QUILT: QUantify, Infer and Label the Thermal Efficiency of Heating and Cooling Residential Homes

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

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

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

ABSTRACT

The European Parliament have set a target of at least a 32.5 % improvement in energy efficiency by 2030. In the EU residential sector, heating and cooling account for half of the energy usage alone, hence there is a need to make homes more thermally efficient. However, there are currently a lack of tools to evaluate the thermal efficiency of homes at scale as previous methods have been intrusive, costly, or inaccurate. More recent approaches have shown that the thermal efficiency of homes can be inferred with non-intrusive methods using only smart-meter and weather data. However, to deploy such algorithms at scale, steps must be taken to quantify in which scenarios it works, and to account for the uncertainty of the inferred thermal efficiency. To address this, we propose QUILT, a non-intrusive, data-driven framework to quantify the thermal efficiency of a home, grounded in physical models of heat flow, which works with heating and cooling systems. A Bayesian model is used to account for the uncertainty associated with data sources and parameters in the thermal model. QUILT is evaluated on real data to demonstrate its feasibility in real world scenarios, and also on simulated data used by the U.S. Department of Energy to evaluate the performance in scenarios not currently captured in real-world datasets. Our experiments quantify the improvements compared to the current state-of-the-art approach by reducing the mean absolute percentage error from 21.6 % to 15.6 % when inferring the heating power loss coefficient (HPLC), which defines how efficiently a building maintains a heated state given the external temperature. Furthermore, the term cooling power loss coefficient (CPLC) is introduced which is the equivalent to HPLC when the building is in a cooled state, and the feasibility of inferring this metric with QUILT is shown. The analysis also highlights the importance of identifying when the heating is on, and how the HPLC or CPLC can be inferred within a week if the occupants are on holiday given the right weather conditions.

CCS CONCEPTS

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

BuildSys '21, November 17–18, 2021, Coimbra, Portugal

ACM ISBN 978-1-4503-9114-6/21/11…$15.00
https://doi.org/10.1145/3486611.3486653

KEYWORDS

Thermal efficiency, smart meters, thermal modelling

ACM Reference Format:
Joe Brown, Alessandro Abate, and Alex Rogers. 2021. QUILT: QUantify, Infer and Label the Thermal Efficiency of Heating and Cooling Residential Homes. In *The 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '21), November 17–18, 2021, Coimbra, Portugal.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3486611.3486653

1 INTRODUCTION

The European Commission has stated that heating and cooling accounts for half of the residential energy consumption of the EU, and approximately 75 % of this is generated by fossil fuel [9]. The European Parliament has set a target of at least a 32.5 % improvement in energy efficiency by 2030 [8]. Hence, reducing the amount of energy required to heat and cool homes will play a significant role in preventing climate change [18–20]. Furthermore, whilst reducing the global carbon footprint, understanding the energy efficiency of individual buildings can help identify how building inefficiency can amplify energy inequality and address preventable fuel poverty caused by poorly insulated buildings and inefficient heating systems. Hence, developing a scalable method to accurately quantify the thermal efficiency of homes is fundamental to understand where the most impactful energy reducing interventions can be made.

The thermal properties of a home can be directly measured with intrusive tests; for instance, the co-heating test is one of the most accurate methods to infer the heating loss coefficient (HLC), which is the power required to maintain a fixed internal temperature given the external temperature [1, 33, 34]. To improve performance and reduce test duration, the P-STAR and the ISABELE methods have been proposed [36, 37]. However, these tests are expensive and require a home to be vacated for up to a week [7, 10]. Alternative approaches to evaluate thermal efficiency have been proposed, such as the QUB method, which infers the U-values, a measure of the thermal transmittance through a surface, of the building elements [26]. However, U-values have been shown to not translate to the actual thermal efficiency of a home [24]. A recent approach deploys a wireless sensor kit to estimate the thermal efficiency of a home in-situ [13]. Since these tests are typically expensive and intrusive, surveys of the features of a home are conducted to evaluate thermal efficiency, however these are often inaccurate [22, 38].

Due to the increasing availability of data from in-home devices, a number of methods have been developed to infer the thermal efficiency of a home non-intrusively, as initially proposed by PRISM [12]. The Deconstruct method uses daily energy usage and weather data to find the heating power loss coefficient (HPLC), a metric that measures how much power needs to be supplied to the heating system in order to maintain the internal temperature, given the external one [5]. Using similar inputs, a smooth maximum function can identify the transition from weather-dependent to weather-independent heat consumption, and infer the HLC [30]. However, using daily data limits the amount of information that can be learned about the house and requires assumptions, such as for how long the heating is on, which can result in overall inaccuracy. Addressing these limitations, SMITE uses half-hourly or hourly data to identify when the heating is on and thus exploits heating periods to infer HPLC, leading to accuracy improvements [3]. Alternative data-driven approaches use internal temperature monitors to collect 10-minute data through nights and use an auto-regressive model to learn the thermal characteristics of a building, however they are not currently widely deployed [31].

Unfortunately, theoretical measurements of thermal efficiency can fail to reflect real-world performance, and the above methods do not quantify when they fail [23]. To address this, a Bayesian framework can be used to capture the uncertainty from the model and the data sources of the inferred thermal efficiency, as the WattScale approach demonstrates [21]. Bayesian models have also been used for heating disaggregation, namely to identify how much energy is used for heating [6]. However, whilst this captures model uncertainty, there may be scenarios where the model fails to infer thermal efficiency accurately, which need to be identified.

In this work we propose QUILT, a non-intrusive, data-driven framework to quantify the thermal efficiency of a home, grounded in physical models of how heat flows through a building. A Bayesian model of the heat flows in a building is used to infer loss coefficients (such as the HPLC), and to account for the uncertainty associated with data sources and parameters in the thermal model, and known distributions of parameters can be used as priors to inform the model. The model can also be adapted when a cooling system is used, and can infer the cooling power loss coefficient (CPLC), which is a cooling equivalent of HPLC. This allows the approach to be deployed in more thermal setups and weather locations. Furthermore, QUILT is implemented in a number of scenarios, which allows us to evaluate the validity of inferring the loss coefficients in common situations, such as when the home is vacated for a holiday, and identify which features cause significant improvements in accuracy when inferring such coefficients.

We evaluate QUILT on real data from the solid wall insulation field trial conducted by the UK Energy Saving Trust [2, 28, 29, 35]. This is a set of seven real households with recorded smart-meter readings and the HLC values calculated by co-heating tests, which allows us to validate the algorithm in the real world. QUILT is non-intrusive, works with existing hardware with half-hourly data, and shows a reduction of the obtained absolute percentage error by 27 % when compared to the current state-of-the-art for inferring the HPLC in buildings with electric heating, reducing the mean absolute percentage error from 21.3 % to 15.6 %. Also, by using Bayesian inference, priors can be used to inform the model, and uncertainty of data sources and parameter contribute to the improved the accuracy.

