



# A grey-box modelling methodology to express home heat-energy usage as statistical distributions — case studies in urban Ireland

Paul Beagon · Fiona Boland

Received: 25 August 2021 / Accepted: 30 April 2022 / Published online: 31 May 2022  
© The Author(s) 2022

**Abstract** Home energy retrofit has recurred in public policy throughout recent decades. However, the savings in energy usage attributable to home retrofit have remained difficult to accurately predict. Occupants cause prediction inaccuracies by varying different factors, especially heating setpoints temperatures and heating patterns. Acting together, such occupant factors result in distributions — not single values — of heat-energy usage, even among similar homes. Datasets of heat-energy distributions can be found by building performance simulation using modern grey-box models. This study presents a methodology to simulate grey-box models of home heating through ranges of heating setpoints and patterns. An entire process to calibrate, validate and simulate at a large scale is described, and then demonstrated using case studies. Grey-box models, written in Modelica

language, can conveniently simulate through large ranges of occupant factors. The case studies exploited this advantage of grey-box models to simulate empirical data on occupant factors. (For instance, empirical data found that home heating setpoints shifted before and after home energy retrofit.) In doing so, the datasets of simulation results enabled the exploration of home heat-energy usage with the normal and Weibull statistical distributions. Additionally, the heat-energy distributions of case-study homes were statistically tested, first for retrofit savings, second for equality to each other and third for equality to an official heat-energy estimate. Results demonstrate that home heat-energy usage, at a large scale, is best expressed as a Weibull distribution not normality. After home energy retrofit, heat-energy usage displays less variation (in general), less skewness, and thus becomes closer to normality. Occupant factors were found to vary home heat-energy usage into distinct distributions, even within similar homes. Therefore, in most case-study homes, heat-energy usage did not equal an official estimate. Finally, shallow retrofit of a modern home in Ireland fails to save heat-energy usage by most occupants.

---

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12053-022-10038-9>.

---

P. Beagon (✉)  
Sustainable Energy Authority of Ireland, 3 Park Place,  
Hatch Street Upper, Dublin 2, Ireland  
e-mail: paul.beagon@seai.ie

F. Boland  
Data Science Centre, Division of Biostatistics  
and Population Health Sciences, Royal College  
of Surgeons in Ireland, 123 St Stephen's Green, Dublin 2,  
Ireland  
e-mail: fionaboland@rcsi.ie

**Keywords** Home energy retrofit · Building performance simulation · Modelica · AixLib library · Weibull distribution · Normal distribution · Distribution goodness-of-fit · Anderson–Darling test

## Introduction

A consensus has emerged on the need for home energy retrofit (HER) and its delivery at large scale (Gupta et al., 2015; Mathiesen et al., 2016; McKenna et al., 2013). Proponents of HER foresee its contribution to different goals: reducing greenhouse gas, combatting energy poverty, — in addition to —, saving energy (DCCAE, 2019; Dixon & Eames, 2013; Fylan et al., 2016; Sdei et al., 2015). Despite these august goals, a gap remains in predicting the energy-usage savings attributable to HER. One prediction approach is to model homes before and after retrofit, and then simulate their performance (Swan & Ugursal, 2009).

Building performance simulation (BPS) can predict home energy usage under different scenarios, but often fail to incorporate heating variations caused by occupants (Huebner et al., 2013; Shipworth et al., 2010). Inaccurate model parameters, such as air changes per hour (ACH) or building fabric properties, further compound the inaccuracies in simulated energy usage. Occupants cause the gap between simulated and measured energy usage for two main reasons during simulation: (1) unrealistic internal temperatures (Jones et al., 2016; Teli et al., 2016) and (2) failure to apply dominant occupancy profiles (Zahiri & Elsharkawy, 2018). Put simply, different occupants operate homes differently.

Occupants also affect HER by causing the *rebound effect*. In the context of home heating, rebound effect is post-retrofit overheating or “temperature take back” (Sorrell et al., 2009), also known as “comfort taking” (Hamilton et al., 2016). Temperature take back occurs when occupants prefer warmer internal temperatures instead of maximising energy and financial savings. In extreme cases, the rebound effect may negate energy efficiency savings, an outcome termed “backfire” where the intervention actually increases overall energy usage (Druckman et al., 2011). Druckman et al. found that the rebound effect results in only a portion of the greenhouse gas (GHG) reductions estimated by engineering calculations. Furthermore, the largest *percentage* rebound hampered potential GHG reductions by low-income households, according to UK analysis (Chitnis et al., 2014). That analysis found that the low-income households were mainly affected by *direct* rebound, whereby occupants use more of an energy service after an intervention decreases its effective price.

In summary, the differences in home operation mean that a single energy-usage value cannot represent every home, even structurally similar homes comprising a neighbourhood. A decade ago, neighbourhoods were proposed as a spatial scale offering the highest potential for large-scale HER (Koch & Girard, 2011; Koch et al., 2012). Koch & Girard defined neighbourhoods as groups of structurally similar homes — proposing that their energy usage is best expressed as *statistical distributions*. This paper now presents a methodology and case studies to calculate the heat-energy usage by structurally similar homes as a statistical distribution.

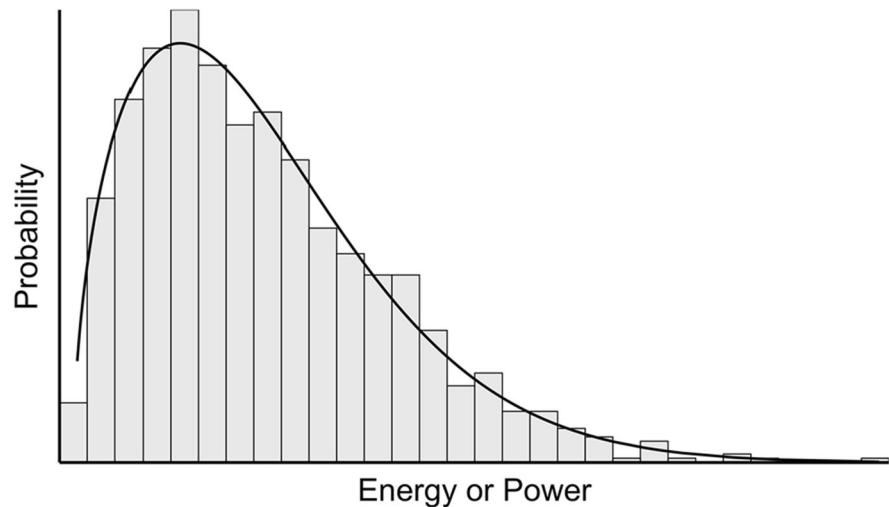
### Statistical distributions of home energy usage

Previous research has expressed home energy usage as statistical distributions (Irwin et al., 1986; Munkhammar et al., 2014). Both research papers describe how meter data was used to find distributions of electricity usage — an energy *vector*. Energy vectors allow transfer, in space and time, a given quantity of energy, hence making it available for use distantly (Orecchini & Santiangeli, 2011). When the energy quantity is used, it provides useful energy services such as heating, cooling, transport and so on. In contrast to the previous research, this study finds the statistical distribution of heat-energy usage for space heating — an individual energy *service*. Furthermore, this study compares heat-energy distributions before and after HER. In contrast, savings in electricity usage are difficult to achieve by HER because of demand from multiple energy services, such as lighting, cooking and supplementary heating (Irwin et al., 1986), and nowadays dishwashers and other appliances (Munkhammar et al., 2014).

Both Irwin et al. and Munkhammar et al. fitted Weibull distributions to positively skewed datasets of electricity usage. Unlike the popular normal distribution, a Weibull distribution covers only positive values in different asymmetric shapes. Figure 1 displays a Weibull distribution fitted to a positively skewed histogram.

Multiple techniques exist to fit a distribution to a dataset of energy values. Munkhammar et al. used the Kolmogorov-Smirnov (KS) test to check the fit of electricity usage to Weibull and log-normal distributions. A similar test by Anderson–Darling (AD) improves the KS test by weighting the values

**Fig. 1** Positively skewed histogram and fitted Weibull distribution



at the distribution tails (Keller, 2011), where non-normality often manifests (Nelson, 2018). Therefore, the AD test is capable of detecting the differences between overlapping normal and Weibull distributions (Engmann & Cousineau, 2011). For these reasons, this study used the AD test for distribution fit. Abbreviations, including KS, AD and others, plus symbols are organised in Table 1.

#### Grey-box models for buildings

Building models can be categorised into three general approaches: white-box, black-box and grey-box. The grey-box approach combines the physics of the white-box approach, and the statistics or machine learning of the black-box approach. Functional labels for the three approaches are “physical”, “machine learning” and “hybrid” respectively (Fouquier et al., 2013). Two decades have passed since a grey-box model was used to represent building envelopes (Déqué et al., 2000). In the meantime, other studies have quantified uncertain parameters in building models using historical time-series data (Bacher & Madsen, 2011; Brastein et al., 2018; Harb et al., 2016; Reynders et al., 2014). This process is known as parameter estimation or “tuning” and is used to calibrate grey-box models.

One type of grey-box building models, called lumped-capacitance, has received much attention and been evaluated (Vivian et al., 2017). That evaluation described how lumped capacitance models simplify






(or lump) a building’s distributed thermal mass into a discrete number of thermal capacitances. Within each model, thermal resistances interconnect the thermal capacitances forming thermal networks; hence, lumped-capacitance is often referred to as “thermal networks”, “resistance–capacitance” or simply “RC”. Henceforth, lumped-capacitance or thermal network models are referred to as “RC models”. RC models are considered grey-box if their parameters are estimated or “tuned” using available weather and energy-use datasets, as opposed to analytical calculation. This study incorporated parameter estimation into its Methodology (section [Calibration procedure and validation](#)).

After model calibration, the system domain of RC models “can be solved analytically, thus avoiding problems of convergence and stability” (Vivian et al., 2017). Vivian et al. concluded that RC models “reliably calculate” overall energy demand but only high-order models, defined by many thermal capacitances, can calculate transient hourly demands. Low-order models are, however, easier to solve during simulation. Easier simulation enables scaling up building simulation to city district (Lauster et al., 2014a, b), or to urban energy in an arid climate (Zekar & El Khatib, 2018). The latter research found that RC models of cooling load provided a “limited loss of accuracy” and “satisfying” level of performance. Nowadays, grey-box models still offer the three main advantages identified two decades ago (Déqué et al., 2000), with minor updates:

**Table 1** Abbreviations and symbols

<b>Abbreviations</b>	
ACH	Air changes per hour
AD	Anderson–Darling, a goodness-of-fit test for statistical distributions
ASHRAE	American Society of Heating, Refrigeration and Air-Conditioning Engineers
BPS	Building performance simulation
CI	Confidence interval, 95% used in this study
CSO	Central Statistics Office
CV(RMSE)	Coefficient of variation of the root mean square error
DASSL	Differential / algebraic system resolver
DEAP	Dwelling Energy Assessment Procedure
EU	European Union
GHG	Greenhouse gas
GOF	Goodness-of-fit
HER	Home energy retrofit
IWEC	International Weather for Energy Calculation
KS	Kolmogorov-Smirnov, a goodness-of-fit test for statistical distributions
kWh	kilowatt hour
MFH	Multi-family house
NMBE	Normalised mean bias error
NREL	National Renewable Energy Laboratory
PID	Proportional-integral-derivative
PDF	Probability distribution function
QQ	Quantile–quantile
RC	Resistance–capacitance
SFH	Single-family house
SEAI	Sustainable Energy Authority of Ireland
Std dev	Standard deviation
TABULA	Typology Approach for Building Stock Energy Assessment
<b>Symbols</b>	
$\alpha_{conv}$	Convective heat transfer coefficient
$\alpha_{rad}$	Radiative heat transfer coefficient
h	Hour
m	Mean of measured data
$m_i$	Measured data, indexed by i
$m_r$	Statistical moment of $r^{\text{th}}$ order
n	Sample size
p	Number of parameters
s	Standard deviation of a sample
$s_i$	Simulated data, indexed by i
$\bar{x}$	Mean average of a dataset x, for example energy use
$\gamma$	Skewness of a statistical distribution

- Limited number of parameters to be found prior to simulation,
  - Fewer equations to solve during simulation, enabling larger spatial scale and
  - Flexibility to detail a model subsystem amongst other model components.
- Modelling home archetypes to calculate energy usage
- Having introduced grey-box building models, definition of the homes represented by the models is necessary. For this study, separate building models must represent the same home in two states: before and

Construction Year Class	Additional Classification	SFH Single Family House	TH Terraced House	MFH Multi Family House	AB Apartment Block
1930 ... 1949	generic	 IE.N.SFH.03.Gen	 IE.N.TH.03.Gen		
1950 ... 1966	generic	 IE.N.SFH.04.Gen	 IE.N.TH.04.Gen		 IE.N.AB.04.Gen

**Fig. 2** TABULA webtool displaying an extract of Ireland’s building matrix. The highlighted cell contains a single-family house archetype TH.03.Gen: a terraced house constructed during 1930–1949 (TABULA, 2016)

after home energy retrofit. Henceforth referred to as *as-built* and *retrofit* states. As already mentioned, homes in the same neighbourhood are structurally similar. Thus, one building archetype can define the structure and fabric of many similar homes forming a neighbourhood.

The archetype approach has been used to understand the *aggregated* impact of energy efficiency policies and technologies (Monteiro et al., 2017; Reinhart & Cerezo Davila, 2016). Similar to this study, Monteiro et al. selected the neighbourhood spatial scale, delineating four neighbourhoods by construction period and major urban intervention. They also identified three stages in archetype generation: *classification*, *parameterisation*, and *modelling*. Regarding the first stage of archetype generation, an existing classification of European homes is now introduced.

