Overheating in English Dwellings: Comparing modelled and monitored large-scale datasets

Phil Symonds[1*], Jonathon Taylor[1], Anna Mavrogianni[1], Mike Davies[1], Clive Shrubsole[1], Ian Hamilton[2], Zaid Chalabi[1,3]

[1]UCL Institute for Environmental Design and Engineering, UCL, London
[2]UCL Energy Institute, UCL, London
[3]London School of Hygiene and Tropical Medicine, London
[*]Corresponding author email: p.symonds@ucl.ac.uk

Acknowledgments

The work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/I02929X/1; and the National Institute for Health Research Health Protection Research Unit at the London School of Hygiene and Tropical Medicine in partnership with Public Health England (PHE), and in collaboration with the University of Exeter, University College London, and the Met Office. We are grateful to the Met Office for making the MetUM available and to the National Centre of Atmospheric Science (NCAS) Computational Modelling Support (CMS) for providing technical support. Built environment input parameters for UHI modelling were provided from the UCL Centre for Advanced Spatial Analysis (CASA) using the GLUD database. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work.

We would also like to thank the Department of Energy and Climate Change (DECC) for allowing us access to the Energy Follow-Up Survey dataset, without which this work would not have been possible.

Overheating in English Dwellings: Comparing modelled and monitored large-scale datasets

Abstract

Monitoring and modelling studies of the indoor environment indicate that there are often discrepancies between simulation results and measurements. The availability of large monitoring datasets of domestic buildings allows for more rigorous validation of the performance of building simulation models derived from limited building information, backed by statistical significance tests and goodness of fit metrics. These datasets also offer the opportunity to test modelling assumptions. This paper investigates the performance of domestic housing models using EnergyPlus software to predict maximum daily indoor temperatures over the summer of 2011. Monitored maximum daily indoor temperatures from the English Housing Survey's (EHS) Energy Follow-Up Survey (EFUS) for 823 nationally representative dwellings are compared against predictions made by EnergyPlus simulations. Due to lack of information on the characteristics of individual dwellings, the models struggle to predict maximum temperatures in individual dwellings and performance was worse on days when the outdoor maximum temperatures were high. This research indicates that unknown factors such as building characteristics, occupant behaviour and local environment makes the validation of models for individual dwellings a challenging task. The models did, however, provide an improved estimate of temperature exposure when aggregated over dwellings within a particular region.

Keywords
Overheating, Housing Stock, Building information modelling (BIM), Validation, Occupant behaviour, EnergyPlus

1. Introduction

The exposure to high indoor temperatures in UK domestic dwellings is of increasing concern due to the established relationship between excess heat and mortality and human performance (Armstrong et al., 2011, ZCH, 2015). The warming climate is expected to lead to an increase in extreme temperature events (Murphy et al., 2009), and energy-efficient building design and retrofit standards may increase the propensity of dwellings to overheat (Mavrogianni et al., 2012, Taylor et al., 2014). Several empirical studies have sought to examine the overheating performance of buildings in the UK by monitoring indoor temperatures. Beizaee et al. (2013) monitored indoor temperatures in 207 dwellings across England, showing detached and pre-1919 homes being significantly cooler than average, and flats and modern dwellings being significantly warmer. A study of 282 homes in Leicester indicated flats had a significantly greater overheating risk than other built forms, while dwellings with solid walls showed a lower risk (Lomas & Kane, 2012).

A larger number of modelling studies (Oikonomou et al. (2012); Hamdy and Hensen (2015); Porritt and Cropper (2010); Peacock, Jenkins, and Kane (2010); Gupta and Gregg (2013)) have examined overheating risk in UK dwellings, in particular the characteristics of dwellings that may influence overheating risk and the adaptations that may be employed to reduce this risk under various climate and occupancy scenarios. Modelling approaches are often used in overheating studies for two main reasons: their ability to model a large number of building variants at a much lower cost than monitoring studies, and also their ability to examine risks under a variety of possible scenarios. Recent large-scale indoor overheating modelling studies in the UK include an investigation of heat-related mortality across the London housing stock (Taylor et al., 2015) and housing modification of heat exposure in Great Britain (Taylor et al., 2016). While the simulation results of these models support empirical findings showing that top-floor flats and highly-insulated buildings may be susceptible to elevated indoor temperatures, there has yet to be a detailed validation of model outputs against a large database of monitored data. Additionally, dwelling stock-level models lack the detailed building and occupancy information that would allow a precise prediction of indoor temperature for individual buildings, but rather rely on the assumption that aggregated overheating estimates would reflect the trends across the building stock itself.

Validation of Dynamic Thermal Building Simulation Software (DTBSS), such as EnergyPlus (US-DoE, 2013), against empirical data is usually only performed for individual buildings or in a test cell environment. Empirical validation of indoor temperatures for building energy simulation programs against monitored data has previously been performed in a number of studies (Zhuang et al, 2010; Royer et al, 2013; Mateus et al, 2014), while combined empirical and inter-model validation of EnergyPlus-predicted indoor temperatures has been performed by Henninger and Witte (2015), Lomas et al (1997), Strachan et al (2015), and Buratti et al (2013). Additionally, calibration of models – or using empirical data to adjust EnergyPlus modelling parameters- has been carried out using sensitivity analyses by Pereira, Bögl and Natschläger (2014) and Roberti, Oberegger, and Gasparella (2015). The validation of DTBSS-derived indoor temperatures is a difficult task due to the large number of parameters relating to occupancy, building construction and microclimate and their associated uncertainties. This makes it hard to draw concrete conclusions about what is causing differences between individual dwellings. The advantage of comparing DTBSS models against large monitored datasets is that outputs averaged over many dwellings can be observed, helping to evaluate the general trends when comparing across various dwelling types and model uncertainty.

The aim of this paper is to evaluate the indoor overheating modelling framework described by Taylor et al. (2015, 2016) and Symonds et al. (2016), which estimates indoor temperature metrics for building variants based on limited available input data from building stock databases. This modelling framework is the basis for ongoing research into the modification of population temperature exposure by the English housing stock under current and future climate and adaptation scenarios - scenarios which cannot be captured by simply creating a statistical model using the EFUS dataset. The evaluation will be performed by comparing monitored indoor temperatures from 823 dwellings in the English Housing Survey's (EHS) Energy Follow-Up Survey (EFUS) (DECC, 2013a) with EnergyPlus modelled indoor temperatures for the same set of dwellings. Overheating trends between modelled and monitored datasets are investigated as a time-series, the relationship between external and indoor temperatures, and by comparing the mean of the daily maximum indoor temperature aggregated by dwelling type. A number of statistical tests and goodness of fit metrics are used to evaluate the model results. The work is, to our knowledge, the first comparison between a large dataset of monitored and modelled dwellings.