Furthermore, simulated data is generated with EnergyPlus using templates from the U.S. Department of Energy's Building Energy Codes Program[1] to create a broader range of scenarios to test QUILT beyond the limited real-world data. In particular, as there are no actual labelled CPLC datasets, the simulated data can be used to validate the feasibility of inferring the CPLC. The experiments across different scenarios highlight a number of key findings, in particular: 1) the importance of only using the data-points when the heating or cooling system is on; 2) holidays provide periods of low-noise data which allow for the HPLC or CPLC to be inferred with a week of data in the right external conditions; and 3) as heating is such a significant source of energy use, energy disaggregation does not have a particularly large effect on accuracy when inferring HPLC or CPLC. By combining improved accuracy with knowledge of when the approach works and fails, whilst being non-intrusive and cheap to deploy, QUILT is a step towards an effective and general method that can be deployed in the real world.

The rest of the paper is formatted as follows. Section 2 introduces the thermal model, and Section 3 extends it to introduce the priors of the parameters in the model, and the modelling choices that can be made. In Section 4 the algorithm used to infer the posterior parameter values is outlined. Section 5 outlines the data that the experiments are performed on and Section 6 outlines the experimental set-up and results. Finally, Section 7 presents actionable insights, before the concluding remarks in Section 8.

2 THERMAL MODELS

The thermal performance of a building can be characterised by the heating loss coefficient, HLC [W/K], a metric representing how much heat energy is transferred through the material of the building given the difference between the internal and external temperatures. However, the actual heat energy in a house is not recorded, so it can be difficult to infer the HLC. Instead, the closest recorded proxy to the heat energy is the energy supplied to the heating system, which is measured by smart meters. Hence, instead of inferring the HLC we focus on inferring the heating power loss coefficient, HPLC [W/K], the HLC scaled by the efficiency of the heating system, and use the HPLC to characterise the steady-state thermal performance.

Furthermore, the cooling power loss coefficient, CPLC [W/K], is introduced, which is the equivalent of the HPLC but for cooling systems. It is the HLC, scaled by the efficiency of the cooling system. To infer the HPLC and CPLC, a thermal model of the home needs to be constructed which accounts for the heat flows of the home.

2.1 Heat Flows

There are a number of sources of heat flow that need to be considered: the flow of heat into the building from heating systems, Q_h, the heat flow through the fabric of the building, Q_{fa}, the heat flow from weather, Q_{we}, the heat flow from occupants, Q_o, and the heat flow from other power sources, Q_b have all been considered in previous models [3, 5]. In this model we also include two more terms to make assumptions in the model explicit. There is a term to account for when the building is not in thermal equilibrium and there is a net heat flow into or out of the building from the previous time-step, $Q_{\Delta t}$, this is typically implicitly assumed in other models. Finally, there is a noise term, σ_Q, to account for unmeasured sources

[1] https://www.energycodes.gov/development/residential/iecc_models

of heat transfer, such as via ventilation due to open windows. The heat flow terms are expressed as follows:

$$Q_h = Q_{fa} + Q_{we} + Q_o + Q_b + Q_{\Delta t} + \sigma_Q. \tag{1}$$

Through the fabric of the building heat flows with respect to the difference between internal and external temperature. The factor by which heat flows in and out given the difference in temperature is the heat loss coefficient (HLC),

$$Q_{fa} = HLC\,(T_{in_t} - T_{out_t}). \tag{2}$$

For the thermal contribution of weather we only consider the contribution from solar irradiance, $I_{sol}[\mathrm{W/m^2}]$ incident on the effective aperture of the house $A_{sol}[\mathrm{m^2}]$, as this is shown to have the biggest impact. Other weather terms have been shown to have an impact although it requires localised and high-frequency data to infer their heat flow contribution, which is not feasible at scale [4, 16]. Hence these contributions are absorbed into the noise term, σ_Q, and the heat flow from the weather factors is represented as,

$$Q_{we} = A_{sol} I_{sol_t}. \tag{3}$$

There is thermal contribution emitted from humans in the building, where each occupant, N_{occ}, typically contributes $\eta_{occ}[\mathrm{W}]$,

$$Q_o = \eta_{occ} N_{occ}. \tag{4}$$

The final term considered is that a proportion, η_b, of power supplied to other appliances, P_b [W], will convert to heat,

$$Q_b = \eta_b P_b. \tag{5}$$

Previous models have failed to explicitly consider that the building may not be in thermal equilibrium and assumed that the internal temperature is fixed due to the heating being on [3, 5]. If the previous time-step has a net flow of heat in or out then this needs to be represented in the heat flow equation. This is represented by a function of the internal temperature at the current and previous time-step, and the thermal mass, M_T, of the building,

$$Q_{\Delta t} = f(T_{in_t}, T_{in_{t-1}}, M_T). \tag{6}$$

To get a state where the thermal properties of the building can be inferred, a state of thermal equilibrium is required. This means a heating or cooling system must be active to ensure the internal temperature is kept at a constant value. If this state is achieved, $Q_{\Delta t} = 0$ and it can be removed from the equation.

2.2 Heating Model

The heat flow into the building from the heating system is given as the power supplied to the heating system, P_h, scaled by the efficiency of the heating system, η_h,

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

Hence, the thermal equilibrium heat flow model is expressed as:

$$\begin{aligned} \eta_h P_{h_t} = {} & HLC\,(T_{in_t} - T_{out_t}) + \eta_{occ} N_{occ_t} + \eta_b P_{b_t} \\ & + A_{sol_t} I_{sol_t} + f(T_{in_t}, T_{in_{t-1}}, M_T) + \sigma_Q. \end{aligned} \tag{8}$$

However, if the efficiency of the heating system is unknown this creates a scenario where each factor might be out by a scaling amount. To address this, we can divide through by the heating efficiency, η_h, and instead learn the ratio of each unknown variable to the heating efficiency. As previously defined, this gives us the HPLC instead,

$$\begin{aligned} P_{h_t} = {} & HPLC\,(T_{in_t} - T_{out_t}) + \hat{\eta}_{occ} N_{occ_t} + \hat{\eta}_b P_{b_t} \\ & + \hat{A}_{sol_t} I_{sol_t} + \hat{f}(T_{in_t}, T_{in_{t-1}}, M_T) + \hat{\sigma}_Q. \end{aligned} \tag{9}$$