### Archetypes from TABULA typology

Project TABULA, conducted during 2009–2012, was formally named “Typology Approach for Building Stock Energy Assessment”. It was the second of three Intelligent Energy Europe projects, including DATA-MINE (2006–2008) and EPISCOPE (2013–2016). A project summary (Loga, 2010), answers the question “What is TABULA about?” as follows:

*The objective ... is to create a harmonised structure for European Building Typologies. Each participating country will on that basis develop a National Building Typology, that is*

*a set of model buildings with characteristic energy related properties. The project focuses on residential buildings ...*

Each National Building Typology produced by TABULA is a “Building Type Matrix” hereinafter “building matrix”, that standardised the classification of residential buildings (Loga et al., 2016). This study used Ireland’s building matrix in section [Home archetypes](#) of the Methodology. Figure 2 displays the standardised matrix axes:

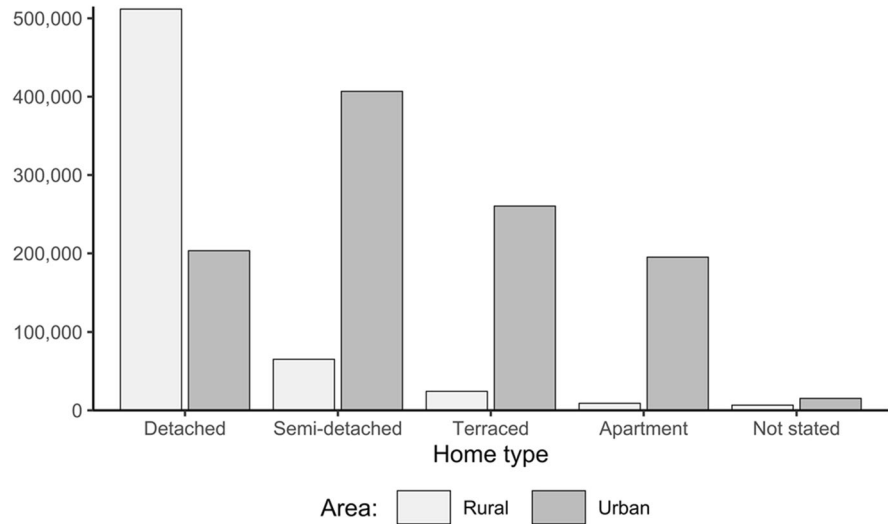
- Building age classes: the construction period of the building.
- Building size classes: clustering the buildings by size, such as single-family houses.

Each home archetype within the building matrix was accompanied by a complete dataset comprising general features such as number of storeys, geometric data of volume and envelope area, thermal properties such as U-values and supply system features such as heat generator efficiency (Loga, 2010). Supply systems pertaining to home energy retrofit were the home heating and cooling systems. TABULA classified heating and cooling systems into three types: (1) initial reference system, termed “as-built” in this study, (2) standard retrofit or (3) advanced retrofit. TABULA archetypes also included energy-use parameters such as air change rates and heating fuels. Simple occupant schedules were assumed by TABULA, thus *not used* in this study because they

**Table 2** Home-type numbers in Ireland, split by rural and urban areas (CSO, 2020). Bed-sit numbers are combined into apartment numbers

Home type	Rural area (number)	Urban area (number)	Total (number)
Detached	511,787	203,346	715,133
Semi-detached	65,150	406,798	471,948
Terraced	24,250	260,319	284,569
Apartment	8997	195,148	204,145
Not stated	6644	15,226	21,870
Total	616,828	1,080,837	1,697,665

**Fig. 3** Numbers of Irish rural and urban home types in 2016. Single-family houses exceed 80% of urban homes



lacked the variation caused by occupants. TABULA assumed a single Irish climate zone but allocates multiple climate zones to larger countries, for instance three zones to Italy.

By focusing on homes, TABULA attempted to predict the impact of energy efficiency measures and select effective strategies for home energy retrofit (Ballarini et al., 2014; Loga et al., 2016). Estimates of building energy relied on a quasi-steady-state calculation over an entire season (TABULA, 2013). Both archetype definitions and energy calculations were published online (TABULA, 2016), and were adopted by other research into Belgium homes (Reynders et al., 2014) and six other countries (Coma et al., 2019).

### Home types in Ireland

The home archetypes defined in TABULA reflect the popularity of single-family houses (SFHs) in Ireland, over multi-family houses (MFHs). Furthermore,

Ireland's 2016 census confirmed the popularity of SFHs and disaggregated their quantities by type: detached, semi-detached and terraced (CSO, 2020).

Disaggregated numbers of home types in Table 2 reveal that detached houses comprised over 80% of rural households, motivating previous research into their retrofit (Ahern et al., 2013). In contrast, the 2016 census reported that urban or "town" households span the three SFH types more evenly, with semi-detached and terraced houses dominating (Fig. 3). The numbers of urban homes in descending order of house type were 406,798 semi-detached, 260,319 terraced and 203,346 detached (CSO, 2020). Given that large-scale retrofit applies to urban areas, this study's case studies focus on the popular semi-detached and terraced home types.

Apartments or other multi-family homes (including bed-sits) constituted 195,148 (under 20%) of all urban households, leaving 15,226 home types not stated (CSO, 2020). Another factor in home energy usage is that Census 2016 recorded an increasing

number of persons per household, reversing the downward trend since 1966 (CSO, 2018). Moreover, increasing household sizes were confined to urban areas.

Turning to the causes of dominant home types, the construction of single-family houses permeated both rural and urban areas of Ireland during the last century. McManus (2011) cited “The typical Irish home, whether in the city or countryside, is a one or two storey, single or semi-detached cottage...” (Pfretzschner, 1965); before remarking that Pfretzschner’s insight described Irish housing in any twentieth century decade as follows:

*The predominance of the standardised three- or four-bedroom, semi-detached or detached house, was not challenged until the 1990s when there was a surge in apartment provision.*

This characterisation of semi-detached and detached homes as “standardised” implies their structural similarity and suitability for classification into home archetypes. Multi-family homes, such as apartments, remained a minor part of Irish twentieth century housing supply. Apartments were perceived as working class, associated with overcrowded tenements, and less popular than houses (McManus, 2011). An official recommendation for the aforementioned “cottages”, i.e. suburban two-storey houses, instead of urban flats appeared in the 1939/43 Report of Inquiry (Rowley, 2015). Thus, purpose-built apartments gradually reached 5% of new builds during the 1980s, only ramping up construction rates during the 1990s (Energy Action, 2014). The resultant 85% growth in apartments during 2002–2016, captured in the census (CSO, 2018), means the average apartment is significantly more modern than the average single-family house.

In addition to older construction age, single-family houses dominate heat-energy usage due to their larger average size and higher specific heat demand. A review of building stock retrofit in Germany forecasted that single-family houses will account for approximately two thirds of heat demand until 2050 (McKenna et al., 2013). Hence, this study selects single-family homes constructed before the 1997–2002 changes in building regulations (Dineen & Ó Gallachóir, 2011), as those most likely to benefit from home energy retrofit. However, the benefits attributable to HER vary because of the variation in energy usage caused

by occupants. For instance, different occupants prefer different internal temperatures controlled by *setpoints*, and warm their homes at different times of the day called *heating patterns*. Since setpoints and heating patterns are critical factors in home heating calculations, they were empirically defined in two Methodology sections: Data sources for simulations, and Heating and internal gains schedules.

#### Calculation of energy savings attributable to HER

As already mentioned, TABULA defined standard and advanced retrofits per archetype that differed in their upgrade of each archetype’s heat generator (TABULA, 2016). A standard retrofit upgraded a boiler, whereas an advanced retrofit replaced a boiler and switched fuel by deployment of an electrical heat pump. To remain independent of fuel type, energy savings attributable to retrofit were measured in sensible heat energy for space heating. This study analyses the *sensible heat energy* used to provide one energy service: *space heating*, before and after a *deep retrofit*. The three preceding terms are now defined:

*Sensible heat energy* or *heat energy* flows as convective energy from heating equipment to maintain internal air at a heating setpoint. For brevity, *sensible heat energy* is hereinafter shortened to *heat energy*.

*Space heating* is the heat energy required to offset the thermal losses across the building envelope by conduction and radiation, as well as air infiltration and ventilation, in an effort to maintain the living space at a comfortable temperature (Swan & Ugursal, 2009).

*Deep retrofit* combines multiple measures to improve the building envelope fabric and upgrade heat generation. Once a building envelope is made consistent, it reduces outside air infiltration and other thermal losses that must be offset by space heating. In contrast, *shallow retrofit* comprises one or two maintenance measures, such as increased boiler efficiency or new glazing. Deep-retrofit designs are often implemented to deliver policy-driven savings in energy or GHG emissions. Policy-driven targets vary around 60% savings (Arcipowska et al., 2017), but have extended to 75% savings (GBPN, 2013).

## Contribution

This study contributes to research by providing a methodology and case study examples to parameterise and evaluate the statistical distributions of heat-energy usage for home space heating. Such distributions of heat-energy usage are particularly relevant to large-scale home energy retrofit, for example at neighbourhood scale. While the case studies are in Ireland, the methodology can be applied internationally, most straightforwardly in countries covered by the TABULA research project. The preceding Introduction covered the points supporting the contribution: the need for home energy retrofit and its large-scale delivery.

In addition, this study contributes to the *research gap* that building energy models typically lack empirical data on internal temperature and heating duration. Furthermore, the same models ignore occupant self-rationing known as the prebound effect (Sunikka-Blank & Galvin, 2012). These findings apply to homes in general (Huebner et al., 2013; Shipworth et al., 2010), and social housing in particular (Jones et al., 2016; Teli et al., 2016; Zahiri & Elsharkawy, 2018). Therefore, empirically sourced heating temperature setpoints and heating patterns are defined in two sections: Data sources for simulations and Heating and internal gains schedules. In total, the aims of this study are fourfold:

- 1) Present a calibration and validation methodology for grey-box models coded in Modelica language, using heat-energy usage data from simulation of equivalent white-box models.
- 2) Convert the heat-energy usage from the simulations into statistical distributions and test their fit to the normal and Weibull distributions.
- 3) Explore the different heat-energy distributions, resulting from different heating setpoints and patterns operated by occupants.
- 4) Use the distributions to calculate mean confidence intervals (CIs) and test the hypothesis that heat-energy usage actually equals an official single value.

The following Methodology covers the selection, calibration, and large-scale simulation of grey-box

models that represent home heating. To simulate different occupant operation, different heating setpoints and heating patterns are defined. Both definitions originate from published empirical data. The resulting distributions and statistics of heat-energy usage by the modelled homes are presented and explored in the Results and Discussion. Finally, the paper closes with Conclusions that include implications and future work.

## Methodology

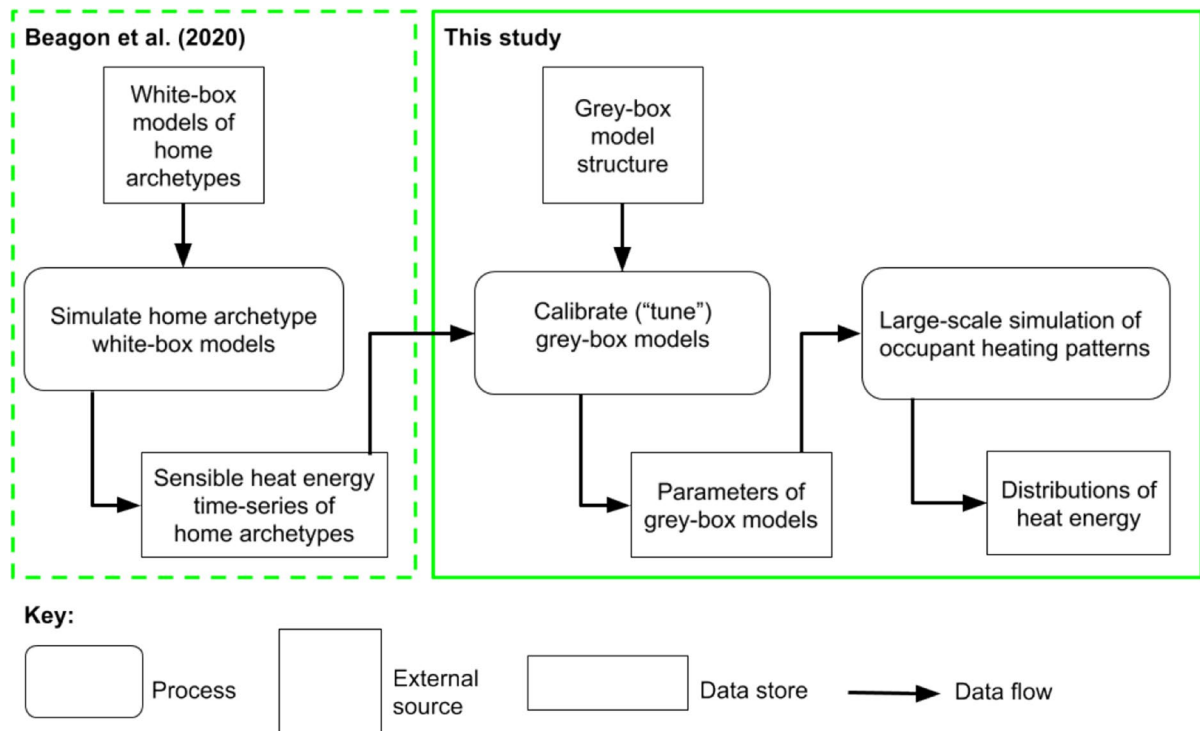
After the structure of a grey-box model was selected, it was calibrated for each home archetype. Calibration of grey-box models relied on the energy usage simulated by white-box models of the same home archetypes. Both white-box and grey-box models were simulated with identical weather files, internal gains, ventilation and infiltration; for more details, see the companion paper (Beagon et al., 2020). That paper details the creation of the white-box models — the first step in the methodology flow undertaken by this study (Fig. 4).

### Home archetypes

Since this study investigated large-scale HER, for example at neighbourhood scale, the methodology selected *urban* home archetypes. As discussed already in section Home types in Ireland, single-family homes dominate Irish housing, primarily as semi-detached and terraced houses in urban areas.

TABULA's typography classified both semi-detached and terraced homes as "terraced"; identified by the letters "TH" within the archetype names (TABULA, 2016). The three archetypes selected as case studies are hereinafter referred to as "TH03", "TH06" and "TH07". The archetypes differed by their construction periods, where TH03 and TH07 are the oldest and youngest respectively (Table 3). All three archetypes shared the same urban area of Dublin, Ireland's capital city. According to the home archetype brochure, TH03 was "very common" in Dublin's 1930s and 1940s building stock, TH06 was "commonly built" in Dublin during 1978–1982 and TH07 was a "very typical house" built in Dublin during the 1980s (Badurek et al., 2014).





