# SMITE: Using Smart Meters to Infer the Thermal Efficiency of Residential Homes


Joe Brown
University of Oxford
joe.brown@cs.ox.ac.uk

Jonathan Chambers
University of Geneva
jonathan.chambers@unige.ch

Alessandro Abate
University of Oxford
alessandro.abate@cs.ox.ac.uk

Alex Rogers
University of Oxford
alex.rogers@cs.ox.ac.uk



## ABSTRACT

Residential homes represent approximately 22 % of global energy use and a large proportion of this is due to space heating. The thermal efficiency of a building is typically evaluated manually via surveys, or via intrusive measurements requiring homes to be vacated for prolonged periods, which can result in great inconvenience and expense. More recently, non-intrusive methods have been developed to infer the thermal efficiency of a home which reduces the time and cost of identifying where interventions, such as installing insulation, will have the greatest impact in reducing heating energy usage and carbon emissions. The insight into thermal energy efficiency can also be used as a tool to help support those who are identified as fuel poor. However, none of the current non-intrusive methods take advantage of the half-hourly smart-meter readings that are presently available. This paper proposes a novel algorithm, SMITE, that detects the time periods of the day where the heating of a home is on for an extended length of time and uses this selected data to infer the heating loss coefficient (HLC) and the heating power loss coefficient (HPLC) of the home. The SMITE method is evaluated on 7 homes where the HLC has been inferred by a co-heating test and compared to a state-of-the-art non-intrusive algorithm for inferring HLC, Deconstruct. Our method shows a significant improvement when there is gas heating, with the mean absolute percentage error (MAPE) between the inferred and the co-heating HLC value reducing from 32.6 % for the Deconstruct method to 12.0 % for the SMITE method. This paper also discusses the merits of using the HPLC (instead of the HLC) as an industry standard for evaluating thermal efficiency.



## CCS CONCEPTS

• **Hardware** → **Temperature monitoring**; **Energy metering**; • **Computing methodologies** → *Machine learning approaches.*

## KEYWORDS

Thermal efficiency, smart meters, thermal modelling





*BuildSys '20, November 18–20, 2020, Virtual Event, Japan*





ACM ISBN 978-1-4503-8061-4/20/11…$15.00
https://doi.org/10.1145/3408308.3427609

**ACM Reference Format:**
Joe Brown, Jonathan Chambers, Alessandro Abate, and Alex Rogers. 2020. SMITE: Using Smart Meters to Infer the Thermal Efficiency of Residential Homes. In *The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20), November 18–20, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3408308.3427609


## 1 INTRODUCTION

Residential homes represent approximately 22 % of global energy consumption, contributing 12 % of total greenhouse gas emission [14, 16]. This is particularly true in colder climates like the United Kingdom, where 65 % of energy use associated with residential homes is due to space heating [3]. Hence, reducing the amount of energy required to heat homes has a significant role in reducing carbon emissions and in mitigating climate change [15]. Whilst reducing global emissions, the UK government and energy companies also have a duty to protect vulnerable consumers who are in fuel poverty; this is defined as those who cannot keep their home warm at a reasonable cost, as per the UK Warm Homes and Energy Conservation Act. As such, developing efficient means to analyse the thermal efficiency of homes is important for both governments and energy providers.

The Energy Performance Certificate (EPC), which evaluates the thermal efficiency of a home by conducting a survey of the physical features, has been shown to be inadequate at estimating energy efficiency by independent investigations [18, 33]. To address this, metrics to evaluate the thermal efficiency of a home have been proposed which are based on recording the external temperature and power required to keep the internal temperature of the home at a fixed internal temperature. The heating loss coefficient (HLC) is one such metric, found in co-heating tests to evaluate the thermal efficiency of homes [2, 27, 28]. However, to control for external factors a co-heating test requires the home to be vacated for days or weeks at a time and a constant internal temperature to be maintained by temporary electric heaters. With 25 million homes in the UK alone, any such intrusive method to evaluate the thermal properties of a home will take significant person-hours and incur notable financial costs [31]. Adaptations have been made to the co-heating test to allow the native heating system to be used, instead of installing temporary electric heaters [10], and alternative approaches to infer HLC have been developed, such as the P-STAR and the ISABELE methods [30, 32]. However, whilst these approaches can show improved performance or reduced testing time, they are still intrusive. The evaluation of the thermal efficiency of homes has also been explored using different metrics, such as U-values, which measure the thermal transmittance through a surface. The QUB method is proposed to quickly calculate the U-values of the building elements in-situ [21]. However, it has been shown that the uncertainty in U-value measurements, and uncertainty in how they translate to the actual thermal efficiency of a home, make them an unreliable metric to use [19].

As such, to achieve a large-scale roll-out of HLC evaluation, there is a need for non-intrusive methods to evaluate the thermal efficiency of homes nationwide, and globally. Original non-intrusive approaches rely on annual predictions of heating fuel and weather, however we are now in a scenario where we have more information thanks to the advent and widespread use of smart meters [11]. An alternative metric to HLC, which can be inferred non-intrusively, is the heating power loss coefficient (HPLC). The HPLC compounds the efficiency of the heating system with the HLC factor. The Deconstruct approach finds this metric using daily smart meter and weather data [6]. This is shown to be effective but has several drawbacks. In particular, it is affected by how long the heating is on during the day. By not separating the length of time the heating is on from the efficiency, the inferred HPLC will vary with the length of time the heating is on. To address this, the Deconstruct approach makes the assumption of a linear relationship between external and internal temperature to infer the average internal temperature for the day [20]. However, as the relationship between internal and external temperature is not known for each home, a national average has to be taken, meaning that each house is treated as if it is the average home and useful information is thus lost. An accurate inference of this parameter for each home, or a method removing this assumption from the model, is required.

To address this need, we propose SMITE, a non-intrusive model to predict HPLC requiring only smart meter and weather data (particularly, external temperature and solar irradiance). Using a hidden Markov model, the periods when the heating is on and the home is in a thermal steady state (referred to as heating-on periods) are identified from half-hourly smart-meter readings. This removes the need to infer the internal temperature, one of the biggest sources of uncertainty in existing approaches. If the heating system efficiency is known, the HLC can also be inferred. This new approach allows a closer replication of co-heating tests, and additionally can be performed retroactively and are magnitudes cheaper to conduct [18]. The HLC and HPLC measure the thermal efficiency of the home directly, unlike alternative methods that are based on surveys, or methods that try to infer the in-situ thermal efficiency from another property, such as the U-values of the walls. As the HLC and HPLC measure actual thermal performance, they are the best suited metric to inform policies to reduce carbon emission, and to identify interventions that could lead to the greatest reduction in carbon emissions associated with energy used to heat homes. This paper presents an algorithm that significantly improves the accuracy that the HPLC and HLC can be non-intrusively inferred - enabling large scale recordings of HPLC or HLC.