2. Methods

This study makes comparisons between simulated and monitored indoor temperatures over the 2011 summer period (May 1st-September 30th). The monitored temperatures were obtained from the EFUS dataset compiled by the Department of Energy and Climate Change (DECC), whilst EnergyPlus was used for the building simulations over the same period. Model performance evaluations focussed on daily maximum indoor temperatures ($T_{\mathrm{Room}}^{\max}$) due to the association of outdoor maximum temperatures with mortality (Armstrong et al., 2011), and their previous use in estimating mortality risk modification by buildings (Taylor et al., 2015).

2.1 Energy Follow-Up Survey Data

EFUS consists of a monitored subset of the dwellings in the 2010/2011 English Housing Survey (EHS) (DCLG, 2011), with sub-hourly temperature measurements in 823 English homes. The EHS is a national survey commissioned by the Department for Communities and Local Government

(DCLG) which has taken place every two years since 2008. The survey consists of a household interview as well as a physical inspection of dwellings. The Energy Follow-Up Survey was initiated by DECC in 2011 (with monitoring running into 2012) as a revisit to a subset of the homes surveyed within the EHS. EFUS consists of more detailed data collected with the aim of understanding the changing patterns in energy use within homes. This is a unique dataset because it includes indoor temperatures (collected over at least 13 months), energy use data, and detailed dwelling information. Both datasets include a large sample of dwellings, designed to be representative of the English housing stock. In this paper, we use the EFUS temperature readings from all 823 dwellings across England.

The details of the monitoring method and data collection are provided elsewhere (DECC, 2013b). We provide here a brief summary. The monitored data were recorded sub-hourly in up to three rooms including living rooms, bedrooms and hallways. The temperature measurements were made using Tiny Tag Transit data loggers (Gemini, Chichester, UK) which have a measurement accuracy of ±0.2 ºC, a resolution of 0.01 ºC and range of -70 ºC to +40 ºC. The memory capacity of the data loggers is 32,000 readings. Data were recorded at 20 minute intervals, as opposed to shorter time intervals, in order to maximise the time frame over which the measurements were made. In order to match the temporal resolution of the simulated output, data from the 20 minute intervals was averaged to obtain an hourly value. The data loggers were installed by EFUS survey interviewers who were instructed to place loggers on internal walls, out of direct sunlight and away from heat sources. Of the 823 houses analysed, 763 dwellings had recordings from three data loggers, whilst 60 dwellings had one or two data loggers installed. Uncertainties and potential biases in the EFUS air temperature measurements are detailed in the EFUS methodology report (DECC, 2013b). Dwellings where any extreme temperature readings were recorded (≤-10 ºC or ≥40 ºC) were removed from further analysis, under the assumption that these readings were from poorly placed or faulty loggers. The remaining sample size used in the analysis was therefore data from 768 living rooms and 772 bedrooms. Data measured in the hallways is not used in this paper.

Each monitored dwelling can be cross-referenced, based on its building code, to the corresponding building and household information in the EHS database. Parameters in the EHS may then be used to inform the building characteristics of EFUS dwellings and provide the required inputs to building simulations (see section 2.2). The Government Office Region (GOR) in which each dwelling was located, was used to assign each dwelling to one of six regions whose climates were modelled in EnergyPlus (see section 2.2.1). Information on the local environment was used to classify the dwelling as being in either a rural, urban, or city location. Building fabric types for the EHS were used alongside building age to calculate *U*-values and building fabric permeability using the UK Governments Standard Assessment Procedure (SAP) for Energy in Buildings (BRE, 2009). Table 1 outlines the characteristics of the buildings and their occupants, monitored within EFUS.

2.2 Building Simulation

EnergyPlus (US-DoE, 2013) is a commonly used building physics simulation tool developed by the United States Department of Energy. It is able to model dynamic indoor temperatures and air movement within the building and can output metrics such as indoor operative temperatures, energy use and relative humidity at user defined time steps. EnergyPlus takes a large amount of user provided information related to the building, the occupants, and the surrounding environment as inputs via an input definition file (".idf"). A Red Hat Enterprise Linux 5 (RHEL5) version of EnergyPlus 8.1 has been used, due to its compatibility on high performance computing facilities. An in-house tool called EnergyPlus Generator2 (EPG2) was used to take information about the monitored buildings from the EHS and create an input file for each monitored building within EFUS.

EnergyPlus files were generated for each of the 823 EFUS dwellings, with their location based on GOR region. Building characteristics which were used as inputs to the model include the dwelling type, whether the dwelling has cavity or solid walls, the glazing fraction[1], the ceiling height, the usable floor area[2] and the *U*-values of the windows, walls, roof and floors (altered by adjusting material thicknesses and thermal conductivities). The layouts of the building archetypes were held constant and were chosen to be representative of the English stock, based on floor plans as described in Symonds et al. (2016). Buildings are shaded by surrounding buildings of the same type (i.e. terraced houses have a row of terraced houses in front and behind them, as well as adjoining dwellings). Information on the orientation of buildings is not provided in the EHS and so this variant was selected at random for each dwelling from a uniform distribution in the range 0º to 360º East of North. Variations within ceiling height, usable floor area, and window size have been incorporated into the model by recalculating the position of surface vertices for each individual building. Further details of the EnergyPlus model development are described in Symonds et al (2016).

2.2.1 Regional Weather Data

EnergyPlus requires the weather data over the course of the simulation period to be specified as an input (".epw") file. Real local weather data from 2011 was required for modelling, however location precision was limited by the fact that EFUS/EHS dwellings are only locatable by region (GOR) rather than a specific geographic location. Since weather files for particular years are not freely available for particular locations in England, they had to be created using raw data from various weather stations. Weather data from the MetOffice Integrated Data Archive System (MIDAS) database (Met Office, 2012) has been used for this purpose. This database contains hourly weather data (including dry bulb and dew point temperatures, wind speed and direction, atmospheric pressure, precipitation, solar radiation, cloud cover) for various locations within the UK and abroad from 1853 up to present day. Jentsch, Bahaj and James (2008) describe methods by which raw weather data can be converted into an EnergyPlus weather file. Global horizontal radiation is the only radiation variable recorded at a limited number of MIDAS weather stations. This meant that the other radiation variables had to be calculated using equations based on geometry (CIBSE, 2002) and the recorded cloud cover. Infra-red radiation from the sky was calculated using cloud cover, dry bulb temperature and vapour pressure following Crawford and Duchon (1999). Illuminance data was calculated from radiation parameters following Perez et al. (1990). The locations of the weather stations used are shown in Figure 1. The choice of weather stations was limited by which locations had a full set of weather observations available in 2011. To fill the gaps in incomplete weather data, it was sometimes necessary to piece together data from two weather stations in relatively close proximity to one another. A full account of the weather data used is given in Table 2.