**Fig. 4** Data flow diagram of modelling and simulation by this study

**Table 3** Summary details of case-study home archetypes, including unique construction year class and building type of either terraced or semi-detached

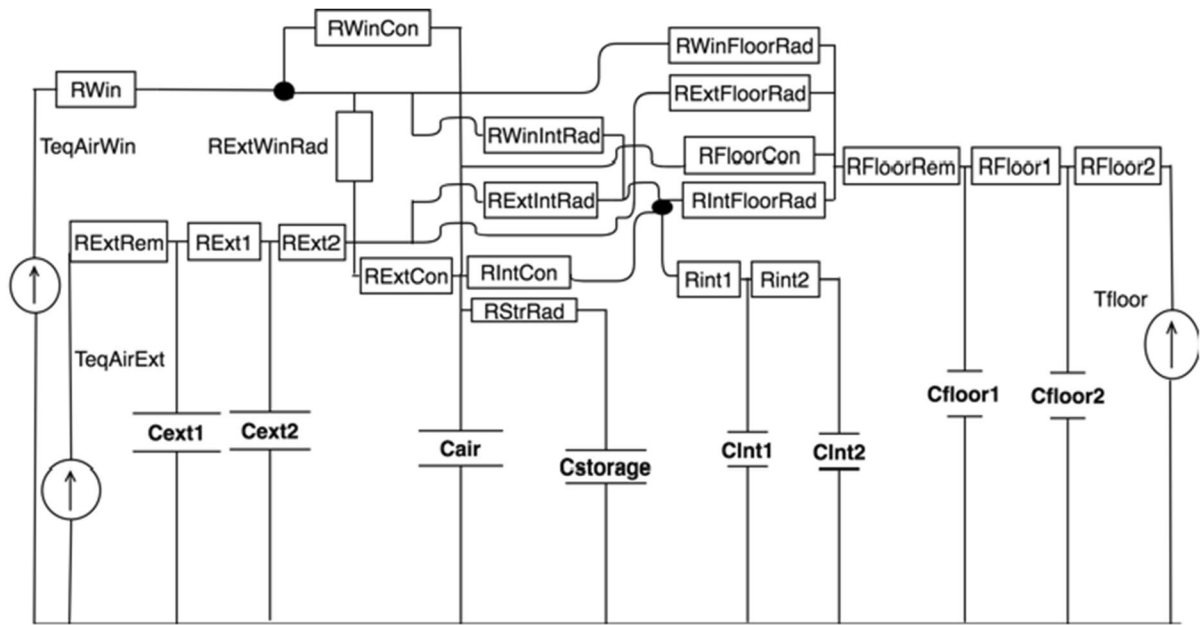
Archetype	Full name	Construction year class	Building type	Additional classification
TH03	IE.N.TH.03.Gen	1930–1949	Terraced	Mass concrete
TH06	IE.N.TH.06.HBlock	1978–1982	Semi-detached	Hollow block
TH07	IE.N.TH.07.Hblock	1983–1993	Terraced	Hollow block

### Selection of Grey-box model

Section [Grey-box models for buildings](#) introduced thermal or RC networks to model building performance. Each grey-box model in this study represented the space heating of a home archetype as one system, containing a RC network to simplify the building fabric subsystem. By simplifying the building fabrics, the grey-box models facilitated the many simulations across a range of heating setpoints and heating patterns. As expected, the range of setpoints and patterns resulted in statistical distributions of heat-energy usage — the main results of this study.

All grey-box models were implemented in Mod-elica code and simulated in the Dymola environment using its default DASSL solver (differential algebraic system resolver). This study used the Mod-elica library *AixLib* and its subsystems (Müller et al., 2016) to implement an existing RC network structure (Lauster et al., 2014a, b). That model separated equivalent air temperatures of wall and windows for each building orientation. Thereby, the effect of solar radiation was more accurately calculated, including long-wave radiation (Lauster et al., 2014a).

Radiative heat traversed the different components of the building fabric. For instance, heat energy transferring from windows to internal walls was impeded



**Fig. 5** Thermal network (RC) subsystem representing the building fabric of a house, adapted from (Lauster et al., 2014a, b). Walls and floors are configured as 3R-2C components

by resistor  $R_{WinIntRad}$ . Specific resistor values are based on a radiative heat transfer coefficient ( $\alpha_{rad}$ ), assumed constant in the simulated temperature range at  $5 \text{ W}/(\text{m}^2\text{K})$ . Concerning convective heat transfer, a coefficient ( $\alpha_{conv}$ ) of  $2.7 \text{ W}/(\text{m}^2\text{K})$  is assumed. Both  $\alpha_{rad}$  and  $\alpha_{conv}$  values were sourced from (Lauster et al., 2014a, b). The  $\alpha_{conv}$  coefficient determines resistances, for example, convective heat transfer at windows  $R_{WinCon}$ . Other resistances and capacitances were specific to the house fabric construction: external wall (ext), internal walls and objects (int), and floor plate (floor).

In order to reduce the over-responsiveness of simulated internal temperatures, lumped capacitances were split in two (Tindale, 1993). For example, the total external wall capacitance was apportioned between  $C_{ext1}$  and  $C_{ext2}$  that interconnected by resistors forming a 3R-2C component. These external wall components were ordered as:  $R_{extRem}$ ,  $C_{ext1}$ ,  $R_{ext1}$ ,  $C_{ext2}$ ,  $R_{ext2}$  (Fig. 5). The resistance (and some capacitance) values of the external wall increased after retrofit. Such increases reflect the fabric retrofit measures of wall insulation and glazing replacement. The next section, Automated model calibration, describes how this study quantified

these RC values by “tuning” grey-box models to datasets of white-box models’ heat-energy usage.

#### Automated model calibration

Model calibration is the process of “tuning” selected model parameters to obtain a better agreement between the model’s simulated or *predicted* behaviour and observed or *measured* behaviour (Dassault Systèmes AB, 2017). In this study, selected parameters of the grey-box models were tuned to match the simulated heat-energy usage by the white-box models.

A *calibrate* task resides in the Dymola optimisation library for Modelica (Elmqvist et al., 2005). This task iterated through the simulation of a grey-box model and compared its energy-usage dataset to those from previous simulations of white-box models (Beagon et al., 2020). By iterating through simulations and comparisons, an optimisation was achieved by automatic re-estimations of selected model parameters. The optimisation objective was to achieve minimum error between the two datasets of heat-energy usage. The selected parameters for tuning were the three

external wall resistors and its two capacitances:  $R_{\text{ext1}}$ ,  $R_{\text{ext2}}$ ,  $R_{\text{extRem}}$ ,  $C_{\text{ext1}}$  and  $C_{\text{ext2}}$  (Fig. 5). After updating the tuned parameters with their new estimated values, the grey-box models were deemed calibrated.

### Model validation indices

Validation of calibrated models used two goodness-of-fit (GOF) indices. The errors between *monthly* heat-energy usage from simulations of white-box and calibrated grey-box models were quantified by both indices. First, the coefficient of variation of the root mean square error (CV(RMSE)) quantified the variation of errors. Second, normalised mean bias error (NMBE) quantified the average of errors. In other words, CV(RMSE) and NMBE quantify the standard deviation and mean of the errors between two datasets (Reddy & Maor, 2006).

Referring to Eqs. (1) and (2),  $\bar{m}$  represents the mean of measured data. Throughout this study, the existing white-box dataset were represented by the “measured” data of month  $i$  ( $m_i$ ), whereas the grey-box model produced the “simulated” data of month  $i$  ( $s_i$ ). Since both models produced monthly results, variable  $n$  equalled twelve. Finally,  $p$  was the number of adjustable model parameters (Robson & McCartan, 2016); however, this study adopted proposed values:  $p=1$  for CV(RMSE) and  $p=0$  for NMBE (Reddy & Maor, 2006).

$$\text{CV(RMSE)} = \frac{1}{\bar{m}} \sqrt{\frac{\sum_{i=1}^n (m_i - s_i)^2}{n - p}} \quad (1)$$

$$\text{NMBE} = \frac{1}{\bar{m}} \sqrt{\frac{\sum_{i=1}^n (m_i - s_i)}{n - p}} \quad (2)$$

ASHRAE<sup>1</sup> Guideline 14 stipulates two necessary GOF validation criteria:  $\text{CV(RMSE)} < 15\%$ , and  $-5\% < \text{NMBE} < 5\%$  (Reddy & Maor, 2006). The larger CV(RMSE) criterion reflects its absolute measurement of all errors (1), whereas the smaller NMBE criterion reflects its cancelling out of positive and negative bias (2).

<sup>1</sup> American Society of Heating, Refrigeration and Air-Conditioning Engineers.

### Calibration procedure and validation

Validated grey-box models met the ASHRAE validation criteria using *monthly* heat-energy usage results from two simulations. One simulation was under a static “steady-state” schedule, whereas the second simulation was under the dynamic DEAP (Dwelling Energy Assessment Procedure) schedule. The steady-state simulations had constant heating, but an unoccupied home. The DEAP simulations had heating and an occupied home during the morning and evening only (07:00–09:00 and 17:00–23:00).

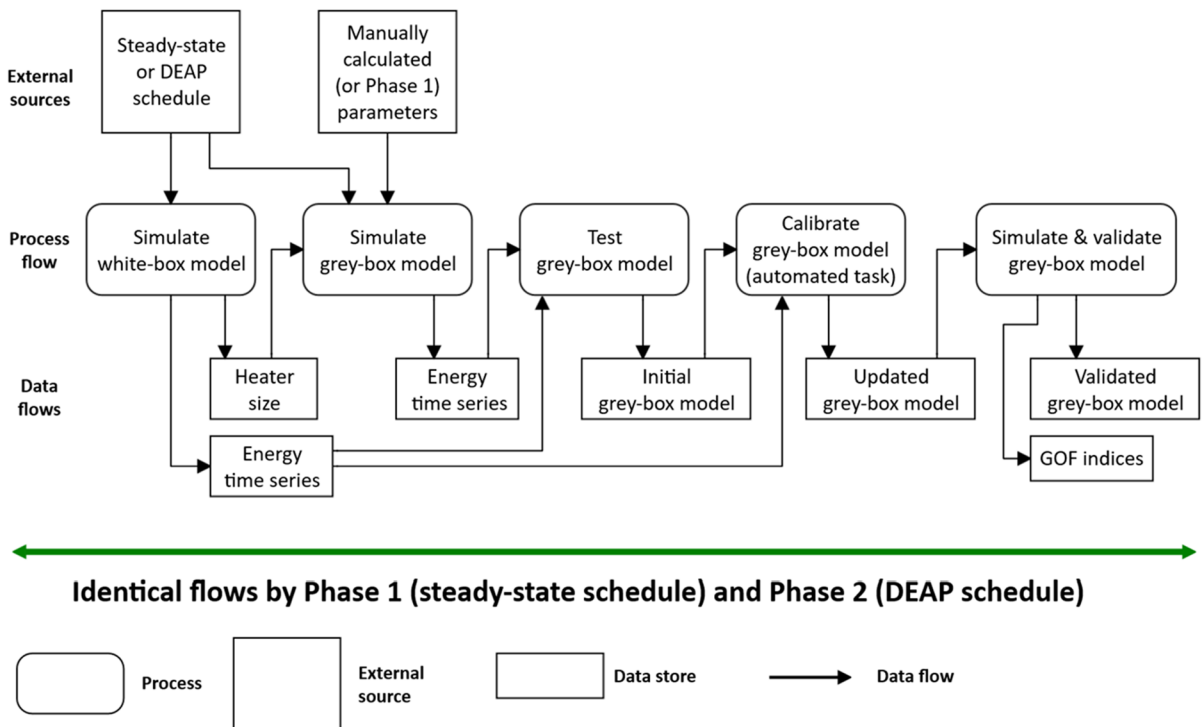
To achieve both validations, the grey-box model calibration procedure comprised two phases. Both phases followed an *identical* data flow (Fig. 6), but each phase calibrated the model against a *different* heating schedule:

- Phase 1: heat-energy usage from simulation under steady-state schedule and
- Phase 2: heat-energy usage from simulation under DEAP schedule.

Home heater sizes in the grey-box models were estimated during previous simulations of the white-box models (Beagon et al., 2020). During calibration procedure, the simulations used the heaters to maintain a 20°C setpoint adopted from TABULA (2013). Before the first simulation of Phase 1, the grey-box models were initialised with manually calculated values of tunable parameters. After the first simulation, the resulting monthly heat-energy usage was tested against that of a white-box model simulated under the same schedule. As expected, simulations with initial parameter values failed validation, and Phase 1 progressed to the automated task labelled “calibrate grey-box model” (Fig. 6).

After the automated calibrate task, the results were used to update the tunable parameters of the grey-box models. The updated models were subsequently re-simulated under both steady-state and DEAP schedules. Each simulation of an updated model calculated its monthly heat-energy usage per schedule. Those results were compared to the monthly results from simulating white-box models under identical schedules. The validation results of as-built and retrofit models appear in Table 4 and Table 5 respectively.

As mentioned, the validation criteria were  $\text{CV(RMSE)} < 15\%$ , and  $-5\% < \text{NMBE} < 5\%$ . Since



**Fig. 6** Calibration procedure and data flow. Identical flows used during both Phase 1 and Phase 2 but differentiated by energy-usage data from different heating schedules

**Table 4** Validation goodness-of-fit indices of calibrated as-built models using monthly heat-energy usage under two heating schedules: (1) steady-state and (2) DEAP

Label	Phase	Steady-state schedule		DEAP schedule	
		CV(RMSE) (%)	NMBE (%)	CV(RMSE) (%)	NMBE (%)
TH03	1	2.51	-0.24	8.08	3.56
TH06	1	0.88	-0.02	4.70	3.54
TH07	1	2.55	-0.47	6.46	4.31

**Table 5** Validation goodness-of-fit indices of calibrated retrofit models using monthly heat-energy usage under two heating schedules: (1) steady-state and (2) DEAP

Label	Phase	Steady-state schedule		DEAP schedule	
		CV(RMSE) (%)	NMBE (%)	CV(RMSE) (%)	NMBE (%)
TH03	1	3.31	-0.07	9.47	8.25
TH06	1	1.48	0.24	22.10	15.11
TH07	1	5.19	4.01	19.04	13.31
TH03	2			3.75	-2.70
TH06	2			7.06	1.76
TH07	2			6.50	0.85

Phase 1 validation failed in the case of retrofit models, Phase 2 became necessary. Phase 2 calibrated grey-box models, initialised with Phase 1 parameters

values, to fit monthly heat-energy usage under the DEAP schedule.

Phase 1 calibration was sufficient for models of as-built archetypes (Table 4). In contrast, models of

**Table 6** Empirical data sources of weather, home archetypes, household heating and occupancy. Descriptions of household data include sample size, location, and home type

Data type	Sample size (n)	Location	Home type	Source
IWEC2	Not applicable	Dublin Airport, Ireland	Not applicable	NREL, 2018
Home archetypes	Not applicable	Dublin, Ireland	Semi-detached and terraced	Badurek et al., 2014
Heating setpoints	49 retrofit, 62 as-built	Plymouth, England	Flats, semi-detached and terraced	Jones et al., 2016
Heating pattern types	Weekday: 43 retrofit, 75 as-built Weekend: 40 retrofit, 74 as-built	Plymouth, England	Flats, semi-detached and terraced	Jones et al., 2016
Heating times	275 (silent on retrofit state)	England nationally	Detached, flats, semi-detached and terraced	Huebner et al., 2015
Occupancy times	32 as-built	London, England	Flats	Zahiri & Elsharkawy, 2018

retrofit archetypes required Phase 2 calibration to achieve validation. The CV(RMSE) and NMBE validation indices improve significantly after Phase 2 as percentage values (Table 5). Note that retrofit archetypes must achieve smaller validation errors in terms of energy values, because retrofitted homes typically use less heat energy.