To benchmark the performance of the SMITE method, the inferred HLC is compared with the Deconstruct approach on a dataset of 7 homes. The dataset contains a year of smart-meter readings for each home and the HLC has been inferred by a co-heating test that is used as the ground truth [4, 23, 29]. The SMITE method shows a significant improvement when there is gas heating, with the mean absolute percentage error (MAPE) between the inferred and the true HLC value reducing from 32.6 % for the Deconstruct method to 12.0 % for the SMITE method. Similarly a MAPE of 38.9 % for the Deconstruct method and 21.3 % is recorded for the SMITE method when the gas and electricity smart-meter readings are combined (to mimic the scenario where a home has electric heating). The results for the SMITE method also show a strong correlation with the measured HLC, with a Pearson's correlation coefficient of 0.91 on gas smart-meter data and 0.70 on gas and electricity smart-meter data indicating this is a viable replacement method for intrusive, expensive evaluations. Furthermore, the real values fall within one standard deviation of the HLC values inferred by the SMITE method for the gas smart-meter data, large misclassifications will lead to engagement with the wrong households for energy-saving interventions, hence it is important that the chosen method has a minimal number of large errors. An experiment is also presented to show the failings of using an approach based on daily data, and how it confounds behaviour of the residents with the thermal efficiency of the home. There is a need for improved solutions to this problem that work in the real world, as shown by the BEIS competition to evaluate the thermal performance of homes using smart-meter readings [8]. To achieve this we are collaborating with a UK energy supplier to industrialise this approach.

The rest of the paper is formatted as follows. Section 2 introduces the modulated heat flow model for a residential home. Section 3 outlines the algorithm implemented to identify the relevant data and infer the HLC and HPLC. The dataset is described in Section 4 and Section 5 presents the results achieved. The impact of this work, potential interventions to reduce carbon emissions and benefits of HPLC over HLC are discussed in Section 6 followed by the conclusions in Section 7.

## 2 MODELS

This section introduces the models used to explain how heat energy flows in and out of a home, and the thermal equilibrium state that needs to be identified to infer HLC. We discuss the practicalities and limitations of thermal modelling in the non-intrusive setting, which leads to the definition of HPLC and how it is inferred. Furthermore, a model is defined to identify the periods when the home is in thermal equilibrium from the smart-meter readings. Comparisons with traditional co-heating tests are made throughout this section to emphasise how it replicates the environment required to conduct a co-heating test, whilst extending it to a non-intrusive situation where occupants are present and there is no control over how the home is being used or heated.

### 2.1 Thermal Model

The HLC [W/K] of a home is a value used to characterise the steady state thermal performance of the building envelope. It is the power required to keep the home at a constant internal temperature, given a lower external temperature. The HLC is inferred through the data obtained from a home and measures how the difference between internal and external temperature, $\Delta T = T_{ex} - T_{in}$ [K], causes the average heat flow into the house i.e. the average heating power used, $Q_{fa}$ [kW], to change [27, 28], given by the equation

$$Q_{fa} = HLC\,\Delta T. \tag{1}$$

This equation concerns a sealed building where the only factors affecting the internal temperature are the external temperature and internal heat sources and the only transfer of heat is through the fabric of the building. In practise, we must also consider ventilation losses, $Q_{ve}$, and also additional sources of heat energy, such as external weather factors providing the home with additional heat energy, $Q_{we}$.

The above terms alone would suffice if the home is unoccupied, however, as the SMITE method is non-intrusive, an occupied home has two additional factors to be considered: (i) heat from the occupants $Q_o$, and (ii) power supplied for tasks other than direct heating which may have a heating effect, $Q_b$ (e.g. cooking, running electrical devices etc). Hence, the total flow of heat energy $Q_h$ into the home is given by

$$Q_h = Q_{fa} + Q_{ve} + Q_{we} + Q_o + Q_b. \tag{2}$$

In this heat flow equation, the term representing the heat flow through the fabric of the building $Q_{fa}$ follows from the initial definition of the co-heating test in Equation (1). The term $Q_{fa}$ is of interest in this work, as it contains the HLC term that we aim to infer. Hence, for the rest of the terms the goal is to either account for them by correctly modelling them, or find a scenario where their contribution is negligible [6]. The first term considered is the ventilation heat flow,

$$Q_{ve} = c_{air}\rho_{air}V_{air}\Delta T, \tag{3}$$

which denotes the transfer of heat energy via the flow of air moving freely between the external and internal environment. In co-heating tests this is accounted for with experiments that track the air movement (e.g., by a tracer gas test or a blowerdoor test) [2]. This heat flow can be modelled as in Equation (3), with $c_{air}$ as the specific heat capacity of air, $\rho_{air}$ as the density of air, and $V_{air}$ as the volume of air that flows between the external and internal environment.

The other factor which effects the internal temperature regardless of if the home is occupied is the external weather factor, $Q_{we}$. For simplicity the only weather contribution considered is the heat flow from solar irradiance,

$$Q_{we} = A_{sol}I_{sol}, \tag{4}$$

defined as the amount of solar irradiance, $I_{sol}$ $[W/m^2]$, incident on the effective aperture of the house $A_{sol}$ $[m^2]$ [13]. Prior studies have shown that other high-level weather factors (i.e. regional wind speed and direction) do not correlate with energy usage, conversely, models with hyper-local data for factors such as wind have demonstrated an impact on internal temperature [7, 13]. With more precise weather data, both regional and with a finer time granularity, more complex ways in which the weather interact with the internal temperature of a home can be modelled, however that is outside the scope of this paper.

There are also the thermal contributions which are caused by occupants. Each person in the home emits heat energy,

$$Q_o = 0.06N_{occ}, \tag{5}$$

where each of the $N_{occ}$ occupants contribute 0.06 kW of heat energy [5]. When the home is occupied, appliances will be used for purposes other than heating and a proportion, $0 \leq \eta_b \leq 1$, of the power used for other appliances, $P_b$, will convert to heat energy,

$$Q_b = \eta_b P_b. \tag{6}$$

## 2.2 Thermal Equilibrium

To infer the HLC, only time periods where the home is in thermal equilibrium are considered. When the home is not in thermal equilibrium there is a difference between the heat energy flowing in and out of the home, resulting in a change in internal temperature. With current approaches the ability to infer the internal temperature from external weather data and smart-meter readings is limited [20]. By restricting the model to only consider times when the home is in thermal equilibrium the flow of heat energy provided by the heating system is equal to the net flow of heat energy out of the home, assuming the thermal mass of the building remains constant. When the heat flows in and out are equal, the heat flow equations for the home is given by:

$$\eta_h P_h = Q_h, \tag{7}$$

which means $P_h$, the power of the heating system to maintain thermal equilibrium, is given by:

$$\begin{aligned} P_h = \frac{1}{\eta_h}(&HLC\,\Delta T + c_{air}\rho_{air}V_{air}\Delta T + A_{sol}I_{sol} \\ &+ 0.06N_{occ} + \eta_b P_b), \end{aligned} \tag{8}$$

where $\eta_h$ is the efficiency of the heating system.

The HLC is only supposed to account for the heat flow through the fabric of the building. In the non-intrusive setting, if the ventilation loss is unaffected by the behaviour of the residents, i.e. it is due to a draft through window sealing rather than due to it being open/closed, then this cannot be separated from the heat flow through the fabric. This type of ventilation heat loss from a property, not due to resident behaviour, is of importance when labeling the thermal efficiency of a home as without an intervention this source of heat loss will remain.