2.2.2 Occupant Behaviour

Occupant behaviour was modelled in EnergyPlus by creating schedules for each occupant within idf input files. We modelled two occupancy types; a family of five who are out during the day and two pensioners who are assumed to stay at home during the day. EFUS homes containing a couple aged

---

[1] Glazing fraction is calculated using the information on the external wall and window areas for the front and the back of the dwelling from the EHS. i.e. glazing fraction = (front window area + back window area) / (front wall area + back wall area).

[2] Useable floor area is defined in the EHS as floor space that can be reasonably used for habitation. It represents all area within the footprint excluding 1) the area under external walls, 2) the are under internal walls, 3) area occupied by staircases. Cupboards, integral balconies, and internal garages are included. Lofts are only included if habitable, with a fixed stair in place to access it.

60 or over with no dependent child(ren) or one person aged over 60 were modelled as pensioner occupants. All other occupancies (defined in Table 1) were modelled as a family. Each occupant was assumed to have a metabolic rate of 100 Watts when awake and 80 Watts when sleeping (Ainsworth et al., 1993). Occupants were assumed to move between rooms in the dwelling depending on the time of day and were able to control various aspects of their environment such as whether or not windows were open. Set schedules were assumed for cooking and the use of electrical appliances, contributing to internal gains. Details on the occupancy schedule and internal gains can be found in Symonds et al. (2016). Variations within occupancy behaviour have been modelled by varying the following:

(i) Annual energy consumption of electrical appliances (MWh). A high consumption of electrical appliances such as kettles and TVs will lead to greater internal gains within the dwelling. The use of electrical appliances is fixed according to set schedules. The power of appliances was varied up (×2) and down (×0.5) to reflect changes in appliance usage. The base-case value of 4 MWh per year was chosen based on findings in DECC (2011c, 2015) reports.

(ii) Window opening temperature thresholds (between May and September). Windows in the bedrooms can be opened between 10 pm at night and 8 am in the morning. In other rooms, pensioners are able to open and close windows all day (8 am – 10 pm). Families can open windows between 8 am and 9 am and 6 pm and 10 pm but close windows during the day when they are assumed to be out.

The simulated baseline and the variations in occupancy parameters are shown in Table 3. These variations in occupancy meant that the 823 buildings were simulated five times. To speed up the simulation runs considerably, the EnergyPlus simulations were run in parallel on high performance computing facilities.

2.3 Statistical assessment of model performance

An assessment of the model's ability to predict the EFUS data was made by comparing the modelled and monitored daily maximum temperatures in the living room ($T_{\mathrm{liv}}^{\mathrm{max}}$) and bedroom ($T_{\mathrm{bed}}^{\mathrm{max}}$). Pearson's correlation coefficient (*r*) and the Root Mean Square Error (*RMSE*) were used to evaluate the performance of the model for each individual dwelling using daily values over the whole summer period. Pearson's correlation coefficient was used to test the level of dependence between the modelled and monitored daily maximum temperatures. It is calculated as:

$$r = \frac{\sum_{t=1}^{n}(x_t - \bar{x}_t)(y_t - \bar{y}_t)}{\sqrt{\sum_{t=1}^{n}(x_t - \bar{x}_t)^2}\,\sqrt{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2}}$$

where *n* is the number of days over the summer period (153), *t* is the day number (i.e. *t* = 1 is May 1st), x_t and y_t are the modelled and monitored daily maximum indoor temperatures, $\bar{x}_t$ and $\bar{y}_t$ are the mean modelled and monitored daily maximum indoor temperatures over the entire summer period. *RMSE* is used to assess the mean deviation between the daily maximum temperature predicted by the model to that of the monitored data. *RMSE* is calculated using the following equation:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_t - y_t)^2}{n}}$$

with *n*, *t*, x_t and y_t defined as above. As there are too many dwellings for us to present the *r* and *RMSE* for each entry, we instead use four performance metrics to evaluate the model performance for dwelling types within regions:

- $\bar{r}$ and $\overline{RMSE}$: the average *r* and *RMSE* across all dwellings of a certain type within a particular region.
- r_{μ} and $RMSE_{\mu}$: the *r* and *RMSE* of the average maximum daily temperatures across dwellings of a certain type within a particular region.

We use r_{μ} and $RMSE_{\mu}$ to determine if a better model performance is observed when attempting to predict aggregated data from dwellings. The model performance was also evaluated as a function of the maximum external temperature ($T_{\mathrm{Ext}}^{\mathrm{max}}$). This was achieved by calculating $\bar{r}$ and $\overline{RMSE}$ using modelled and monitored data from days where the external temperature at the regional weather station (described in section 2.2.1) exceeded certain thresholds. The external temperature threshold ($T_{\mathrm{Ext}}^{\mathrm{max}}$ thresh.) was varied in the range 10-30 ºC in increments of 2 ºC.

The models prediction of the mean of the maximum indoor temperature over the whole summer period was evaluated with respect to various characteristics of dwellings and their geographic location (GOR). Two tailed t-tests were used to test the statistical significance of the level of agreement between the modelled and monitored mean maximum temperatures (Mean $T_{<\mathrm{Room}>}^{\mathrm{max}}$). The null hypothesis being tested was that the means of the modelled and monitored daily maximum temperatures are equal. P-values were calculated to test the predictions of the models allowing the null hypothesis to be accepted or rejected at a certain level of statistical significance.

3. Results

3.1 Comparisons of daily maximum temperatures in living rooms

Time-series comparisons between the modelled and monitored maximum daily temperatures for all dwellings in the dataset over the 2011 summer period (May-September) have been performed. The metrics ($\bar{r}$, $\overline{RMSE}$, r_{μ} and $RMSE_{\mu}$) described in section 2.3 were used to assess model performance. Table 4 gives all $\bar{r}$, $\overline{RMSE}$, r_{μ} and $RMSE_{\mu}$ values for the various dwelling types within each of the six modelled regions. The standard errors on the $\bar{r}$ and $\overline{RMSE}$, $\sigma_{\bar{r}}$ and $\sigma_{\overline{RMSE}}$, defined as the standard deviation divided by the square root of number of dwellings, are also shown. The standard errors indicate the spread in the mean performance of the models. The $\bar{r}$ and $\overline{RMSE}$ calculated across all dwellings within the dataset are 0.52 and 2.66 ºC, respectively. When comparing average daily temperatures across all dwellings by type within all regions, values of 0.82 and 1.37 ºC are achieved for r_{μ} and $RMSE_{\mu}$, respectively. These values were found by finding the dwelling weighted average of the r_{μ} and $RMSE_{\mu}$ values in Table 4. This indicates that an improvement in model performance can be achieved by grouping dwellings together by type for a particular region.