#### Data sources for simulations

With the RC values of the building fabric subsystem now known, the modelling effort focused on the subsystem controlled by occupants: *the internal thermal zone*. Different occupant-operated setpoints and heating patterns result in different heat-energy usage calculated by simulations. Large variations in heat-energy usage are expressed as statistical distributions — a key result of this study.

Recent literature provided empirical sources to enable simulation of heating patterns and setpoints (Table 6). A review of gas or oil-fired central heated homes found thermostat settings from 13.1 to 27.3°C — an “enormous” range (Shipworth et al., 2010). Although the mean and median temperatures were both 21°C, the sample data was positively skewed with 40% at or above 22°C.

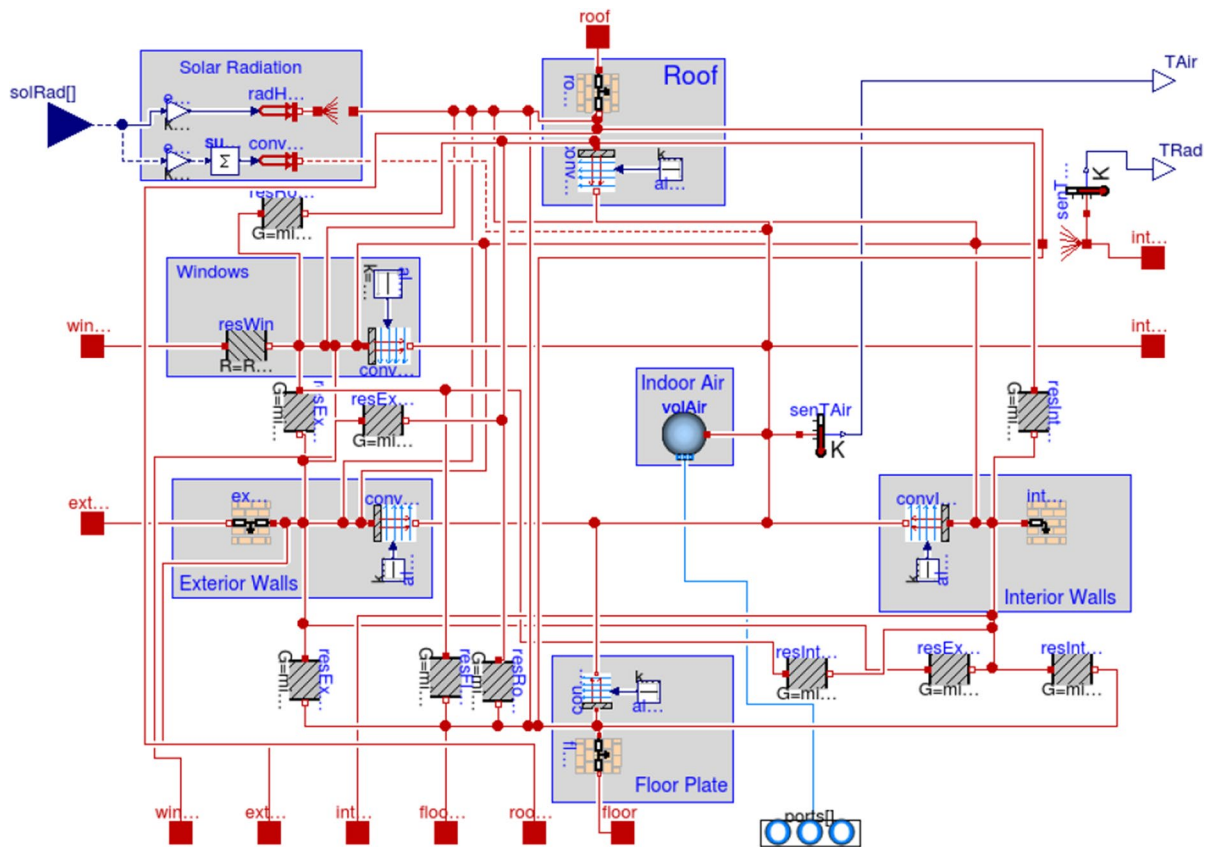
Social housing in England displayed large variations in heating setpoints, even among 62 as-built and 49 retrofit homes analysed (Jones et al., 2016). For instance, occupants in retrofit homes chose setpoints from a distribution of higher temperatures ( $\bar{x} = 21.7^\circ\text{C}$ , std dev. = 3.2°C,  $n = 49$ ) than occupants

of as-built homes ( $\bar{x} = 20.2^\circ\text{C}$ , std dev. = 3.3°C,  $n = 62$ ). Notably, the same research referred to “thermal upgrade” of wall insulation as opposed to “retrofits”, indicating that primary heat generation (probably boilers) remained unchanged.

Considering the variation in occupant heating preferences, this study adopted setpoint temperatures, heating durations and schedules from available empirical sources, that themselves differ in sample size, home type and location (Table 6). In addition to setpoint distributions, three heating patterns emerged in social housing (Jones et al., 2016). “Double” (morning and evening) was the most common (45.9%), followed by “single” (24.3%) and “constant” 24 h (20.3%). Triple heating periods occurred in under 10% of weekdays and thus were neglected by this study.

Occupants also require ventilation to maintain good health. As explained in the companion paper (Beagon et al., 2020), the simulations maintained the minimum 0.4 ACH recommendation during occupancy periods (Ramos et al., 2015; Wargocki, 2018). Where lower levels of air permeability are achieved, Part L regulations require purpose provided ventilation (DECLG, 2011). Therefore, the simulations assumed that retrofit homes operated demand controlled ventilation (DCV) simultaneous with occupant-caused internal gains.

Finally, heating and ventilation operated upon a single internal zone within all models. This study’s AixLib model of a single zone building appears in (Fig. 7); it displays the four building elements of



**Fig. 7** Modelica model of four building elements enclosing one indoor thermal zone labelled “Indoor Air volAir” (IEA Annex 60, 2017)

exterior walls, interior walls, floor plate and roof (IEA Annex 60, 2017). A proportional-integral-derivative (PID) controller, recommended in the AixLib library (RWTH EBC, 2020), tracked the temperature setpoint of the internal zone. Controller parameters enabled activation of heating when internal air temperature dropped 0.5°C below setpoint, accompanied by limitation of short cycling by one-minute off-time. This controller was easily configured, closely tracked every setpoint, and limited short cycling (where a heat generator turns on and off too frequently).

#### Heating and internal gains schedules

As listed in Table 6 data sources, research on 275 homes throughout England plotted four heating schedules as daily time-series: “Steady rise”, “Flat line”, “Two peak” and “Steep rise” (Huebner et al.,

2015). Simulation heating times were extracted from the time-series plots by Huebner et al., (2015) and other research (Jones et al., 2016). Both sources lengthened weekend heating durations which this study also incorporated into simulation schedules (Table 7).

The Steep rise time-series plot increased temperature from the morning throughout the day, peaking “before 21:00”, specifically 20:30 in its diurnal plot (Huebner et al., 2015). Since Steep rise descended to the lowest temperature of 16°C among all four time-series plots, it resembled the single heating pattern identified by Jones et al. (2016). Occupants significantly increased the duration of the single heating pattern after retrofit, compared to the double pattern. Durations of the *single* heating pattern were lengthened: weekdays 4.9 h to 8.4 h, and weekends 5.4 h to 8.6 h (Table 7). Weekday and weekend durations

**Table 7** Heating patterns schedules by time of day and hours (h) duration. Each pattern comprises four schedules, split by as-built and retrofit state, and weekday and weekend. Times adapted from Huebner et al., 2015; Jones et al., 2016

Pattern	As-built		Retrofit	
	Weekday	Weekend	Weekday	Weekend
Constant	00:00–24:00 (24 h)	00:00–24:00 (24 h)	00:00–24:00 (24 h)	00:00–24:00 (24 h)
Single	15:30–20:25 (4.9 h)	15:15–20:35 (5.4 h)	13:20–21:45 (8.4 h)	13:15–21:50 (8.6 h)
Double	06:40–08:10 & 16:40–20:35 (5.4 h)	07:00–09:00 & 16:30–20:35 (6.1 h)	06:40–08:10 & 16:20–20:50 (6.0 h)	07:00–09:00 & 16:20–20:30 (6.2 h)

of the *double* pattern lengthened marginally after retrofit: weekday 5.4 h to 6.0 h, and weekend 6.1 h to 6.2 h.

Longer heating durations after retrofit indicated that the occupants had been rationing their home heating. While occupants ought to be comfortable in their homes, the lengthening of heating durations and other “comfort taking” may have reduced or negated any savings in heat energy. Occupant “comfort taking” and the resulting rebound effect and potential “backfire” have been outlined in the Introduction.

Finally, internal gains caused by occupants were already calculated (Beagon et al., 2020), and their schedules lengthened at weekends for the single and double heating patterns in Table 8. Schedules of internal gains under the constant and single heating patterns were derived from the dominant occupancy patterns defined for a London tower block by Zahiri and Elsharkawy (2018). The dominant occupancy patterns emerged from 32 responses to the 108 questionnaires distributed to every tower block flat. This sample size of 32 exceeded the 25 necessary for valid correlation tests according to Oppenheim (1992), whom Zahiri & Elsharkawy cited as justification.

Regarding the double heating pattern, Marshall et al. (2016) provided detailed times of internal gains by an applicable “Working Family”. This presented study used a double heating pattern where the morning heating coincided with internal gains (Table 7, Table 8). At the weekend, the double heating pattern and associated internal gains did not start until 07:00, as extracted from time use surveys (Buttitta et al., 2017).

#### Characterising heat-energy distributions

With the heating patterns and internal gains defined (Table 7 and Table 8), simulations then calculated

heat-energy usage across two ranges of heating setpoints. As described in section [Data sources for simulations](#), setpoint distribution parameters differed between as-built ( $\bar{x} = 20.2^\circ\text{C}$ , std dev. =  $3.3^\circ\text{C}$ ) and retrofit homes ( $\bar{x} = 21.7^\circ\text{C}$ , std dev. =  $3.2^\circ\text{C}$ ). An R function used the parameters of the two setpoint distributions to randomly generate separate sets of as-built or retrofit setpoints. Normal distributions were assumed. Each of the two sets contained 300 sample setpoints ( $n = 300$ ), to enable distribution fitting (Keller, 2011). Note that  $n = 300$  did *not* imply that each neighbourhood comprises over 300 similar homes.

The 300 as-built setpoints were applied to all as-built models, then simulated under the as-built schedules of each heating pattern. Similarly, the 300 retrofit setpoints were applied to all retrofit models, then simulated under the retrofit schedules of each heating pattern. A total of 18 models represented every combination of three archetypes, three heating patterns, and two states of as-built and retrofit. Thus, the 18 models were each simulated under 300 heating setpoints to produce 5400 values of yearly heat-energy usage.

All 300 simulations of an archetype and heating pattern combination, modelled in either as-built or retrofit state, calculated one value of heat-energy usage. After the simulations, the 300 usage values

**Table 8** Simulated internal gains schedules by heating pattern. Schedules apply to both as-built and retrofit home archetypes

Pattern	Weekday	Weekend
Constant	08:00–22:00	08:00–22:00
Single	06:30–22:00	08:30–22:30
Double	06:40–08:30 & 16:20–22:15	07:00–10:30 & 15:30–22:30

**Table 9** Heat-energy usage by combination of archetype and heating pattern, in as-built and retrofit states. Sample statistics comprise the mean ( $\bar{x}$ ), standard deviation (s) and skewness ( $\gamma$ )

Pattern	Label	As-built			Retrofit		
		$\bar{x}$ (kWh)	s (kWh)	$\gamma$	$\bar{x}$ (kWh)	s (kWh)	$\gamma$
Constant	TH03	9654	4236	0.462	4437	2151	0.507
Constant	TH06	11,553	4653	0.415	6201	2637	0.423
Constant	TH07	6215	2913	0.502	6992	2686	0.363
Single	TH03	5079	2371	0.510	3155	1667	0.601
Single	TH06	4184	1878	0.697	4357	1798	0.411
Single	TH07	3035	1656	0.696	4050	1789	0.490
Double	TH03	7016	2804	0.345	3903	1694	0.382
Double	TH06	7275	2656	0.382	4805	1848	0.355
Double	TH07	4893	2110	0.426	5328	1881	0.303

formed a dataset representing that archetype, heating pattern and state. Each energy-usage dataset was itself a distribution, characterised by sample statistics of mean ( $\bar{x}$ ) standard deviation (s) and skewness ( $\gamma$ ). Skewness, a dimensionless statistic, indicates a distribution's shape and symmetry. A perfectly symmetrical distribution was measured at zero skewness. Positive or negative skewness values indicated a longer right or left tail respectively. Multiple skewness formulas existed in literature (Tabor, 2010; Versluis & Straetmans, 2015); the former publication defined 11 skewness formulas and the latter 14. This study adopted the “formal” or “textbook” formula; the Fisher-Pearson third moment coefficient of skewness (3), unadjusted due to the large sample  $n=300$  (Joanes & Gill, 1998; NIST, 2013).

$$\text{Skewness} : \gamma = \frac{m_3}{\sqrt{m_2^3}} [-] \text{ where } m_r = \frac{1}{n} \sum (x_i - \bar{x})^r \quad (3)$$

A highly skewed distribution requires  $\gamma > 1$ , a moderately skewed distribution  $0.5 \leq |\gamma| \leq 1$ , and an approximately or “fairly” symmetrical distribution  $|\gamma| < 0.5$  (Brown, 2016; Bulmer, 1979). All calculations and plots were executed in R statistical language. Function `skewness(x, type=1)` calculated the textbook skewness (Meyer, 2020), and `ggplot2` functions plotted the graphs (Beagon, 2020; Wickham, 2016).

As introduced in section [Statistical distributions of home energy usage](#), the Anderson–Darling (AD) technique was selected to test fit of the heat-energy datasets to statistical distributions. The AD test adopted a null hypothesis that sample values arose

from a common statistical distribution, specified by name or as another dataset. Smaller AD statistics and larger  $p$ -values indicated no evidence of deviation from the hypothesised common distribution. Thus, AD test statistics *reduced* with increasing goodness-of-fit (GOF) with the hypothesised common distribution. This study uses the AD test function within the R `kSamples` package (Scholz & Zhu, 2019). As a `kSamples` function, it tested the hypothesis that  $k$  number of independent samples arose from a common continuous distribution.