Furthermore, to infer the HLC the heating system efficiency, $\eta_h$, must be known for the home or estimated based on manufacturer guidelines of the typical heating system efficiency. As Equation (8) shows, the HLC and heating system efficiency, $\eta_h$, are compounded which means a non-intrusive method cannot separate the two values when only considering periods where the home has the heating on and is in thermal equilibrium.

## 2.3 Heating Power Loss Coefficient (HPLC)

Since in general the heating system efficiency cannot be separated from HLC, it makes sense to develop non-intrusive approaches to infer the HPLC instead. The HPLC is defined as the rate of supplied power loss to maintain the home at a constant internal temperature, where the supplied power is the average energy consumption to heat the home (i.e. energy input to the heating system).

With the less strict definition of HPLC compared to HLC, the ventilation loss is split into its two sources: behavioural ventilation loss, i.e. when heat is lost through open windows; and home ventilation loss, i.e. when heat is lost through the frame of windows. These two sources are separated by the volume of air lost by each source: $V_{air_b}$ is the volume of air which leaves the house due to behavioural causes, whereas $V_{air_d}$ as the volume of air leaving due to the structure of the home. The ventilation loss related to the home is absorbed into the HPLC term as it cannot be separated from the heat transfer through the fabric of the building. The HPLC is then expressed as follows,

$$HPLC = \frac{1}{\eta_h}(HLC + c_{air}\rho_{air}V_{air_d}), \tag{9}$$

and its relation with the power supplied to the home is given by

$$P_h = HPLC\,\Delta T + \frac{1}{\eta_h}(c_{air}\rho_{air}V_{air_b}\Delta T + A_{sol}I_{sol} + 0.06N_{occ} + \eta_b P_b). \tag{10}$$

It can be assumed that when the building is in thermal equilibrium there is no loss due to the behavioural ventilation.

### 2.4 Detection of Heating Periods

Co-heating tests are conducted with the heating continuously on. To replicate the conditions required for a co-heating test non-intrusively, the heating-on periods of the day need to be identified. The heating-on periods of the day can be recorded by devices such as smart thermostats, however they are yet to be adopted at a large scale. As such, until smart thermostats or similar devices are widely adopted, the heating-on periods of the day need to be inferred via smart-meter readings and local weather data. The SMITE method uses a hidden Markov model to identify these heating-on periods.

As space heating accounts for 65 % of energy consumption in homes, it is assumed that energy consumption at any time $t$ can be split into two modes: (i) when heating is off ($h_{on}^{(t)} = 0$), and (ii) when heating is on ($h_{on}^{(t)} = 1$) [3]. The power used to heat the home at time step t is denoted by $P_h^{(t)}$, and the power used for all other appliances at time step t by $P_b^{(t)}$. The $\xi$ terms represent the noise associated with each scenario and is assumed to follow a Gaussian distribution. Then, $P_{all}^{(t)}$, the total power used by the home at time $t$, can be expressed as a function of the heating mode,

$$\mathbb{P}(P_{all}^{(t)}|h_{on}^{(t)}) \sim \begin{cases} P_b^{(t)} + \xi_1 & \text{if } h_{on}^{(t)} = 0, \\ P_h^{(t)} + P_b^{(t)} + \xi_2 & \text{if } h_{on}^{(t)} = 1. \end{cases} \tag{11}$$

The state of the heating system at any particular time $t$ is unknown, however based on prior knowledge of how a heating system is used and how it contributes to the power output, a model can be constructed to infer when the heating is on or off. As per Equation (11), the likelihood of the power recorded given each state of the heating system can be inferred, $\mathbb{P}(P_{all}^{(t)}|h_{on}^{(t)})$. A temporal component to heating systems makes them more likely to stay in the same state (on or off) for consecutive half-hour periods, rather than fluctuating between the two. This is encompassed by the transition probabilities between the hidden states representing the probability of the heating switching between on or off, as $\mathbb{P}(h_{on}^{(t)}|h_{on}^{(t-1)})$. This scenario can be appropriately modelled with a hidden Markov model and the expression for the joint probability is given as

$$\mathbb{P}(h^{(1:T)}, P_{all}^{(1:T)}) = \prod_t \mathbb{P}(P_{all}^{(t)}|h_{on}^{(t)})\mathbb{P}(h_{on}^{(t)}|h_{on}^{(t-1)}), \tag{12}$$

where $\mathbb{P}(h_{on}^{(t)}|h_{on}^{(t-1)})$ is defined by a transition matrix, and the emission function $\mathbb{P}(P_{all}^{(t)}|h_{on}^{(t)})$ is defined in Equation (11). The joint probability is taken across all time steps, $t \in 1...T$, where T is the number of time steps (T=48 for a day of half-hour periods). By maximising the joint probability in Equation (12) it is possible to find the most likely heating pattern for each day, identifying the heating-on periods. The model is outlined in Figure 1.

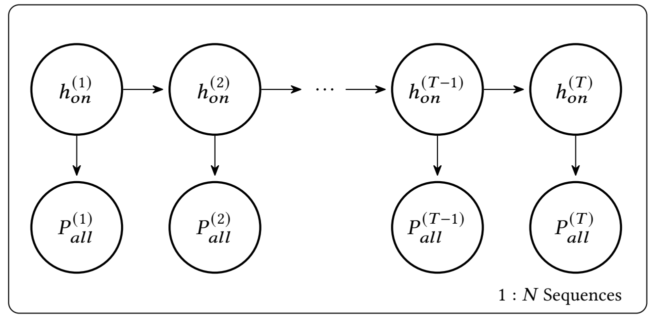


**Figure 1: The hidden Markov model used to identify the periods when heating is on or off when the half-hourly average power is observed. This is performed for all N sequences of consecutive smart-meter readings that have passed through the filtering process. The transition matrix between hidden states and the parameters of the Gaussian distributions representing the emission probabilities are learned via the Baum-Welch algorithm.**

## 3 THE SMITE METHOD

The goal is to construct a post-hoc experiment to infer HPLC (and HLC when heating system efficiency is available) by extracting a sample of data from the smart meter where the home is in thermal equilibrium and the energy used to heat the home is known. To infer the HPLC, a three-step method is proposed: (i) identify the data points that have minimal factors affecting the internal temperature (other than the heating system and external temperature); (ii) identify the heating-on periods; and (iii) infer the HPLC and HLC from the selected data points. The aim is to reduce the model to a situation where the only parameter that needs to be inferred is the HPLC, this process is outlined in Figure 2.

The intrusive data collection process in the co-heating test is replaced by smart-meter readings and regional weather data, and is a process that can be easily scaled across the country for energy companies as it requires no additional hardware installations. A smart meter records energy usage every 30 minutes compared to being recorded every 40 minutes for a co-heating test. The similar frequency of recording of the smart meter means that the granularity of the data will not be a performance bottleneck when compared to the co-heating test.

### 3.1 Identification of Usable Data

Sampling is performed to reduce noise and remove terms that cannot be accurately monitored in a non-intrusive setting. With the current limited ability to model the thermal effects of factors such as solar irradiance these steps are taken to identify and remove periods when these factors have a significant contribution to the heat flow into the home.