When looking at particular dwelling types within regions, the lowest $\bar{r}$ was observed for converted flats in the North East (0.24), whilst the highest is for high-rise flats in the North East (0.73). The highest $\overline{RMSE}$ is for bungalows in the North West (4.36 ºC), whilst the lowest is for low-rise flats in the South West (1.62 ºC). Figure 2(a-d) show some time-series comparisons of the monitored and modelled daily maximum living room temperatures for the best and worst performing models in terms of $\bar{r}$ and $\overline{RMSE}$ with more than 10 dwellings. The median properties of these dwellings are given in Table A1. The modelled time-series were produced using the base-case occupancy behaviour defined in Table 3. The time-series distributions plot the mean of the EnergyPlus prediction and the EFUS data for particular dwelling types within each of the six regions. The 95% confidence intervals (CI) indicate a large spread in the modelled and monitored data. The confidence interval is, in general, slightly wider for the monitored than the modelled data. This could be due to the fact that variations within occupant behaviour and local climate are inherently accounted for in the data but are less so in models.

The performance of EnergyPlus models was also evaluated as a function of the external temperature. Figure 3(a-d) show the relationship between the external temperature threshold and the mean of the maximum living room temperature on days where the external temperature threshold was exceeded, for the best and worst performing models (with more than 10 dwellings). $\bar{r}$ and $\overline{RMSE}$ are displayed at the bottom of the plot alongside the number of days on which external temperature thresholds were exceeded. The prediction of EnergyPlus models tend to diverge from the EFUS data at higher external temperatures with the modelled prediction overestimating internal temperatures. This is reflected in the $\bar{r}$ and $\overline{RMSE}$ values.

3.2 Statistical significance of mean maximum temperature predictions

An evaluation of the statistical significance of the differences between EnergyPlus predictions and the EFUS data was performed using two tailed t-tests. These tests were able to determine whether the mean of the daily maximum temperatures predicted by the models is in agreement with the monitored data. P-values were calculated for EnergyPlus models categorised by dwelling type and by the GOR of the dwelling under base-case and the variations in simulated occupant window opening and electrical consumption behaviour shown in Table 3. Table 5 displays the p-values of the t-tests comparing the means of the simulation results produced by EnergyPlus and the counterpart EFUS data. The null hypothesis, that the two means are compatible with one another at the 95% confidence level, was accepted in four out of the eight base-case comparisons of living room temperatures. This indicates that some of the modelling assumptions may need refining. The models tend to perform worse when predicting bedroom temperatures, where only the high-rise flat, which is a relatively small sample of dwellings (9) can be accepted at the 95% confidence level.

Figure 4 shows a box and tail plot comparing the mean maximum temperature for various dwelling types over the summer period in the living room simulated under base-case occupancy. The number of dwellings and the p-values are shown at the bottom of the plot. The results of both the monitoring and modelling seem to support previous findings (Beizaee et al., 2013), and suggest that higher average temperatures are observed in flats. Detached homes are observed to be least prone to higher average temperatures during the summer. Figure 5 shows a box and tail plot comparing modelled and monitored mean maximum temperatures across the nine GOR regions. The results indicate that, as expected, homes in London and the South East are most prone to overheating, whilst homes in the North and West are least prone to overheating. The p-values are shown at the bottom of the plot and indicate that models for six of the nine regions can be accepted at the 95% confidence level. Models for the North West, West Midlands, and South East fail to satisfy the null hypothesis at 95% confidence. This indicates that the weather stations used to model these regions are not representative of the average weather for the region in question.

4. Discussion

The use of rich monitored datasets such as EFUS provides an excellent opportunity to thoroughly evaluate the predictions made using the EnergyPlus building physics software. The direct validation of EnergyPlus (EP) simulations against monitored data is a challenging task. This has been demonstrated to a certain extent by the results of our analyses. Although we had a relatively large sample of dwellings (~800) we did not have all the relevant information on individual dwellings (e.g. orientation, occupancy behaviour, local weather information). This meant that we had to make assumptions on what some of the input parameters entering the EnergyPlus models should be. Past sensitivity analyses have shown that some of the missing information, such as occupancy behaviour, is important (Mavrogianni et al., 2014). This means that we are not in a position where we are able validate the model against monitored data for individual dwellings. We can however, achieve more valid predictions when modelling collectively for a groups of dwelling types or those in the same geographic region. Naturally, this resulted on occasions in large CIs. Although the central estimates of the EnergyPlus simulations were mostly within the respective CIs, this poses some difficulty in interpreting the goodness of the model in situations with large CIs.

Pearson's correlation coefficients, *RMSE*s, and t-tests have been used to evaluate the performance of model at predicting maximum indoor temperatures. The mean correlation, $\bar{r}$, between the daily maximum living room temperature predicted by EnergyPlus and that in the monitored EFUS dataset averaged across all dwellings was 0.52. The $\overline{RMSE}$ between the modelled and monitored data for all dwellings was found to be 2.66 ºC with the performance getting worse on days when external temperatures were higher. The results highlight the difficulties in predicting maximum temperatures for individual dwellings. An improved prediction can be made when estimating the average of the maximum indoor temperature for groups of dwelling types within particular regions. In this case it was possible to achieve a correlation of 0.82 and *RMSE* of 1.37 ºC. These values were calculated by taking the dwelling weighted averages of r_μ and $RMSE_\mu$ from Table 4. This suggests that in 68.2% of cases (since *RMSE* represents one standard deviation), we are able to predict the average maximum indoor temperature for a given dwelling type in a region to within approximately ±6%.

Figure 4 shows that a modest agreement between the modelled and the monitored data is observed for the various types of built form. Some improvements in the agreement between data measurements and simulation were observed when using an alternative occupant variation from the base-case for some of the dwelling types. The statistical significance hypothesis tests suggest that occupant behaviour may have a strong influence on the agreements between simulations and measurements. This could reflect adaptive occupant behaviour adopted to prevent high temperatures in more overheating-prone dwelling types. It may also be easier to keep windows open for longer in some dwelling types such as high-rise flats without security fears. The box plots also show that the simulated variance in indoor temperatures is less than those in the monitored data. This can be explained by the fact that occupant behaviour and environmental factors such as building shading and the UHI are being held constant in models.