## Results

### Statistics of heat-energy usage

The heat-energy usage statistics of all nine combinations of heating pattern and home archetype appear in Table 9. Only five of the combinations, however, show decreased mean heat-energy usage after retrofit. Four combinations increased mean heat-energy usage, including all three TH07 heating patterns. Among the *shallow* TH07 retrofits, heat-energy savings by building fabric upgrades were counteracted by higher occupant heating demands.

Table 10 displays the decreases and increases in mean heat-energy usage as positive and negative savings respectively. Paired  $t$ -tests were performed on each combination's heat-energy usage between their as-built and retrofit states. The null hypothesis of no difference between as-built and retrofit states was rejected with a  $p$ -value below 0.001 for all combinations of archetype and heating pattern.



**Table 10** Savings in mean heat-energy usage by combination of archetype and heating pattern. The mean value is accompanied by its 95% confidence interval (CI)

Heating pattern	Archetype label	Mean heat-energy saving (kWh)	CI of mean heat-energy saving (kWh)
Constant	TH03	5217	4981, 5454
Constant	TH06	5352	5213, 5581
Constant	TH07	-777	-805, -747
Single	TH03	1924	1843, 2004
Single	TH06	-173	-192, -153
Single	TH07	-1014	-1034, -996
Double	TH03	3113	2987, 3239
Double	TH06	2470	2397, 2562
Double	TH07	-436	-463, -408

Based on the mean values in Table 10, the single heating pattern fails to save heat energy in archetype TH06, and the aforementioned TH07. An explanation is that occupants extend the weekday duration of the single heating pattern by three and half hours after retrofit (section [Heating and internal gains schedules](#)). The short duration of the single pattern, however, means that it minimises the heat-energy usage by *all* archetypes in both as-built and retrofit states.

#### Exploring distributions of heat-energy usage

Having calculated the statistics from each dataset of heat-energy usage, the datasets were then analysed as probability density distributions. As mentioned, each dataset of 300 energy-usage values represents one combination of archetype and heating pattern in as-built or retrofit state. These distributions were explored using quantile–quantile plots, goodness-of-fit (GOF) evaluation and hypothesis testing.

#### Quantile–quantile plots

A quantile–quantile (QQ) plot is an initial exploration of a dataset's fit to normality. All QQ plots appear in this paper's Supplementary Material. The plots display a dataset of heat-energy usage, alongside the theoretical normal distribution parameterised from energy-usage dataset. Every plot's horizontal axis displays Z-scores representing the distance between an

**Table 11** Anderson–Darling tests of fit between heat-energy usage datasets and normal distribution datasets. Results comprise test statistics and asymptotic  $p$ -values. All sample sizes were  $n = 300$ 

Pattern	Label	As-built		Retrofit	
		AD statistic	$p$ -value	AD statistic	$p$ -value
Constant	TH03	0.671	0.585	0.807	0.476
Constant	TH06	0.515	0.732	0.538	0.709
Constant	TH07	0.769	0.504	0.413	0.836
Single	TH03	0.914	0.405	1.180	0.273
Single	TH06	1.500	0.175	0.509	0.737
Single	TH07	1.590	0.157	0.759	0.512
Double	TH03	0.414	0.835	0.534	0.712
Double	TH06	0.494	0.753	0.401	0.848
Double	TH07	0.605	0.644	0.334	0.911

energy-usage value and the dataset mean. Since the distance is expressed in standard deviations, the range from  $-3$  to  $+3$  captures 99.7% of the theoretical normal distribution. All QQ plots approximate to normality at central values but diverge at their tails. The largest proportional divergence at the tails manifests under the single heating pattern.

#### Goodness-of-fit to normal distributions

The Anderson Darling (AD) technique, described in section [Characterising heat-energy distributions](#), was used to test the fit between two 300-sample datasets. The first dataset contained the simulation heat-energy usage and the second dataset contained 300 quantiles of the theoretical normal distribution. The mean and variance parameters of that theoretical normal distribution were calculated from the simulation dataset. Results of the AD tests appear in Table 11, comprising the test statistics and the  $p$ -values. Despite deviations from normality seen in the QQ plots, there is insufficient evidence to reject the null hypothesis of normality at a standard  $p$ -value of 0.05.

#### Goodness-of-fit to Weibull distributions

Weibull distributions, introduced in section [Statistical distributions of home energy usage](#), have been used to represent home energy use at different spatial scales (Irwin et al., 1986; Munkhammar et al., 2014). Furthermore, the Weibull distribution is the most popular

**Table 12** Mean ( $\bar{x}$ ) and 95% confidence interval (CI) of heat-energy usage, by combination of archetype and heating pattern in as-built and retrofit states. The  $p$ -value results from the hypothesis that mean heat-energy use equals 4600 kWh

Archetype label	Heating pattern	As-built		Retrofit	
		$\bar{x}$ and CI (kWh)	$p$ -value	$\bar{x}$ and CI (kWh)	$p$ -value
TH03	Constant	9654 (9173, 10 136)	<0.001	4437 (4192, 4681)	0.190
TH03	Single	5079 (4810, 5348)	<0.001	3155 (2966, 3345)	<0.001
TH03	Double	7016 (6697, 7334)	<0.001	3903 (3710, 4095)	<0.001
TH06	Constant	11 553 (11 025, 12 082)	<0.001	6201 (5901, 6500)	<0.001
TH06	Single	4184 (3971, 4397)	<0.001	4357 (4152, 4561)	0.020
TH06	Double	7275 (6973, 7577)	<0.001	4804 (4594, 5015)	0.056
TH07	Constant	6215 (5884, 6546)	<0.001	6992 (6686, 7297)	<0.001
TH07	Single	3035 (2847, 3224)	<0.001	4050 (3847, 4254)	<0.001
TH07	Double	4893 (4653, 5133)	0.017	5328 (5115, 5542)	<0.001

model of wind energy evaluation (Wais, 2017). Wind energy analysts often estimate Weibull parameter by the “graphical method”, that is mathematically similar to a linear regression technique.

Applying the graphical method to the energy-usage datasets enabled estimation of the Weibull parameters of shape and scale. Shape and scale parameters were estimated for each combination of archetype and heating pattern, in both as-built and retrofit states (Table 13). The larger AD  $p$ -values in both as-built and retrofit states (Table 13) indicate that heat-energy distributions fit closer to Weibull distributions compared to normality (Table 11).

Hypothesis test that mean heat-energy usage equals an official estimate

The hypothesis test was a one sample  $t$ -test against a population average ( $\mu$ ). It tested the null hypothesis that the difference between each distribution of heat-energy usage and  $\mu = 4600$  kWh was zero. The 4600-kWh value reflects the SEAI estimate that a terraced, “efficient” home demands 4.6 MWh/year heat energy (SEAI, 2015). Confidence intervals (CI) were calculated at 95% for all combinations of archetypes and heating patterns, in as-built and retrofit state (Table 12).

In all but two combinations of retrofit archetype and heating pattern,  $p$ -values show sufficient evidence to reject the null hypothesis at the 5% significance level. In conclusion, the null hypothesis that mean

heat-energy usage equals 4600 kWh is rejected by all archetypes in as-built state, and seven of the nine archetypes in retrofit state, at the 0.05 significance level.

## Discussion

### Discussion of heat-energy statistics

The statistics of the heat-energy usage are now discussed in the sequence of constant, single and double heating patterns.

#### Constant heating pattern

As expected, the largest savings in heat-energy usage attributable to retrofit were achieved under the constant heating pattern. Within the constant pattern, the mean values of heat-energy usage decreased under deeper retrofits of archetypes TH03 and TH06 but increased after the shallow retrofit of TH07, causing positive and negative energy savings respectively (Table 10). Thus, the retrofit of TH07 actually “backfired” by actually increasing the archetype’s mean heat-energy usage. As described in the Introduction, cases where occupant “comfort taking” exceeded savings in energy usage were termed “backfire” by (Druckman et al., 2011). Specific to the constant pattern, the TH07 backfire entailed a 10% increase in mean (95% CI) heat-energy usage: from

**Table 13** Heat-energy usage expressed as Weibull parameters accompanied by the goodness-of-fit  $p$ -value to the Weibull distribution. The Anderson–Darling (AD) test quantifies the goodness-of-fit  $p$ -value, while CV(RMSE) and NMBE indices quantify the goodness-of-fit error

Archetype label	Heating pattern					Retrofit				
	As-built									
	Shape	Scale (kWh)	AD $p$ -value	CV(RMSE) (%)	NMBE (%)	Shape	Scale (kWh)	AD $p$ -value	CV(RMSE) (%)	NMBE (%)
TH03	Constant	2.51	10 817	0.998	2.85	2.20	4985	0.999	2.49	0.144
TH03	Single	2.35	5694	0.997	3.05	1.96	3550	0.999	3.00	-0.104
TH03	Double	2.77	7844	0.999	2.34	2.50	4378	0.999	2.29	0.153
TH06	Constant	2.77	12 910	0.993	2.88	2.58	6947	0.999	2.54	0.170
TH06	Single	2.62	4666	0.542	6.29	2.67	4878	0.996	2.75	0.141
TH06	Double	3.14	8088	0.908	3.37	2.89	5364	0.993	2.66	0.137
TH07	Constant	2.30	6977	0.999	2.61	2.92	7800	0.993	2.69	0.168
TH07	Single	1.94	3402	0.999	3.55	2.51	4535	0.991	3.16	0.293
TH07	Double	2.54	5483	0.999	2.65	3.22	5921	0.972	2.72	0.148

6215 (5884, 6546) kWh to 6992 (6686, 7297) kWh (Table 12).

Under the constant heating pattern, variations of all archetypes’ heat-energy usage reduced from as-built to retrofit state. The variations, expressed as ranges of standard deviation (s), reduced from asbuilt state 2913–4653 kWh to retrofit state 2151–2686 kWh (Table 9). As expected, the largest reductions in standard deviation were achieved by deeper retrofits of older archetypes: TH03 – 2085 kWh and TH06 – 2016 kWh, whereas the shallow retrofit of the youngest archetype TH07 reduced standard deviation by merely – 227 kWh.

Despite their reductions in variation after retrofit, the heat-energy distributions of TH03 and TH06 slightly increased positive skew. Retrofit TH03 entered the moderate skewness range  $0.5 \leq |\gamma| \leq 1$ . The as-built and retrofit skewness ranges were 0.415–0.502 and 0.363–0.507 respectively (Table 9). In Appendix Fig. 9, heat-energy distributions under the constant pattern displayed the manifestations of positive skew — extending right-hand tails over high values and diverging from the symmetry that characterises a normal distribution.

### Single heating pattern

The smallest heat-energy savings were achieved under the single heating pattern, because of the pattern’s longer heating durations after retrofit (section [Heating and internal gains schedules](#)). Under the single pattern, only the deepest retrofit of archetype TH03 delivered savings in mean heat-energy usage. Archetypes TH06 and TH07 actually increased heat-energy usage, shown as negative savings (Table 10). A shallow retrofit of TH07, compounded by longer and warmer heating durations after retrofit, resulted in the study’s largest retrofit backfire. Specific to the single pattern, the TH07 backfire entailed a one-third increase in mean (95% CI) heat energy: from 3035 (2847, 3224) kWh to 4050 (3847, 4254) kWh (Table 12).

Under the single heating pattern, the smallest reductions in variation from as-built to retrofit state occurred. Variations, expressed as ranges of standard deviation, reduced from as-built state 1656–2371 kWh to retrofit state 1667–1798 kWh (Table 9). Within the pattern, retrofit of the oldest archetype

TH03 again achieved the largest reduction in standard deviation, from 2371 to 1667 kWh.

In Appendix Fig. 10, heat-energy distributions under the single pattern displayed the variation and skewness across archetypes. Overall, the distributions' skewness ranges in the as-built state (0.510–0.696) or retrofit state (0.411–0.601) were moderately skewed or mainly “fairly” symmetrical respectively (Table 9). However, the skewness of the TH03 distribution increased from 0.510 to 0.601, although remained within the moderate range  $0.5 \leq |\gamma| \leq 1$ , while the younger TH06 and TH07 archetypes reduced their distribution skewness. The post-retrofit increase in TH03 skewness reflected the non-linear increases in heat-energy usage to maintain the highest heating setpoints.

### Double heating pattern

Energy savings achieved under the double heating pattern exceeded those under the single pattern, but undershot those under the constant pattern (Table 10). Like the constant heating pattern, savings in mean heat-energy usage are achieved by deeper retrofits of older archetypes TH03 and TH06. The shallow retrofit of TH07 consistently failed to save heat energy under all heating patterns. Specific to the double pattern, the TH07 backfire entailed a 9% increase in mean (95% CI) heat energy: from 4893 (4653, 5133) kWh to 5328 (5115, 5542) kWh (Table 12).

The variations in heat-energy distributions, expressed as ranges of standard as-built deviation, reduced from as-built state 2110–2804 kWh to retrofit state 1694–1881 kWh (Table 9). As expected, retrofit of the oldest archetype TH03 achieved the largest reduction in standard deviation from 2804 to 1694 kWh.

In Appendix Fig. 11, heat-energy distributions under the double pattern displayed the variation and skewness across archetypes. All archetype distributions met the “fairly” symmetrical definition where  $0.0 \leq |\gamma| \leq 0.5$ . The as-built and retrofit skewness ranges were 0.345–0.426 and 0.303–0.382 respectively (Table 9). These variation and skewness statistics manifested in heat-energy distributions that resemble normal

distributions and peaked over mean values, especially after retrofit (Fig. 11).

### Discussion of heat-energy distribution fitting

The fit of the heat-energy distributions with normal and Weibull distributions are now discussed.

#### Quantile–quantile plots

It is important to note that all heat-energy usage results are, by their nature, positive. Therefore, positive values of heat-energy usage must diverge from any negative values extrapolated by the *lowest* Z-scores of a QQ plot. The QQ plots assumed a *theoretical* normal distribution. For instance, as-built QQ plots under the constant pattern extended into negative values because of large standard deviations with the energy-usage datasets. Note that standard deviation is both a measure of dispersion, and a parameter of the theoretical normal distribution. To summarise, as-built heat-energy usage appeared to diverge further from normality, because of their larger standard deviations compared to retrofit heat-energy usage.

