurs when the reduction of energy demand facilitates the acquisition of more devices; it is also related to individual behaviours that intensify the heating energy demand.

In terms of behaviours, it results computationally expensive to represent highly disaggregated individual activities among the housing stock. Nevertheless, a more succinct process of calibration of household profiles might be achieved. For example, a multi-layered characterisation of housing stock behaviours might involve: a) a first layer, employing the same sample units of the surveys, that quantifies household expenditure and disaggregated heating system; b) a second layer, being a typological sub-sample of the survey to characterise household appliances, including their specifications; and c) a third layer, being a sub-sample based on surveyors' expertise, to quantify disaggregated indoor activities. Although thee process have been developed in the past, they have been tackled in isolation. Their consideration within the same survey would benefit the calibration and functionality provided to HSEM.

Appliances also suffer from over-assumptions and their representation in surveys is not balanced. For example, the characterisation of household appliances in five groups (wet, cold, brown, lighting, cooking) might be used to enhance surveys, as following: 1) despite based on surveyors' expertise to fill information on these groups, a process of calibration with energy bills might be used to increase the accuracy of datasets. 2) their representation might be benefited from smart meters and more intense in-situ calibration 3) splitting appliances into groups might benefit surveys to incorporate evolution in the ownership of appliances.

What else?

In addition to the improvements on modelling approaches and data sources, the strategies to decarbonise the UK housing stock require significant deployments on three perspectives: fuel pureness (mix), structure supply (decentralisation), and fuel use (efficiency).

In this regard, it should be noted that this Thesis is not questioning the (historical) dependency on fossil fuels, nor the cost associated with their demand. Indeed, though the fuel mix has changed over the last forty years, this change has not significantly altered the proportion of fossil fuels in the national fuel mix. For example, the generation of electricity made a substantial change by initially replacing coal with gas in the 1990s. Nowadays, gas and renewable sources have completely replaced the use of coal in power plants. The carbon emissions associated to gas, as compared with coal, are less intense, yet there is carbon in them. Realistically, the national fuel mix is quite unlikely to nullify the use of gas in the short-term, which currently constitutes 80% of the supplied fuel for the domestic sector.

In addition to real changes in the national fuel mix, the decentralisation of the fuel supply has resulted advantageous. This has boosted the development of low-carbon power plants, the installation of solar panels, and the investment on district heating networks. Unfortunately, the cost-benefit for such changes has received a limited amount of incentives (e.g. early Feed-In Tariffs (FIT) and sporadic photovoltaic projects) and so their growth has been slow and partially successful.

In terms of the efficiency of household appliances and systems, not only direct rebound effects need to be considered, but also further effects related to the evolution of appliances. For example, the market penetration is not willing to reduce its actual rate which, in turn, actively increases the ownership of appliances in the domestic sector. Accordingly, technical improvements are highly counteracted by adopted lifestyles and levels of comfort normality. The problem of an increase in residential end-use energy intensity can be illustrated with the case of TVs. Forty years ago, dwellings had one low-efficient TV (120W), which demanded more electricity, but also provided more heat to the indoor air. Today, in average, households own more than one TV or one larger screen. These devices have considerably improved in their efficiencies, yet they now deliver less heat. Their modelling needs to address all these changes. Their ownership has evolved in a way that, now, households have decentralised the use of a TV, adding other devices, such as computers, tablets and mobiles, to satisfy similar needs and to complement some others. Changes on both the devices and their use can be highly diverse and very quick, which carries on a high level of uncertainty for their energy impact estimations. Here the potential to incorporate demand-response strategies to improve autonomy in the stock energy supply opens a further avenue for future work, in order to better represent household appliances.

It is therefore important to quantify the energy usage for household appliances, and to capture their evolution in the terms of their ownership and efficiency. This will, undoubtedly, benefit the reduction of energy demand by employing more accurate supporting data. However, in this respect, it is worthy of mention that not only householders must shape the energy demand reduction, but also the government must play an active role in the adoption of these measures, as these reductions have a sound impact in the alleviation of peak demands in power plants. Such improvements in the operations are as important as the reductions in the end-use sector.

More incentives, subsides and regulations are required to improve energy demand and, consequently, to decarbonise the housing stock. First and foremost, It must be considered at all times that the energy-related problem of housing (i.e. fuel poverty, health issues and discomfort) cannot be attributed to householders. They use dwellings to satisfy basic activities, to shelter, to find comfort, to gather, to rest and to entertain. The strategies to improve the housing stock cannot be supported by timid regulations. Yet, all these decisions need to be reviewed, questioned and improved constantly, since only this way the long term strategy of decarbonisation will be successful.

7.3 Contribution to Knowledge

This Thesis engages with different methods and techniques to represent energy usage in the UK housing stock. It also aims at identifying the analysis of data sources employed to achieve this. On the one hand, it has been pointed out the importance of a comprehensive review on the elements that make up HSEMs, but on the other hand, it has been elaborated on the combination of many attributes in previous studies to consolidate a national decarbonisation platform. This means that a theoretical revision is extended, but complemented with the application of further evaluation processes, not only highlighting the areas that require more attention, but also indicating those areas of opportunity where multi-participative and active collaboration is suitable. For this, significant changes on the ways the platform is constructed and distributed, are required. The contributions to knowledge affect different components and branches in the simulation platform, where it has been identified that major attention is needed. They also exemplify the ability to improve the study of analysis to decarbonise the UK housing stock. The main contributions of this work are summarised in the following list:

i/ harmonisation of datasets describing c14,000 housing and household archetypes. This is also used to indicate how survey data might be updated and complemented;

ii/ development of programming functions to convert raw data from survey data (e.g. EHS) to ready-to-use parameters;

iii/ sensitivity analysis to reduce the dimensions of this stock of archetypes, and to indicate where special attention should be given to the development of surveys and evaluation techniques;

iv/ recontextualisation of a generic representation of dwelling types for evaluating attachment conditions, and for scaling better characterised elements of the stock;

v/ preparation of input files to the dynamic simulation program *EnergyPlus* describing each retained archetypal house;

vi/ inclusion of household-related profiles to improve the consideration of socio-demographics in a dynamic simulation environment;

vii/ development of a modular platform to centralise Housing Stock Energy Models (HSEMs). This is to test a set of tasks on the same platform (sensitivity, archetype modelling, stock analysis, agency);

viii/ exploration of different methods for optimising the use of a fit-for-purpose dataset in iterative simulations;

ix/ scripts to control the execution of these simulations on a High Performance Computing (HPC); and to extract and analyse the simulation results;

x/ scripts to modify initial models and to repeat simulations to analyse specific decarbonisation scenarios (regarding occupant's behaviours, building fabric and building systems). This is to test the uppermost potential of specific interventions;

xi/ development and testing of prototype Agent Based Modelling (ABM) framework to predict the probable uptake of interventions and their impacts on future carbon emissions, and to test the likely impact of specific interventions.

7.4 Directions for Future Research

Because the platform is scalable and adaptable, there is significant potential to adapt the platform to different stocks of houses. Not only to study more detailed regions but also to study the housing stock in different countries. Additionally, there are significant areas for improvement, especially those related to feedbacks, either from households or from individuals.

The following list summarises the main directions for future research:

- verification, validation and calibration of models with empirical data, to identify areas for improvement of both algorithms and sources of information;
- sensitivity analysis on weather profiles, to clearly identify the influence of the assumption related to the environment (i.e. heat island effects, solar radiation, overheating risks);
- stochastic behaviour modelling, including demand-response models, to identify the relevance of more disaggregated information regarding household appliances and energy-systems;
- further development of a modelling platform capable to capture investment behaviours, and behavioural practices in response to policy measures;
- calibration of models with more specific geographic information, so that this can be linked with other surveys and cadastral databases;
- simulation of spatio-temporal conditions and representation of how these evolve in response to changes in behaviour;
- inclusion of decentralised fuel supply, particularly to identify the relevance of renewable and the impact on the national energy supply;
- inclusion of more granular data regarding health and comfort, to also incorporate transdisciplinary collaboration with more specific tools regarding inhabitants, and to provide with more tools to support uptakes;
- projections of future population (fragmented households, more households), fuel mix (renewable and decentralised systems) and climate change (heating demand, health issues);

- integration of an ABM framework to study emulated impact of policies and highly specific responses of different subsets of the population;
- dynamic updates as new data becomes available (flexible information, cloud databases, centralised standards and efficiency labels); and
- application of EnHub to a different housing stock.

A Standards and Regulations

The following list is mainly derived from the *CIBSE Guide A* [83] cross-references

Table A.1: List of Standards by Module

Building: *for opaque elements*	
BS EN 12664 (2001)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Dry and moist products of low and medium thermal resistance
BS EN 12667 (2000)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Products of high and medium thermal resistance
BS EN 12939 (2001)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Thick products of high and medium thermal resistance
BS EN 1745 (2002)	Masonry and masonry materials. Methods for determining design thermal values
BS EN 1745 (2012)	Masonry and masonry materials. Methods for determining thermal properties
BS EN ISO 10456 (2000)	Building materials and products – Procedures for determining declared and design thermal values
BS EN ISO 10456 (2007)	Building materials and products – Hygrothermal properties – Tabulated design values and procedures for determining declared and design thermal values
BS EN ISO 12524 (2000)	Building materials and products. Hygrothermal properties. Tabulated design values
BS EN ISO 13789 (1999)	Thermal performance of buildings - Transmission heat loss coefficient. Calculation method
BS EN ISO 13789 (2007)	Thermal performance of buildings – Transmission and ventilation heat transfer coefficients – Calculation method
BS EN ISO 6946 (1997)	Building components and building elements – Thermal resistance and thermal transmittance – Calculation method
BS EN ISO 6946 (2007)	Building components and building elements – Thermal resistance and thermal transmittance – Calculation method
BS EN ISO 7345 (1996)	Thermal insulation. Physical quantities and definitions
ISO 8302 (1991)	Thermal insulation. Determination of steady-state thermal resistance and related properties. Guarded hot plate apparatus
ISO 9164 (1989)	Thermal insulation - Calculation of space heating requirements for residential buildings
Building: *for windows and doors*	
BS EN ISO 12567-1 (2000)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method
BS EN ISO 12567-1 (2010)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method
BS EN ISO 12567-2 (2005)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method – Roof windows and other projecting windows
BS EN 673 (1998)	Glass in building. Determination of thermal transmittance (U-value). Calculation method

Continued on next page

Table A.1 – continued from previous page

BS EN 673 (2011)	Glass in building. Determination of thermal transmittance (U-value). Calculation method
BS EN ISO 10077-1 (2000)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Simplified method
BS EN ISO 10077-1 (2006)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. General
BS EN ISO 10077-2 (2003)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Numerical method for frames
BS EN ISO 10077-2 (2012)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Numerical method for frames
Building: *for grounds*	
BS EN ISO 13370 (1998)	Thermal performance of buildings – Heat transfer via the ground – Calculation methods
BS EN ISO 13370 (2007)	Thermal performance of buildings – Heat transfer via the ground – Calculation methods
Loads: *internal gains*	
BS EN 1745 (2002)	Masonry and masonry materials. Methods for determining design thermal values
BS EN 1745 (2012)	Masonry and masonry materials. Methods for determining thermal properties
BS EN 28996 (1994)	Ergonomics. Determination of metabolic heat production
BS EN ISO 7730 (1995)	Moderate thermal environments. Determination of the PMV and PPD indices and specification of the conditions for thermal comfort
BS EN ISO 8990 (1996)	Thermal insulation. Determination of steady-state thermal transmission properties. Calibrated and guarded hot box
BS EN ISO 9346 (1996)	Thermal insulation. Mass transfer. Physical quantities and definitions
BS EN ISO 13370 (1998)	Thermal performance of buildings – Heat transfer via the ground – Calculation methods
BS EN ISO 12572 (2001)	Hygrothermal performance of building materials and products. Determination of water vapour transmission properties
BS EN ISO 9920 (2003)	Ergonomics of the thermal environment. Estimation of the thermal insulation and evaporative resistance of a clothing ensemble
BS EN ISO 8996 (2004)	Ergonomics of the thermal environment — Determination of metabolic rate
BS EN ISO 7730 (2005)	Ergonomics of the thermal environment. Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria
BS EN ISO 9346 (2007)	Hygrothermal performance of buildings and building materials — Physical quantities for mass transfer — Vocabulary
BS EN ISO 9920 (2007)	Ergonomics of the thermal environment. Estimation of thermal insulation and water vapour resistance of a clothing ensemble
BS EN ISO 9920 (2009)	Ergonomics of the thermal environment. Estimation of thermal insulation and water vapour resistance of a clothing ensemble
Loads: *external gains*	
BS EN 1745 (2002)	Masonry and masonry materials. Methods for determining design thermal values
BS EN 1745 (2012)	Masonry and masonry materials. Methods for determining thermal properties
ISO 9164 (1989)	Thermal insulation - Calculation of space heating requirements for residential buildings
ISO 8302 (1991)	Thermal insulation. Determination of steady-state thermal resistance and related properties. Guarded hot plate apparatus
BS EN ISO 10211-1 (1996)	Thermal bridges in building construction. Calculation of heat flows and surface temperatures
BS EN ISO 6946 (1997)	Building components and building elements – Thermal resistance and thermal transmittance – Calculation method
BS EN ISO 13786 (1999)	Thermal performance of building components. Dynamic thermal characteristics. Calculation method
BS EN ISO 13789 (1999)	Thermal performance of buildings - Transmission heat loss coefficient. Calculation method
BS EN 12667 (2000)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Products of high and medium thermal resistance
BS EN ISO 10077-1 (2000)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Simplified method
BS EN ISO 12567-1 (2000)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method
BS EN 12664 (2001)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Dry and moist products of low and medium thermal resistance
BS EN 12939 (2001)	Thermal performance of building materials and products. Determination of thermal resistance by means of guarded hot plate and heat flow meter methods. Thick products of high and medium thermal resistance