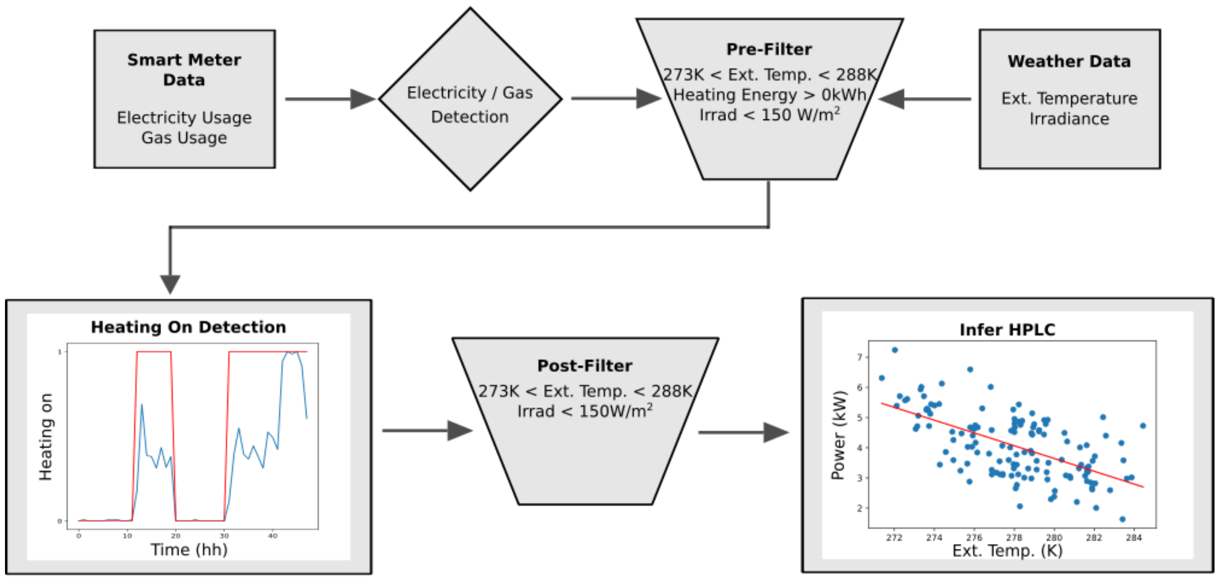


**Figure 2: This flowchart outlines the steps of the SMITE method. The boxes in the upper right and left corners give required input data. A check is performed on the smart-meter readings to infer if the home has gas or electric heating then the pre-filter removes days with high solar irradiance and external temperatures outside the range of accepted values. The cleaned data is used to infer the heating-on periods. The filters are applied again on the heating-on periods to ensure they do not coincide with a moment of high solar irradiance or external temperature. The HPLC is then inferred on the remaining data. The reason the gradient is negative is due to the external temperature, not temperature difference being plotted as shown in Equation (13).**

*3.1.1 Data Cleaning.* The first step is to remove days with corrupt data, missing data and outlier energy readings. These outlier days are identified by calculating a Z-score to identify days where the smart-meter readings are in excess of what is realistically expected and days where 0 kWh of energy are recorded as it is either a faulty reading or provides no information. If there is an hour or more of consecutive readings that are missing, corrupt or outliers, the data for the day is split into two sequences. However, if it is detected that the heating-on period begins or ends immediately before or after, it is discarded as there is no certainty that the full period is accurately recorded. If there is a single missing weather observation with recordings immediately before and after, the missing value is inferred linearly.

*3.1.2 Gas or Electric Heating Detection.* After the data is cleaned whether the source of heating energy is gas or electricity is identified. As 65 % of the energy usage of the average home is for heating, the heating energy source can be identified with 100% accuracy on the dataset used by identifying which source consumes the most energy through the winter months.

*3.1.3 Temperature Filter.* To ensure there is sufficient difference between the external and internal temperature, days with an average external temperature above 15°C (283.15 K) are removed. Further to this, a filter to exclude extremely cold days is added, if the average external temperature is below 0°C (273.15 K) it is also removed. This avoids situations where the heating system is running at maximum power for the duration of the heating-on period and the building does not reach a thermal steady state.

*3.1.4 Solar Irradiance Filter.* All days with an average solar irradiance above 150 W/m$^2$ are filtered out to reduce the heat energy flow into the home from solar irradiance. This value is selected based on empirical results on the dataset and may be altered for different scenarios. Methods have been proposed to estimate the effects of solar irradiance on the temperature of a home, however it introduces modelling uncertainty which is unnecessary if there is enough data after the filtering process to infer the thermal efficiency of a home [12].

*3.1.5 Repeated Filters.* The temperature and solar irradiance filters are applied twice. First against the average for the whole day; and again, on the average of the heating-on periods once they are identified. This filters out the days with large sources of external heat energy when detecting the heating-on periods and afterwards to avoid scenarios where there are spikes in temperature or solar irradiance during a heating-on period.

## 3.2 Detection of Periods when Heating is On

In order to detect the heating-on periods a hidden Markov model is used, as outlined in Figure 1, with emissions affected by an additive Gaussian distribution. The only observations required are the half-hourly values of average power from the smart-meter readings. The hidden binary state represents whether the heating is on or off. Using the filtered data, each complete day (and the segments of a day split around unusable data) is used as a sequence and the Baum-Welch algorithm is applied across all sequences to maximise the joint probability distribution in Equation (12) by finding the most likely parameters for the Gaussian emission distributions and transition matrix [1].

Once the parameters are inferred, the Viterbi algorithm is applied to find the most likely heating state for each time step of each day [1]. All the identified sequences of at least two hours of consecutive heating are extracted as a heating-on period. The first hour of each heating-on period is ignored to allow the house time to reach thermal equilibrium, and the average of the remaining power and external temperature recordings in the heating-on period are used to infer the HPLC.

Notice that as part of the steady state assumption, it is assumed that the internal temperature is kept constant, which is true for the heating-on periods after the first hour is omitted, however the exact value is unknown. Under this assumption, the actual value of the internal temperature does not particularly matter as it is a fixed value which will only provide a constant shift, assuming the home is always heated to the same temperature. The value of interest, the HPLC, is not affected by this as it is inferred through the gradient of the heating power against the external temperature.

## 3.3 Inference of the HPLC

After identifying the heating-on periods for a home, the average power and external temperature for each period is found. The process of finding these periods is outlined in Figure 2. By considering Equation (10), it shows that the rate the average power used to heat the house changes with the external temperature gives the HPLC, if the other terms are accounted for.

It is assumed that there is the same number of occupants in the home for all of the heating periods, this causes a constant shift in the energy usage and as the gradient is the value of interest it can be ignored. Note that across the data used in this study the average power supplied to heat a home is 6.5kW, the contribution of heat energy per occupant is $< 1\%$. Similarly, power for other purposes has been shown to not correlate with the external temperature and can be treated as noise [7]. Furthermore, it is assumed that there is no behavioural ventilation loss during heating-on periods as they are assume to be identified by the outlier filter. Finally, the solar filter ensures there is only a minimal contribution from the solar irradiance and this is also absorbed into a noise term, meaning we can express the relation between the heating power and external temperature as

$$P_h = -HPLC\ T_{ex} + C_p + \xi_h. \tag{13}$$

To infer the HPLC, linear regression is used to find the expected value of the HPLC and the variance is calculated from the external temperature and heating power data points. Note the negative sign in front of the HPLC as the external temperature is used instead of the difference between internal and external temperature and $C_p$ represents constant contribution of other sources of heat energy.