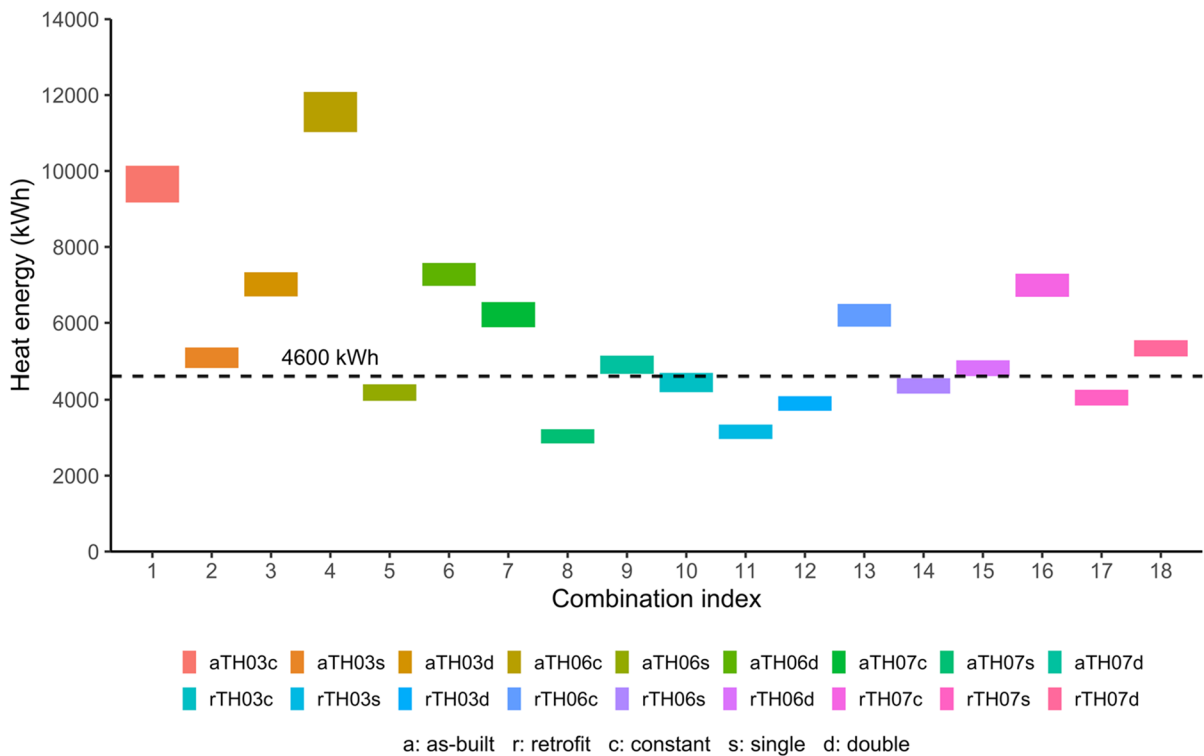
#### Exploring the goodness-of-fit to normal distributions

The QQ plots displayed evidence of divergence from normality at the distribution tails. While it remained relatively small, the divergence was also reflected in distribution statistics. For instance, as-built archetypes heated under the single pattern resulted in the largest AD statistic (Table 11), and the most skewed plots (Appendix Fig. 9).

All AD test *p*-values were too large to reject the null hypothesis of a common normal distribution at the significance value of 0.05. However, as seen in the QQ plots, there were deviations from normality and the *p*-values were inconsistent, ranging from 0.157 to 0.911. Therefore, AD tests were repeated in case each dataset fitted better with its theoretical Weibull distribution.

#### Exploring the goodness-of-fit to Weibull distributions

Weibull shape parameters range from 1.94 to 3.22 (Table 13), and a majority of 13 achieved the



**Fig. 8** Mean (95% CI) of heat-energy usage by combinations of archetype and heating pattern in as-built and retrofit states. Across the three heating patterns, retrofit TH06 approaches the 4600-kWh official estimate (dashed line)

log-normal distribution criteria of shape value  $\geq 2.5$  (NCSS LLC, 2019). The double heating pattern was the only pattern that was consistently expressed as a log-normal distribution of heat-energy usage. Nonetheless, the pattern’s highest shape value of 3.22 revealed a significant shortfall below the 3.6 value of a normal distribution.

While four of the largest shape parameters manifest under the double heating pattern, a well-defined association between shape parameters and heating patterns fails to appear. Retrofitting the archetypes does not consistently reduce, increase or transform the shape parameters of heat-energy distributions. As expected, Weibull scale parameters reduce after retrofit, due to the savings in heat-energy usage.

The goodness-of-fit achieved by each Weibull distribution is measured by model calibration indices: CV(RMSE) and NMBE. Almost universally, these indices improve slightly after retrofit. All

heating combinations, as-built and retrofit, achieve both model calibration standards:  $CV(RMSE) \leq 5\%$  and  $-5\% \leq NMBE \leq 5\%$ . Weibull distributions can be seen to maximise these fitting errors at their plot peaks (Appendix Figs. 9, 10 and 11).

Hypothesis that mean heat-energy usage equals an official estimate

Finally, the discussion reviews which one of the mean values of heat-energy usage approaches the official estimate. All the confidence intervals (CI) are *dis-jointed*, highlighting how occupants cause *different* mean values of heat-energy usage in the *same* home archetype. Most of the CIs exclude the official 4600-kWh estimate of mean heat-energy usage (Fig. 8).

Two combinations in *retrofit* state contain 4600 kWh within their mean 95% CIs: TH03 constant heating pattern 4192–4681 kWh, and TH06 double

heating pattern 4594–5015 kWh (Table 12). After retrofit, the semi-detached TH06 reduces its heat energy to become the closest archetype to the 4600 kWh value. Under the constant, single and double patterns, retrofitted TH06 used three different means of heat energy: 6201 (5901, 6500) kWh, 4357 (4152, 4561) kWh and 4804 (4594, 5015) kWh.

Overall, archetypes in the retrofit state diverge less from the official 4600 kWh value, than those in the as-built state (Fig. 8). Considering all heating patterns, the retrofit TH06 archetype of a semi-detached house comes closest to the official estimate. Therefore, there is evidence that the official estimate is representative of “efficient” medium homes — especially retrofit semi-detached houses, as opposed to low-performing homes.

#### Discussion of study limitations

Certain limitations existed in the models and in the predictive application of their results. Starting with the models themselves, the heating controller assumed heating setpoints were maintained throughout the year, even after the heating season. Occupants may have behaved otherwise, and completely deactivated space heating during the summer months.

While the Modelica models did incorporate increased air infiltration due to temperature differentials, they lacked a realistic response to wind speeds. In general, the models also assumed the correct installation and operation of demand controlled ventilation.

This study ignored uncommon occupant heating patterns, for example the triple heating pattern. A mass transition to working from home is covered by the constant heating pattern. Any evaluation of large-scale home energy retrofit must re-assess the popularity of heating patterns, especially for working families.

The weather file used by all simulations comprised selected days from previous years. Therefore, simulations exclude extreme weather conditions that are more likely in the future. Conventionally, homes located in cool temperate climates, such as Ireland’s, lacked built-in air conditioning. Certain climate change outcomes would make air-conditioning

necessary for human comfort, necessitating a holistic reassessment of home energy usage.

#### Conclusions

This paper defined a methodology to calibrate and realistically simulate grey-box models representing home heating. By incorporating occupant variables of heating setpoints and heating patterns, the methodology provides a distribution of heat-energy usage instead of a single value. Three case studies, of home archetypes in a cool temperate climate, demonstrated the results of the methodology.

Regarding the results of the methodology, all datasets of home heat-energy usage formed distributions that were positively skewed instead of normality. Positively skewed distributions cluster most values around their lower left tail, while the distribution’s right tail extends longer — pulled rightward by a few high values.

After retrofit however, most distributions were classified as fairly symmetric, having a skewness value below 0.5, with a minority remaining moderately skewed, having a skewness value between 0.5 and 1.0. There is no evidence of deviations from the Weibull distributions parameterised from the heat-energy dataset. However there exists, some evidence of deviation from log-normal distributions and evidence to reject fitting to normal distributions. Other remarks on the heat-energy usage distributions include:

- Heat-energy usage distributions are *different* between heating patterns of the *same* home archetype, both as-built and retrofit.
- Only deepest retrofit, implemented on the oldest archetype TH03, appears to consistently reduce the mean average of heat energy use.
- Retrofit of a higher performing home tends to increase the distribution symmetry but significant variance persists.
- These large variances in home heat-energy values require Weibull distributions that cannot extend into negative values.

- Positive skewness, albeit reduced, persists in the heat-energy usage of retrofitted homes.

Turning to the hypothesis that average heat-energy usage by an “efficient medium” home in Ireland equals a single value of 4600 kWh/year: this study finds evidence contrary to this hypothesis. Of the 18 heat-energy datasets, only two produced a mean value confidence interval that overlapped with the hypothesised average value. Two retrofit archetypes TH03 and TH06, each under a different heating pattern, produce a mean 95% confidence interval containing 4600 kWh/year. Since “double” is the most popular heating pattern, retrofit TH06 double heating pattern is the more likely source of a 4600 kWh/year average than retrofit TH03 under a constant heating pattern. TH06 is a semi-detached home constructed during 1978–1982, while TH03 is a terraced home constructed during 1930–1949.

#### Implications and future work

The hypothesis results imply that an official estimate of heat demand by a medium-sized, efficient home archetype (SEAI, 2015), does *not* equal the mean heat-energy usage by most combinations of the studied home archetypes and heating patterns (Fig. 8). The reason is that heat-energy usage by each combination is itself a distribution, characterised by its own distinct average.

Evidence does exist that the estimated 4600 kWh of heat demand by efficient “medium” homes, defined by SEAI as terraced homes, is relevant. Heat-energy usage by retrofitted (and arguably efficient) archetype

TH06 approaches the official estimate. However, TH06 represents a semi-detached home a home type that fits SEAI’s third, and final, category labelled “large” (SEAI, 2015).

In conclusion, evidence exists that the range of nine archetypes used by SEAI need expansion in order to accurately represent the Irish housing stock (SEAI, 2015). Higher resolution in home archetypes is needed, especially among the terraced and semi-detached homes that are popular in urban areas. Given their similar floorspace but different envelope, terraced and semi-detached homes warrant their own archetypes. Such archetypes should be further subdivided by as-built and retrofit state, to assist homeowners in selecting the optimum combination of fabric retrofit and heat pump. Elements of this proposed work will form part of SEAI’s National Heat Study: a “comprehensive assessment of the potential for efficient heating and cooling in Ireland” (SEAI, 2021).

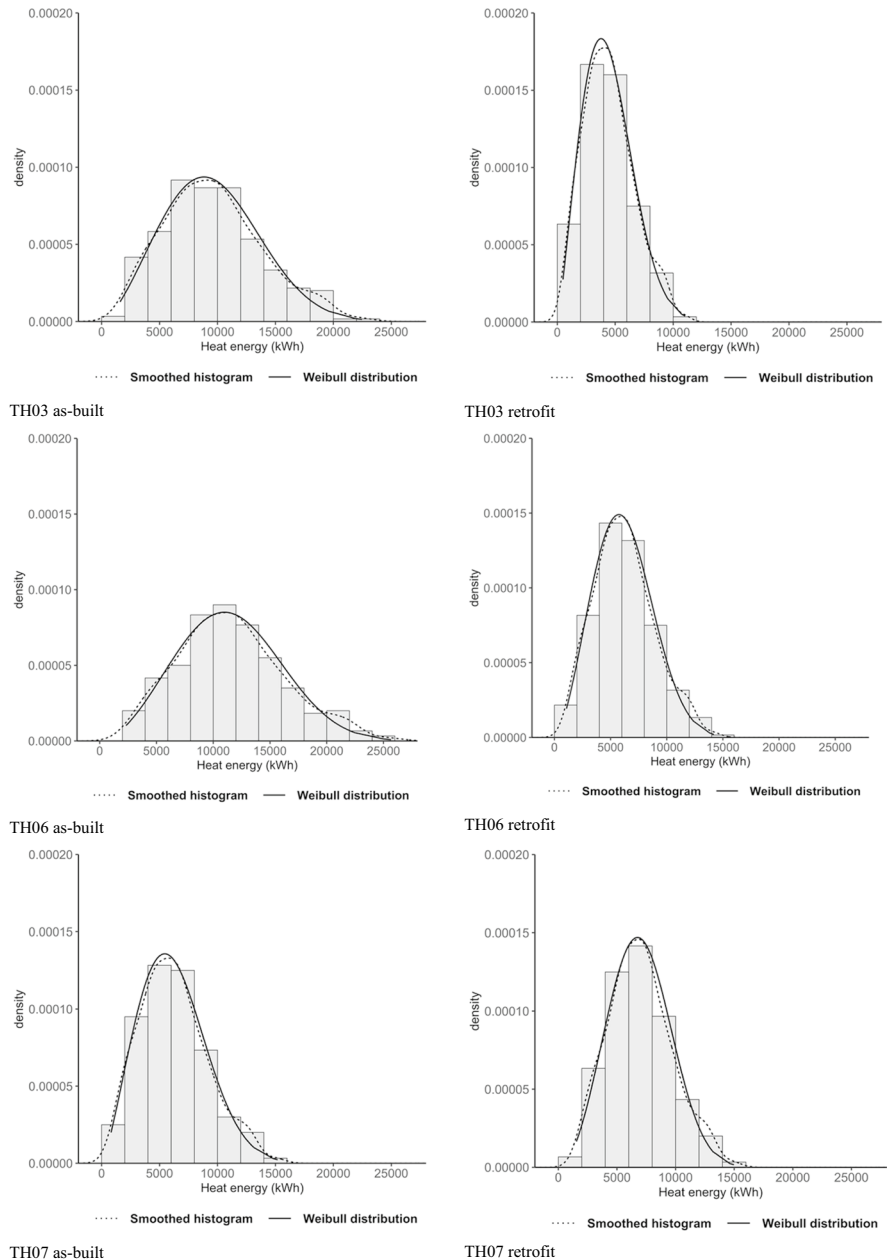
Looking forward, work could extend from home archetype’s usage of heat energy to secondary energy. Research questions could be (1) what are the secondary-energy savings attributable to boiler upgrades or replacement by heat pumps? and (2) what is the energy usage across a neighbourhood of home archetypes before and after retrofit? Any new neighbourhood analysis would require new proportions of heating patterns, given the pandemic-induced shift to home working. Realistic distributions of secondary and primary energy usage — and savings attributable to *large-scale* home energy retrofit — could then be calculated.