Continued on next page

Table A.1 – continued from previous page

BS EN ISO 10211-2 (2001)	Thermal bridges in building construction. Heat flows and surface temperatures
BS EN ISO 13788 (2001)	Hygrothermal performance of building components and building elements. Internal surface temperature to avoid critical surface humidity and interstitial condensation. Calculation methods
BS EN ISO 10077-2 (2003)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Numerical method for frames
BS EN ISO 12567-2 (2005)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method – Roof windows and other projecting windows
BS EN ISO 10077-1 (2006)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. General
BS EN ISO 10211 (2007)	Thermal bridges in building construction. Heat flows and surface temperatures — Detailed calculations
BS EN ISO 13370 (2007)	Thermal performance of buildings – Heat transfer via the ground – Calculation methods
BS EN ISO 13786 (2007)	Thermal performance of building components. Dynamic thermal characteristics. Calculation method
BS EN ISO 6946 (2007)	Building components and building elements – Thermal resistance and thermal transmittance – Calculation method
BS EN ISO 12567-1 (2010)	Thermal performance of windows and doors – Determination of thermal transmittance by hot box method
BS EN ISO 10077-2 (2012)	Thermal performance of windows, doors and shutters. Calculation of thermal transmittance. Numerical method for frames
BS EN ISO 7345 (1996)	Thermal insulation. Physical quantities and definitions
BS EN 410 (1998)	Glass in buildings. Determination of luminous and solar characteristics of glazing
BS EN 410 (2011)	Glass in buildings. Determination of luminous and solar characteristics of glazing
Energy Performance	
BS EN ISO 13790 (2008)	Energy performance of buildings — Calculation of energy use for space heating and cooling
CIBSE Guide B (2001)	Heating, ventilating, air conditioning and refrigeration
ASHRAE Handbook (2009)	Fundamentals, S-I Edition - Principles and data used in the HVAC industry

B Statistical Concepts

Correlation

In statistics, two variables may be associated either by chance or by a degree of dependance. When these variables are associated by the latter, and can be described in terms of a linear relationship, then they present a level of correlation.

Pearson's correlation coefficient (r) is a measure of the strength of the association between the two variables. The correlation coefficient ranges between -1 and +1 and quantifies the direction and strength of the linear association between the two variables. The formula for the sample correlation coefficient is:

$$r = \frac{Cov(x,y)}{\sqrt{S_x^2 \cdot S_y^2}} \tag{B.1}$$

where $Cov(x,y)$ is the covariance of x and y, which represents the deviations of two variables from their respective means, and it is defined as:

$$Cov(x,y) = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{n-1} \tag{B.2}$$

S_x^2 and S_x^2 are the sample variances of x and y, defined as:

$$S_x^2 = \frac{\sum(X-\bar{X})^2}{n-1} \tag{B.3}$$

$$S_y^2 = \frac{\sum(Y-\bar{Y})^2}{n-1} \tag{B.4}$$

The variances of x and y measure the variability of the x scores and y scores around their respective sample means. The covariance measures the variability of the (x,y) pairs around the mean of x and mean of y, considered simultaneously.

Distance

In order to minimise the resulting overlap on the assumption of a linear correlation between the variables, a distance value can be derived for each of the variables and their simultaneous association. The distance between two points is defined by $d = |b-a|$; in other words, it is the length of a straight line drawn between them. If both a and b are defined in terms of a reference value, $o = (0,0)$ or in an euclidean domain, then they can be defined as vectors and thus form the relationship:

$$d^2 = a^2 + b^2 - 2 \cdot ab \cdot cos(\theta) \tag{B.5}$$

But because the correlation of a and b is constant, regardless of the length of the vectors, these can be normalised and, consequently, r can be expressed as $r = cos(\theta)$, and $d(a,b) = \sqrt{\sum(a_i - bi)^2}$. Therefore, the distance can be expressed as Equation B.6 or, inversely, the correlation in terms of the squared distance between them, as in Equation B.7.

$$d = \sqrt{2 \cdot (1-r)} \tag{B.6}$$

$$r = 1 - \frac{1}{2} \cdot d^2 \tag{B.7}$$

The Five-number Summary

To promptly provide a summary of a numeric variable [102], a set of five descriptive statistics is formed with the following terms (or percentiles):

1. minimum value (or smallest observation),
2. lower quartile or first quartile,
3. median (or middle quartile),
4. upper quartile or third quartile,
5. maximum value (or largest observation).

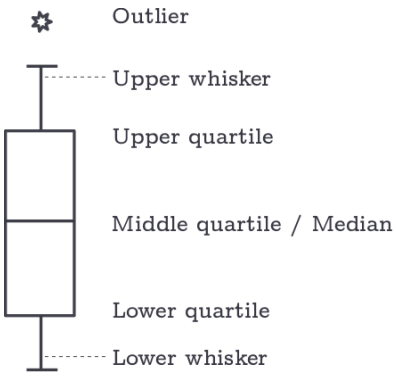
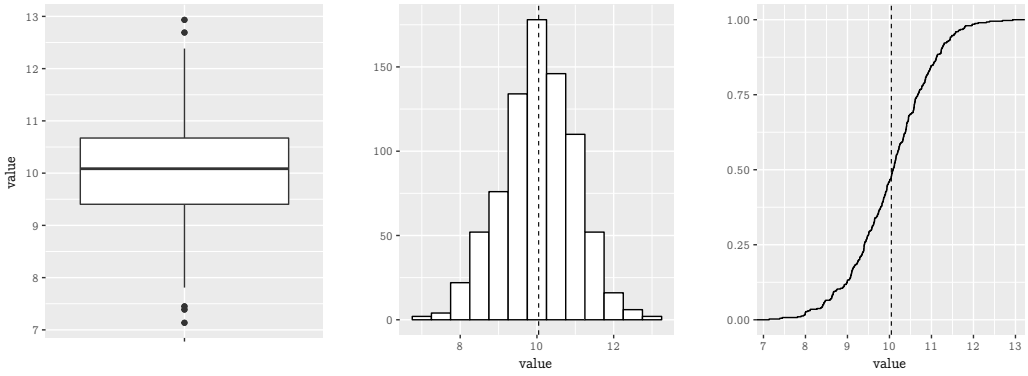


Figure B.1: Boxplot interpretation

In here, the median value gives a measure of the centre of the data; the minimum and maximum values provide the range of the data, and determine the presence of outliers; the lower and upper quartiles provide the spread of the data, especially when compared to the minimum, maximum, and median. These two values define an interquartile range that encapsulates 50% of the data. Furthermore, these five numbers conform the basis of the box-and-whisker plot ((see Figure B.1). The plot labels the limits of the data as whiskers, which represent the ranges for ± 25% of the data, excluding outliers.

The spread of the data may be also represented via histograms and Cumulative Distribution Functions (CDFs). Figure B.2 provides equivalent representations of the same distribution of data.



(a) boxplot *(b)* histogram *(c)* CDF

Figure B.2: Equivalent plotting of distributions

Cluster and Dendrograms

By way of demonstration, the following example describes the process of forming dendrogram based on the *distance* among points. Here, these points represent a cartesian space limited by x and y variables, but these could be replaced by n dimensions. Furthermore, the values for the axis could result from a previous analysis of *proximity*, which could be used to identify direct clusters in the studied dataset.

	x	y
a	40	50
b	60	20
c	30	30
d	80	70
e	70	30
f	50	50
g	60	80

(a) Data

	a	b	c	d	e	e	e
a	-						
b	36.1	-					
c	22.4	31.6	-				
d	44.7	53.9	64.0	-			
e	36.1	**14.1**	40.0	41.2	-		
f	**10.0**	31.6	28.3	36.1	28.3	-	
g	36.1	60.0	58.3	22.4	51.0	31.6	-

(b) Resemblance Matrix of Distances

Table B.1: Scatterplot (Dendogram)

The first step consists of plotting the dataset defined by Table B.1a. The second step consists in estimating the distances among the points in Figure B.3. This results in a correlation matrix (also known as resemblance matrix) defined by B.1b.

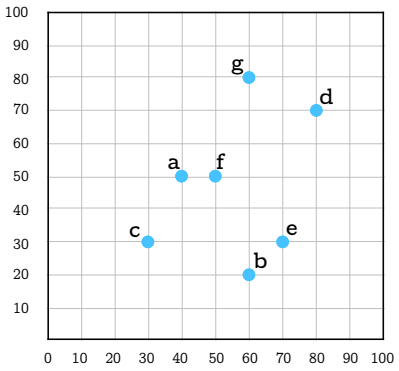


Figure B.3: Scatterplot: dendograms development

The third step consists in identifying the nearest non-adjacent points in the dataset. Table B.1b reveals that $a-f$ are the nearest points, followed by $b-e$. Then, both pairs are synthesised, to form a couple of new points (see Figure B.4).

The fourth step considers these new points and performs the similar evaluation of distances. Here, the nearest points are those between $af-c$ and $d-g$ (see Figure B.5).

The fifth step identifies the nearest point between $abcef$ and dg (see Figure B.6).

Finally, the remaining points are used to form the trunk of the tree. Noteworthy, it must be considered that the approach used in the example may be altered and the structure of the tree, as well as the criteria to stop the formation, can be adjusted as required. Figure B.7 illustrates the generated dendrogram, which reveals (graphically) the closest proximity of points in the dataset.

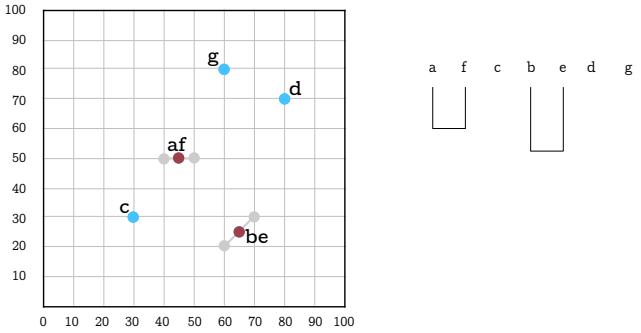


Figure B.4: Scatterplot: dendograms development (B)

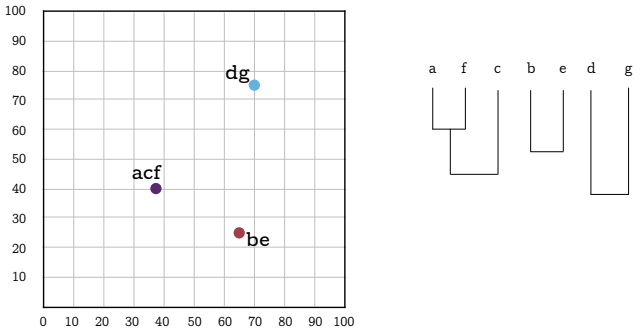


Figure B.5: Scatterplot: dendograms development (C)

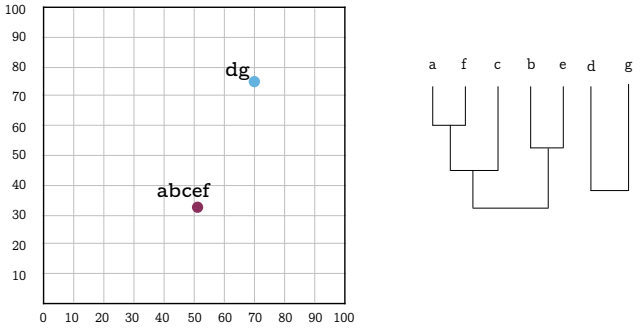


Figure B.6: Scatterplot: dendograms development (D)

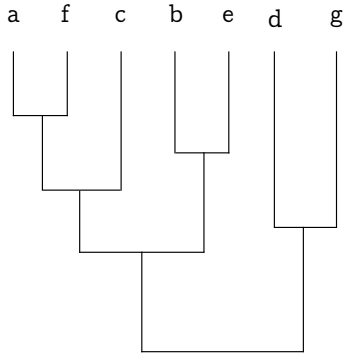


Figure B.7: Scatterplot: dendograms development (E)

C Modelling Process: Composition and Simulation

By way of demonstration, one of the most common typologies in the housing stock—despite its high energy demand—is selected to explain the simulation process (see Table C.1).

Housing Code	1750
Number of Dwellings	2022
SAP Age band	1850-1899
Tenure Type	Private owner occupied
Dwelling Type	Semi-detached
Region	London
Wall	Solid brick
Window	Single-glazed (wood sash)
Heating system	Gas central heating
DHW	Condensing combi

Table C.1: Archetypes Parameters / simple subset selection

By following the approach described in both Figure 4.1 and Table 4.1, this exemplary archetype is represented as a two-floor dwelling with an attachment on one side—represented in Figure C.1. The glazing ratio is adjusted according to the value provided by the EHS; this also explains the slightly unrealistic configuration of windows in both front and back façades.

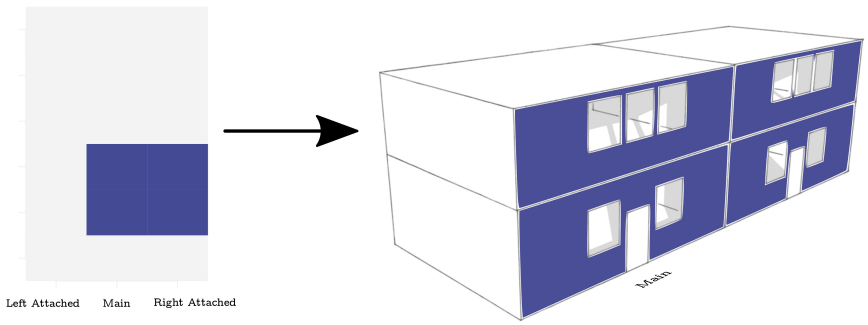


Figure C.1: Volumetric representation of a semi-detached house

Script C.1 exemplifies the assignment of materials to the idf format, as a function of the R core script. The function receives information from the survey dataset, and compares the value with the corresponding library of materials.