Whilst the HPLC is the metric of interest in this paper, for comparisons against existing approaches the HLC is required. This is calculated from the HPLC by multiplying it by the standard efficiency of a boiler, $\eta_h$, of 0.85, which is the mean boiler efficiency found in a Department of Energy and Climate Change study [24].

# 4 DATASET

To conduct this experiment the ideal dataset consists of homes where the HLC has been inferred via a co-heating test on a vacated house as a ground truth, the heating system efficiency is known to infer the HPLC, half-hourly smart-meter recordings for gas and electricity usage for at least a year whilst the home is occupied, and half-hourly weather data, particularly external temperature and solar irradiance. This presents a major difficulty for obtaining a large scale dataset as each home in the dataset requires an expensive, intrusive and time-consuming co-heating test prior to being occupied and monitored for a year. One alternative is to use simulated data, however modelling the thermal properties of a home is still an open research challenge. As such, inferring the HLC or HPLC on simulated data would only demonstrate the ability to fit to the model used to generate data rather than correctly model the real world scenario [26].

To address this challenge, the dataset selected for our experiments is a subset of homes from the solid wall insulation field trial originally conducted by the UK Energy Saving Trust and then processed and anonymised by another study [4, 23, 25, 29]. This data was collected with the intent to evaluate the effectiveness and experience of customers who decided to install solid wall insulation, however the study recorded the data required for our experiments. For each home the HLC was measured after the installation of the solid wall insulation along with half-hourly smart-meter readings of gas and electricity usage for an extended time period after. Unfortunately, HPLC or the heating system efficiency were not reported hence the HLC must be used to evaluate the performance of the SMITE method. Furthermore, the relevant weather data for each home is available as an anonymised location of each home was provided and the data could be collected from the nearest weather station provided by the Met Office [22].

To ensure that each home in the subset has the data required to infer the HLC it must have enough days recorded after filtering out the corrupted data and be within 15km of the nearest weather station in the Met Office dataset. From the solid wall dataset 14 homes have sufficient data and, when combined with the Met Office dataset, 7 of these homes are within the required proximity of a weather station. The 7 homes fitting the requirements use gas heating. Whilst our experiments show the feasibility of the SMITE method on the available data, there is a pressing need for publicly available large-scale datasets to enable research into evaluating the thermal efficiency of homes.

# 5 EXPERIMENTS

To evaluate the SMITE method it is implemented on the 7 homes selected from the solid wall insulation field trial dataset and compared to the ground truth HLC calculated by a co-heating test [4, 23, 25, 29]. For comparison, the Deconstruct method is also implemented and the two are compared across a number of metrics.

## 5.1 Experimental Setup

The SMITE method is implemented as described in Section 3 and the co-heating inferred HLC values are used as the ground truth.

| | | | Gas meter readings | | | | Gas and electric combined meter readings | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Co-heating | | Deconstruct | | SMITE | | Deconstruct | | SMITE | |
| HLC [W/K] | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| house 1 | 304 | 10.1 | 252 | 28.0 | 264 | 45.4 | 271 | 41.1 | 186 | 43.5 |
| house 2 | 223 | 6.3 | 244 | 14.6 | 248 | 31.2 | 258 | 14.7 | 235 | 34.2 |
| house 3 | 198 | 11.7 | 501 | 33.8 | 177 | 59.7 | 517 | 35.4 | 195 | 64.9 |
| house 4 | 351 | 6.3 | 436 | 40.2 | 387 | 42.6 | 433 | 63.8 | 524 | 47.2 |
| house 5 | 214 | 7.1 | 235 | 30.8 | 211 | 45.1 | 302 | 37.9 | 239 | 71.4 |
| house 6 | 189 | 7.4 | 209 | 37.8 | 148 | 51.2 | 205 | 40.5 | 218 | 70.9 |
| house 7 | 189 | 6.6 | 195 | 31.0 | 219 | 64.2 | 206 | 32.3 | 240 | 90.1 |

**Table 1: The mean and standard deviation of the HLC values inferred from the Deconstruct and SMITE approaches are presented alongside the HLC values inferred for each home from the co-heating tests (ground truth). The standard deviation was not provided for the co-heating tests, instead it has been inferred based on the known uncertainty of co-heating tests [17].**

| | Gas heating | | | | Gas and electric combined meter readings | | | |
|---|---|---|---|---|---|---|---|---|
| | Deconstruct | | SMITE | | Deconstruct | | SMITE | |
| | % Diff | Z-score | % Diff | Z-score | % Diff | Z-score | % Diff | Z-score |
| house 1 | -17.1 | -1.86 | -13.4 | -0.90 | -11.0 | -0.81 | -38.9 | -2.72 |
| house 2 | 9.4 | 1.44 | 11.4 | 0.81 | 15.6 | 2.38 | 5.5 | 0.36 |
| house 3 | 53.0 | 8.97 | -10.5 | -0.35 | 160.8 | 9.00 | -1.7 | -0.05 |
| house 4 | 24.3 | 2.12 | 10.2 | 0.84 | 23.2 | 1.28 | 49.2 | 3.66 |
| house 5 | 10.0 | 0.69 | -1.3 | -0.06 | 41.4 | 2.34 | 11.6 | 0.35 |
| house 6 | 10.9 | 0.55 | -21.7 | -0.80 | 9.0 | 0.42 | 15.5 | 0.41 |
| house 7 | 3.2 | 0.20 | 15.7 | 0.46 | 8.8 | 0.52 | 26.8 | 0.56 |
| **Average** | 27.7 | 1.73 | **-1.3** | **0.00** | 35.4 | 2.16 | **9.7** | **0.37** |
| **Average of absolute values** | 32.6 | 2.26 | **12.0** | **0.60** | 38.9 | 2.39 | **21.3** | **1.15** |

**Table 2: The table shows the percentage difference between the non-intrusive methods and the ground truth, and the Z-score gives how many standard deviations away the ground truth is from the mean of the inferred value for each non-intrusive method. The average of absolute values gives the average magnitude of the percentage differences (MAPE) and Z-scores for each method in each setting. The SMITE approach shows a significant reduction across all measurements. The average without taking the absolute of each value shows that there is less bias in the SMITE method, whilst the Deconstruct method over-predicts.**

The standard deviation values for the co-heating test are not reported and are hence inferred based on the known uncertainty in co-heating HLC measurements [17].

For the Deconstruct method the only parameter changed is the solar irradiance filter from 50 W/m$^2$ to 150 W/m$^2$ as this minimised the MAPE of the Deconstruct method on this dataset. The SMITE method uses the same filters as Deconstruct to find the viable days. Then the heating-on periods are found using the proposed detection algorithm and the filters are re-applied.