### Appendix. Distributions of heat-energy usage

The following distributions plots display how probability density distributions of heat-energy usage by home archetypes fit to Weibull distributions. Separate plots are presented for archetypes in both

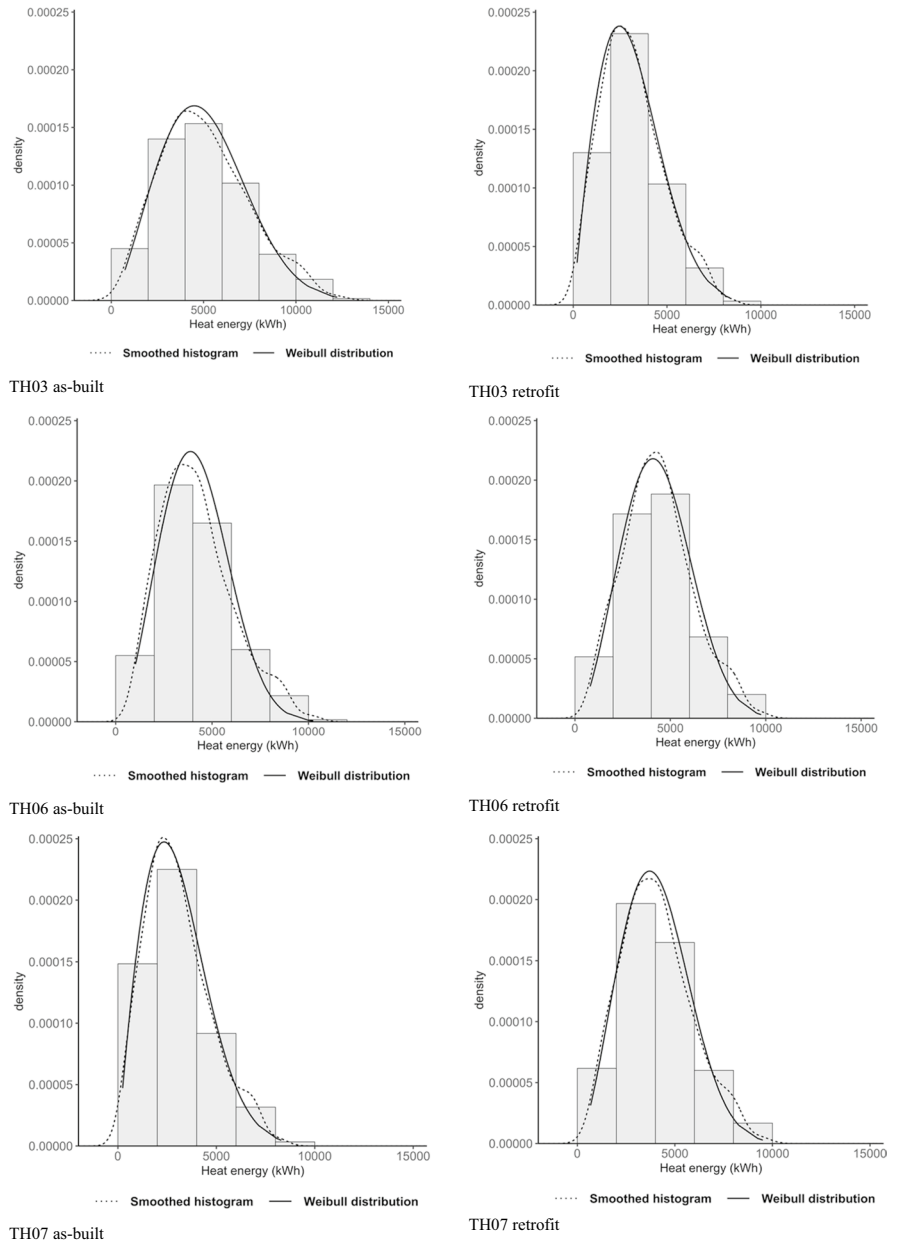
as-built and retrofit states. Heat energy is the energy type that warms homes' internal air to setpoint temperatures at scheduled times. Occupants vary heating setpoints and schedules, thus varying heat-energy usage across the same archetype. Savings in heat energy are attributable to the fabric-only retrofits analysed in this study.

**Fig. 9** Distributions of yearly heat-energy usage by as-built and retrofit archetypes under *constant* heating pattern. The three distributions are histogram bars, smoothed histogram density and Weibull density

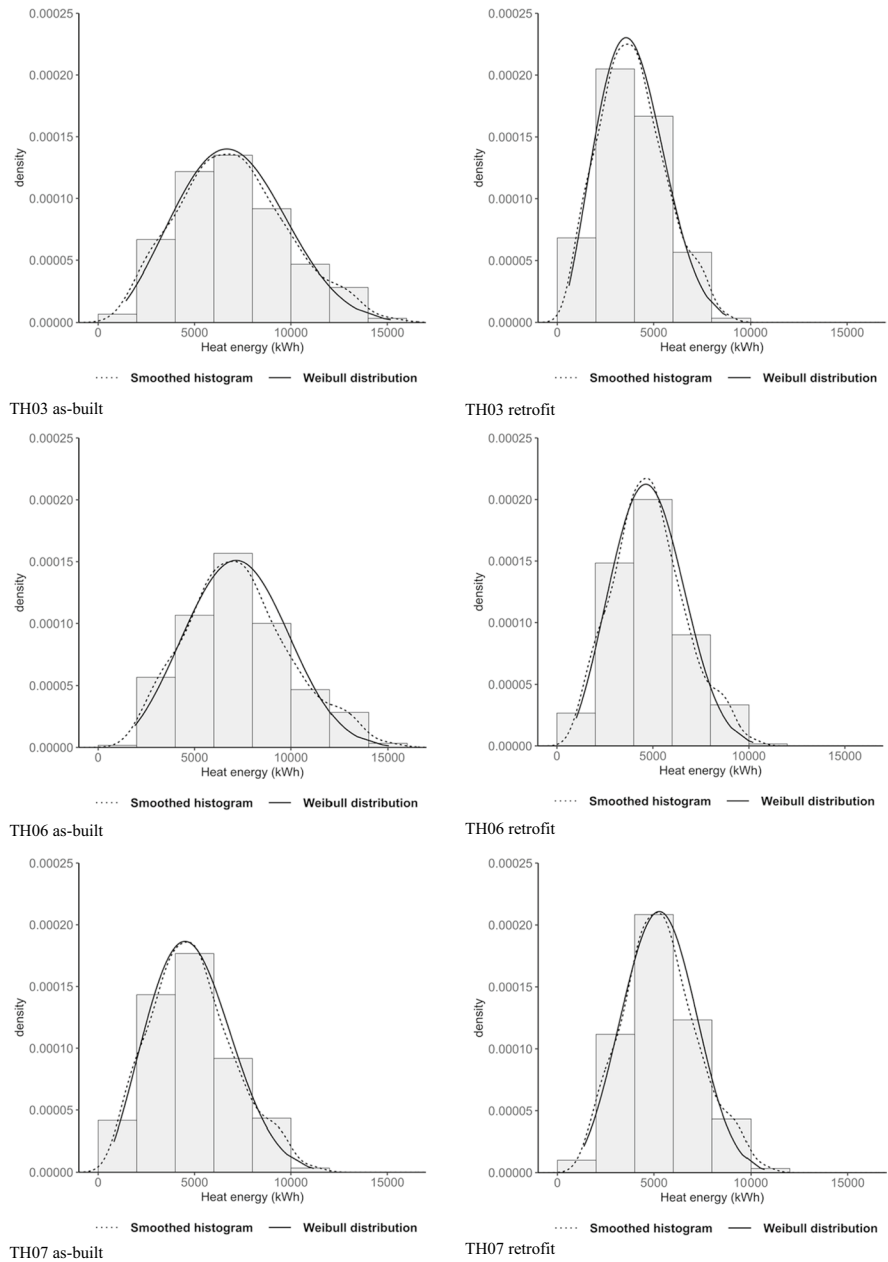




**Fig. 10** Distributions of yearly heat-energy usage by as-built and retrofit archetypes under *single* heating pattern. The three distributions are histogram bars, smoothed histogram density and Weibull density



**Fig. 11** Distributions of yearly heat-energy usage by as-built and retrofit archetypes under *double* heating pattern. The three distributions are histogram bars, smoothed histogram density and Weibull density



**Acknowledgements** The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Science Foundation Ireland.

**Author contribution** Paul Beagon designed the research, acquired the data (by modelling), drafted and revised the manuscript and interpreted the results (energy retrofit practices). Fiona Boland analysed (statistically), interpreted the results, revised manuscript and supervised the research.

**Funding** Paul Beagon was funded by Science Foundation Ireland (SFI) under the SFI Strategic Partnership Programme Grant Number SFI/15/SPP/E3125.

**Data availability** The datasets generated owned by Paul Beagon are available for publication if Journal wishes.

**Declarations**

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahern, C., Griffiths, P., & O'Flaherty, M. (2013). State of the Irish housing stock—Modelling the heat losses of Ireland's existing detached rural housing stock & estimating the benefit of thermal retrofit measures on this stock. *Energy Policy*, 55, 139–151. <https://doi.org/10.1016/J.ENPOL.2012.11.039>
- Arcipowska, A., Rapf, O., Faber, M., Fabbri, M., Tigchelaar, C., Boermans, T., Surmeli-Anac, N., Pollier, K., Dal, F., Sebi, C., & Karásek, J. (2017). *Support for setting up an observatory of the building stock and related policies*. [https://ec.europa.eu/energy/studies\\_main/final\\_studiessup-port-setting-observatory-building-stock-and-related-policies\\_en](https://ec.europa.eu/energy/studies_main/final_studiessup-port-setting-observatory-building-stock-and-related-policies_en).
- Bacher, P., & Madsen, H. (2011). Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7), 1511–1522. <https://doi.org/10.1016/j.enbuild.2011.02.005>
- Badurek, M., Hanratty, M., Sheldrick, B., & Stewart, D. (2014). *Building Typology Brochure Ireland*. [http://episcope.eu/fileadmin/tabula/public/docs/brochure/IE\\_TABULA\\_TypologyBrochure\\_EnergyAction.pdf](http://episcope.eu/fileadmin/tabula/public/docs/brochure/IE_TABULA_TypologyBrochure_EnergyAction.pdf).
- Ballarini, I., Corgnati, S. P., & Corrado, V. (2014). Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project. *Energy Policy*, 68, 273–284. <https://doi.org/10.1016/J.ENPOL.2014.01.027>
- Beagon, P. (2020). *distFit: An R package to investigate goodness-of-fit with normal and Weibull distributions*. <https://zenodo.org/record/4060239>.
- Beagon, P., Boland, F., & Saffari, M. (2020). Closing the gap between simulation and measured energy use in home archetypes. *Energy and Buildings*, 224. <https://doi.org/10.1016/j.enbuild.2020.110244>.
- Brastein, O. M., Perera, D., Pfeifer, C., & Skeie, N. (2018). Parameter estimation for grey-box models of building thermal behaviour. *Energy and Buildings*, 169, 58–68. <https://doi.org/10.1016/j.enbuild.2018.03.057>
- Brown, S. (2016). *Measures of shape: Skewness and kurtosis*. <http://brownmath.com/stat/shape.htm#Skewness>.
- Bulmer, M. G. (1979). *Principles of Statistics*. Dover Publications.
- Buttitta, G., Neu, O., Turner, W. J. N., & Finn, D. (2017). Modelling household occupancy profiles using data mining clustering techniques on time use data. *Building Simulation 2017*. <http://www.buildingsimulation2017.org/>.
- Chitnis, M., Sorrell, S., Druckman, A., Firth, S. K., & Jackson, T. (2014). Who rebounds most? Estimating direct and indirect rebound effects for different UK socioeconomic groups. *Ecological Economics*, 106, 12–32. <https://doi.org/10.1016/j.ecolecon.2014.07.003>
- Coma, J., Maldonado, J. M., de Gracia, A., Gimbernat, T., Botargues, T., & Cabeza, L. F. (2019). Comparative analysis of energy demand and CO2 emissions on different typologies of residential buildings in Europe. *Energies*, 12(12), 2436. <https://doi.org/10.3390/EN12122436>
- CSO. (2018). *Census of Population 2016 - Profile 1 Housing in Ireland*. <https://www.cso.ie/en/releasesandpublications/ep/p-cp1hii/cp1hii/od/>.
- CSO. (2020). *Private Households in Permanent Housing Units*. <https://data.cso.ie/table/E1008>.
- Dassault Systèmes AB. (2017). *Dymola Dynamic Modelling Laboratory - User Manual Volume 2*.
- DCCAE. (2019). *Climate Action Plan To Tackle Climate Breakdown*. <https://assets.gov.ie/10206/d042e174c16546ca14f39242fb07d22.pdf>.
- DECLG. (2011). *Technical Guidance Document L - Conservation of Fuel & Energy - Dwellings (2011) | Department of Housing, Planning, Community and Local Government*. <http://www.housing.gov.ie/housing/building-standards/tgd-part-l-conservation-fuel-and-energy/technical-guidance-document-tgd-0>.
- Déqué, F., Ollivier, F., & Poblador, A. (2000). Grey boxes used to represent buildings with a minimum number of geometric and thermal parameters. *Energy and Buildings*, 31(1), 29–35. [https://doi.org/10.1016/S0378-7788\(98\)00074-7](https://doi.org/10.1016/S0378-7788(98)00074-7)
- Dineen, D., & ÓGallachóir, B. P. (2011). Modelling the impacts of building regulations and a property bubble on residential space and water heating. *Energy and Buildings*, 43(1), 166–178. <https://doi.org/10.1016/j.enbuild.2010.09.004>
- Dixon, T., & Eames, M. (2013). Scaling up: The challenges of urban retrofit. *Building Research & Information*, 41(5), 499–503. <https://doi.org/10.1080/09613218.2013.812432>
- Druckman, A., Chitnis, M., Sorrell, S., & Jackson, T. (2011). Missing carbon reductions? Exploring rebound and backfire effects in UK households. *Energy Policy*, 39(6), 3572–3581. <https://doi.org/10.1016/j.enpol.2011.03.058>
- Elmqvist, H., Olsson, H., Mattsson, S. E., Brück, D., Schweiger, C., Joos, D., & Otter, M. (2005). Optimization for design and parameter estimation. In Gerhard Schmitz (Ed.), *4th International Modelica Conference* (pp. 255–266). <http://www.modelica.org/events/Conference2005/>.
- Energy Action. (2014). *TABULA: Ireland - Country Page*. <http://episcope.eu/building-typology/country/ie/>.