Pseudo-code C.1 Assignment of materials

```
fn_material_ExtWall <- function(type, width, data){
  #  type : material type from survey data
  # width : wall width from survey data
  #  data : reference [protected library]

  matLbls <- c("",
               "!- Name","!- Roughness",
               "!- Thickness {m}",
               "!- Conductivity {W/m-K}",
               "!- Density {kg/m3}",
               "!- Specific Heat {J/kg-K}")

  myMat <- as.data.frame(data[[20]][,type])
  names(myMat) <- "Name"
  NoLayers <- length(myMat[,1]) - sum(myMat[,1] == "" )
  myMat <- merge(myMat, data[[19]], by.x = "Name", by.y = "Name", sort = F)

  myName <- c(as.character(unlist(data[[20]][1,type], recursive = F,
                                  use.names = F), paste (sep = "")))
  myMatLayer <- list(myName,myMat)

  Xtot <- sum(myMatLayer[[2]][3])
  Xnew <- width
  myMatLayer[[2]][3] <- myMatLayer[[2]][3]*(Xnew/Xtot)
  myMatLayer[[2]][3] <- round(myMatLayer[[2]][3],3)

  myMatLayer[[3]] <- as.data.frame(c(cbind(c(as.character(myMatLayer[[1]]))),
                                     cbind(c(as.character(myMatLayer[[2]][,1])))))
    names(myMatLayer[[3]]) <- "Construction"
    myMatLayer[[3]] <- t(myMatLayer[[3]])
    myMatLayer[[3]] <- cbind(c("Construction"), myMatLayer[[3]])
    myMatLayer[[3]] <- t(myMatLayer[[3]])
    myMatLayer[[3]] <- as.data.frame(cbind(myMatLayer[[3]],c(rep(c(","),
                                       each=length(myMatLayer[[3]])-1), c(";"))))
    myMatLayer[[3]]$V <- paste(myMatLayer[[3]]$Construction,myMatLayer[[3]]$V2,sep="")
    myMatLayer[[3]]$Construction <- NULL
    myMatLayer[[3]]$V2 <- NULL

  Xmy <- myMatLayer[[2]]
    Xmy$V3 <- matLbls
    Xmy$V0 <- c("",rep("     ",6))
    Xmy <- paste(Xmy$V0,Xmy$V,Xmy$V3,sep="")
  myMatLayer[[3]] <- as.character(Xmy)

  myMatLayer[[1]][2,] <- "Exterior Wall,"
    Xmz <- myMatLayer[[1]]
    Xmz$V3 <- c("!- Name",
                "!- Outside Layer",
                rep("!- Layer",length(Xmz$V)-2))
    Xmz$V0 <-  c("",rep("     ",D_lim-1))
    Xmz <- paste(Xmz$V0,Xmz$V,Xmz$V0,Xmz$V3,sep="")

  myMatLayer[[1]] <- as.character(Xmz)
  myMatLayer[[2]] <- myMatLayer[[4]] <- "!- "

  return(myMatLayer)
}
```

A similar process is applied to parse information related to heating system parameters, appliance properties, and occupancy values.

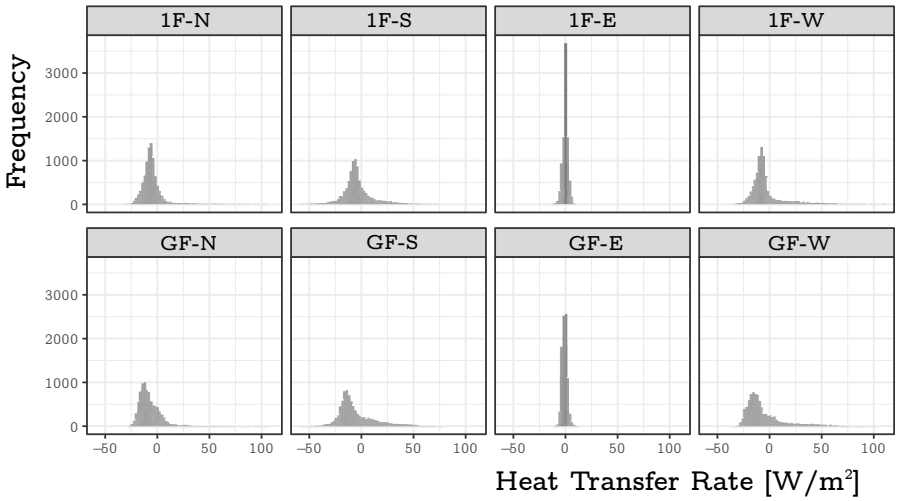


Figure C.2: Convective surface transfer rate for the two-floor dwelling

The heat loss parameter, which provides an indication of the building leakiness, is 5.13 W/m^2K. This high value results from the properties of the envelope, which in this case, are related to an old dwelling. Figure C.2 illustrates the overall performance of the surrounding walls, and their representation in the simulation. It can also be seen that the transfer rate in conjoined walls (GF-E and 1F-E) is significantly low.

As described in Section 4.1.1, the integration and processing of information into the platform, derived from different datasets and catalogues, is implemented using the open-source statistical software `R`, and the dynamic simulation of dwellings is performed using the `EnergyPlus` engine. Script C.2 describes the communication between `R`, `EnergyPlus` and the shell.

Pseudo-code C.2 EP script

```
runEnergyPlus <- function(idf_toRun, NameOut, epw, myOS){
  path.run <- system('pwd', intern = T)
  writeLines(idf_toRun,"myModel/ModelToRun.idf")

  if (myOS == "unix") {
     inpfilename <- paste("-r ",normalizePath(path.run),"/myModel/ModelToRun.idf", sep="")
     # define location of idf (schdule needs to be in the same location)
     exename <- paste(normalizePath(Path.EnergyPlus), "/energyplus -x", sep="");
     # find /usr/bin/runenergyplus file
     outname <- paste("-d ",normalizePath(path.run), "/myModel/Output/", sep="");
   } else {
     inpfilename <- paste(normalizePath(getwd()),"\\ModelToRun", sep="")
     # define location of idf (schdule needs to be in the same location)
     exename <- "path_to_\\EnergyPlus\\EnergyPlusV8-1-0\\RunEPlus";
     # find RunEPlus.bat file // on cmd >> dir /b /s .\\*RunEPlus.bat
   }
   weathername <- paste("-w ",fnEPWeatherPathFinder(),epw,".epw",sep='');
   # assign weathe file - needs to be stored on the EP/Weatherfile folder
   command <- paste(exename,weathername,outname, inpfilename);
   rm(exename,inpfilename,weathername);
     ptm <- proc.time()
  system(command, wait = F, intern = T)
     proc.time() - ptm

  invisible(EP_TryOut)
}
```

Similarly, the inclusion of addition tools is performed by implementing such a communication. Scripts C.3 and C.4 outline the steps applied to perform the SA based on the model archetype, which follows a systematic process.

Pseudo-code C.3 OAT script

```
modelToUse=$1
sizeToUse=$2
typeEval="$3"

OS="`uname`"
case $OS in
  'Linux')
    OS='Linux' ;;
  'Darwin')
    OS='Mac' ;;
  *) ;;
esac

#[0] : Help  ============================================================
if [ "$1" == "-h" ]; then
  echo "
  Usage: ./myScript/`basename $0` id s

  id [model] :: reference model to analyse
   s  [size] :: size of the sample
  "
  exit 0
fi

#[1] : Population  ============================================================
  commandPop="./myScript/Run_SAnU_OATa.sh $modelToUse $sizeToUse $typeEval"
  eval $commandPop
  wait

#[2] : Simulation  ============================================================
  commandSim="./myScript/Run_SAnU_OATb.sh $modelToUse"
  eval $commandSim
  wait

#[3] : Analysis  ============================================================
  commandAna="./myScript/Run_SAnU_OATc.sh $modelToUse $typeEval"
  eval $commandAna
  wait

#[-] : End
echo ' parametrics / evaluation / analysis '
```

Pseudo-code C.4 OAT section

```
fnDefineDensity <- function(DataSrc, id, vars.rqstd, var.avlbl, blocs){
    fnExtractDens <- function(tmp.id){
      tmp.name <- sub("density.","",tmp.id)
      vars.fact.Use <- list()
      if(any(grep(tmp.name, vars.rqstd))){
        tmp.density <- abs(rnorm(blocs, mean = mean(DataSrc[,tmp.name]),
        sd = sd(DataSrc[,tmp.name])))
        tmp.density <- ifelse(tmp.density>=0,tmp.density,0)
        tmp.density <- tmp.density[order(tmp.density)]
      }else{
        tmp.density <- as.numeric(DataSrc[id,tmp.name])
      }
      vars.fact.Use[[tmp.id]] <- tmp.density;
      assign(tmp.id, as.vector(unlist(vars.fact.Use)), envir = .GlobalEnv);
      invisible(vars.fact.Use)
    }
    lapply(var.avlbl, fnExtractDens)
  }
```

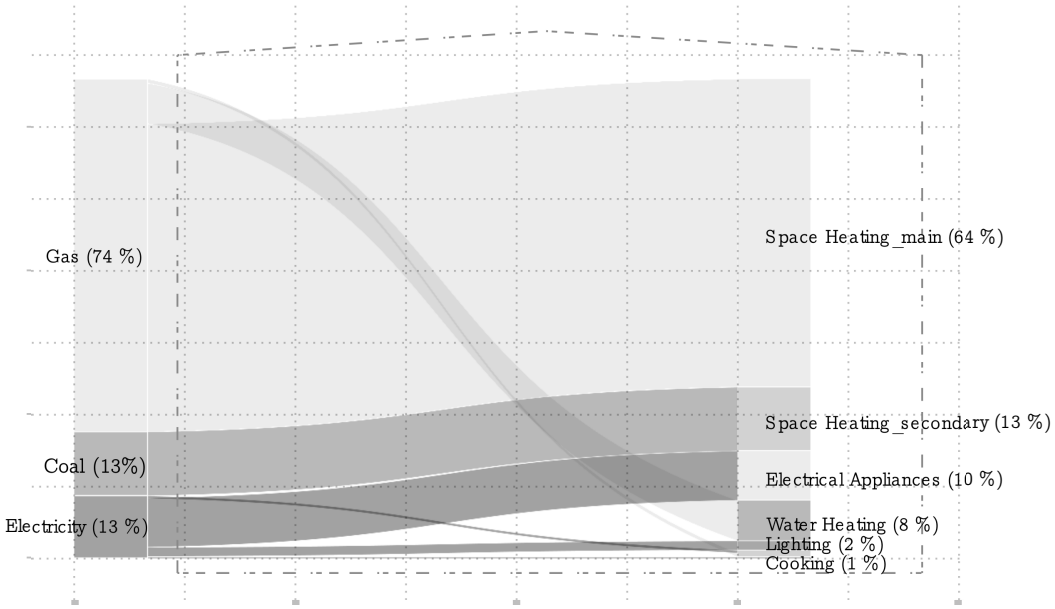


Figure C.3: Synthesised energy share for the semi-detached house

The simulation, as mentioned in 4.1.4, is performed for an entire calendar year. The energy demand for this archetype is 61,167.19 kWh. Three quarters of this demand are attributable to gas, as outlined in Figure C.3. The total carbon emissions associated are 24,655.43 kgCO2. Further, the secondary heating system for this archetypes is coal-based, which increases the total carbon footprint. The average temperature (i.e. for the two zones) is 16.4 C. Figure C.4 shows that, along the year, this temperature is below typical comfort temperature levels.

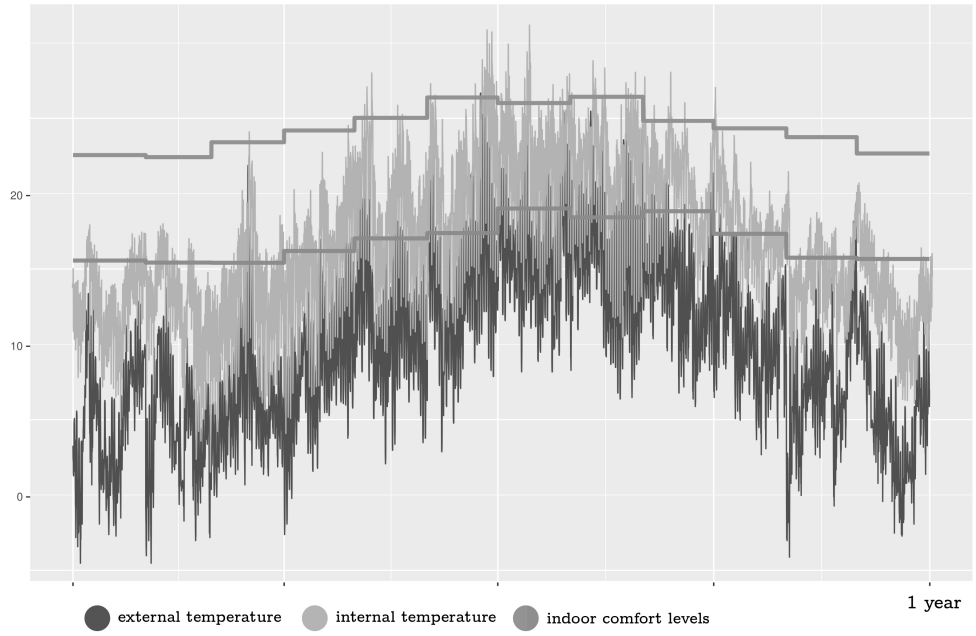


Figure C.4: Relation of temperatures for the semi-detached house

Approximately, the construction of a volumetric archetype takes about 3 seconds, and the simulation about 20 seconds. Therefore, there are some applications where the whole process can be performed without further hardware requirements, and some other applications where the distributed computing is necessary.