As all the 7 homes in the dataset have gas heating, the experiments are set up to show that the approach works when metering of the gas usage is recorded. Another scenario is constructed, adding the gas and electricity smart-meter readings together, as the combined meter readings are what would be expected when a home only has electric utilities and no supply of gas. The purpose of this is to demonstrate that the SMITE method works in a scenario similar to what is expected in a home that has electric heating.

As previously discussed, the HPLC is the HLC multiplied by the heating system efficiency. Hence without any insight into each house an assumption has to be made on the heating system efficiency to infer the HLC from the HPLC.

## 5.2 Evaluation Metrics

To evaluate the performance of the SMITE method a selection of metrics are used to demonstrate how accurately it can infer the HLC of a home, compared to the ground truth. As a benchmark, the Deconstruct method is implemented and a comparison of the results shows the significant performance improvement that the SMITE method provides. A number of metrics are implemented to compare the approaches. The mean absolute percentage error (MAPE) calculates the absolute error as a percentage of the real HLC, measuring how accurate the predictions are and is given by,

$$\mathrm{MAPE} = \frac{1}{N} \sum_{n=1}^{N} \left| \frac{\mathrm{H}_{real}^{(n)} - \mathrm{H}_{pred}^{(n)}}{\mathrm{H}_{real}^{(n)}} \right|,$$

where the HLC value inferred through the co-heating test for house n is $H^{(n)}_{real}$, and $H^{(n)}_{pred}$ refers to the HLC value inferred through the SMITE or Deconstruct method. The Z-score is used to measure how the uncertainty of the prediction captures the ground truth, measuring how many standard deviations away the ground truth HLC is from the predicted HLC. The $\sigma$ term denotes the standard deviation inferred through linear regression for home n with the SMITE or Deconstruct method,

$$\text{Z-score} = \frac{H^{(n)}_{pred} - H^{(n)}_{real}}{\sigma^{(n)}_{H_{pred}}}.$$

Also, the Pearson's correlation coefficient (PCC) is used to measure how the predicted values correlate with the ground truth,

$$\rho\left(\mathbf{H}_{real}, \mathbf{H}_{pred}\right) = \frac{\text{cov}\left(\mathbf{H}_{real}, \mathbf{H}_{pred}\right)}{\sigma_{\mathbf{H}_{real}}\sigma_{\mathbf{H}_{pred}}},$$

where The bold $\mathbf{H}_{real}$ and $\mathbf{H}_{pred}$ represent the vector of all the HLC values inferred via the method in question. whilst the $\sigma$ with a bold subscript represents the standard deviation across the vector of HLC values. The 'cov' function finds the covariance between the two vectors.

## 5.3 Evaluation

The SMITE method shows a significant performance improvement compared to the Deconstruct method when using gas smart-meter data and the combined gas and electricity smart-meter readings to infer the HLC as shown in Table 2. The accuracy across all settings makes the SMITE method the state-of-the-art non-intrusive approach. Additionally, It can be seen in Table 1 that the SMITE method also has a larger variance than the Deconstruct method. This is expected as the SMITE method uses the average of shorter periods of data and whilst there is more variance in the data it removes a modelling bias that is present in the Deconstruct approach.

First we consider the setting where gas smart-meter readings are used. There is a significant gain in accuracy as the MAPE decreases from 32.6% for the Deconstruct method to 12.0% for the SMITE method. This performance increase should be expected as by identifying the heating-on periods the data used to infer the HLC is more similar to the data collected in a co-heating test. Furthermore, the reduction in Z-score when moving from the Deconstruct method to the SMITE method shows that the predicted values and the uncertainty inferred through the SMITE method better capture the true value. For the SMITE method the ground truth HLC is within 1 standard deviation of the inferred HLC for all homes. The consistency of the accurate predictions is valuable in this setting as incorrect predictions may result in costly interventions for the wrong homes. On the point of reducing cost, the SMITE method is shown to have a strong correlation with the intrusive co-heating test, with a PCC of 0.91 compared to a PCC of 0.35 for the Deconstruct method, meaning the SMITE method is a viable replacement for a co-heating test as it achieves similar results whilst bringing the benefits of being non-intrusive.

In the setting with combined energy recording, a slight drop in performance is expected, as there are two sources of energy usage combined making it more difficult to separate out the energy used for heating. There is again a significant improvement in accuracy when moving from the Deconstruct method to the SMITE method with MAPE of 38.9% and 21.3% respectively. There is also a similar decrease in Z-score, from 2.39 to 1.15 and the PCC increases from 0.31 for the Deconstruct method to 0.70 for the SMITE method. This shows that the SMITE method improves on the current approaches to infer HLC non-intrusively in all settings and as algorithms to disaggregate heating energy usage from other uses of energy improve we can expect the performance to be in line with what is achieved in the gas-only setting.

Considering the average, rather than the average of absolute values, highlights how the SMITE method removes a modelling error and source of bias, by using half-hourly rather than daily data. The average percentage difference is 27.7% for the Deconstruct method compared to -1.3% the SMITE method. This is most likely caused by the fact that using daily data rather than half-hourly confounds the duration of heating with the HLC, an issue discussed further in Section 6.3, causing an over-estimate of the HLC. By removing this source of bias with the SMITE approach the predicted HLC values are more accurate and inferences made over large samples of homes are more likely to be accurate compared to the Deconstruct method.

To gain insight into the results we consider the HLC inferred for a few individual houses. There is a large performance drop for Houses 1, 4 and 7 when switching to the combined gas and electricity smart-meter readings. This can partly be attributed to all three houses having high electricity usage of other appliances. By improving heating energy disaggregation, this source of inaccuracy can be reduced. Considering the SMITE method for house 3 where the HLC is inferred from gas data two potential issues arise: (i) the reduction in MAPE when in the setting with the combined gas and electricity usage data indicates there may be a source of electric heating as well as gas heating in this house, which is currently unaccounted for; (ii) The gas data has systematic noise, indicating the presence of another system using gas energy, perhaps with improved disaggregation the performance on this house could be improved. Finally, looking at the Deconstruct approach on house 3, this is where there is the biggest failure caused by not separating the duration of heating from the thermal efficiency of a home. The following section addresses this point and demonstrates how it is correctly handled by SMITE by considering half-hourly smart-meter readings instead of daily.

# 6 DISCUSSION

Following the successful experimental results, three concepts are discussed: (i) The merits of using HPLC instead of HLC and what is the most appropriate metric for evaluating the thermal efficiency of a home given the limitation of a non-intrusive setting; (ii) detecting when heating is used and what further insights into the home this can give to guide energy usage related policy; and (iii) the shortcomings of approaches using daily energy usage and how considering only daily data causes multiple factors to be mixed up.