The models were seen to perform poorly for bedrooms, which suggests that occupancy schedules need revising. Occupants may, for example, spend more time in bedrooms or open windows differently to what is currently modelled and hence occupant behaviour, in this regard, may be having more influence on the indoor temperature. Figure 3 indicated that the models perform less well at high external temperatures. This could be explained by occupants making preventative measures which are not modelled, such as leaving doors and windows open for longer or keeping curtains closed. Due to the complexity of deriving regional climate files, some GOR regions were combined. The weather data used in simulations was generally taken from rural areas such as air fields rather than urban areas, and therefore do not include temperature increments due to the Urban Heat Island (UHI). The weather stations chosen may not be representative of the locations of the dwellings in that particular region. The lack of information about the EFUS dwelling locations makes this assumption hard to test.

The differences between the predictions of EnergyPlus and the EFUS indoor temperature measurements can be explained by several factors including 1) a lack of information on building occupant behaviour, 2) the simulated regional climates used in the EnergyPlus models differing from the actual local climates to which the monitored dwellings are exposed, 3) a lack of information about local shading, surrounding terrain, or building orientation in the EFUS/EHS datasets, 4) the uncertainty in the model inputs including inferred building characteristics calculated using SAP (Francis et al., 2014), and 5) biases and uncertainties which may be present in the monitored EFUS data, for example due to sensor exposure to radiation heat transfer. The models are able to account for variations in some occupant behaviours, such as temperature thresholds for window-opening, and the internal gains produced by electrical equipment. However, other important behavioural-related variables that may contribute to internal gains, such as cooking schedules, were held constant in the simulations.

Future work will focus on improving model performance. We are seeking to gain secure access to the EHS which will enable us to locate dwellings more precisely. This will allow more localised weather to be used in the models. A metamodelling framework has been developed by Symonds et al. (2016) based on EnergyPlus outputs, which enables overheating estimates to be rapidly calculated for individual dwellings. This framework will be used to optimise the input variables for individual buildings, to enable model calibration. The metamodel will also help to quantify the uncertainties in the building fabric characteristics, such as *U*-values and permeability, of the surveyed EHS dwellings and also to investigate the association between built forms and occupant behaviours. Different ranges and distributions of occupant behaviour could also be tested for particular dwelling and household types. Future EnergyPlus validation work could also incorporate energy measurements made as part of EFUS.

In the absence of making dedicated measurements (which are very costly) whose purpose is to validate large scale housing models such as EnergyPlus, researchers will always be confronted with using data whose purpose was not to validate their models. Nevertheless, the existence of such large datasets has provided us with the opportunity to evaluate the performance of EnergyPlus models at the building stock level at a scale not previously seen before.

5. Conclusion

In this paper, the results of building simulations carried out using EnergyPlus were compared to temperature data recorded in 823 homes over the summer of 2011. Six regional climates within England were modelled using weather data compiled from MIDAS weather stations. The model prediction of the maximum daily temperatures within individual dwellings correlated to the data at a level of about 0.5 with an average *RMSE* of 2.66 ºC found. When modelling daily maximum temperatures across aggregate dwellings within particular regions the performance improved by nearly a factor of two, with 0.82 and 1.37 ºC calculated for r_{μ} and $RMSE_{\mu}$, respectively. The results of the statistical hypothesis tests against model predictions, indicate that some of the modelling assumptions may need modifying. Particularly in bedrooms and in relation to occupant behaviour and local weather conditions. The availability of the large scale EFUS dataset has allowed a thorough evaluation of the performance of the models against empirical data. The large number unknowns relating to building characteristics, occupants and local climate makes it difficult to draw concrete conclusions about which model assumptions are right and which are wrong. This emphasises the need for further research into the tuning and uncertainty analysis of model inputs and assumptions using a metamodeling technique.

References

Ainsworth B.E., Haskell, W. Leon A.S., Jacobs D.R. Jr, Montoye H.J., Sallis J.F., Paffenbarger, R.S. Jr. 1993. Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*. 25(1):71-80.

Armstrong, B.G. Chalabi, Z. Fenn, B. Hajat, S. Kovats, S. Milojevic, A. Wilkinson, P. 2011. Association of mortality with high temperatures in a temperate climate: England and Wales. *Journal of epidemiology and community health*, 65(4), pp.340–5.

Beizaee, A., Lomas, K.J. & Firth, S.K., 2013. National survey of summertime temperatures and overheating risk in English homes. *Building and Environment*, 65(null), pp.1–17.

BRE, 2009. The Government's Standard Assessment Procedure for Energy Rating of Dwellings, Watford, UK: *Building Research Establishment*.

Buratti, C., Moretti, E., Belloni, E. and Cotana, F., 2013. Unsteady simulation of energy performance and thermal comfort in non-residential buildings. *Building and Environment, 59*, pp.482-491.

CIBSE, CIBSE Guide J: Weather, Solar and Illuminance data, The Chartered Institution of Building Services Engineers, London, 2002.

CIBSE. 2003. British Standards: Heating systems in buildings. Method for calculation of the design heat load." *BSi* (12831).

Crawford, T.M., Duchon. C.E. An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. Journal of Applied Meteorology, 38 (4) (1999), pp. 474–480.

DCLG. 2011. The English Housing Survey. *Department for Communities and Local Government* (76).

DECC, 2013a. Energy Follow-Up Survey 2011: Report 2: Mean household temperatures, London, UK.

DECC, 2013b. Energy Follow-Up Survey 2011: Report 11: Methodology, London, UK.

DECC, 2013c. Energy Follow-Up Survey 2011: Report 9: Domestic appliances, cooking & cooling equipment, London, UK.

DECC. 2015. Energy Consumption in the UK. Chapter 3: Domestic energy consumption in the UK between 1970 and 2014.

Francis, G. N. Li. Smith, A.Z.P., Biddulph, P. Hamilton, I. Lowe, R. Mavrogianni, A. Oikonomou, E. Raslan, R. Stamp, S. Stone, A. Summerfield, A.J, Veitch, D. Gori, V. Oreszczyn, T. Solid-wall U -values: heat flux measurements compared with standard assumptions. *Building Research & Information* (43:2):238-252

Gemini Data Loggers Limited, Chichester, UK. http://www.geminidataloggers.com/.

Gupta, R., and M. Gregg. 2013. Preventing the overheating of English suburban homes in a warming climate. *Building Research & Information* (41): 281-300.

Hamdy, M., and J. Hensen. 2015. Ranking of Dwelling types in terms of Overheating Risk and Sensitivity to Climate Change. *International Conference of the International Building Performance Simulation Association (BS2015)*, 7-9 December 2015, Hyderabad, India.

Henninger, R.H. and Witte, M.J., 2012. EnergyPlus Testing with Building Thermal Envelope and Fabric Load Tests from ANSI/ASHRAE Standard 140-2011.

Jentch M., Bahaj A., James P. Climate change future proofing of buildings generation and assessment of building simulation weather files. *Energy Build.*, 40 (12) (2008), pp. 2148–2168.