Pseudo-code C.5 HPC job script

```
#!/bin/bash

#OPTIONS FOR PBS PRO ==========================================

#PBS -P HPCA-012345-PGR
# This specifies the project associated

#PBS -l walltime=02:20:00
# This specifies the job should run for no longer than 2.2 hours

#PBS -l select=8:ncpus=4:mem=4Gb
# This specifies the job needs 8 chunks, with 4 CPU core, and 4Gb of RAM (memory)

#PBS -o out2job
# This send your output to the file rather than the standard filename

#OPTIONS FOR PBS PRO ==========================================

cd /home/ezxgas/Project/BETA

./myScript/Run_BASH_BakeOffArray.sh 1 14951 32

echo "using EP+! -- idf simulator"
```

To end, Script C.5 exemplifies a script-job to request resources in the HPC facility and perform the simulation. Essentially, the scripts consists of a set of instructions which are given to the external environment. Interestingly, rather than a sophisticated script, this represents an old-fashioned approach to control a computing machine.

* More details on the EnHub repository are given in the provided *List of Publications*.

D Unified Modelling Language

The UML architecture shows the internal logic of the modules, and can be presented using different types of diagram such as class, use-case, state, sequence, component, or deployment. These diagrams can be easily expanded and complemented. EnHub implements a UML class diagram type, focussed on the characteristics of the modules and their relationships. The diagram is based on Mhalas' implementation [156] and complemented with BREDEM parameters [95].

Here, each class contains attributes and methods—reported below. The former contain state components and the latter the functionality of a class. Both attributes and methods define their elements, also known as modifiers, as public (+), private (−) or protected (#). Classes are linked by an inheritance relation. Such a relation is interpreted as a dependency, ownership, or association [177].

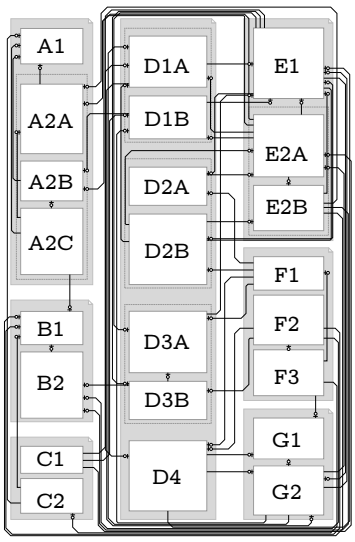
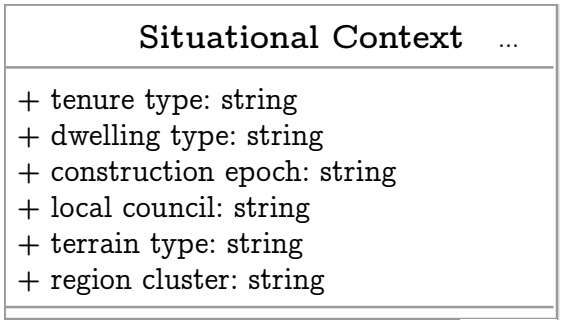
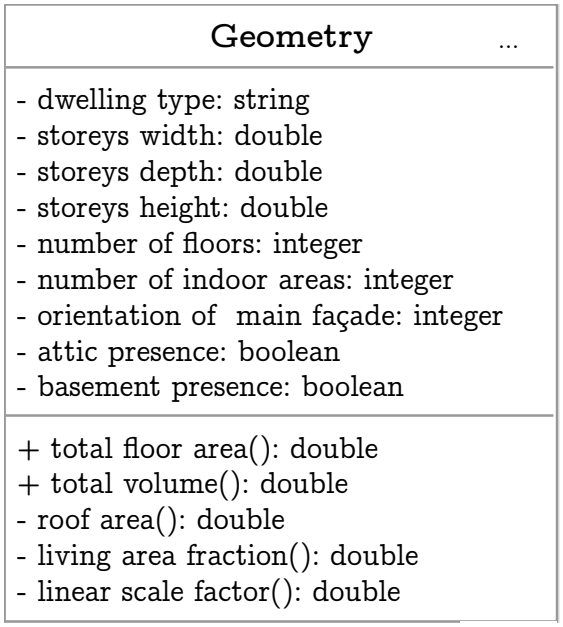


Figure D.1: UML Schematic

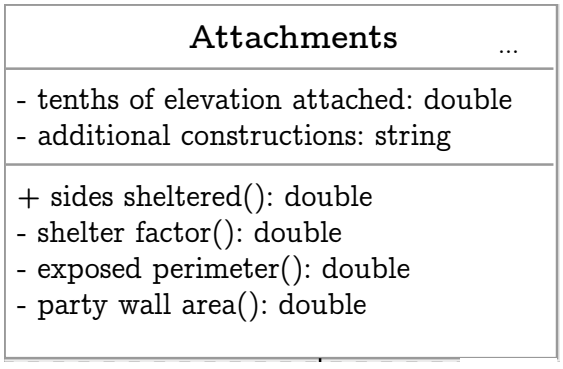
Below are summarised each of the components outlined in Figure D.1. These correspond to the modules described in Figure 4.3 from Section 4.