2.3 Cooling Model

Similarly, if the external temperature is greater than the internal temperature, the flow of energy to cool the building from the cooling system is given as the power supplied to the cooling system, P_c, scaled by the efficiency of the cooling system, η_c,

$$Q_c = \eta_c P_c. \tag{10}$$

Then to construct the thermal equilibrium heat flow model from Equation (1), Q_h is replaced with Q_c to represent the cooling situation we are now in. Then the equation is similar to Equation (8), except the flow of heat due to the difference in internal and external temperature is inwards rather than outwards, causing the sign to change so we have $(T_{out_t} - T_{in_t})$ instead of $(T_{in_t} - T_{out_t})$ also the signs of the other terms are turned negative as the energy provided by the cooling system is now being used to cool rather than heat,

$$\begin{aligned} \eta_c P_{c_t} = {} & HLC\,(T_{out_t} - T_{in_t}) - \eta_{occ} N_{occ_t} - \eta_b P_{b_t} \\ & - A_{sol_t} I_{sol_t} - f(T_{in_t}, T_{in_{t-1}}, M_T) - \sigma_Q. \end{aligned} \tag{11}$$

Similar to the scenario with heating, the terms are re-scaled by the efficiency of the cooling system, η_c, and the ratio of each variable is instead learned, this gives us a scenario where the CPLC can be inferred,

$$\begin{aligned} P_{c_t} = {} & CPLC\,(T_{out_t} - T_{in_t}) + \tilde{\eta}_{occ} N_{occ_t} - \tilde{\eta}_b P_{b_t} \\ & - \tilde{A}_{sol_t} I_{sol_t} - \tilde{f}(T_{in_t}, T_{in_{t-1}}, M_T) - \tilde{\sigma}_Q. \end{aligned} \tag{12}$$

3 BAYESIAN THERMAL MODEL

The theoretical model of the heat flow is constructed, as in Equation (9) for heating, and cooling in Equation (12). However, in practice some terms are not fully observed and may have to be inferred, which introduces sources of uncertainty and inaccuracy. As such, the model needs to reflect the practicalities of gathering the data required to infer the HPLC or CPLC and the uncertainty and inaccuracy in different scenarios needs to be quantified. It is assumed that all parameters are independently distributed.

3.1 Fixed Internal Temperature

When the heating is on and it is sufficiently cold outside, it is reasonable to assume that there is a significant difference between the internal temperature T_{in_t} and the external temperature T_{out_t}, such that the heat flow through the fabric of the building, Q_{fa}, is large relative to all other terms, except for the heat flow from the heating system. This provides a scenario where it is easier to infer the HPLC. Furthermore, when the heating is on the internal temperature is relatively fixed, and the net heat flow into or out of the building from the previous time-step, $Q_{\Delta T}$, can be absorbed into the noise term, σ_Q. To demonstrate the importance of exclusively considering heating-on periods, Figure 1 shows how taking daily data leads to a gradient (which represents the HPLC) that is different to the ground truth, whereas hourly data is noisy and splits into different usage modes.

Hence, there are two ways in which modelling error can occur for the internal temperature, T_{in}, errors in detecting when the heating is on, and errors in identifying the temperature that the thermostat is set to. If a smart thermostat is installed in the home, the internal temperature can be directly observed, t_{in}, with a small measurement noise. A smart thermostat would also provide a clear indication of when the heating is on. However, if a smart thermostat is not present, the internal temperature must be inferred and a model must be used to infer when the heating is on. A uniform distribution across expected thermostat temperatures can be used as a prior for the internal temperature, and it can be assumed that this value will be the same for all heating periods, thus in summary

$$T_{in} \sim \begin{cases} \mathcal{N}(t_{in}, 0.5) & \text{if } t_{in} \text{ observed,} \\ U(19, 23) & \text{otherwise.} \end{cases} \tag{13}$$

3.2 External Temperature and Weather

With regards to the external temperature, it is assumed that it is known and the only uncertainty required to account for is the discrepancy between the temperature recording at the nearest weather station, t_{out}, and the home,

$$T_{out} \sim \mathcal{N}(t_{out}, 0.5). \tag{14}$$

Similarly, it can be assumed that that solar irradiance is observed with a small amount of measurement uncertainty due to discrepancy between weather station recordings of solar irradiance,i_{sol}, and the solar irradiance incident on the house. A gamma distribution is used to represent this to avoid negative values as follows,

$$I_{sol} \sim \text{gamma}(\mu = i_{sol}, \sigma = 10). \tag{15}$$

The contribution of heat flow from solar irradiance is how much is incident on the aperture of the home. The solar aperture for a building can not be larger than the size of the sun-facing walls, however it also depends on the wall material, size of windows, and the structure of the building, hence the prior has a large amount of uncertainty [11, 25]. A gamma distribution is also used to represent this to avoid negative values,

$$A_{sol} \sim \text{gamma}(\mu = 10, \sigma = 10). \tag{16}$$

3.3 Power

If there is separate metering for heating and cooling power, then they are observed as p_h or p_c, respectively, alongside the power used for other appliances, p_b, and only measurement noise needs to be considered. However, if this is not the case, they need to be disaggregated and inferred as $\hat{p}_h$ or $\hat{p}_c$, and $\hat{p}_b$. The errors in the disaggregation process can propagate through and effect the inferred value of the HPLC,

$$P_h \sim \begin{cases} \mathcal{N}(p_h, \sigma_{p_h}) & \text{if } p_h \text{ observed,} \\ \mathcal{N}(\hat{p}_h, \sigma_{\hat{p}_h}) & \text{otherwise.} \end{cases} \tag{17}$$

$$P_c \sim \begin{cases} \mathcal{N}(p_c, \sigma_{p_c}) & \text{if } p_c \text{ observed,} \\ \mathcal{N}(\hat{p}_c, \sigma_{\hat{p}_c}) & \text{otherwise.} \end{cases} \tag{18}$$

$$P_b \sim \begin{cases} \mathcal{N}(p_b, \sigma_{p_b}) & \text{if } p_b \text{ observed,} \\ \mathcal{N}(\hat{p}_b, \sigma_{\hat{p}_b}) & \text{otherwise.} \end{cases} \tag{19}$$

Where the prior over the variance is given as,

$$\begin{aligned} \sigma_{p_h} &\sim |\mathcal{N}(0, 500)|, \sigma_{\hat{p}_h} \sim |\mathcal{N}(0, 1000)|, \\ \sigma_{p_c} &\sim |\mathcal{N}(0, 500)|, \sigma_{\hat{p}_c} \sim |\mathcal{N}(0, 1000)|, \\ \sigma_{p_b} &\sim |\mathcal{N}(0, 500)|, \sigma_{\hat{p}_b} \sim |\mathcal{N}(0, 1000)|. \end{aligned} \tag{20}$$