## 6.1 Non-Intrusive Thermal Efficiency Metrics

In this paper the metric used for comparisons across methods is HLC as this was the only relevant metric labelled in the dataset

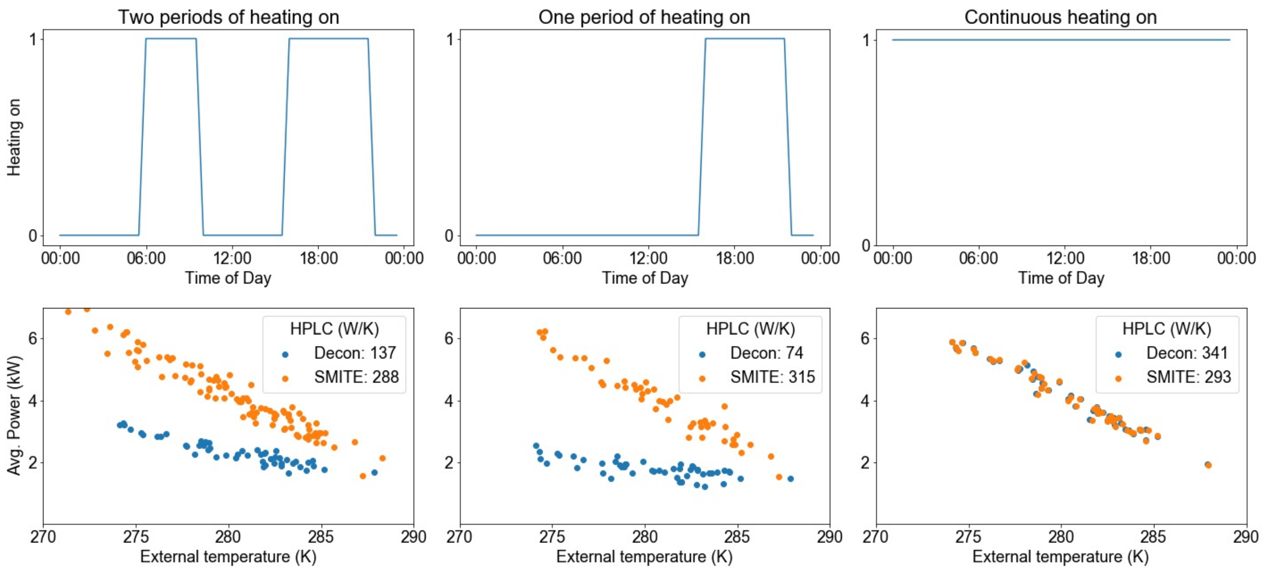


**Figure 3: The Plots demonstrate how the length of time heating is on affects the HPLC inferred with the daily approach. The top row shows their respective heating patterns for a 24 hour period whilst the bottom row shows the average power and external temperature for each period recorded by each method. The inferred HPLC remains consistent for the half-hour approach, However, the daily approach confounds heating usage behaviour with the homes thermal efficiency. Note that the HPLC for the Deconstruct approach is scaled by a factor to attempt to account for the internal temperature, hence the value does not match the gradient.**

used. To infer the HLC it requires an assumption to be made about the efficiency of the heating system. However, a more appropriate metric may be HPLC as it alleviates the requirement of this assumption and measures the compounded value of HLC and heating system efficiency instead, resulting in a more accurate evaluation when performing non-intrusive inference of the thermal efficiency of homes.

In the absence of sensors recording real-time data in the home, such as smart meters or smart thermostats, measuring the HLC of a house with an intrusive thermal efficiency measurement is required - such as a co-heating test. When the HLC is measured, specific procedures can be taken to ensure that only the efficiency of the building material are measured. However, with the advent of smart meters this creates a significant paradigm shift in how the thermal properties of a house can be measured and hence it may make sense to change the metric that is used. The HPLC may be a more appropriate metric to track in this new data-rich scenario. Whilst the HPLC compounds the HLC with heat loss through ventilation and heating system efficiency, it can be inferred through cheap non-intrusive techniques, which means it can be deployed at a scale that is not feasible for HLC.

Finding the HPLC instead of HLC could identify which homes are currently the most inefficient at heating, and where the biggest impact from heating efficiency interventions could be made. In some cases, it may be more important to replace a heating system than to provide insulation. As any intervention would require engagement with the homeowner, the heating system efficiency can be measured, and then the HLC can be inferred with this new information.

## 6.2 Detection of Periods when Heating is On

In the process of finding the HLC, the SMITE method detects periods when the heating is on. When evaluating which homes require intervention, this information may also prove useful as studies have shown that presenting consumers with insights into their energy behaviour can have a significant impact on their energy usage [9, 34]. Homes with a high HPLC and extended heating periods may be identified as those that could financially benefit the most from energy saving interventions, thus making such interventions more financially viable. On the other side, homes with a high HPLC but short heating periods may be flagged as households potentially in fuel poverty, as the short heating periods may be due to financial constraints and the house may not be sufficiently heated. Furthermore, as mentioned in Section 3.1.3, this method can be extended to detect periods when the heating system is at full capacity for extended periods of time, which can suggest that a thermal equilibrium is never reached. This may prove to be useful when identifying households in fuel poverty with an inefficient heating system. Whilst the metrics from this paper alone will not suffice for these decisions, they can be a valuable tool when combined with wider socioeconomic factors.

## 6.3 Issues with Daily Energy Usage

A major issue with the daily approach is that it confounds the length of time the heating is on with the thermal efficiency of the home (HPLC). Generally speaking, the heating is on for longer intervals when the weather is colder. As expected, if the heating is on for longer periods of the day, the average heating energy usage for the day is greater and if the length of time the heating is on is correlated with the external weather, this has an effect on the inferred HPLC value. By not separating the length of time the heating is on and the efficiency, the predicted HPLC will vary with the time heating is on, a behavioural factor, which should not be affecting the HPLC. This is a short-coming addressed by taking the half-hourly smart-meter readings in the SMITE method.

To illustrate the point, a synthetic dataset is created. External temperature values are taken from the Met Office dataset and the energy usage when heating is on is generated by using Equation (13). The plots in Figure 3 demonstrate the point, as the heating schedule changes, the daily approach infers varying HPLC values, whilst the half-hourly approach consistently infers it correctly.

## 7 CONCLUSIONS

In this paper, an improved approach to infer the HLC non-intrusively, using only smart-meter readings and local weather data, is proposed. To evaluate the new SMITE method, a dataset of 7 homes is used. Using the HLC inferred through a co-heating test as the ground truth, the SMITE method achieves a MAPE of 12.0 % on the gas data with the ground truth HLC falling within one standard deviation of the predicted HLC for all homes. This is compared to a MAPE of 32.6 % for an existing alternative approach (Deconstruct). In the setting with combined gas and electricity smart-meter readings the SMITE method achieves a MAPE of 21.3 % whilst the Deconstruct method attains a MAPE of 38.9 %. With the roll-out of smart meters and the ease of access to weather data, the SMITE method can be used at large scale and low cost. We are working with a UK energy supplier to implement the proposed solution, to provide feedback to their customers about the thermal efficiency of their home.

The SMITE framework proposed is flexible, making future improvements to the model very easy. It can be adapted to account for additional information from sensors i.e. occupancy detection, localised solar irradiance measurements using solar panel readings, or internal temperature readings from smart thermostats. Furthermore, as the underlying models improve (cf. thermal equilibrium detection), specific components of the algorithm can be replaced with state-of-the-art ones, thanks to its modular structure.

## REFERENCES


[1] David Barber. 2012. *Bayesian reasoning and machine learning.* Cambridge University Press.