- Engmann, S., & Cousineau, D. (2011). Comparing distributions: The two-sample Anderson – Darling test as an alternative to the Kolmogorov – Smirnov test. *Journal of Applied Quantitative Methods*, 6(May), 1–17. [http://www.jaqm.ro/issues/volume-6,issue-3/1\\_engmann\\_cousineau.php](http://www.jaqm.ro/issues/volume-6,issue-3/1_engmann_cousineau.php).
- Fouquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23, 272–288. <https://doi.org/10.1016/j.rser.2013.03.004>
- Fylan, F., Glew, D., Smith, M., Johnston, D., Brooke-Peat, M., Miles-Shenton, D., Fletcher, M., Aloise-Young, P., & Gorse, C. (2016). Reflections on retrofits: Overcoming barriers to energy efficiency among the fuel poor in the United Kingdom. *Energy Research and Social Science*, 21, 190–198. <https://doi.org/10.1016/j.erss.2016.08.002>
- GBPN. (2013). *What is a deep renovation definition?* <https://tools.gbpn.org/reports/what-deep-renovation-definition-3>.
- Gupta, R., Gregg, M., Passmore, S., & Stevens, G. (2015). Intent and outcomes from the Retrofit for the Future programme: Key lessons. *Building Research & Information*, 43(4), 435–451. <https://doi.org/10.1080/09613218.2015.1024042>
- Hamilton, I. G., Summerfield, A. J., Shipworth, D., Steadman, J. P., Oreszczyn, T., & Lowe, R. J. (2016). Energy efficiency uptake and energy savings in English houses: A cohort study. *Energy and Buildings*, 118, 259–276. <https://doi.org/10.1016/j.enbuild.2016.02.024>
- Harb, H., Boyanov, N., Hernandez, L., Streblov, R., & Müller, D. (2016). Development and validation of grey-box models for forecasting the thermal response of occupied buildings. *Energy and Buildings*, 117, 199–207. <https://doi.org/10.1016/j.enbuild.2016.02.021>
- Huebner, G. M., McMichael, M., Shipworth, D., Shipworth, M., Durand-Daubin, M., & Summerfield, A. (2013). The reality of English living rooms - A comparison of internal temperatures against common model assumptions. *Energy and Buildings*, 66, 688–696. <https://doi.org/10.1016/j.enbuild.2013.07.025>
- Huebner, G. M., McMichael, M., Shipworth, D., Shipworth, M., Durand-Daubin, M., & Summerfield, A. J. (2015). The shape of warmth: Temperature profiles in living rooms. *Building Research & Information*, 43(2), 185–196. <https://doi.org/10.1080/09613218.2014.922339>
- IEA Annex 60. (2017). *Thermal zone with four elements for exterior walls, interior walls, floor plate and roof*. [https://www.iea-annex60.org/releases/modelica/1.0.0/help/Annex60\\_ThermalZones\\_ReducedOrder\\_RC.html#Annex60.ThermalZones.ReducedOrder.RC](https://www.iea-annex60.org/releases/modelica/1.0.0/help/Annex60_ThermalZones_ReducedOrder_RC.html#Annex60.ThermalZones.ReducedOrder.RC).
- Irwin, G. W., Monteith, W., & Beattie, W. C. (1986). Statistical electricity demand modelling from consumer billing data. *IEE Proceedings C Generation, Transmission and Distribution*, 133(6), 328. <https://doi.org/10.1049/ip-c.1986.0048>
- Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society Series d: The Statistician*, 47(1), 183–189. <https://doi.org/10.1111/1467-9884.00122>
- Jones, R. V., Fuertes, A., Boomsma, C., & Pahl, S. (2016). Space heating preferences in UK social housing: A socio-technical household survey combined with building audits. *Energy and Buildings*, 127, 382–398. <https://doi.org/10.1016/j.enbuild.2016.06.006>
- Keller, P. (2011). *Six Sigma Demystified* (2nd ed.). McGraw-Hill.
- Koch, A., & Girard, S. (2011). Urban neighbourhoods -- An intermediate scale for the assessment of energy performance of buildings. *Eceee 2011 Summer Study*, 1377–1385.
- Koch, A., Girard, S., & McKoen, K. (2012). Towards a neighbourhood scale for low- or zero-carbon building projects. *Building Research & Information*, 40(4), 527–537. <https://doi.org/10.1080/09613218.2012.683241>
- Lauster, M., Remmen, P., Fuchs, M., Teichmann, J., Streblov, R., & Müller, D. (2014a). Modelling long-wave radiation heat exchange for thermal network building simulations at urban scale using Modelica. *10th International Modelica Conference*, 125–133. <https://doi.org/10.3384/ECP14096125>.
- Lauster, M., Teichmann, J., Fuchs, M., Streblov, R., & Mueller, D. (2014b). Low order thermal network models for dynamic simulations of buildings on city district scale. *Building and Environment*, 73, 223–231. <https://doi.org/10.1016/j.buildenv.2013.12.016>
- Loga, T. (2010). *Project Information: Typology Approach for Building Stock Energy Assessment* (p. 2). <http://www.buildup.eu/sites/default/files/link-files/tabula-info.pdf>.
- Loga, T., Stein, B., & Diefenbach, N. (2016). TABULA building typologies in 20 European countries — Making energy-related features of residential building stocks comparable. *Energy and Buildings*, 132, 4–12. <https://doi.org/10.1016/j.enbuild.2016.06.094>
- Marshall, E., Steinberger, J. K., Dupont, V., & Foxon, T. J. (2016). Combining energy efficiency measure approaches and occupancy patterns in building modelling in the UK residential context. *Energy and Buildings*, 111, 98–108. <https://doi.org/10.1016/j.enbuild.2015.11.039>
- Mathiesen, B. V., Drysdale, D., Lund, H., Paardekooper, S., Ridjan, I., Connolly, D., Thellufsen, J. Z., & Jensen, J. S. (2016). *Future green buildings*. Department of Development and Planning, Aalborg University. [http://vbn.aau.dk/files/234005850/Future\\_Green\\_Buildings\\_A\\_key\\_to\\_cost\\_effective\\_sustainable\\_energy\\_systems\\_ENGLISH.pdf](http://vbn.aau.dk/files/234005850/Future_Green_Buildings_A_key_to_cost_effective_sustainable_energy_systems_ENGLISH.pdf).
- McKenna, R., Merkel, E., Fehrenbach, D., Mehne, S., & Fichtner, W. (2013). Energy efficiency in the German residential sector: A bottom-up building-stock-model-based analysis in the context of energy-political targets. *Building and Environment*, 62, 77–88. <https://doi.org/10.1016/j.buildenv.2013.01.002>
- McManus, R. (2011). Suburban and urban housing in the twentieth century. *Proceedings of the Royal Irish Academy*, 111C, 253–286. <https://doi.org/10.3318/PRIAC.2011.111.253>
- Meyer, D. (2020). *Skewness function (Package e1071)*. R Documentation. <https://www.rdocumentation.org/packages/e1071/versions/1.7-4/topics/skewness>.
- Monteiro, C. S., Pina, A., Cerezo, C., Reinhart, C., & Ferrão, P. (2017). The use of multi-detail building archetypes in urban energy modelling. *Energy Procedia*, 111, 817–825. <https://doi.org/10.1016/j.egypro.2017.03.244>

- Müller, D., Lauster, M., Constantin, A., Fuchs, M., & Remmen, P. (2016). AIXLIB - An open-source Modelica library within the IEA-EBC Annex 60 framework. *BauSIM 2016*, 3–6. <http://www.iea-annex60.org/downloads/2016-bausim-aixlib.pdf>.
- Munkhammar, J., Ryden, J., & Widen, J. (2014). Characterizing probability density distributions for household electricity load profiles from high-resolution electricity use data. *Applied Energy*, 135, 382–390. <https://doi.org/10.1016/j.apenergy.2014.08.093>
- NCSS LLC. (2019). *Data simulation: Weibull distribution*. [https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Data\\_Simulation.pdf](https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Data_Simulation.pdf).
- Nelson, L. S. (2018). The Anderson-Darling Test for Normality. *Journal of Quality Technology*, 30(3), 298–299. <https://doi.org/10.1080/00224065.1998.11979858>
- NIST. (2013). *Measures of skewness and kurtosis*. Engineering Statistics Handbook. <https://itl.nist.gov/div898/handbook/eda/section3/eda35b.htm>.
- NREL. (2018). *Weather Data by Region - Ireland*. [https://energyplus.net/weather-region/europe\\_wmo\\_region\\_6/IRL/](https://energyplus.net/weather-region/europe_wmo_region_6/IRL/).
- Oppenheim, A. N. (1992). *Questionnaire design, interviewing, and attitude measurement*. Pinter Publishers.
- Orecchini, F., & Santiangeli, A. (2011). Beyond smart grids - The need of intelligent energy networks for a higher global efficiency through energy vectors integration. *International Journal of Hydrogen Energy*, 36(13), 8126–8133. <https://doi.org/10.1016/j.ijhydene.2011.01.160>
- Pfretzschner, P. A. (1965). *The dynamics of Irish housing*. Institute of Public Administration Dublin.
- Ramos, N. M. M., Almeida, R. M. S. F., Curado, A., Pereira, P. F., Manuel, S., & Maia, J. (2015). Airtightness and ventilation in a mild climate country rehabilitated social housing buildings – What users want and what they get. *Building and Environment*, 92, 97–110. <https://doi.org/10.1016/j.buildenv.2015.04.016>
- Reddy, T., & Maor, I. (2006). *Procedures for reconciling computer-Calculating results with measured energy data*.
- Reinhart, C. F., & Cerezo Davila, C. (2016). Urban building energy modeling - A review of a nascent field. *Building and Environment*, 97, 196–202. <https://doi.org/10.1016/j.buildenv.2015.12.001>
- Reynders, G., Diriken, J., & Saelens, D. (2014). Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals. *Energy and Buildings*, 82, 263–274. <https://doi.org/10.1016/j.enbuild.2014.07.025>
- Robson, C., & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings* (4th Edition). Wiley.
- Rowley, E. (2015). The Architect, the planner and the bishop: The shapers of ‘Ordinary’ Dublin, 1940–60. *FOOT-PRINT*, 9(2), 69–88. <https://doi.org/10.7480/footprint.9.2.865>
- RWTH EBC. (2020). *PID controller with anti-windup, hysteresis and timer to prevent short cycling*. <https://github.com/RWTH-EBC/AixLib/blob/development/AixLib/Controls/Continuous/PIDHysteresisTimer.mo>.
- Scholz, F., & Zhu, A. (2019). *Package “kSamples”: K-Sample Rank Tests and their Combinations*. CRAN. <https://cran.r-project.org/package=kSamples>.
- Sdei, A., Gloriant, F., Tittlein, P., Lassue, S., Hanna, P., Beslay, C., Gournet, R., & McEvoy, M. (2015). Social housing retrofit strategies in England and France: A parametric and behavioural analysis. *Energy Research & Social Science*, 10, 62–71. <https://doi.org/10.1016/j.erss.2015.07.001>
- SEAI. (2015). *Replacing oil boilers with air-source heat pumps: Household economics and system-wide impacts*. <https://www.seai.ie/resources/publications/Replacing-oil-boilers-with-heat-pump-household-economics-and-system-wide-impacts-Summary-document.pdf>.
- SEAI. (2021). *National Heat Study*. <https://www.seai.ie/data-and-insights/national-heat-study/>.
- Shipworth, M., Firth, S. K., Gentry, M. I., Wright, A. J., Shipworth, D. T., & Lomas, K. J. (2010). Central heating thermostat settings and timing: Building demographics. *Building Research & Information*, 38(1), 50–69. <https://doi.org/10.1080/09613210903263007>
- Sorrell, S., Dimitropoulos, J., & Sommerville, M. (2009). Empirical estimates of the direct rebound effect: A review. *Energy Policy*. <https://doi.org/10.1016/j.enpol.2008.11.026>
- Sunikka-Blank, M., & Galvin, R. (2012). Introducing the rebound effect: The gap between performance and actual energy consumption. *Building Research & Information*, 40(3), 260–273. <https://doi.org/10.1080/09613218.2012.690952>
- Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8), 1819–1835. <https://doi.org/10.1016/j.rser.2008.09.033>
- Tabor, J. (2010). Investigating the investigative task: Testing for skewness an investigation of different test statistics and their power to detect skewness. *Journal of Statistics Education*, 18(2), 1–13. <https://doi.org/10.1080/10691898.2011.11889611>
- TABULA. (2013). *TABULA calculation method – Energy use for heating and domestic hot water*. [http://episcopes.eu/fileadmin/tabula/public/docs/report/TABULA\\_CommonCalculationMethod.pdf](http://episcopes.eu/fileadmin/tabula/public/docs/report/TABULA_CommonCalculationMethod.pdf).
- TABULA. (2016). *Tabula WebTool*. <http://webtool.buildingtypology.eu/>.
- Teli, D., Dimitriou, T., James, P., Bahaj, A., Ellison, L., & Waggott, A. (2016). Fuel poverty-induced ‘rebound effect’ in achieving the anticipated carbon savings from social housing retrofit. *Building Services Engineering Research and Technology*, 37(2), 176–193. <https://doi.org/10.1177/0143624415621028>
- Tindale, A. (1993). Third-order lumped-parameter simulation method. *Building Services Engineering Research and Technology*, 14(3), 87–97. <https://doi.org/10.1177/014362449301400302>
- Versluis, C., & Straetmans, S. (2015). Skewness measures for the Weibull distribution. *SSRN Electronic Journal*, 1–16. <https://doi.org/10.2139/ssrn.2590356>.
- Vivian, J., Zarrella, A., Emmi, G., & De Carli, M. (2017). An evaluation of the suitability of lumped-capacitance models in calculating energy needs and thermal behaviour of buildings. *Energy and Buildings*, 150, 447–465. <https://doi.org/10.1016/j.enbuild.2017.06.021>

- Wais, P. (2017). Two and three-parameter Weibull distribution in available wind power analysis. *Renewable Energy*, *103*, 15–29. <https://doi.org/10.1016/j.renene.2016.10.041>
- Wargocki, P. (2018). The effects of ventilation in homes on health. *International Journal of Ventilation*, *12*(2), 101–118. <https://doi.org/10.1080/14733315.2013.11684005>
- Wickham, H. (2016). *ggplot2 — Elegant graphics for data analysis* (2nd ed.). Springer. [https://doi.org/10.1007/978-3-319-24277-4\\_4](https://doi.org/10.1007/978-3-319-24277-4_4).
- Zahiri, S., & Elsharkawy, H. (2018). Towards energy-efficient retrofit of council housing in London: Assessing the impact of occupancy and energy-use patterns on building performance. *Energy and Buildings*, *174*, 672–681. <https://doi.org/10.1016/j.enbuild.2018.07.010>
- Zekar, A., & El Khatib, S. (2018). Development and assessment of simplified building representations under the context of an urban energy model: Application to arid climate environment. *Energy and Buildings*, *173*, 461–469. <https://doi.org/10.1016/j.enbuild.2018.04.030>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.