3.4 HPLC & CPLC

The HPLC and CPLC are unobserved parameters that must be inferred to identify how thermally efficient the home is. The prior for the HPLC is based on the distribution of HPLC values from a previous study of 541 homes [5]. However, the CPLC is a newly introduced term so we base the prior on the fact that the buildings will have the same properties, however cooling systems are typically more efficient,

$$\begin{gathered} HPLC \sim \mathcal{N}(HPLC_\mu, HPLC_\sigma), \\ HPLC_\mu \sim \mathcal{N}(400, 200),\ HPLC_\sigma \sim |\mathcal{N}(0, 300)|. \end{gathered} \tag{21}$$

$$\begin{gathered} CPLC \sim \mathcal{N}(CPLC_\mu, CPLC_\sigma), \\ CPLC_\mu \sim \mathcal{N}(350, 200),\ CPLC_\sigma \sim |\mathcal{N}(0, 300)|. \end{gathered} \tag{22}$$

Finally, to account for uncertainty in the model there is a general noise term,

$$\sigma_Q \sim \mathcal{N}(0, 300). \tag{23}$$

4 ALGORITHM

To infer the HPLC or CPLC, filters are applied to identify data-points where the thermal contributions from hard-to-model terms are small, and then Bayesian inference is performed on the probabilistic model constructed in Section 3 to find the posterior distributions of the HPLC or CPLC.

4.1 Data Selection

If there is enough data it is possible to remove the data-points where there are high contributions from difficult to model sources of heat flow [3, 5]. If the solar irradiance is large, it can create a source of uncertainty as the linear model used can fail to capture the true dynamics of how it converts to heat energy, hence data-points with solar irradiance over 50 W/m^2 can be removed [15]. Furthermore, if the external temperature is too low the difference between the internal and external temperature can be too small to infer the HPLC, and the heating will typically not be on, hence data-points with external temperatures below 15°C can be removed. Also, if the external temperature is too cold the heating system may run at maximum capacity without maintaining thermal equilibrium, hence external temperatures below 0°C are filtered out.

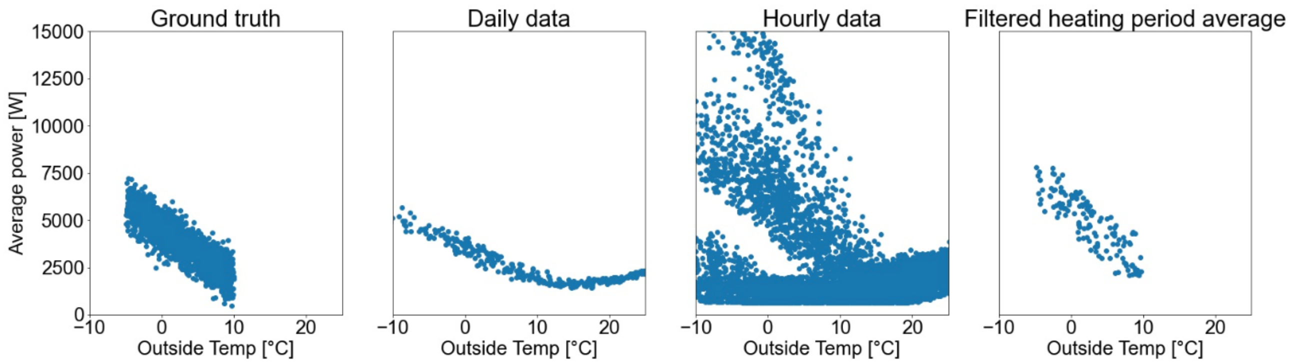

Figure 1: The data used to infer the ground truth is presented alongside the daily average energy usage, hourly energy usage and the average energy usage of each heating period. The daily data has little variance as the energy usage is averaged across the day, however the gradient does not accurately reflect the HPLC as it is influenced by the duration that the heating is on. The hourly data is noisy due to variance of energy usage between hours and different modes of energy usage (heating off, heating on and stable internal temperature, and heating up the building from a colder state). Using the average energy usage for each heating period ensures that the correct data is collected for inferring the HPLC.

Scenario	hardware	Observed	Unobserved
Daily data	smart-meter (daily readings)	$T_{out}, I_{sol}, P_{all}$	HPLC, T_{in}, A_{sol}, η_b
Hourly data	smart-meter	$T_{out}, I_{sol}, P_{all}$	HPLC, T_{in}, A_{sol}, η_b
Week holiday data	smart-meter	$T_{out}, I_{sol}, P_h, P_b$	T_{in} HPLC, A_{sol}, η_b
Filtered hourly data	smart-meter	$T_{out}, I_{sol}, P_{all}$	HPLC, T_{in}, A_{sol}, η_b
Heating-on data	smart-meter, smart-thermostat	$T_{out}, I_{sol}, P_{all}$	T_{in} HPLC, A_{sol}, η_b
averaged filtered heating period data	smart-meter, smart-thermostat	$T_{out}, I_{sol}, P_{all}$	P_h, P_b, T_{in} HPLC, A_{sol}, η_b
averaged filtered heating period data & heating energy known	smart-meter, smart-thermostat, energy disaggregation	$T_{out}, I_{sol}, P_{all}, T_{in}$	P_h, P_b HPLC, A_{sol}, η_b

Table 1: The scenarios of data procurement created with the EnergyPlus simulator, and the corresponding variables which are observed or unobserved in the thermal models presented in Equation (9) and Equation (12).

The reverse arguments can be made with regards to external temperature for the CPLC. If the external temperature is too low the difference between the internal and external temperature will not be large enough, hence data-points with an external temperature below 25°C are filtered out. And if the external temperature is too high, the cooling system may run at maximum capacity without maintaining thermal equilibrium, hence external temperatures above 35°C are filtered out. As high external temperature is correlated with high solar irradiance, this typically relies on the cooling system running at night and occasional days when there is high external temperature and low solar irradiance.

4.2 Inference

Bayesian inference is used to find the posterior distribution of the parameters in Equation (9) for heating, and in Equation (12) for cooling. The prior distributions for the parameters and the data source measurement uncertainty are introduced in Section 3, and they are used to inform the posterior distributions alongside the data observed in each scenario, which is outlined in Table 1. To generate the posterior distribution of the unknown variables, in particular the HPLC and CPLC, a Hamiltonian Monte Carlo (HMC) algorithm is used to propose samples that are distributed according to the target distribution, where the samples are generated using the No-U-Turn Sampler (NUTS) to improve the efficiency of the HMC [14]. The construction of the Bayesian model is outlined in Algorithm 1, and it has been implemented using PyMC3 [32].