[2] G. Bauwens and S. Roels. 2014. Co-Heating Test: A State-of-the-Art. *Energy and Buildings* 82 (2014), 163–172.

[3] BEIS. 2019. Department for Business, Energy and Industrial Strategy, Energy Consumption in the UK (ECUK). https://www.gov.uk/government/statistics/energy-consumption-in-the-uk. (2019). Online; accessed 21 July 2020.

[4] S Birchall, C Pearson, and R Brown. 2011. Solid Wall Insulation Field Trial-Baseline Performance of the Property Sample. *Bracknell: BSRIA* (2011).

[5] BRE. 2010. Building Research Establishment, SAP 2009: The Government's Standard Assessment Procedure for Energy Rating of Dwellings. https://www.bre.co.uk/filelibrary/SAP/2009/SAP-2009_9-90.pdf. Online; accessed 21 July 2020.

[6] J. Chambers and T. Oreszczyn. 2019. Deconstruct: A Scalable Method of As-Built Heat Power Loss Coefficient Inference For UK Dwellings Using Smart Meter Data. *Energy and Buildings* 183 (2019), 443–453.

[7] Jonathan David Chambers. 2017. *Developing a rapid, scalable method of thermal characterisation for UK dwellings using smart meter data.* Ph.D. Dissertation. UCL (University College London).

[8] J Crawley, E McKenna, V Gori, and T Oreszczyn. 2020. Creating Domestic Building Thermal Performance Ratings Using Smart Meter Data. *Buildings & Cities* (2020).

[9] Sarah Darby. 2006. The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays* (2006).

[10] David Farmer, David Johnston, and Dominic Miles-Shenton. 2016. Obtaining the heat loss coefficient of a dwelling using its heating system (integrated coheating). *Energy and Buildings* 117 (2016), 1–10.

[11] Margaret F Fels et al. 1986. PRISM: an introduction. *Energy and Buildings* 9, 1-2 (1986), 5–18.

[12] F. Hollick, V. Gori, and C. Elwell. 2020. Thermal Performance of Occupied Homes: A Dynamic Grey-Box Method Accounting for Solar Gains. *Energy and Buildings* 208 (2020), 109669.

[13] Z. Huifen, Y. Fuhua, and Z. Qian. 2014. Research on the Impact of Wind Angles on the Residential Building Energy Consumption. *Mathematical Problems in Engineering* (2014).

[14] IEA. 2018. International Energy Agency, World energy statistics and balances. https://www.iea.org/reports/world-energy-balances-2019. Online; accessed 21 July 2020.

[15] IPCC. 2018. Intergovernmental Panel on Climate Change, 2018: Global Warming of 1.5°C. https://www.ipcc.ch/sr15/. Online; accessed 21 July 2020.

[16] Climate Change IPCC. 2014. Mitigation of Climate Change. https://www.ipcc.ch/report/ar5/wg3/. *Contribution of working group III to the fifth assessment report of the intergovernmental panel on climate change* (2014). Online; accessed 21 July 2020.

[17] Richard Jack. 2015. *Building diagnostics: practical measurement of the fabric thermal performance of houses.* Ph.D. Dissertation. Loughborough University.

[18] Scott Kelly, Doug Crawford-Brown, and Michael G Pollitt. 2012. Building performance evaluation and certification in the UK: Is SAP fit for purpose? *Renewable and Sustainable Energy Reviews* 16, 9 (2012), 6861–6878.

[19] Francis GN Li, AZP Smith, Phillip Biddulph, Ian G Hamilton, Robert Lowe, Anna Mavrogianni, Eleni Oikonomou, Rokia Raslan, Samuel Stamp, Andrew Stone, et al. 2015. Solid-wall U-values: heat flux measurements compared with standard assumptions. *Building Research & Information* 43, 2 (2015), 238–252.

[20] Jennifer A Love. 2014. *Understanding the interactions between occupants, heating systems and building fabric in the context of energy efficient building fabric retrofit in social housing.* Ph.D. Dissertation. UCL (University College London).

[21] Johann Meulemans, Florent Alzetto, David Farmer, and Christopher Gorse. 2017. QUB/e: A Novel Transient Experimental Method for in situ Measurements of the Thermal Performance of Building Fabrics. In *Building Information Modelling, Building Performance, Design and Smart Construction.* Springer, 115–127.

[22] Met Office. 2020. Met Office MIDAS Open: UK Land Surface Stations Data (1853-current). Centre for Environmental Data Analysis. http://catalogue.ceda.ac.uk/uuid/dbd451271eb04662beade68da43546e1. Online; accessed 21 July 2020.

[23] T Oreszczyn, J Chambers, V Gori, P Biddulph, I Hamilton, and Cliff Elwell. 2015. How solid is our knowledge of solid walls? Comparing energy savings through three different methods. In *Proceedings of CISBAT 2015 international conference on future buildings and districts-sustainability from nano to urban scale-vol 1*, Vol. 2015. EPFL Solar Energy and Building Physics Laboratory, 107–112.

[24] G Orr, T Lelyveld, S Burton, and I Summerfield. 2009. In-situ monitoring of efficiencies of condensing boilers and use of secondary heating. *Energy Savings Trust Report* (2009).

[25] Daniel Prince. 2014. Solid Wall Insulation Study. *Energy Saving Trust* (2014).

[26] Yoshiaki Sakakura. 2014. Household power consumption simulator with compact representation of occupant behaviors. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm).* IEEE, 170–175.

[27] R. Sonderegger, P. Condon, and M. Modera. 1980. *In-Situ Measurements of Residential Energy Performance Using Electric Co-Heating.* Technical Report. California Univ., Berkeley (USA). Lawrence Berkeley Lab.

[28] RC Sonderegger and MP Modera. 1979. Electric co-Heating: A Method For Evaluating Seasonal Heating Efficiencies and Heat Loss Rates in Dwellings. (1979).

[29] G Stevens and J Bradford. 2013. Do U-value insulation? England's field trial of solid wall insulation. In *ECEEE Summer Study Proceedings*, Vol. 2013. 1269–1280.

[30] K. Subbarao. 1988. *PSTAR: Primary and Secondary Terms Analysis and Renormalization: A Unified Approach to Building Energy Simulations and Short-Term Monitoring.* Technical Report. Solar Energy Research Inst., Golden, CO (USA).

[31] AJ Summerfield, T. Oreszczyn, J. Palmer, I. Hamilton, and R. Lowe. 2015. Comparison of Empirical And Modelled Energy Performance Across Age-Bands of Three-bedroom Dwellings in the UK. *Energy and Buildings* 109 (2015), 328–333.

[32] Simon Thébault and Rémi Bouchié. 2018. Refinement of the ISABELE method regarding uncertainty quantification and thermal dynamics modelling. *Energy and Buildings* 178 (2018), 182–205.

[33] USwitch. 2020. Energy Performance Certificate. https://www.uswitch.com/energy-saving/guides/energy-performance-certificates/. Online; accessed 21 July 2020.

[34] Yan Zhang, Xuemei Bai, Franklin P Mills, and John CV Pezzey. 2018. Rethinking the role of occupant behavior in building energy performance: A review. *Energy and Buildings* 172 (2018), 279–294.