(a) A1

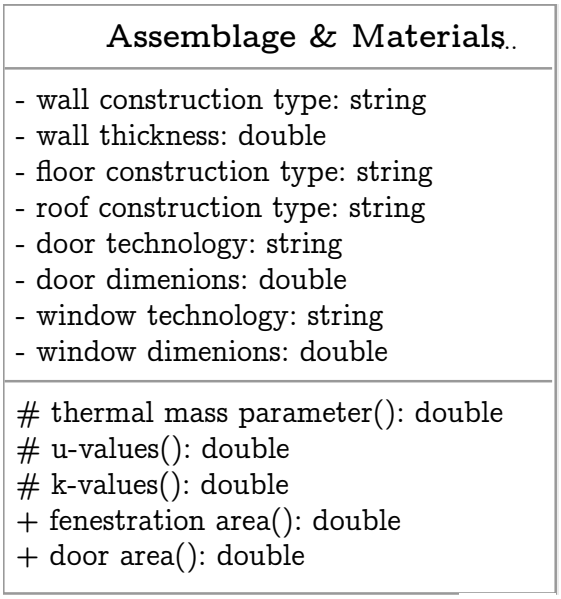


(b) A2A

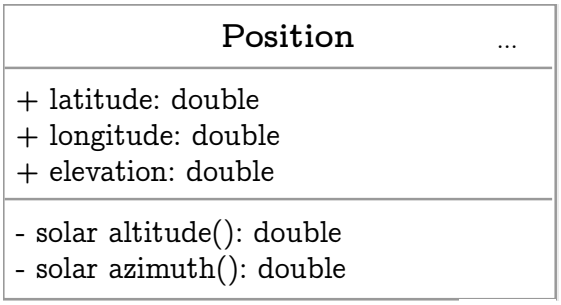


(c) A2B

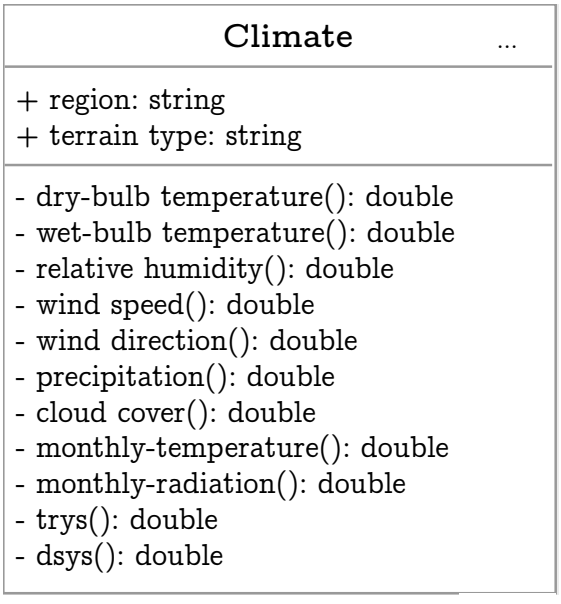
Figure D.2: UML Sections



(d) A2C

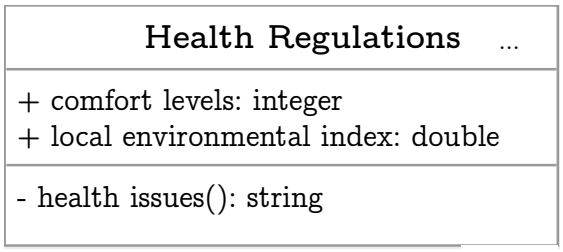


(e) B1

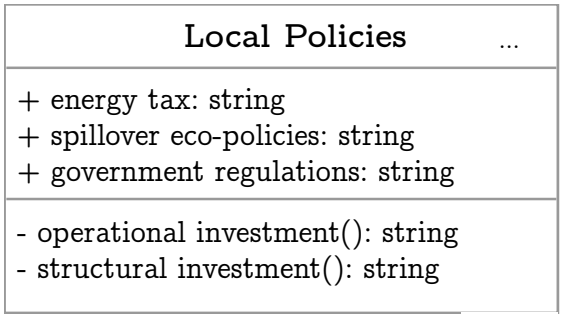


(f) B2

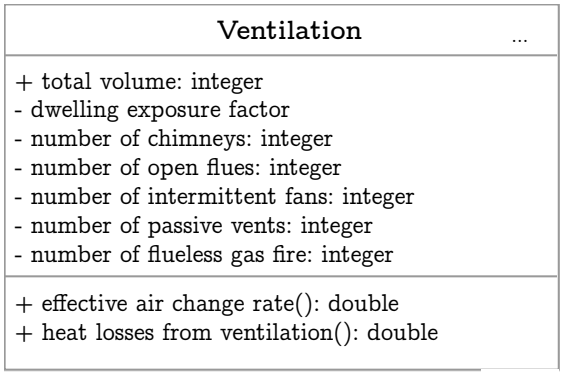
Figure D.2: UML Sections – continued



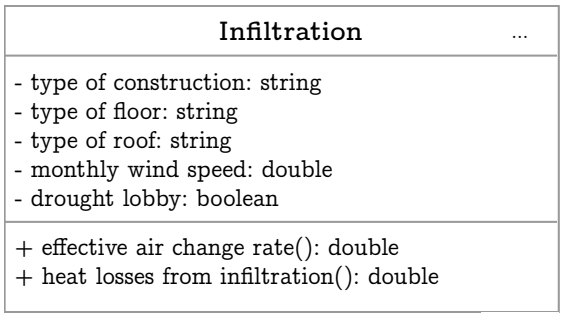
(g) C1



(h) C2

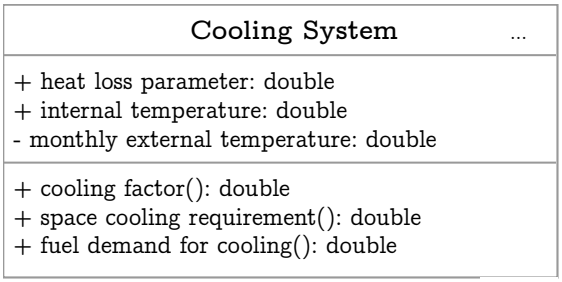


(i) D1A

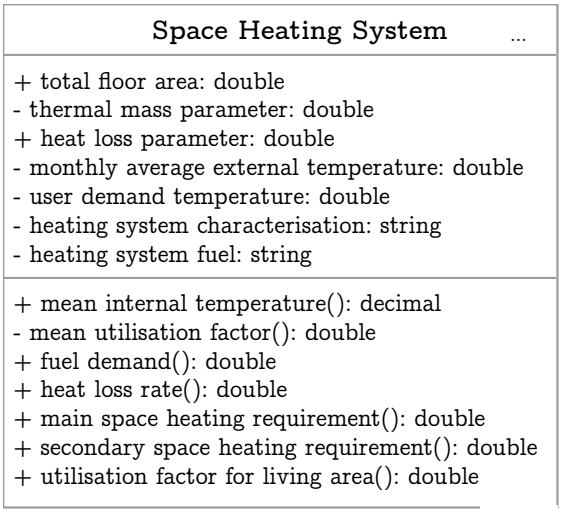


(j) D1B

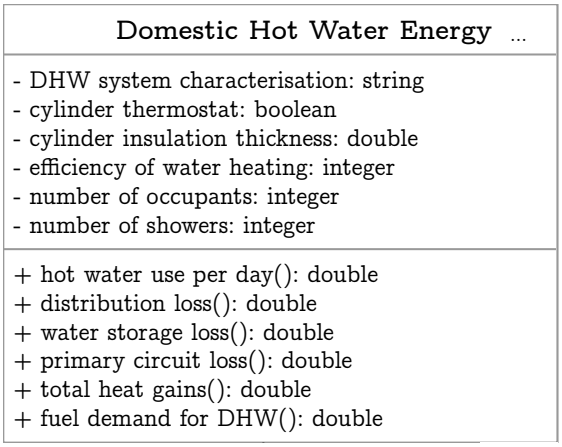
Figure D.2: UML Sections – continued



(k) D2A

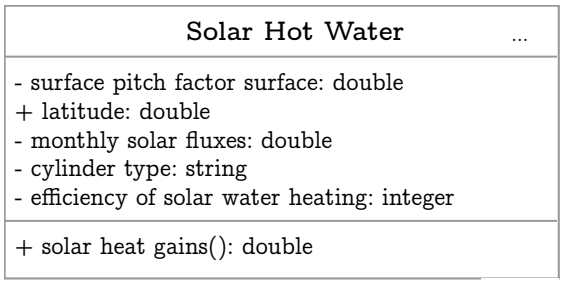


(l) D2B

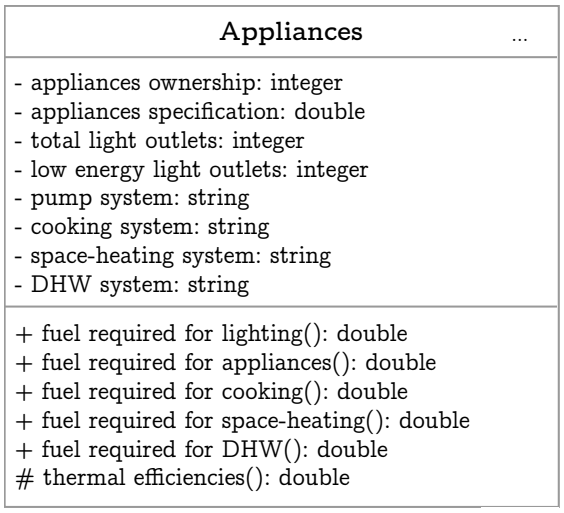


(m) D3A

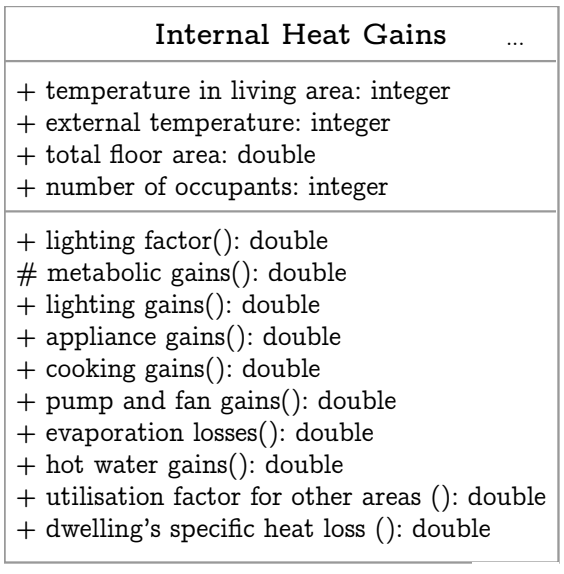
Figure D.2: UML Sections – continued



(n) D3B

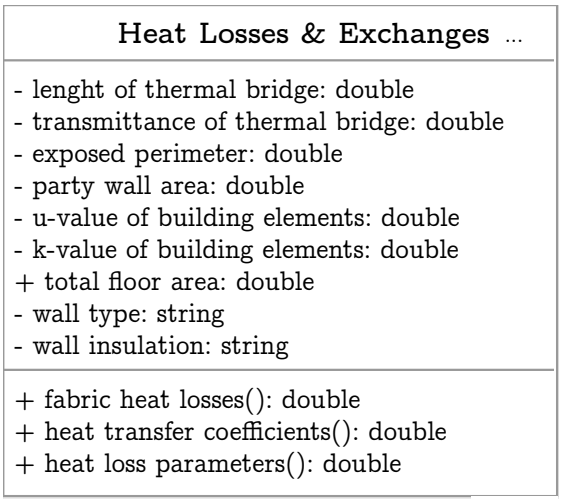


(o) D4

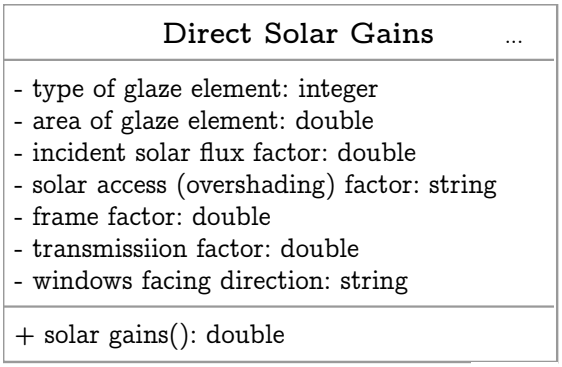


(p) E1

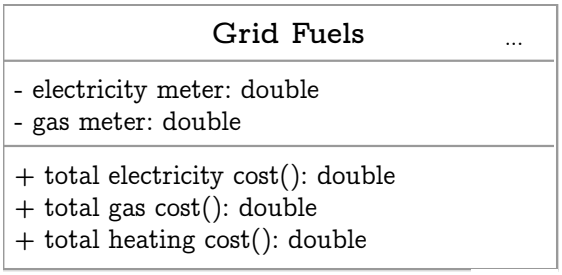
Figure D.2: UML Sections – continued



(q) E2A

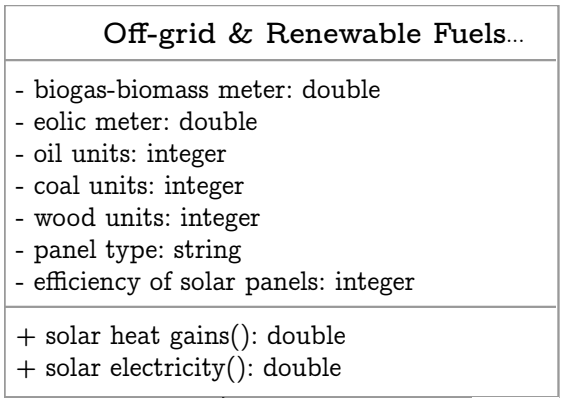


(r) E2B

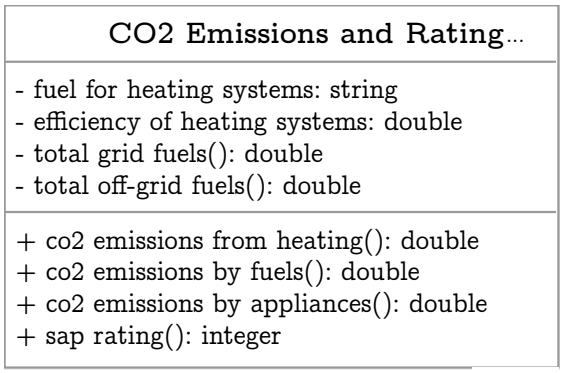


(s) F1

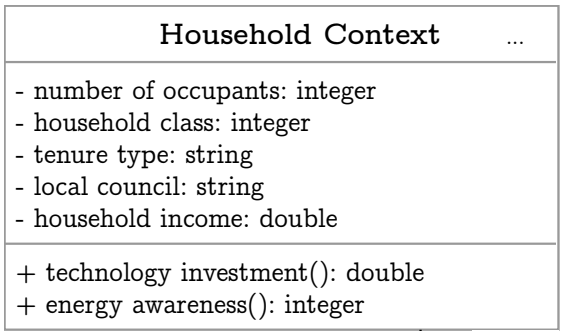
Figure D.2: UML Sections – continued



(t) F2

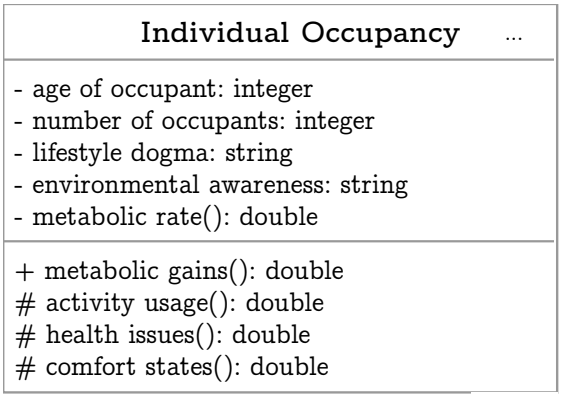


(u) F3



(v) G1

Figure D.2: UML Sections – continued



(w) G2

Figure D.2: UML Sections – continued

E Glossary

A

active participation

This refers to ability to share the platform, by providing sufficient information (and requirements) to validate and edit each module. In this way, each set of modules may be employed and adapted according to the purposes of the modellers, however maintaining the main functionality of the platform.

agent

A discrete entity with its own goals and behaviours. Such an entity employs a combination of parameters to define their unique state (or condition) and therefore reacts according to their current environment. An agent can be static or dynamic, and it can also evolve in time.

agent-based model

A technique for studying behaviour among a group of agents, for assessing its own situation and responding appropriately. In other words, this process simulates the actions and interactions of autonomous agents.

C

calibration

This is defined as the process of determining or rectifying a parameter. In sum, a process of calibration determines the likelihood—with their corresponding adjustments—to represent a real-world phenomenon with numerical results.

comfort

This is defined as the condition of mind that expresses satisfaction with the thermal environment, although assessed by subjective evaluation. The main factors that influence thermal comfort are those that determine heat gain and loss (e.g. metabolic rate, clothing insulation, air temperature, mean radiant temperature, air speed and relative humidity). Other factors are those related to individual expectations and adopted lifestyles.

complexity

This property exists when the number of relationships and interconnections between elements represented in a system or model, creates multiple dependencies within inputs and internal variables. A complex system also exists when the representation of physical phenomena depends on the unique response to a set of inputs by generating processes of organisation, emerging dependencies and unique conditions.

D

decision-making

The cognitive process resulting in the selection of a belief or a specific action among several alternative possibilities.

dynamic simulation

This mode stands for the analysis of time-dependant variables in a given representation of energy balance. In building analysis domain, this mode is used to represent the resulting behaviour from the interaction between building services and fabric components. Due to the consideration of time, dynamic simulation is helpful to provide energy demand in an appropriate timeframe.

E

encapsulation

The process of enclosing a message or signal in a set of codes which allow transfer across a modelling platform. It is also defined as the provision of an interface to simplify access for the user. In other words, it is a standard method of communication within the platform.

F

framework

This represents a layered structure indicating the main components and attributes of programs and their stage in the general algorithm. Here it is defined the main relationships and dependencies of programs. Therefore, it also represents a template for the development of software applications or components.

fuel poverty

A classic definition states this condition when a household spends more than 10 per cent of its income on energy fuels. Although there is debate on the number of factors that may affect household expenditure, this is an indicator for the housing stock in terms of how affordable is to reach indoor thermal comfort.

H

household

A household represents one person living alone or a group of people (not necessarily related) living at the same address who share cooking facilities and share a living room or sitting room or dining area.

M

modelling

The purposeful abstraction of reality, resulting in the formal specification of a conceptualisation and underlying assumptions and constraints.

modelling classes

These are defined as specific functions, routines and attributes upon which models can be built. As a result of such selection of properties, the adopted approach can be different, despite pursuing the same goal. The class defines the nature in the connection between the abstraction and the phenomena. Therefore, a modelling class is defined under different domains and attributes.

O

open modularity

This refers to the ability of providing open-access to a standardised platform. In this way, it is possible to validate and edit each module, by maintaining the functionality within the platform.

P

platform

This represents the underlying components of a system, on which application programs can be executed and linked. It also defines a standard around which a system can be developed.

policy

This is defined as a written method of action to guide, manage and determine present and future decisions. In short, a policy represents the principle of action(s) proposed and applied by an organisation or an individual.

policy outcome

In practice, this is defined as the resulting conditions when a particular policy is used. It tends to be a quantifiable consequence of the application of a policy.

R

rebound effect

The resulting outcome in energy demand when changes in behaviour offset energy savings from improvements and energy conservation measures. More specifically, the rebound effects can be distinguished in three types: a) *direct effects* - savings from improvements are translated into more intense use of energy in other systems; b) *indirect effects* - savings reflected in energy-related bills are used to acquire other products and services; c) *economy-wide effects* - these reflect socio-technical transitions that affect changes in the economy and energy-related patterns and demands.

S

scalability

This refers to the ability of a computer application to continue to function well when it is changed in size, in scope and in detail.

simulation

The representation of the behaviour or characteristics of one system through the use of another system, especially a computer program designed for the purpose.

steady-state

These are defined as those models in which their conditions are not affected by time and the phenomena is represented as a system in equilibrium. In some cases, the representation of physical phenomena may assume reduced periods of time in which different conditions are assumed, however these do not reflect fluctuation of processes in the models.

V

validation

This is defined as the process to make valid or proper. In brief, a process of validation determines how adequate is the abstraction of the real phenomena. For this, a comparison between of the simulated results and empirical values is typically performed.

verification

This is defined as the process to confirm or corroborate values, against expected ones. In practice, a process of verification identifies the ability to develop an abstraction (i.e. developing a numerical model, building a software algorithm, processing a survey), and to accurately reflect the specifications of the chosen algorithm.

W

workflow

The defined series of tasks and stages within an algorithm to produce outcomes. A workflow usually conveys an explicit representation of the necessary links within a process.

Bibliography

[1] W. Abrahamse and L. Steg, *How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings?*, Journal of Economic Psychology, 30 (2009), pp. 711–720.

[2] I. Ajzen, *The theory of planned behavior*, Orgnizational Behavior and Human Decision Processes, 50 (1991), pp. 179–211.

[3] G. Anandarajah, N. Strachan, P. Ekins, R. Kannan, and N. Hughes, *Pathways to a Low Carbon Economy: Energy Systems Modelling*, tech. rep., UK Energy Research Centre, London, 2009.

[4] F. Anstett-collin, J. Goffart, T. Mara, and L. Denis-vidal, *Sensitivity analysis of complex models : Coping with dynamic and static inputs*, Reliability Engineering and System Safety, 134 (2015), pp. 268–275.

[5] ASHRAE (American Society of Heating Refrigerating and Air-Conditioning Engineers), *2009 ASHRAE Handbook: Fundamentals, I-P Edition*, no. 404, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., 2009.