5 DATASET

Two datasets are used to evaluate the performance of QUILT in different situations. The first dataset used is a selection of houses from the solid wall insulation field trial conducted by the UK Energy Saving Trust [2, 28, 29, 35]. For each home in the dataset, the HLC was measured and half-hourly electricity and gas meter readings were provided for a year with the house in normal use. Unfortunately, the HPLC can not be inferred as the heating system efficiency is not provided, hence for these experiments the results will be compared on HLC with an assumed thermal efficiency of 0.84. Also, there are no cooling systems installed, so there is no CPLC to infer. The weather data of each house is available from the nearest weather station provided by the Met Office [27].

5.1 Simulated Data

To identify the scenarios in which the HPLC or CPLC can be accurately inferred, the EnergyPlus simulator is used to generate a

Algorithm 1: Inferring HPLC with QUILT. The CPLC can be inferred with minor adjustments to align the algorithm with Equation (12)

input: t_{out}, i_{sol}, $((p_h, p_b) \vee (\hat{p}_h, \hat{p}_o))$, optional: t_{in}
$T_{out} \sim \mathcal{N}(t_{out}, 0.5)$
$I_{sol} \sim \text{gamma}(\mu = i_{sol}, \sigma = 10)$
$HPLC_\mu \sim \mathcal{N}(400, 200)$, $HPLC_\sigma \sim |\mathcal{N}(0, 300)|$
$HPLC \sim \mathcal{N}(HPLC_\mu, HPLC_\sigma)$
$A_{sol} \sim \text{gamma}(\mu = 10, \sigma = 10)$
$\sigma_Q \sim |\mathcal{N}(0, 500)|$
if t_{in} *observed* **then**
  $T_{in} \sim \mathcal{N}(t_{in}, 0.5, \text{Observed} = t_{in})$
else
  $T_{in} \sim \mathcal{U}(19, 23)$
end
if p_h *observed* **then**
  $p_{h_\sigma} \sim |\mathcal{N}(0, 50)|$, $p_{b_\sigma} \sim |\mathcal{N}(0, 50)|$
  $P_h \sim \mathcal{N}(p_h, p_{h_\sigma})$
  $P_b \sim \mathcal{N}(p_b, p_{b_\sigma})$
else
  $p_{h_\sigma} \sim |\mathcal{N}(0, 200)|$, $p_{b_\sigma} \sim |\mathcal{N}(0, 200)|$
  $P_h \sim \mathcal{N}(\hat{p}_h, p_{h_\sigma})$
  $P_b \sim \mathcal{N}(\hat{p}_b, p_{b_\sigma})$
end
$H \leftarrow P_h - (HPLC(T_{in} - T_{out}) + A_{sol}I_{sol} + \eta_b P_b)$
$\mathcal{N}(H, \sigma_Q, \text{Observed} = 0)$
Sample()

number of different energy usage scenarios on a set of buildings. By using the EnergyPlus simulator, it enables the evaluation of the performance of QUILT in scenarios where no real world data exists, and provides exploratory results which may be beneficial in targeting future data collection efforts.

The simulated data is created using the residential prototype building models produced by the Pacific Northwest National Laboratory (PNNL) which is used for the U.S. Department of Energy's Building Energy Codes Program[2]. The simulations are run in a variety of climate settings and energy usage scenarios to evaluate the performance of QUILT and benchmark methods to identify when the algorithm can reliably infer the HPLC or CPLC. By simulating different energy usage scenarios on the same set of buildings it allows for direct comparisons to be made.

The scenarios that are simulated are presented in Table 1, they are selected to represent different set-ups of data collection which are possible with different hardware to identify what data is required to accurately infer HPLC and CPLC. This paper assumes the use of hardware to acquire the heating-on periods and energy disaggregation however research has shown this can be acquired non-intrusively from smart-meter data.

The datasets selected validate the performance of the algorithm on real data and show the potential of it working on a broader range of scenarios with the simulated data. However there is a huge need for publicly available, large-scale datasets of thermal efficiency in homes as better understanding of building thermal efficiency can lead to significant reductions in global energy usage.

[2]https://www.energycodes.gov/development/residential/iecc_models

6 EXPERIMENTS

The experiments demonstrate that QUILT outperforms the current state-of-the-art approaches. Furthermore, the experiments identify the data and models required for the HPLC and CPLC to be accurately inferred. The experiments are implemented on both the real, and simulated data described in Section 5.

6.1 Experimental Setup

For the seven real houses, the HLC values are known and the uncertainty is inferred from the known accuracy of co-heating tests. On the simulated dataset, the HPLC and CPLC are inferred with a scenario that is set up to replicate a co-heating test and the uncertainty in this process is quantified. Then the EnergyPlus simulations are run a number of times to create the scenarios outlined in Table 1 for each building in each climate. In each scenario, unless specified otherwise, the schedules of energy usage within the house and occupancy follow the guidelines of the PNNL prototype residential buildings. For the real houses the results are taken from the SMITE paper as it is performed on the same dataset [3], and for the simulated data both methods are re-implemented.

6.2 Evaluation Metrics

To reflect the broad range of applications that QUILT can be used for, a number of evaluation metrics are used. The mean absolute percentage error (MAPE) is used to calculate the absolute error as a percentage of the true HPLC or CPLC value, which is given by,

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

Whilst the MAPE does not account for uncertainty it is a good benchmark for how accurate the algorithm is when the most likely predicted value is selected, which may be how it is used in practise. To evaluate how the uncertainty in the predictions are capturing the true value, we use the Z-score, to measure how many standard deviation (σ) away the predicted HPLC or CPLC is from the true HPLC or CPLC,

$$\text{Z-score} = \frac{H_{pred} - H_{real}}{\sigma_{H_{pred}}}.$$

6.3 Evaluation

QUILT shows improvement in performance compared to the previous state-of-the-art approach, SMITE, in the real world settings, with an overview of the inferred HLC values by all methods available in Table 2. The evaluation metrics are presented in Table 3, where most significantly the MAPE for the inferred HLC compared to the co-heating test is reduced from 21.3 % to 15.6 % in the setting with the combined gas and electricity smart-meter readings, used to represent a scenario with electric heating. There was also a marginal improvement from 12.0 % for the SMITE approach to 11.5 % by QUILT with only gas-meter readings.

	Gas meter readings								Gas and electric combined meter readings					
	Co-heating		Deconstruct		SMITE		QUILT		Deconstruct		SMITE		QUILT	
HLC [W/K]	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
house 1	304	10.1	252	28.0	264	45.4	**246**	**28.4**	271	41.1	186	43.5	**159**	**29.1**
house 2	223	6.3	244	14.6	248	31.2	**212**	**20.7**	258	14.7	235	34.2	**179**	**24.9**
house 3	198	11.7	501	33.8	177	59.7	**220**	**32.8**	517	35.4	195	64.9	**170**	**42.4**
house 4	351	6.3	436	40.2	387	42.6	**299**	**23.7**	433	63.8	524	47.2	**373**	**32.5**
house 5	214	7.1	235	30.8	211	45.1	**215**	**28.6**	302	37.9	239	71.4	**196**	**42.1**
house 6	189	7.4	209	37.8	148	51.2	**188**	**29.7**	205	40.5	218	70.9	**175**	**41.5**
house 7	189	6.6	195	31.0	219	64.2	**244**	**33.2**	206	32.3	240	90.1	**208**	**54.9**

Table 2: The HLC values inferred via Deconstruct, SMITE and QUILT for the seven real households alongside the ground-truth HLC inferred via a co-heating test.