Lomas, K. Eppel, H., Martin, C. Bloomfield, D. 1997. Empirical validation of building energy simulation programs. *Energy and Buildings* (26): 253-275

Lomas, K. & Kane, T., 2012. Summertime temperatures in 282 UK homes: thermal comfort and overheating risk. In *Proceedings of 7th Windsor Conference: The changing context of comfort in an unpredictable world*.

Mateus, N.M., Pinto, A. and da Graça, G.C., 2014. Validation of EnergyPlus thermal simulation of a double skin naturally and mechanically ventilated test cell. *Energy and Buildings, 75*, pp.511-522.

Mavrogianni, A. Wilkinson, P. Davies, M. Biddulph, P. Oikonomou, E. 2012. Building characteristics as determinants of propensity to high indoor summer temperatures in London dwellings. *Building and Environment*, 55(null), pp.117–130.

Mavrogianni, A. M. Davies, J. Taylor, Z. Chalabi, P. Biddulph, E. Oikonomou, P. Das, B. Jones., 2014. The impact of occupancy patterns, occupant-controlled ventilation and shading on indoor overheating risk in domestic environments. *Building and Environment*, 78, pp.183–198.

Met Office, 2012. Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). *NCAS British Atmospheric Data Centre*.

Murphy, J. et al., 2009. *UKCP09 Climate change projections*, Exeter.

Oikonomou, E., M. Davies, A. Mavrogianni, P. Biddulph, P. Wilkinson, and M. Kolokotroni. 2012. Modelling the relative importance of the urban heat island and the thermal quality of dwellings

for overheating in London. *Building and Environment* (57): 223-38.
Peacock, A. D., D. P. Jenkins, and D. Kane. 2010. Investigating the potential of overheating in UK dwellings as a consequence of extant climate change. *Energy Policy* (38: 32);77-88.
Perez, R., Ineichen, P., Seals, R., Michalsky, J., Stewart, R. Modelling daylight availability and irradiance components from direct and global irradiance. Solar Energy, 44 (5) (1990), pp. 271–289.
Porritt, S., and P. Cropper. 2010. Heat wave adaptations for UK dwellings and introducing a retrofit toolkit. *International Journal of Disaster Resilience in the Built Environment* (4:3): 269-286.
Roberti, F., Oberegger, U.F. and Gasparella, A., 2015. Calibrating historic building energy models to hourly indoor air and surface temperatures: Methodology and case study. *Energy and Buildings*, *108*, pp.236-243.
Royer, S., Bressan, M., Thil, S. and Talbert, T., 2013, August. Modelling of a multi-zone building and assessment of its thermal behaviour using an energy simulation software. In *CASE* (pp. 735-740).
Stephen, R., 2000. Airtightness in UK Dwellings, Watford.
Strachan, P., Svehla, K., Heusler, I. and Kersken, M., 2015. Whole model empirical validation on a full-scale building. *Journal of Building Performance Simulation*, pp.1-20.
Symonds, P., J. Taylor, A. Mavrogianni, M. Davies, I. Hamilton, S. Vardoulakis, C. Heaviside, H. Macintyre. 2016. Development of an England-wide indoor overheating and air pollution model using articial neural networks. *Journal of Building Performance Simulation*.
Taylor, J. M. Davies, A Mavrogianni, C. Shrubsole, I. Hamilton, P. Das, B. Jones, E. Oikonomou, and P. Biddulph. 2016. Mapping indoor overheating and air pollution risk modification across Great Britain: A modelling study. *Building and Environment*.
Taylor, J., P. Wilkinson, M. Davies, B. Armstrong, Z. Chalabi, A. Mavrogianni, P. Symonds, E. Oikonomou, and S. I. Bohnenstengel. 2015. Mapping the effects of Urban Heat Island, housing, and age on excess heat-related mortality in London. *Urban Climate*, 14, pp.517–528.
Taylor, M. et al. 2014. Preventing Overheating. Good Homes Alliance.
US-DoE. 2013. EnergyPlus V8.1. http://apps1.eere.energy.gov/buildings/energyplus/.
ZCH. 2015. Overheating in Homes - The Big Picture. Zero Carbon Hub.
Zhuang, C.L., Deng, A.Z., Chen, Y., Li, S.B., Zhang, H.Y. and Fan, G.Z., 2010. Validation of veracity on simulating the indoor temperature in PCM light weight building by EnergyPlus. In *Life System Modeling and Intelligent Computing* (pp. 486-496). Springer Berlin Heidelberg.

		Number of households	**% of total households**
House type	Detached	143	17
	Semi-detached	244	30
	Mid-terrace	124	15
	End-terrace	83	10
	Bungalow	101	12
	Converted Flat	15	2
	Low-rise Flat	103	13
	High-rise Flat	10	1
Government Office Region (GOR)	North East	57	7
	North West	130	15
	Yorkshire and Humber	106	13
	East Midlands	79	9
	West Midlands	71	9
	East of England	112	14
	Greater London	62	8

	South East	125	15
	South West	81	10
Surrounding environment	City: population > 10k	639	78
	Urban: town, fringe or village	160	19
	Rural: hamlet and isolated dwellings	24	3
Household type	Couple, no dependent child(ren)	319	39
	Couple with dependent child(ren)	179	22
	Lone parent with dependent child(ren)	45	5
	Other multi-person households	40	5
	One person under 60	94	11
	One person aged 60 or over	146	18

Table 1 – Characteristics of the dwelling types, building locations and occupants within the Energy Follow-Up Survey (EFUS) dataset.

Region	Abbrev.	Weather data used
North East and Yorkshire and Humber	NE	Bramham in West Yorkshire was used for all weather data except cloud cover, visibility and atmospheric pressure where the Bingley weather station was used.
North West	NW	Ringway weather station at Manchester airport was used for weather observations. Hulme Library in Greater Manchester was used for solar and precipitation observations.
East Midlands and East of England	EM	Wittering Airfield, Cambridgeshire was used for all weather observations.
West Midlands	WM	The University of Birmingham weather station in Winterbourne was used for all observations except cloud cover, visibility, atmospheric pressure, and wind speed and direction where Birmingham Airport was used.
South East and Greater London	SE	Kenley Airfield, Greater London was used for all weather observations.
South West	SW	Dunkeswell Airfield, Devon was used for all weather except solar and precipitation observations where Exeter airport was used.

Table 2 – Summary of weather data obtained from the Met Office Integrated Data Archive System (MIDAS) used to create the weather files used in EnergyPlus simulations.

	Symbol	Unit	Base-case	Upward variation	Downward variation
Annual electrical appliance energy consumption	E_{use}	MWh	4	8	2

Window opening temperature threshold	W_{open}	ºC	22	26	18

Table 3 – Base-case occupancy behaviour for modelled annual electrical energy consumption and window opening threshold. Upward and downward variations (relative to the base-case) in window opening and electrical appliance usage were also modelled for all dwellings. The window opening temperature threshold relates to the temperature within the room at which windows are opened.