[6] S. Y. Auyang, *Foundations of complex-system theories : in economics, evolutionary biology, and statistical physics*, Cambridge University Press, Cambridge, 1998.

[7] R. M. Axelrod, *The complexity of cooperation*, Princeton University Press, 1997.

[8] K. J. Baker and R. M. Rylatt, *Improving the prediction of UK domestic energy-demand using annual consumption-data*, Applied Energy, 85 (2008), pp. 475–482.

[9] T. Baker and K. Jenkins, *The Domestic Energy Sub-Model in MDM-E3*. 2007.

[10] C. S. Bale, N. J. McCullen, T. J. Foxon, A. M. Rucklidge, and W. F. Gale, *Harnessing social networks for promoting adoption of energy technologies in the domestic sector*, Energy Policy, 63 (2013), pp. 833–844.

[11] I. Ballarini, S. P. Corgnati, and V. Corrado, *Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project*, Energy Policy, 68 (2014), pp. 273–284.

[12] J. Banks, ed., *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, John Wiley & Sons, Inc., 1998.

[13] B. Boardman, S. Darby, G. Killip, M. Hinnells, C. N. Jardine, J. Palmer, and G. Sinden, *40 % House Project Background material A. United Kingdom Domestic Carbon Model. Description, method and analysis.*, tech. rep., Environment Change Institute, Oxford University, 1996.

[14] ———, *40% House report*, tech. rep., Environment Change Institute, Oxford University, 2005.

[15] B. Boardman, D. Favis-Mortlock, M. Hinnells, K. Lane, G. Milne, J. Palmer, E. Small, V. Strang, and J. Wade, *Domestic Equipment and Carbon Dioxide Emissions: Second Year Report*, tech. rep., Energy and Environment Programme, Environmental Change Unit, University of Oxford, 1995.

[16] C. Böhringer and T. F. Rutherford, *Combining bottom-up and top-down*, Energy Economics, 30 (2008), pp. 574–596.

[17] A. Booth and R. Choudhary, *Displaying Uncertainty in Energy Savings for Large-Scale Building Refurbishments*, in Retrofit 2012, Cambridge, 2012.

[18] A. Booth, R. Choudhary, and D. Spiegelhalter, *Handling uncertainty in housing stock models*, Building and Environment, 48 (2012), pp. 35–47.

[19] V. Brechling and S. Smith, *The pattern of Energy Efficiency Measures amongst Domestic Households in the UK*, tech. rep., Institute for Fiscal Studies, London, UK, 1992.

[20] ———, *Household Energy Efficiency in the UK*, Fiscal Studies, 15 (1994), pp. 44–56.

[21] J. Buckley and S. King-Hele, *What is weighting?*, tech. rep., UK Data Service, University of Essex and University of Manchester, 2014.

[22] P. C. Cacciabue, *Modelling and simulation of human behaviour in system control*, Springer, Berlin, 1998.

[23] G. Calleja Rodríguez, A. Carrillo Andrés, F. Domínguez Muñoz, J. M. Cejudo López, and Y. Zhang, *Uncertainties and sensitivity analysis in building energy simulation using macroparameters*, Energy and Buildings, 67 (2013), pp. 79–87.

[24] Cambridge Architectural Research Limited, *DEMScot 2 : Scotland's Housing Carbon Model The extensions*, Tech. Rep. April, Cambridge Architectural Research Limited, 2010.

[25] J. Chapman, *Data Accuracy and Model Reliability*, in BEPAC Conference, Canterbury, UK, 1991.

[26] J. Chapman, *Agent based modelling of occupants' behaviours for building simulation and control*, phd thesis, University of Nottingham, 2017.

[27] J. Chapman, R. J. Lowe, and R. Everett, *The Pennyland Project*, tech. rep., Milton Keynes, 1985.

[28] J. Chapman, P.-O. Siebers, and D. Robinson, *Multi-agent stochastic simulation of occupants in buildings*, Journal of Building Performance Simulation.

[29] Chartered Institution of Building Services Engineers (CIBSE), *Degree-days: theory and application. TM41: 2006.*, tech. rep., The Chartered Institution of Building Services Engineers, London, UK, 2006.

[30] J. Chen, J. E. Taylor, and H. H. Wei, *Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation*, Energy and Buildings, 47 (2012), pp. 515–524.

[31] V. Cheng and K. Steemers, *Modelling domestic energy consumption at district scale: A tool to support national and local energy policies*, Environmental Modelling & Software, 26 (2011), pp. 1186–1198.

[32] CIBSE (Chartered Institution Building Services Engineers), *Building energy and environmental modelling*, CIBSE applications manual ; AM11: 1998, Chartered Institution of Building Services Engineers, London :, 1998.

[33] J. A. Clarke, *Energy simulation in building design*, Taylor & Francis Ltd, Oxford, 2001.

[34] J. A. Clarke, C. M. Johnstone, S. Ghauri, J. M. Kim, and P. Tuohy, *The EDEM methodology for Housing Upgrade Analysis, Carbon and Energy Labelling and National Policy Development*. 2012.

[35] L. Collins, S. Natarajan, and G. J. Levermore, *Climate change and future energy consumption in UK housing stock*, Building Service Engineering, 31 (2010), pp. 75–90.

[36] Committee on Climate Change, *Building a low-carbon economy: the UK's contribution to tackling climate change*, Tech. Rep. December, 2008.

[37] ———, *The Fourth Carbon Budget Reducing emissions through the 2020s*, Carbon, (2011), pp. 1–18.

[38] ———, *The Fifth Carbon Budget: The next step towards a low-carbon economy*, Tech. Rep. November, 2015.

[39] V. Corrado and I. Ballarini, *Refurbishment trends of the residential building stock: Analysis of a regional pilot case in Italy*, Energy and Buildings, 132 (2016), pp. 91–106.

[40] B. D. B. Crawley and L. K. Lawrie, *EnergyPlus : Energy Simulation Program*, Ashrae, 42 (2000), pp. 49–56.

[41] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, and J. Glazer, *EnergyPlus : creating a new-generation building energy simulation program*, Energy and Buildings, 33 (1994), pp. 443–457.

[42] A. Dagoumas and T. Barker, *Pathways to a low-carbon economy for the UK with the macro-econometric E3MG model*, Energy Policy, 38 (2010), pp. 3067–3077.

[43] P. Das, C. Shrubsole, B. Jones, I. G. Hamilton, Z. Chalabi, M. Davies, A. Mavrogianni, and J. Taylor, *Using probabilistic sampling-based sensitivity analyses for indoor air quality modelling*, Building and Environment, 78 (2014), pp. 171–182.

[44] M. S. de Wit, *Identification of the important parameters in thermal building simulation models*, Journal of Statistical Computation and Simulation, 57 (1997), pp. 305–320.

[45] M. S. de Wit and G. Augenbroe, *Analysis of uncertainty in building design evaluations and its implications*, Energy and Buildings, 34 (2002), pp. 951–958.

[46] S. A. Delre, W. Jager, T. H. A. Bijmolt, and M. A. Janssen, *Will it spread or not? the effects of social influences and network topology on innovation diffusion*, Journal of Product Innovation Management, 27 (2010), pp. 267–282.

[47] Department for Communities and Local Government (DCLG), *English Housing Survey : Housing Stock Data*, tech. rep., Colchester, Essex, 2011.

[48] ———, *English Housing Survey: HOMES report*, tech. rep., Department for Communities & Local Government, 2011.

[49] ———, *English Housing Survey: Technical Advice Note - Sampling and Weighting: 2011-2012 Update*, tech. rep., 2013.

[50] Department for Environment Food and Rural Affairs (DEFRA), *Climate Change: The UK Programme 2006*, tech. rep., 2006.

[51] ———, *Public Understanding of Sustainable Energy Consumption in the Home*, Tech. Rep. November, Brook Lyndhrust, London, 2007.

[52] ———, *2012 Guidelines to Defra/DECC GHG Conversion Factors for Company Reporting*, tech. rep., DEFRA/DECC, 2012.

[53] Department of Energy & Climate Change and Building Research Establishment, *Energy Follow Up Survey, 2011*, tech. rep., UK Data Service, 2016.

[54] Department of Energy & Climate Change (DECC), *The Government's Standard Assessment Procedure for Energy Rating of Dwellings*, Tech. Rep. March, Building Research Establishment, Watford, 2009.

[55] ———, *Energy Flow Chart 2011*, tech. rep., Department of Energy and Climate Change, 2011.

[56] ———, *Appendix S : Reduced Data SAP for existing dwellings*, in The Government's Standard Assessment Procedure for Energy Rating of Dwellings, vol. 91, Watford, 2012.

[57] ———, *The Government's Standard Assessment Procedure for Energy Rating of Dwellings*, Tech. Rep. October, Building Research Establishment, Watford, 2012.

[58] ———, *Digest of United Kingdom Energy Statistics 2013*, Crown Copyright, London, 2013.

[59] ———, *Domestic energy consumption in the UK between 1970 and 2012*, tech. rep., Department of Energy and Climate Change, 2013.

[60] ———, *Energy Consumption in the United Kingdom (ECUK)*, Tech. Rep. July, Department of Energy and Climate Change, London, 2013.

[61] ———, *Energy Consumption in the United Kingdom (ECUK) : domestic data tables*, 2013.

[62] ———, *Energy Flow Chart 2012*, tech. rep., Department of Energy and Climate Change, 2013.

[63] ———, *Summary of Analysis using the National Energy Efficiency Data-Framework: Part I Domestic Energy Consumption*, (2013).

[64] ———, *Local authority carbon dioxide emissions estimates 2012*, Tech. Rep. June, 2014.

[65] P. E. Dodds, *Integrating housing stock and energy system models as a strategy to improve heat decarbonisation assessments*, Applied Energy, 132 (2014), pp. 358–369.

[66] P. Dolan, M. Hallsworth, D. Halpern, D. King, and I. Vlaev, *Influencing behaviour through public policy*, tech. rep., UK Cabinet Office & Institute for Government, London, 2010.

[67] S. Dorofeev and P. Grant, *Statistics for Real-Life Sample Surveys: Non-Simple-Random Samples and Weighted Data*, Cambridge University Press, Cambridge, 2006.

[68] A. Druckman and T. Jackson, *Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model*, Energy Policy, 36 (2008), pp. 3167–3182.

[69] B. Edmonds, *Simplicity Is not Truth-indicative*, World Scientific, Liverpool, 2007, ch. 5, pp. 65–80.

[70] Episcope, *Building Typology Brochure England*, Tech. Rep. September, Episcope, 2014.

[71] R. Everett, A. Horton, and J. Doggart, *Linford Low Energy Houses*, tech. rep., Milton Keynes, 1985.

[72] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*, Wiley, 2011.

[73] B. Falissard, *Focused Principal Component Analysis: Looking at a Correlation Matrix With a Particular Interest in a Given Variable*, Journal of Computational and Graphical Statistics, 8 (1999), pp. 906–912.

[74] S. K. Firth, K. J. Lomas, and A. J. Wright, *Targeting household energy efficiency measures using sensitivity analysis*, Building Research & Information, 38 (2010), p. 31.

[75] E. Foda, D. Allinson, K. J. Lomas, and D. Loveday, *Using Uk Green Deal Assessment Data Of A Stock Of Dwellings To Run A Batch Of Building Energy Models*, in Building Simulation and Optimization Conference, no. Xml, 2014.

[76] V. Gaigalis and R. Skema, *Analysis of fuel and energy transition in Lithuanian households sector and its sustainable development in compliance with the EU policy*, Renewable and Sustainable Energy Reviews, 37 (2014), pp. 273–287.

[77] N. Gilbert, J. C. Hawksworth, and P. A. Swinney, *An Agent-Based Model of the English Housing Market*, Association for the Advancement of Artificial Intelligence, (2009), pp. 30–35.

[78] S. L. Graham, M. Snir, and C. A. Patterson, eds., *Getting Up to Speed: The Future of Supercomputing*, vol. 170, The National Academies Press, 2010.

[79] R. Gupta, *Applying CO2 reduction strategies to existing UK dwellings using GIS-based modelling : a case study in Oxford*. 2006.

[80] R. Gupta and M. Gregg, *Using Urban Energy Modelling To Rapidly Assess Potential For Climate Change Mitigation And Adaptation Of UK Homes*, in Building Simulation and Optimization Conference, London, 2014.

[81] P. Gyberg and J. Palm, *Influencing households' energy behaviour—how is this done and on what premises?*, Energy Policy, 37 (2009), pp. 2807–2813.

[82] V. Haines, K. Lomas, T. Murray, R. Ian, N. Tracy, and G. Monica, *How Trends in Appliances Affect Domestic CO2 Emissions : A Review of Home and Garden Appliances*, Tech. Rep. March 2010, Department of Energy and Climate Change, 2010.

[83] T. Halsey, *CIBSE Guide A: Environmental Design*, tech. rep., 2010.

[84] D. M. Hamby, *A review of techniques for parameter sensitivity analysis of environmental models.*, Environmental monitoring and assessment, 32 (1994), pp. 135–54.

[85] I. Hamilton, J. Milner, Z. Chalabi, P. Das, B. Jones, C. Shrubsole, M. Davies, and P. Wilkinson, *Health effects of home energy efficiency interventions in England: a modelling study.*, BMJ open, 5 (2015), p. e007298.

[86] I. G. Hamilton, D. Shipworth, A. J. Summerfield, P. J. Steadman, T. Oreszczyn, and R. J. Lowe, *Uptake of energy efficiency interventions in English dwellings*, Building Research & Information, 42 (2014), pp. 1–21.