	Gas meter readings						Gas and electric combined meter readings					
	Deconstruct		SMITE		QUILT		Deconstruct		SMITE		QUILT	
	% Diff	Z-score	% Diff	Z-score	% Diff	Z-score	% Diff	Z-score	% Diff	Z-score	% Diff	Z-score
house 1	-17.1	-1.86	-13.4	-0.90	**-19.3**	**-2.07**	-11.0	-0.81	-38.9	-2.72	**-47.7**	**-4.96**
house 2	9.4	1.44	11.4	0.81	**-5.0**	**-0.54**	15.6	2.38	5.5	0.36	**-18.7**	**-1.68**
house 3	53.0	8.97	-10.5	-0.35	**11.0**	**0.66**	160.8	9.00	-1.7	-0.05	**-11.8**	**-0.56**
house 4	24.3	2.12	10.2	0.84	**-14.8**	**-2.20**	23.2	1.28	49.2	3.66	**6.2**	**0.71**
house 5	10.0	0.69	-1.3	-0.06	**0.1**	**0.05**	41.4	2.34	11.6	0.35	**-4.7**	**-0.22**
house 6	10.9	0.55	-21.7	-0.80	**-0.1**	**-0.04**	9.0	0.42	15.5	0.41	**-9.8**	**-0.41**
house 7	3.2	0.20	15.7	0.46	**29.0**	**1.66**	8.8	0.52	26.8	0.56	**10.3**	**0.37**
Average	27.7	1.73	-1.3	0.00	**0.00**	**-0.35**	35.4	2.16	9.7	0.37	**-10.9**	**-0.96**
Average of absolute values	32.6	2.26	12.0	0.60	**11.5**	**1.03**	38.9	2.39	21.3	1.15	**15.6**	**1.27**

Table 3: The percentage difference and Z-score between the inferred HLC values and the ground truth values. The metrics indicate an improvement in performance by QUILT, particularly in the setting with gas and electricity meter readings combined.

	MAPE (%)	
Data collection & model	HPLC	CPLC
SMITE	18.7	-
Daily data	46.7	-
Hourly data	42.0	-
Filtered hourly data	42.0	-
Heating on data	64.5	-
Averaged filtered heating period data	**15.8**	**18.9**
Averaged filtered heating period data & heating energy known	**15.6**	**17.4**
Week holiday data (all)	24.9	30.3
Week holiday data (>100 data-points)	**13.1**	**16.1**

Table 4: The MAPE of the inferred HPLC and CPLC in different scenarios from the simulated dataset.

A significant change between SMITE and QUILT is the use of a Bayesian model to allow for prior knowledge to inform the inferred HLC value, and to more accurately account for additional terms in the thermal model. This is demonstrated in Figure 2 which shows the HLC inferred by both methods for House 4 from the real dataset. Here the HLC is more accurately inferred by QUILT than SMITE, as the new approach accounts for terms beyond the energy usage and the external temperature mean. Furthermore, the usage of a Bayesian approach means that the prior knowledge of how the HLC values are distributed can inform the inferred HPLC for each house. Figure 2 also shows that there is a marginal correction when solar irradiance is accounted for, as it is worth remembering in the filtering process that data points with >50 W/m^2 of solar irradiance are removed, hence a small correction is expected.

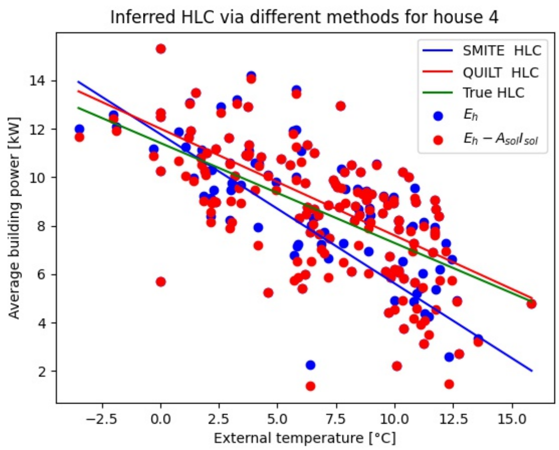

Figure 2: How the HLC inferred by SMITE and QUILT relates to the data. Notice that there is only a marginal shift due to the solar irradiance, mainly due to the larger values of solar irradiance being filtered out, and that the prior information helps inform QUILT to better match the true HLC.

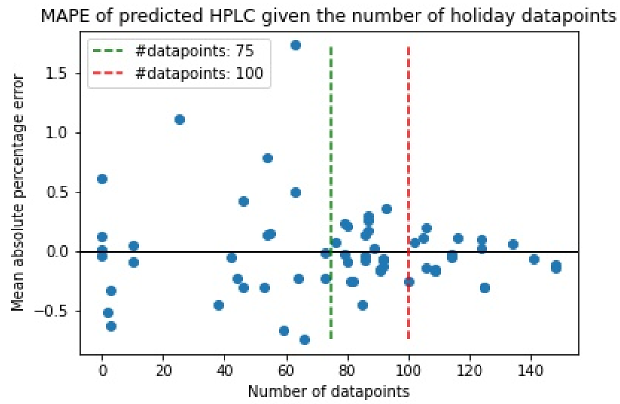

Figure 3: A random week from winter is selected as the holiday period for each house. The average MAPE is 24.9 % across all houses. However, when there are over 100 data-points the MAPE is 13.1 % which is the more accurate than all other scenarios. This shows that inferring the HPLC is viable for periods when the homeowner is on holiday if the external conditions allow for enough valid data-points.

In order to compare the accuracy of the inferred HPLC in different scenarios with different algorithms, the EnergyPlus simulator is used to create 60 building scenarios, and the MAPE of the HPLC and CPLC is calculated under different data collection techniques and with different models. The experiments demonstrate the importance of the different steps taken when inferring HPLC and CPLC, and highlight how it is more accurate than the existing state-of-the-art SMITE in a broad range of simulated settings. Across all houses QUILT shows a reduction in MAPE to 15.6 %, compared to SMITE which achieves a MAPE of 18.7 %. When inferring the CPLC, there was instead a slightly larger MAPE. This was expected, as with a cooling system there are terms in the thermal equilibrium model that are hard to model, e.g. the higher solar irradiance and the presence of open windows. This leads to less data collected, or noisier data-points. Beyond the improved accuracy, the simulated data experiments also allow for a number of scenarios to be tested to identify where QUILT works. The MAPE for the inferred HPLC and CPLC in these different scenarios are presented in Table 4.