Built Form	**Region**	**N**	$\bar{r}$	$\sigma_{\bar{r}}$	$\overline{RMSE}$	$\sigma_{\overline{RMSE}}$	r_{μ}	$RMSE_{\mu}$
Terrace	NW	32	0.36	0.06	3.31	0.26	0.78	1.74
	NE	35	0.42	0.05	3.37	0.3	0.85	1.39
	WM	13	0.45	0.07	3.39	0.47	0.8	1.4
	EM	37	0.53	0.03	2.28	0.19	0.86	0.83
	SE	43	0.52	0.03	2.63	0.19	0.8	1.32
	SW	17	0.49	0.05	2.37	0.21	0.7	1.51
Semi-detached	NW	49	0.53	0.04	2.82	0.19	0.86	1.73
	NE	50	0.47	0.04	2.71	0.16	0.86	1.23
	WM	29	0.56	0.04	2.31	0.17	0.87	0.94
	EM	53	0.57	0.03	2.34	0.15	0.86	1.24
	SE	37	0.59	0.03	2.66	0.24	0.86	1.3
	SW	14	0.58	0.06	2.89	0.33	0.83	1.07
Detached	NW	12	0.43	0.09	2.66	0.24	0.81	1.85
	NE	21	0.57	0.04	2.8	0.29	0.84	1.42
	WM	13	0.52	0.04	2.86	0.24	0.83	1.58
	EM	40	0.64	0.03	2.4	0.16	0.89	1.09
	SE	33	0.62	0.03	2.81	0.21	0.85	1.29
	SW	20	0.55	0.05	2.33	0.17	0.87	1.43
Bungalow	NW	12	0.42	0.09	4.36	0.91	0.77	3.41
	NE	21	0.56	0.04	2.26	0.2	0.88	0.77
	WM	3	0.6	0.02	2.55	0.53	0.74	2.38
	EM	37	0.54	0.03	2.47	0.19	0.8	1.57
	SE	14	0.61	0.03	2.65	0.32	0.83	1.67
	SW	10	0.61	0.04	2.51	0.47	0.8	2.03
Converted Flat	NW	2	0.5	0.14	2.22	0.49	0.6	1.95
	NE	3	0.24	0.18	3.95	1.5	0.43	2.75
	WM	2	0.52	0.05	1.85	0.33	0.48	2.1
	SE	8	0.47	0.06	2.72	0.43	0.74	2.39
Low-rise Flat	NW	13	0.41	0.06	2.85	0.46	0.77	1.37
	NE	22	0.47	0.05	2.42	0.2	0.86	0.98
	WM	6	0.53	0.05	2.35	0.4	0.75	1.28
	EM	13	0.59	0.06	1.95	0.3	0.87	0.94
	SE	34	0.49	0.03	2.72	0.23	0.78	1.02
	SW	10	0.53	0.05	1.62	0.17	0.77	0.67
High-rise Flat	NW	2	0.7	0.02	2.73	0.55	0.75	2.5
	NE	3	0.73	0.06	3	0.62	0.81	1.83
	SE	4	0.41	0.05	3.13	0.35	0.67	2.42

Table 4 – Pearson correlation coefficients ($\bar{r}$) and Root Mean Square Errors ($\overline{\text{RMSE}}$) for maximum daily living room temperatures predicted by EnergyPlus models to the EFUS data. The results presented above are averages across the number of dwellings (i.e. sample size) shown in column, N. The standard error for both ($\bar{r}$) and $\overline{RMSE}$ are also presented to indicate the range

in the model performance. We also present r_μ and $RMSE_\mu$ which indicate the performance of the model when the mean maximum temperatures are averaged across dwelling types within regions.

		p-values				
Built form	Room	Base-case	$E_{\text{use}}^{\text{up}}$	$E_{\text{use}}^{\text{down}}$	$W_{\text{open}}^{\text{up}}$	$W_{\text{open}}^{\text{down}}$
Semi-detached	Bedroom	<0.001	<0.001	<0.001	<0.001	<0.001
	Living room	<0.001	<0.001	<0.001	<0.001	<0.001
Detached	Bedroom	0.001	<0.001	0.002	<0.001	<0.001
	Living room	0.574	0.27	0.151	0.186	<0.001
Bungalow	Bedroom	<0.001	<0.001	<0.001	<0.001	0.002
	Living room	<0.001	<0.001	0.001	<0.001	0.003
Mid-terrace	Bedroom	<0.001	<0.001	<0.001	<0.001	0.343
	Living room	0.024	<0.001	0.064	<0.001	<0.001
End-terrace	Bedroom	0.026	0.005	0.09	<0.001	<0.001
	Living room	0.78	0.067	0.4	0.013	<0.001
Low-rise flat	Bedroom	<0.001	<0.001	<0.001	<0.001	<0.001
	Living room	0.932	0.176	0.364	<0.001	<0.001
Converted flat	Bedroom	0.013	0.008	0.017	0.003	0.684
	Living room	0.01	0.002	0.027	<0.001	0.998
High-rise flat	Bedroom	0.592	0.713	0.531	0.895	0.083
	Living room	0.852	0.585	0.999	0.432	0.406

Table 5 – P-value results from the independent two sample t-tests between EFUS data and the EnergyPlus prediction of the mean maximum daily temperatures over the 2011 summer (May-Sept). Results are shown for the living room and bedroom. T-tests assumed non-equal variances between the two data samples.