[87] I. G. Hamilton, P. J. Steadman, H. Bruhns, A. J. Summerfield, and R. J. Lowe, *Energy efficiency in the British housing stock: Energy demand and the Homes Energy Efficiency Database*, Energy Policy, 60 (2013), pp. 462–480.

[88] V. S. K. V. Harish and A. Kumar, *A review on modeling and simulation of building energy systems*, Renewable and Sustainable Energy Reviews, 56 (2016), pp. 1272–1292.

[89] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag New York, 2009.

[90] M. He, T. Lee, S. Taylor, S. K. Firth, and K. Lomas, *Coupling a Stochastic Occupancy Model To Energyplus To Predict Hourly Thermal Demand of a Neighbourhood*, in 14th International Conference of IBPSA, 2015, pp. 2101–2108.

[91] M. He, T. Lee, S. Taylor, S. K. Firth, and K. J. Lomas, *Dynamic modelling of a large scale retrofit programme for the housing stock in the North East of England*, in Proceedings of the 2nd International Conference in Urban Sustainability and Resilience, 2014.

[92] P. Heiselberg, H. Brohus, A. Hesselholt, H. Rasmussen, E. Seinre, and S. Thomas, *Application of sensitivity analysis in design of sustainable buildings*, Renewable Energy, 34 (2009), pp. 2030–2036.

[93] J. Henderson, *A review of the relationship between floor area and occupancy in SAP*, Tech. Rep. April, 2009.

[94] ———, *Review of auxiliary energy use and the internal heat gains assumptions in SAP*. 2009.

[95] J. Henderson and J. Hart, *BREDEM 2012 – A technical description of the BRE Domestic Energy Model*, tech. rep., 2012.

[96] H. Herring, *Is Britain a Third World country?*, Energy Policy, 22 (1994), pp. 779–787.

[97] M. Hinnells, B. Boardman, S. J. Darby, G. Killip, and R. Layberry, *Transforming UK homes : achieving a 60 % cut in carbon emissions by 2050*, in European Council for an Energy Efficient Economy, no. June, 2007, pp. 1105–1109.

[98] M. Hinnells, B. Boardman, and R. Layberry, *The UK Housing Stock 2005 to 2050: Assumptions used in Scenarios and Sensitivity Analysis in UKDCM2*, Tech. Rep. March, Environmental Change Institute - University of Oxford, Oxford, 2007.

[99] T. Hinton and J. Thumim, *An analysis of D3 in DECC's energy system models*, tech. rep., Department of Energy & Climate Change, 2014.

[100] HM Government, *The Building Regulations 2010: L1A Conservation of fuel and power in new dwellings*, tech. rep., Department for Communities and Local Government, London, 2010.

[101] ———, *The Building Regulations 2010: L2A Conservation of fuel and power in new buildings other than dwellings*, tech. rep., Department for Communities and Local Government, London, 2010.

[102] D. C. Hoaglin, F. Mosteller, and J. W. Tukey, eds., *Understanding Robust and Exploratory Data Analysis*, Wiley, 2000.

[103] D. Holt and D. Elliot, *Methods of Weighting for Unit Non-Response*, The Statistician, 40 (1991), pp. 333–342.

[104] T. Hong, H. Sun, Y. Chen, S. C. Taylor-Lange, and D. Yan, *An occupant behavior modeling tool for co-simulation*, Energy and Buildings, 117 (2015), pp. 272–281.

[105] C. J. Hopfe and J. L. Hensen, *Uncertainty analysis in building performance simulation for design support*, Energy and Buildings, 43 (2011), pp. 2798–2805.

[106] P. J. Huber and E. M. Ronchetti, *Robust Statistics*, vol. 523, Wiley, 2009.

[107] G. M. Huebner, I. Hamilton, Z. Chalabi, D. Shipworth, and T. Oreszczyn, *Explaining domestic energy consumption – The comparative contribution of building factors, socio-demographics, behaviours and attitudes*, Applied Energy, 159 (2015), pp. 589–600.

[108] M. Hughes, P. Armitage, J. Palmer, and A. Stone, *Converting English Housing Survey Data for Use in Energy Models*, Tech. Rep. January, Cambridge Architectural Research Ltd. and University College London, 2012.

[109] M. Hughes, J. Palmer, V. Cheng, and D. Shipworth, *Sensitivity and uncertainty analysis of England's housing energy model*, Building Research & Information, 41 (2013), pp. 156–167.

[110] M. Hughes, J. Palmer, and P. Pope, *A Guide to The Cambridge Housing Model*, Tech. Rep. November, Department of Energy & Climate Change (DECC), 2013.

[111] IEA-ECBCSP (International Energy Agency Energy Conservation in Buildings and Community Systems Programme), *Stock aggregation, in Annex 31 – Energy related environmental impact of buildings*, tech. rep., IEA, 2005.

[112] R. Ihaka and R. Gentleman, *R: A Language for Data Analysis and Graphics*, Journal of Computational and Graphical Statistics, 5 (1996), pp. 299–314.

[113] International Energy Agency, *Calculation Methods to Predict Energy Savings in Residential Buildings*, in IEA Energy Conservation in Buidings And Community Systems Programm, 1983.

[114] ———, *Mapping the energy future : energy modelling and climate change policy*, OECD Paris, 1998.

[115] International Organization for Standardization, *ISO 9164:1989 Thermal insulation - Calculation of space heating requirements for residential buildings*, 1989.

[116] ———, *BS EN ISO 13790:2008 Energy performance of buildings — Calculation of energy use for space heating and cooling*, 2008.

[117] S. Jaboob, *Stochastic Modelling of Occupants' Activities and Related Behaviours*, phd thesis, University of Nottingham, 2015.

[118] J. N. R. Jeffers, *Two Case Studies in the Application of Principal Component Analysis*, Journal of the Royal Statistical Society. Series C: Applied Statistics, 16 (1967), pp. 225–236.

[119] D. Jenkins, A. Peacock, P. Banfill, D. Kane, V. Ingram, and R. Kilpatrick, *Modelling carbon emissions of UK dwellings – The Tarbase Domestic Model*, Applied Energy, 93 (2012), pp. 596–605.

[120] T. Jensen and É. Chappin, *Towards An Agent-Based Model On Co-Diffusion Of Technology And Behavior: A Review*, in 28th European Conference on Modelling and Simulation, F. Squazzoni, F. Baronio, C. Archetti, and M. Castellani, eds., Ecms, may 2014, pp. 782–788.

[121] W. S. Jevons, *The coal question*, A. M. Kelley, New York, 1865.

[122] D. Johnston, *A physically-based energy and carbon dioxide emission model of the UK housing stock*, phd thesis, Leeds Metropolitan University, 2003.

[123] D. Johnston, R. Lowe, and M. Bell, *An exploration of the technical feasibility of achieving CO2 emission reductions in excess of 60% within the UK housing stock by the year 2050*, Energy Policy, 33 (2005), pp. 1643–1659.

[124] B. Jones, P. Das, Z. Chalabi, M. Davies, I. Hamilton, R. Lowe, A. Mavrogianni, D. Robinson, and J. Taylor, *Assessing uncertainty in housing stock infiltration rates and associated heat loss: English and UK case studies*, Building and Environment, 92 (2015), pp. 644–656.

[125] B. Jones, P. Das, Z. Chalabi, M. Davies, I. Hamilton, R. Lowe, J. Milner, I. Ridley, C. Shrubsole, and P. Wilkinson, *The effect of party wall permeability on estimations of infiltration from air leakage*, International Journal of Ventilation, 12 (2013), pp. 17–29.

[126] B. M. Jones, R. J. Lowe, M. Davies, Z. Chalabi, P. Das, and I. Ridley, *Modelling uniformly porous facades to predict dwelling infiltration rates*, Building Services Engineering Research and Technology, 35 (2013), pp. 408–416.

[127] P. Jones, J. Patterson, and S. Lannon, *Modelling the built environment at an urban scale—Energy and health impacts in relation to housing*, Landscape and Urban Planning, 83 (2007), pp. 39–49.

[128] R. Kannan and N. Strachan, *Modelling the UK residential energy sector under long-term decarbonisation scenarios: Comparison between energy systems and sectoral modelling approaches*, Applied Energy, 86 (2009), pp. 416–428.

[129] M. Kavgic, A. Mavrogianni, D. Mumovic, A. J. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic, *A review of bottom-up building stock models for energy consumption in the residential sector*, Building and Environment, 45 (2010), pp. 1683–1697.

[130] M. Kavgic, D. Mumovic, A. J. Summerfield, Z. Stevanovic, and O. Ecim-Djuric, *Uncertainty and modeling energy consumption: Sensitivity analysis for a city-scale domestic energy model*, Energy and Buildings, 60 (2013), pp. 1–11.

[131] J. Keirstead, M. Jennings, and A. Sivakumar, *A review of urban energy system models: Approaches, challenges and opportunities*, Renewable and Sustainable Energy Reviews, 16 (2012), pp. 3847–3866.

[132] S. Kelly, *Do homes that are more energy efficient consume less energy?: A structural equation model of the English residential sector*, Energy, 36 (2011), pp. 5610–5620.

[133] S. Kelly, D. Crawford-Brown, and M. G. Pollitt, *Building performance evaluation and certification in the UK: Is SAP fit for purpose*, Renewable and Sustainable Energy Reviews, 16 (2012), pp. 6861–6878.

[134] J. P. C. Kleijnen, *Experimental Design for Sensitivity Analysis, Optimization and Validation of Simulation Models*, John Wiley & Sons, Inc., 1998, ch. 6, pp. 173—-223.

[135] M. Kurant and P. Thiran, *Layered complex networks*, Physical Review Letters, 96 (2006), pp. 1–4.

[136] R. H. Landau, M. J. Paez, and C. C. Bordeianu, *A Survey of Computational Physics: Introductory Computational Science*, Princeton University Press, 2008.

[137] E. P. LBNL, *EnergyPlus Version 8.6 Documentation: Engineering Reference*, tech. rep., Lawrence Berkeley National Laboratory, 2006.

[138] T. Lee and R. Yao, *Incorporating technology buying behaviour into UK-based long term domestic stock energy models to provide improved policy analysis*, Energy Policy, 52 (2013), pp. 363–372.

[139] T. Lee, R. Yao, and P. Coker, *An analysis of UK policies for domestic energy reduction using an agent based tool*, Energy Policy, 66 (2014), pp. 267–279.

[140] T. Loga, B. Stein, and N. Diefenbach, *TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable*, Energy and Buildings, 132 (2016), pp. 4–12.

[141] K. J. Lomas and H. Eppel, *Sensitivity analysis techniques for building thermal simulation programs*, Energy and Buildings, 19 (1992), pp. 21–44.

[142] S. W. Lorimer, *A housing stock model of non-heating end-use energy in England verified by aggregate energy use data*, Energy Policy, 50 (2012), pp. 419–427.

[143] B. G. Louis and F. Özmizrak, *A Knowledge-Based Approach for the Validation of Simulation Models : The Foundation*, ACM Transactions on Modelling and Computer Simulation, 6 (1996), pp. 76–98.

[144] R. Lowe, *Technical options and strategies for decarbonizing UK housing*, Building Research and Information, 35 (2007), pp. 412–425.

[145] L. Lutzenhiser, H. Hu, M. Moezzi, A. Levenda, and J. Woods, *Lifestyles, Buildings and Technologies: What Matters Most?* 2012.

[146] C. Macal and M. North, *Agent-based modeling and simulation*, Winter Simulation Conference (WSC), (2009), pp. 86–98.

[147] I. Macdonald and P. Strachan, *Practical application of uncertainty analysis*, Energy and Buildings, 33 (2001), pp. 219–227.

[148] S. M. Macgill and B. Sheldrick, *Monergy: qualifying imperfect measures of need and of performance*, Environment & Planning C: Government & Policy, 6 (1988), pp. 209–223.

[149] P. S. Mallaburn and N. Eyre, *Lessons from energy efficiency policy and programmes in the UK from 1973 to 2013*, Energy Efficiency, 7 (2014), pp. 23–41.

[150] I. Mansouri, M. Newborough, and D. Probert, *Energy consumption in UK households: Impact of domestic electrical appliances*, Applied Energy, 54 (1996), pp. 211–285.

[151] E. Marshall, J. K. Steinberger, V. Dupont, and T. J. Foxon, *Combining energy efficiency measure approaches and occupancy patterns in building modelling in the UK residential context*, Energy and Buildings, 111 (2016), pp. 98–108.

[152] A. Mastrucci, O. Baume, F. Stazi, and U. Leopold, *Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam*, Energy and Buildings, 75 (2014), pp. 358–367.

[153] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, Chapman and Hall/CRC, 2nd ed., 1989.

[154] N. J. McCullen, A. M. Rucklidge, C. S. E. Bale, T. J. Foxon, and W. F. Gale, *Multiparameter Models of Innovation Diffusion on Complex Networks*, SIAM Journal on Applied Dynamical Systems, 12 (2013), pp. 515–532.

[155] E. McKenna, M. Krawczynski, and M. Thomson, *Four-state domestic building occupancy model for energy demand simulations*, Energy and Buildings, 96 (2015), pp. 30–39.

[156] A. Mhalas, *Assessment and decision support for energy performance improvement of dwellings: framework and prototype development*, phd thesis, Teesside University, 2013.

[157] A. Mhalas, M. Kassem, T. Crosbie, and N. Dawood, *A visual energy performance assessment and decision support tool for dwellings*, Visualization in Engineering, 1 (2013), p. 7.

[158] D. C. Mikulecky, *Complexity Science As An Aspect Of The Complexity Of Science*, World Scientific, Liverpool, 2007, ch. 3, pp. 30–52.