6.3.1 Filtering, Clustering and Disaggregation. To identify what data is required and which models need to be used to accurately infer the HPLC, the MAPE in a number of different simulated scenarios are compared. As a starting point, when hourly energy usage data is used with no filters, the MAPE is 70.6 %, which is unsurprising as steps need to be taken to get the correct data. When filters are applied to only take data-points when the external temperature is low enough that the heating could feasibly be on, and the solar irradiance is small such that it doesn't have a significant impact on the thermal energy in the building,the MAPE decreases to 46.4 %, this is however still inaccurate and not useful in practise.

As a next step, only using hourly data when the heating is detected to be on gives a worse MAPE of 68.6 %. This counter-intuitive result is caused by the fact that when the heating is first on, the building is not in thermal equilibrium and excess energy is required to heat up the building before it reaches thermal equilibrium. This can be seen in how the data splits into 3 modes below 15°C in the plot of hourly data in Figure 1. Hence, the reason why the filtered-only data appears to be doing better is because two sources of error are slightly cancelling each other out. This indicates that the heating periods of data need careful treatment.

Hence, the start of the heating period is removed to account for the period where the building is not in thermal equilibrium, and the average is taken for the energy usage, external temperature and solar irradiance for the duration of the rest of the heating period. By using these values to infer the HPLC, the MAPE significantly reduces to 15.8 %. By applying the correct filters and models to select the data, it ensures the data being input into the model aligns with the modelling assumptions that are made (i.e. that the building is in thermal equilibrium). Also, by taking the average across the heating period it reduces the variation in data-points caused by energy usage for other appliances.

With existing disaggregation techniques, the energy usage for heating can be identified and used in the algorithm. However, this did not provide an increase in MAPE when inferring the HPLC, as when the heating is on it typically consumes over 80 % of the energy used by the house and a significant proportion of other energy usage for appliances is relatively constant across time-steps. Hence energy usage disaggregation did not provide a significant increase in accuracy when inferring HPLC on the simulated data and other steps made a much more significant impact. However, it is worth noting that this may be a limitation of the simulated data.

6.3.2 Using Holiday Data. The experiments show that it is feasible to infer the HPLC of a home accurately from holiday data when there are enough data-points collected. Whilst the average MAPE across all situations is 24.9 %, for houses with over 100 data-points the MAPE reduces significantly to 13.1 %. This demonstrates that a viable method of inferring the HPLC is to monitor the energy usage whilst the building is unoccupied for a short duration. However this comes with the caveat that there must be enough data-points that are accepted by the filter stage, i.e. there needs to be at least 100 hours of temperatures below 15°C and with solar irradiance below 50 W/m^2. Figure 3 outlines the relation between the MAPE and the number of data-points. Further real-world experiments may identify that this can be inferred to a sufficient level of accuracy with fewer data-points, and it may be possible with 100 half-hour data-points instead, which would half the amount of time the house needs to be in eligible conditions through the holiday. Furthermore, real world experiments could explore how low the temperature on the thermostat can be set whilst still accurately inferring the HPLC, to minimize the energy used in this HPLC labelling process. These results show that using data when the home is unoccupied can provide accurate inference of the HPLC with approximately a week-worth of data with the right conditions. As such, this can prove to be a useful tool when there is not enough data to infer the HPLC via other methods, or when the inferred HPLC via the other methods is inaccurate.

6.3.3 Cooling and CPLC. QUILT accurately infers the CPLC with a MAPE of 17.4 % compared to 15.6 % when inferring HPLC. Since the thermal model is fundamentally the same when the building has a cooling system or heating system running, it should not be a surprise that the accuracy levels are similar. The slightly worse performance can be attributed to the fact that the cooling system typically uses less power, and there is typically more solar irradiance and windows opened when a cooling system is operating which means there are less viable data-points and more noise in the remaining data-points.

When the occupants are on holiday the CPLC was inferred with a MAPE of 16.1 % when there was enough data. This is a useful insight as it is more likely that the occupants will take a holiday over the summer and there is potential to learn the HLC if the cooling system efficiency is known. Further work could explore scenarios beyond holiday periods where the house is completely empty and identify scenarios where heating and cooling systems are run in situations with low noise and solar irradiance. For example it is common for people to run a cooling system through the night when there is typically no solar irradiance, the number of occupants is fixed and energy usage for other applications is low. This reiterates the need for more labelled real-world data to help identify these scenarios where the HPLC and CPLC can be accurately inferred non-intrusively.

7 ACTIONABLE INSIGHTS

QUILT can be used to inform energy companies, governing bodies and homeowners of the thermal efficiency of buildings which can lead to energy saving interventions. Studies have shown that the energy efficiency of a building can increase its price, providing an economic incentive beyond energy savings, which often do not cover the cost of the energy-saving interventions [17]. Hence, correctly labelling the thermal efficiency of buildings is important to align environmental impact with financial incentives. This is now achievable non-intrusively, accurately, and at low cost thanks to the development of QUILT.

The experimental results highlight how effective QUILT can be at inferring HPLC and CPLC, which works with smart-meter and local weather data. Our experiments have demonstrated the feasibility on real-world data and on a broader set of scenarios using simulated data. However, steps need to be taken to operationalise this research in the real world. There is a need for more real world data on the thermal efficiency of buildings, and such efforts of creating real world datasets are underway[3]. Deploying QUILT on larger real-world datasets is an important next step to demonstrate how robust it is and to identify scenarios where adaptations are needed to work beyond the datasets it has been tested on.

If QUILT is deployed at scale, the operators of the algorithm can combine it with other sources of data to identify the most suitable interventions. Houses with high heating energy usage and low HPLC may benefit from learning to reduce their thermostat temperature, and how more careful heating scheduling can lead to reduced energy usage. Whilst survey data has been shown to not be useful for inferring the thermal efficiency, it does identify what potential interventions can be made to improve insulation, and combined with an accurate HPLC measurement more targeted interventions can be made to the least efficient buildings with the highest HPLC. Furthermore, if QUILT is deployed widely, it will uncover a large number of inefficient heating systems, which can be replaced by modern solutions such as heat pumps.

If the evaluation of QUILT on a broader range of real homes identifies it to be inaccurate in certain scenarios then steps can be taken to improve the data collection process. For example, smart thermostats could remove this potential source of inaccuracy by providing information of when the heating is on, and what the internal temperature is. As discussed in Section 6, identifying the energy usage for heating does not provide any noticeable benefit when identifying HPLC in our simulated scenarios, however real world scenarios may be more complex, meaning this could still be a valuable tool.