Appendices

Region	Built Form	Wall type	N	Wall U-value ($W/m^2/K$)	Roof U-value ($W/m^2/K$)	Window U-value ($W/m^2/K$)	Floor U-value ($W/m^2/K$)	Perme-ability ($m^3/m^2/s$)	Glazing Fraction	Ceiling height (m)	Useable Floor Area (m^2)
SW	Lowrise	Cavity	10	0.64	2.3	2.76	0.23	10.05	0.25	2.3	47.51
	Lowrise	Solid	1	1.14	2.3	2.76	1.23	8.18	0.34	2	55.96
	Semi	Cavity	12	0.5	0.22	2.76	0.73	16.05	0.2	2.4	81.5
	Semi	Solid	3	2.1	0.28	4.03	0.72	21.06	0.29	2	130.73
	Bungalow	Cavity	8	0.5	0.36	2.76	0.65	10.81	0.22	2.4	88.19
	Bungalow	Solid	2	0.36	0.23	2.76	0.76	12.44	0.26	2.3	61.66
	Terrace	Cavity	15	0.64	0.28	2.76	0.69	14.87	0.3	2.4	68.79
	Terrace	Solid	7	2.1	0.5	2.76	0.68	16.93	0.23	2.4	90.62
	Detached	Cavity	18	0.5	0.33	2.76	0.78	17.45	0.2	2.3	138.2
	Detached	Solid	5	2.1	0.39	2.76	0.81	18.03	0.17	2	128.6
WM	Lowrise	Cavity	5	0.64	0.39	2.76	0	10.02	0.22	2.3	50.79
	Lowrise	Solid	1	2	2.3	2.76	0.95	9.3	0.2	2.3	57.91
	Semi	Cavity	16	0.5	0.28	2.76	0.7	17.42	0.28	2.4	80.01
	Semi	Solid	14	2.1	0.2	2.76	0.73	16.46	0.22	2.4	78.97
	ConvertedFlat	Solid	2	2.1	2.3	3.39	0.41	13.69	0.15	2.65	37.39
	Bungalow	Cavity	3	0.5	0.22	2.76	0.68	13.03	0.2	2.4	55.96
	Terrace	Cavity	9	0.5	0.39	2.76	0.69	16.32	0.23	2.4	70.82
	Terrace	Solid	8	2.1	0.33	2.76	0.75	18.04	0.29	2.45	76
	Detached	Cavity	11	0.5	0.33	2.76	0.73	17.91	0.2	2.4	133.34
	Detached	Solid	2	1.31	0.36	3.39	0.78	15.43	0.13	2.25	119.35
SE	Lowrise	Cavity	33	0.5	2.3	2.76	0	10.74	0.31	2.4	45.86
	Lowrise	Solid	4	2.05	2.3	2.76	0	13.41	0.34	2.45	67.17
	Semi	Cavity	29	1.6	0.33	2.76	0.71	16.52	0.26	2.4	88.97
	Semi	Solid	10	2.1	0.39	2.76	0.74	18.72	0.27	2.65	108.44
	ConvertedFlat	Cavity	2	1.3	2.3	2.76	0.72	9.69	0.25	2.4	51.81
	ConvertedFlat	Solid	6	2.1	2.3	4.03	0.63	10.85	0.19	2.6	60.71
	Bungalow	Cavity	12	0.5	0.39	2.76	0.65	12.54	0.28	2.4	85.3
	Bungalow	Solid	3	1.73	0.46	2.76	0.61	10.87	0.31	2.4	73.88
	Terrace	Cavity	29	0.64	0.39	2.76	0.61	15.53	0.33	2.4	77.19
	Terrace	Solid	22	2.1	0.39	2.76	0.62	17.97	0.29	2.6	91.66
	Detached	Cavity	30	0.5	0.28	2.76	0.75	19.38	0.21	2.4	153.67
	Detached	Solid	3	2.1	0.39	2.76	0.79	30	0.18	2.4	257.06
	Highrise	Cavity	1	1.6	2.3	2.76	0	16.86	0.29	2.4	53.72
	Highrise	Solid	3	1.74	2.3	2.76	0	18.36	0.32	2.4	51.12
EM	Lowrise	Cavity	12	0.5	0.86	2.76	0	9.57	0.21	2.3	46.53
	Lowrise	Solid	1	0.42	0.24	2.76	0	18.17	0.46	2.4	64.56
	Semi	Cavity	33	0.5	0.39	2.76	0.74	16.36	0.25	2.4	85.43
	Semi	Solid	22	2.1	0.39	2.76	0.74	16.99	0.27	2.5	88.11
	Bungalow	Cavity	33	0.5	0.28	2.76	0.65	11.45	0.23	2.4	61.94
	Bungalow	Solid	6	1.64	1.04	2.76	0.78	10.91	0.22	2.35	60.45
	Terrace	Cavity	28	0.5	0.28	2.76	0.68	15.47	0.23	2.3	75.75
	Terrace	Solid	15	2.1	0.33	2.76	0.74	15.73	0.29	2.6	66.96
	Detached	Cavity	30	0.46	0.36	2.76	0.75	17.96	0.2	2.3	134.1
	Detached	Solid	10	2.1	0.39	3.39	0.79	20.28	0.21	2.3	130.13
	Highrise	Solid	1	2.1	2.3	4.03	0	13.69	0.25	2.6	43.93
NE	Lowrise	Cavity	21	0.5	2.3	2.76	0	10.27	0.23	2.4	50.22
	Lowrise	Solid	1	2.1	2.3	2.76	0.46	9.47	0.25	2.5	57.21
	Semi	Cavity	46	0.5	0.33	2.76	0.72	17.02	0.24	2.4	88.83
	Semi	Solid	6	2.05	0.45	2.76	0.72	17.22	0.26	2.45	108.44
	ConvertedFlat	Solid	3	2.1	2.3	4.03	0	15.12	0.25	2.6	53.24
	Bungalow	Cavity	22	0.5	0.22	2.76	0.66	11.53	0.19	2.4	70.91
	Terrace	Cavity	25	0.5	0.28	2.76	0.61	15.97	0.25	2.4	71.32
	Terrace	Solid	15	2.1	0.33	2.76	0.58	18.13	0.29	2.7	75.77
	Detached	Cavity	21	0.42	0.22	2.76	0.76	18.41	0.19	2.4	136.73
	Highrise	Cavity	2	1.05	2.3	2.76	0	16.4	0.29	2.35	54.83

	Highrise	Solid	1	1.73	1.46	2.76	0	17.06	NA	2.3	56.58
NW	Lowrise	Cavity	13	0.5	2.3	2.76	0	10.26	0.27	2.3	60.26
	Lowrise	Solid	1	1.73	2.3	2.76	0.68	8.85	0.12	2.3	43.01
	Semi	Cavity	46	0.5	0.33	2.76	0.72	16.43	0.26	2.45	84.44
	Semi	Solid	7	2.1	0.5	2.76	0.72	17.12	0.32	2.6	99.93
	Converted Flat	Cavity	2	1.3	1.4	2.76	0.35	10.73	0.22	2.6	54.93
	Bungalow	Cavity	12	0.5	0.22	2.76	0.68	11.07	0.19	2.4	74.28
	Terrace	Cavity	21	0.5	0.28	2.76	0.62	16.65	0.37	2.5	80.73
	Terrace	Solid	13	2.1	0.28	2.76	0.59	16.28	0.29	2.7	78.54
	Detached	Cavity	10	0.46	0.25	2.76	0.79	17.49	0.19	2.4	121.47
	Detached	Solid	3	2.1	0.39	2.76	0.84	23.04	0.16	2.5	136.41
	Highrise	Cavity	2	1.6	2.3	2.76	0	15.97	0.2	2.35	48.13

Table A1 – Median building characteristics for the dwelling types within particular regions for the full dataset of 823 dwellings.