[159] J. Milner, C. Shrubsole, P. Das, B. Jones, I. Ridley, Z. Chalabi, I. Hamilton, B. Armstrong, M. Davies, and P. Wilkinson, *Home energy efficiency and radon related risk of lung cancer: modelling study*, BMJ (Clinical research ed.), 348 (2014), p. f7493.

[160] B. Minasny and A. B. McBratney, *A conditioned Latin hypercube method for sampling in the presence of ancillary information*, Computers and Geosciences, 32 (2006), pp. 1378–1388.

[161] L. Mitton, H. Sutherland, and M. Weeks, *Microsimulation Modelling for Policy Analysis: Challenges and Innovations*, 2000.

[162] M. D. Morris, *Factorial Sampling Plans for Preliminary Computational Experiments*, Technometrics, 33 (1991), pp. 161–174.

[163] S. Natarajan and G. J. Levermore, *Predicting future UK housing stock and carbon emissions*, Energy Policy, 35 (2007), pp. 5719–5727.

[164] S. Natarajan, J. Padget, and L. Elliott, *Modelling UK domestic energy and carbon emissions: an agent-based approach*, Energy and Buildings, 43 (2011), pp. 2602–2612.

[165] N.Diefenbach, T.Loga, B.Stein, A.Arcipowska, A.Rakušček, M.ŠijanecZavrl, K.B.Wittchen, J.Kragh, N.Altmann-MavaddatJ., J.Hulme, J.Riley, E.Dascalaki, C.BalarasD., D.Cuypers, M.VanHolm, V.Corrado, I.BallariniT., T.Vimmr, M.Hanratty, B.Sheldrick, C.Roarty, T.Csoknyai, G.Szendrő, S.Hrabovszky-Horváth, L.Ortega, D.Serghides, M.Katafygiotou, N.Nieboer, F.Filippidou, U.Rochard, S.Shanthirabalan, H.Brattebø, M.Jovanovic-Popovic, and D.Ignjatovic, *Energy Performance Indicators for Building Stocks*, Tech. Rep. April 2013, Institut Wohnen und Umwelt GmbH, 2014.

[166] S. Nicol, C. Beer, and C. Scott, *The Age and Construction of English Housing*, IHS BRE Press, 2014.

[167] I. Nikolic and J. Kasmire, *Theory*, in Agent-Based Modelling of Socio-Technical Systems, K. H. van Dam, I. Nikolic, and Z. Lukszo, eds., Springer Netherlands, 2013, ch. 2, p. 284.

[168] M. Nye, L. Whitmarsh, and T. J. Foxon, *Sociopsychological perspectives on the active roles of domestic actors in transition to a lower carbon electricity economy*, Environment and Planning A, 42 (2010), pp. 697–714.

[169] Object Management Group, *OMG Unified Modeling Language*, tech. rep., 2015.

[170] Office for National Statistics, *Harmonised Concepts and Questions for Social Data Sources Introduction to Harmonised Standards*, tech. rep., 2011.

[171] ———, *Labour Force Survey User Guide*, 1 (2011), pp. 0–146.

[172] M. G. Oladokun and I. A. Odesola, *Household energy consumption and carbon emissions for sustainable cities – A critical review of modelling approaches*, International Journal of Sustainable Built Environment, 4 (2015), pp. 231–247.

[173] T. Oreszczyn and R. Lowe, *Challenges for energy and buildings research: objectives, methods and funding mechanisms*, Building Research & Information, 38 (2009), pp. 107–122.

[174] P. Owen, *The Rise of the Machines*, Tech. Rep. June, Energy Saving Trust, 2006.

[175] ———, *The Elephant in the Living Room: How Our Appliances And Gadgets Are Trampling The Green Dream*, tech. rep., Energy Saving Trust, 2011.

[176] ———, *Powering the Nation: Household electricity-using habits revealed*, tech. rep., Energy Saving Trust/DECC/DEFRA, London, 2012.

[177] M. Page-Jones and L. L. Constantine, *Fundamentals of Object-Oriented Design in UML*, Addison-Wesley Professional, Boston, US, 2000.

[178] J. Palmer and I. Cooper, *United Kingdom Housing Energy Fact File*, tech. rep., Department of Energy and Climate, 2012.

[179] ———, *United Kingdom Housing Energy Fact File*, tech. rep., Department of Energy and Climate, 2013.

[180] J. Palmer and N. Terry, *Powering the Nation 2: Electricty use in homes, and how to reduce it*, tech. rep., Energy Saving Trust/DECC/DEFRA, London, 2014.

[181] J. Palmer, N. Terry, S. Firth, T. Kane, D. Godoy-shimizu, and P. Pope, *Energy use at home : models, labels and unusual appliances*, Tech. Rep. October, Cambridge Architectural Research Ltd, Loughborough University, 2013.

[182] J. Palmer, A. Tillson, and P. Armitage, *Comparing the Cambridge Housing Model against the National Energy Efficiency Data-Framework and Meter Readings*, Tech. Rep. July, 2013.

[183] Parliament of the United Kingdom, *Climate Change Act 2008*, tech. rep., 2008.

[184] I. Preston, V. White, J. Thumim, and T. Bridgeman, *Distribution of Carbon Emissions in the UK: Implications for Domestic Energy Policy*, Tech. Rep. March, Centre for Sustainable Energy, 2013.

[185] R Core Team, *R: A Language and Environment for Statistical Computing*, 2016.

[186] P. Rees, D. Martin, and P. Williamson, eds., *The census data system*, vol. 35, John Wiley and Sons Ltd, Chichester, 2003.

[187] I. Richardson, M. Thomson, and D. Infield, *A high-resolution domestic building occupancy model for energy demand simulations*, Energy and Buildings, 40 (2008), pp. 1560–1566.

[188] D. Robinson, N. Filchakova, J. H. Kämpf, A. Rasheed, and J.-L. Scartezzini, *Towards an evolutionary model of city sustainability*, in Proceedings of the Holcim Forum, Beijing, China, 2007, pp. 17–24.

[189] Royal Institution of Chartered Surveyors (RICS) and Department for Communities and Local Government (DCLG), *Energy Efficiency and Value Project*, Tech. Rep. March, London, UK, 2010.

[190] A. Saltelli, *Global sensitivity analysis : the primer*, John Wiley, Chichester, 2008.

[191] A. Saltelli, K. Chan, and E. Scott, *Sensitivity Analysis*, Wiley, 2008.

[192] A. Saltelli, S. Tarantola, and F. Campolongo, *Sensitivity Analysis as an Ingredient of Modeling*, Statistical Science, 15 (2000), pp. 377–395.

[193] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity analysis in practice: a guide to assessing scientific models*, Wiley, 2004.

[194] A. Sancho-Tomás, M. Sumner, and D. Robinson, *A generalised model of electrical energy demand from small household appliances*, Energy and Buildings, 135 (2017), pp. 350–366.

[195] N. Schwarz and A. Ernst, *Agent-based modeling of the diffusion of environmental innovations — An empirical approach*, Technological Forecasting and Social Change, 76 (2009), pp. 497–511.

[196] E. R. Scoffham, *The shape of British housing*, London, 1984.

[197] Scottish Government Social Research (SGSR), *Modelling Greenhouse Gas Emissions from Scottish Housing: Final Report*, tech. rep., Scottish Government Social Research/Cambridge Architectural Research / Cambridge Econometrics / Roger Talbot & Associates Ltd. / Alembic Research, 2009.

[198] L. Shorrock and J. Dunster, *The physically-based model BREHOMES and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock*, Energy Policy, 25 (1997), pp. 1027–1037.

[199] L. D. Shorrock, *A detailed analysis of the historical role of energy efficiency in reducing carbon emissions from the UK housing stock — ECEEE*, Tech. Rep. 1, 2003.

[200] L. D. Shorrock and B. R. Anderson, *A Guide to the Development of BREDEM*, BRE information paper, Building Research Establishment, 1995.

[201] L. D. Shorrock, G. Henderson, and J. H. F. Bown, *BREHOMES: A Physically Based Model of the Energy Use of the United Kingdom Housing Stock*, tech. rep., 1991.

[202] L. D. Shorrock, J. Henderson, and J. Utley, *Reducing carbon emissions from the UK housing stock*, BRE Press, 2005.

[203] E. Shove, *Comfort, cleanliness and convenience: the social organization of normality*, Oxford, 2003.

[204] A. S. Silva and E. Ghisi, *Uncertainty analysis of the computer model in building performance simulation*, Energy and Buildings, 76 (2014), pp. 258–269.

[205] B. M. Sopha, C. a. Klöckner, and E. G. Hertwich, *Exploring policy options for a transition to sustainable heating system diffusion using an agent-based simulation*, Energy Policy, 39 (2011), pp. 2722–2729.

[206] B. M. Sopha, C. A. Klöckner, and E. G. Hertwich, *Adoption and diffusion of heating systems in Norway: Coupling agent-based modeling with empirical research*, Environmental Innovation and Societal Transitions, 8 (2013), pp. 42–61.

[207] S. Sorrell, *Improving the evidence base for energy policy: The role of systematic reviews*, Energy Policy, 35 (2007), pp. 1858–1871.

[208] G. Sousa, B. M. Jones, P. A. Mirzaei, and D. Robinson, *A review and critique of UK housing stock energy models, modelling approaches and data sources*, Energy and Buildings, (2017), pp. 66–80.

[209] C. Spitz, L. Mora, E. Wurtz, and A. Jay, *Practical application of uncertainty analysis and sensitivity analysis on an experimental house*, Energy and Buildings, 55 (2012), pp. 459–470.

[210] P. Steadman, S. Evans, and M. Batty, *Wall area, volume and plan depth in the building stock*, Building Research & Information, 37 (2009), pp. 455–467.

[211] P. Steadman, I. Hamilton, and S. Evans, *Energy and urban built form: an empirical and statistical approach*, Building Research & Information, 42 (2014), pp. 17–31.

[212] R. K. Strand, D. B. Crawley, C. O. Pedersen, R. J. Liesen, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, and D. E. Fisher, *EnergyPlus : A New-Generation Energy Analysis and Load Calculation Engine for Building Design Introduction : What is EnergyPlus ?*, in ACSA Technology Conference, vol. 2, NREL, 2000.

[213] A. J. Summerfield, R. Lowe, and T. Oreszczyn, *Two models for benchmarking UK domestic delivered energy*, (2010), pp. 37–41.

[214] A. J. Summerfield, R. J. Lowe, H. R. Bruhns, J. A. Caeiro, J. P. Steadman, and T. Oreszczyn, *Milton Keynes Energy Park revisited: Changes in internal temperatures and energy usage*, Energy and Buildings, 39 (2007), pp. 783–791.

[215] L. G. Swan and V. I. Ugursal, *Modeling of end-use energy consumption in the residential sector: A review of modeling techniques*, Renewable and Sustainable Energy Reviews, 13 (2009), pp. 1819–1835.

[216] S. Taylor, D. Allinson, S. Firth, and K. Lomas, *Dynamic Energy Modelling of UK Housing: Evaluation of Alternative Approaches*, 13th Conference of International Building Performance Simulation Association, (2013), pp. 745–752.

[217] N. Terry, *Energy and Carbon Emissions: The Way We Live Today*, UIT Cambridge, 2011.

[218] W. Tian, *A review of sensitivity analysis methods in building energy analysis*, Renewable and Sustainable Energy Reviews, 20 (2013), pp. 411–419.

[219] M. A. Tovar, *The structure of energy efficiency investment in the UK households and its average monetary and environmental savings*, Energy Policy, 50 (2012), pp. 723–735.

[220] UK Government, *Building Act 1984*, 1984.

[221] J. I. Utley and L. D. Shorrock, *Domestic Energy Fact File*, tech. rep., BRE Housing, 2003.

[222] D. Vickers and P. Rees, *Creating the UK National Statistics 2001 output area classification*, Journal of the Royal Statistical Society. Series A: Statistics in Society, 170 (2007), pp. 379–403.

[223] D. Virk and M. Eames, *CIBSE Weather Files 2016 release: Technical Briefing and Testing*, tech. rep., Chartered Institution of Building Services Engineers (CIBSE), 2016.

[224] I. Vlaev and P. Dolan, *From changing cognitions to changing the context : a dual-route model of behaviour change*. 2009.

[225] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442.

[226] M. Wetter and P. Haves, *A modular building controls virtual test bed for the integration of heterogeneous systems*, IBPSA-USA Journal, 3 (2008), pp. 69–76.

[227] U. Wilke, F. Haldi, J. L. Scartezzini, and D. Robinson, *A bottom-up stochastic model to predict building occupants' time-dependent activities*, Building and Environment, 60 (2013), pp. 254–264.

[228] A. J. Wright, *What is the relationship between built form and energy use in dwellings?*, Energy Policy, 36 (2008), pp. 4544–4547.

[229] P. Wyatt, *A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England*, Energy Policy, 60 (2013), pp. 540–549.

[230] R. Yao and K. Steemers, *A method of formulating energy load profile for domestic buildings in the UK*, Energy and Buildings, 37 (2005), pp. 663–671.

[231] T. Zhang, P.-O. Siebers, and U. Aickelin, *A three-dimensional model of residential energy consumer archetypes for local energy policy design in the UK*, Energy Policy, 47 (2012), pp. 102–110.

[232] Y. Zhang, *Use jEPlus as an efficient building design optimisation tool*, in CIBSE ASHRAE Technical Symposium, no. April, London, 2012, pp. 1–12.

[233] F. Zhao, S. H. Lee, and G. Augenbroe, *Reconstructing building stock to replicate energy consumption data*, Energy and Buildings, 117 (2016), pp. 301–312.

[234] J.-P. Zimmermann, M. Evans, J. Griggs, N. King, L. Harding, P. Roberts, and C. Evans, *Household Electricity Survey A study of domestic electrical product usage*, tech. rep., Oxford, 2012.