Furthermore, if the energy usage data for the house is too noisy, or there is not enough data, a targeted data collection process may be beneficial. The results show that the HPLC can be inferred if the house is unoccupied, and there are at least 100 viable half-hour readings where the solar irradiance and temperature filters are satisfied. This requires the heating system to be on for the duration of the holiday. In real world experiments, different thermostat set-points could be explored to minimise the cost of inferring the HPLC. Also, as inferring the HPLC is a function of the difference between internal and external temperature, if there is not enough variation in external temperature the thermostat set-point can be varied.

8 CONCLUSIONS

In this paper we have proposed QUILT, a Bayesian approach for inferring the HPLC and CPLC of buildings non-intrusively using only smart-meter readings and weather data. The new approach improves on the state of the art, as demonstrated by a number of experiments on real data, and on simulated data to represent

[3]https://www.gov.uk/guidance/smart-meter-enabled-thermal-efficiency-ratings-smeter-innovation-programme

scenarios with unavailable real-world data. Furthermore, the experiments conducted explore the importance of different features when inferring HPLC and CPLC, and identify that correctly clustering heating periods is of great importance, that the HPLC and CPLC can also be inferred when the occupants are on holiday if there is enough valid data points, and that energy disaggregation for heating and cooling does not make a big impact on the accuracy of inferred HPLC and CPLC. These findings emphasise how important a successful smart-meter roll-out is to enable accurate inference of the thermal efficiency of buildings across the country, and provide governments and energy suppliers the tools they need to help improve the thermal efficiency of buildings.

REFERENCES

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

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

[3] J. Brown, J. Chambers, A. Abate, and A. Rogers. 2020. SMITE: Using Smart Meters to Infer the Thermal Efficiency of Residential Homes. In *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation.* 21–30.

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

[5] 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.

[6] F. Culière, L. Leduc, and A. Belikov. 2020. Bayesian model of electrical heating disaggregation. In *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*. 25–29.

[7] C. Deb, LV Gelder, M. Spiekman, G. Pandraud, R. Jack, and R. Fitton. 2021. Measuring the heat transfer coefficient (HTC) in buildings: A stakeholder's survey. *Renewable and Sustainable Energy Reviews* 144 (2021), 111008.

[8] European Commission. 2018. Directive (EU) 2018/2002 of the European Parliament and of the Council of 11 December 2018 amending Directive 2012/27/EU on energy efficiency (Text with EEA relevance.). https://eur-lex.europa.eu/eli/dir/2018/2002/oj. *Official Journal of the European Union* (2018). Online; accessed 09 July 2021.

[9] European Commission. 2019. Commission Delegated Regulation (EU) 2019/826 of 4 March 2019 amending Annexes VIII and IX to Directive 2012/27/EU of the European Parliament and of the Council on the contents of comprehensive assessments of the potential for efficient heating and cooling. https://eur-lex.europa.eu/eli/reg_del/2019/826/oj. *Official Journal of the European Union* (2019). Online; accessed 09 July 2021.

[10] D. Farmer, D. Johnston, and D. 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] A. Fateh, D. Borelli, A. Spoladore, and F. Devia. 2019. A state-space analysis of a single zone building considering solar radiation, internal radiation, and PCM effects. *Applied Sciences* 9, 5 (2019), 832.

[12] M. Fels. 1986. PRISM: an introduction. *Energy and Buildings* 9, 1-2 (1986), 5–18.

[13] M. Frei, C. Deb, Z. Nagy, I. Hischier, and A. Schlueter. 2020. Building Energy Performance Assessment Using an Easily Deployable Sensor Kit: Process, Risks and Lessons Learned. *Frontiers in Built Environment* 6 (2020), 222.

[14] MD Hoffman, A. Gelman, et al. 2014. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15, 1 (2014), 1593–1623.

[15] 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.

[16] 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).

[17] M. Hyland, RC Lyons, and S. Lyons. 2013. The value of domestic building energy efficiency—evidence from Ireland. *Energy economics* 40 (2013), 943–952.

[18] IEA. 2018. International Energy Agency, World energy statistics and balances. https://www.iea.org/reports/world-energy-balances-2019. Online; accessed 09 July 2021.

[19] IPCC. 2018. Intergovernmental Panel on Climate Change, 2018: Global Warming of 1.5°C. https://www.ipcc.ch/sr15/. Online; accessed 09 July 2021.

[20] 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 09 July 2021.

[21] S. Iyengar, S. Lee, D. Irwin, P. Shenoy, and B. Weil. 2021. WattScale: A Data-driven Approach for Energy Efficiency Analytics of Buildings at Scale. *ACM Transactions on Data Science* 2, 1 (2021), 1–25.

[22] S. Kelly, D. Crawford-Brown, and M. 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.

[23] E. Lambie and D. Saelens. 2020. Identification of the building envelope performance of a residential building: A case study. *Energies* 13, 10 (2020), 2469.

[24] F. Li, AZP Smith, P. Biddulph, IG Hamilton, R. Lowe, A. Mavrogianni, E. Oikonomou, R. Raslan, S. Stamp, A. Stone, et al. 2015. Solid-wall U-values: heat flux measurements compared with standard assumptions. *Building Research & Information* 43, 2 (2015), 238–252.

[25] C. Marino, A. Nucara, and M. Pietrafesa. 2017. Thermal comfort in indoor environment: Effect of the solar radiation on the radiant temperature asymmetry. *Solar Energy* 144 (2017), 295–309.

[26] J. Meulemans, F. Alzetto, D. Farmer, and C. 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.

[27] Met Office. 2021. 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 09 July 2021.

[28] T. Oreszczyn, J. Chambers, V. Gori, P. Biddulph, I. Hamilton, and C. 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.

[29] D. Prince. 2014. Solid Wall Insulation Study. *Energy Saving Trust* (2014).

[30] C. Rasmussen, P. Bacher, D. Calì, HA Nielsen, and H. Madsen. 2020. Method for scalable and automatised thermal building performance documentation and screening. *Energies* 13, 15 (2020), 3866.

[31] JP Real, C. Rasmussen, R. Li, K. Leerbeck, OM Jensen, KB Wittchen, and H Madsen. 2021. Characterisation of thermal energy dynamics of residential buildings with scarce data. *Energy and Buildings* 230 (2021), 110530.

[32] J. Salvatier, TV Wiecki, and C. Fonnesbeck. 2016. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* 2 (2016), e55.

[33] 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.

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

[35] 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.

[36] 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).

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

[38] USwitch. 2020. Energy Performance Certificate. https://www.uswitch.com/energy-saving/guides/energy-performance-certificates/. Online; accessed 09 July 2021.