

# Mapping the Pitfalls in the Characterisation of the Heat Loss Coefficient from On-Board Monitoring Data Using ARX Models

Marieline Senave<sup>1,2,3,\*</sup>, Glenn Reynders<sup>1,2,3</sup>, Behzad Sodagar<sup>4</sup>, Stijn Verbeke<sup>2,3,5</sup>, and Dirk Saelens<sup>1,3</sup>

<sup>1</sup>KU Leuven, Department of Civil Engineering, Building Physics Section, Belgium

<sup>2</sup>VITO, Unit Smart Energy and Built Environment, Belgium

<sup>3</sup>EnergyVille, Cities in Transition Section, Belgium

<sup>4</sup>University of Lincoln, School of Architecture and the Built Environment, UK

<sup>5</sup>University of Antwerp, Applied Engineering, EMIB, Belgium

## Abstract

Several studies have demonstrated the capability of data-driven modelling based on on-site measurements to characterise the thermal performance of building envelopes. Currently, such methods include steady-state and dynamic heating experiments and have mainly been applied to scale models and unoccupied test buildings. Nonetheless, it is proposed to upscale these concepts to characterise the thermal performance of in-use buildings based on on-board monitoring (OBM) devices which gather long-term operational data (e.g. room temperatures, gas and electricity consumption...). It remains, however, to be proven whether in-use data could be a cost-effective, practical and reliable alternative for the dedicated tests whose more intrusive measurements require on-site inspections. Furthermore, it is presently unclear what the optimal experimental design of the OBM would be and which data analysis methods would be adequate. This paper presents a first step in bridging this knowledge gap, by using on-board monitoring data to characterise the overall heat loss coefficient (HLC) [W/K] of an occupied, well-insulated single-family house in the UK. With the aid of a detailed building physical framework and specifically selected data subsets a sensitivity analysis is carried out to analyse the impact of the measurement set-up, the duration of the measurement campaign and the applied data analysis method. Although the exact HLC of the building is unknown and no absolute errors could hence be calculated, this paper provides a new understanding of the decisions that have to be made during the process from design of experiment to data analysis. It is demonstrated that such judgements can lead to differences in the mean HLC estimate of up to 89.5 %.

*Keywords: Characterisation, Building Energy Performance, HLC, On-board Monitoring, Data Analysis, Sensitivity, Uncertainty*

---

## 1. Introduction

In response to growing environmental concerns and given the fact that the residential sector accounts for a major share of the final energy consumption in the European Union (25.4 %, [1]), governments have developed increasingly stringent policies with regard to the energy performance of new and renovated dwellings. The European Energy Performance of Buildings Directive (EPBD) recast impels its member states, amongst other things, to set minimum requirements for the energy performance of (1) a newly constructed or thoroughly renovated building as a whole; (2) the building elements that form part of the building envelope and (3) the technical building systems, and to regularly update them [2].

The performance metrics are mainly determined using quasi-steady-state calculations which are (necessarily) based on building plans, theoretical performance indicators of materials and technical systems, as well as default assumptions on occupant behaviour. However, research shows that the

---

\* Corresponding author. E-mail address: [marieline.senave@kuleuven.be](mailto:marieline.senave@kuleuven.be)

results of these calculations, but also of dynamic simulations, may significantly differ from the actual as-built performance, a problem often referred to as the 'performance gap'. Several authors [3–8] reported, for instance, considerable differences between the theoretical and actual energy use in dwellings, which could among other things lead to false assumptions about energy-saving measures. Potential causes of these deviations are (1) invalid simplifications in the prediction models; (2) differences between assumed and actual material properties; (3) workmanship issues; (4) incorrect modelling or malfunctioning of building services and (5) the influence of user behaviour. The first three sources of the gap already happen on the building element level [9,10] and propagate further to the building level, as signalled by Gupta & Dantsiou [11], Johnston et al. [12], Marshall et al. [13] and Wingfield et al. [14,15].

In this context, characterisation methods which adequately combine on-site measurements with statistical modelling are put forward, for instance by the IEA EBC Annex 58 project [16]. These methods could prove useful in supporting a range of applications including certification for new or renovated constructions, the evaluation of the effectiveness of applied energy-saving measures and the identification of the potential for demand side flexibility.

One of the performance indicators that is interesting to determine is the heat loss coefficient (HLC). This steady-state characteristic of the building envelope describes the amount of heating power [W] that is needed to sustain a temperature difference of 1 K between the interior and exterior environment. The most commonly applied on-site measurement to assess the HLC is the coheating test [17]. The data, collected during the (quasi-)steady-state experiment, need to be further analysed to obtain the HLC. Bauwens [18] demonstrates how Auto-Regressive Models with exogenous inputs (ARX models), adopted from the field of system identification, can be used for this purpose. ARX models [19] are transfer function models that are able to describe and predict a building's dynamic behaviour by factoring in a number of past observations from so-called 'input' and 'output' variables. The models base their parametrization on statistical correlations present in the data, disregarding any prior knowledge of the physical processes underpinning the behaviour. Hence, their estimated model coefficients cannot directly be related to actual physical properties, which explains their classification as 'black-box models'.

In the course of the years, several dynamic heating experiments have been suggested as an alternative to the coheating test to shorten the measurement duration. These include the Quick U-value of Buildings (QUB) test [20], the ISABELE method [21,22] and the Short Term Energy Monitoring (STEM) using PSTAR [23]. Nevertheless, these methods, and the current established characterisation methods in general, are still difficult to apply on large scale, since they deploy costly experimental set-ups. Therefore, they are categorised as being intrusive, since they require unoccupied buildings.

Meanwhile, a trend to equip buildings with relatively cheap data collection systems such as wireless temperature sensors and 'smart' energy meters can be observed. In its revision of the EPBD, which forms part of the Clean Energy for All Europeans Package, the European Commission also encourages the collection of data on the measured energy consumption of buildings and the use of smart meters and building automation and control systems [24,25]. This increasing uptake of sensors and Internet of things (IoT) technology in buildings raises the question whether on-board monitoring of the indoor climate and energy use of real in-use buildings could form a worthy alternative for the dedicated tests [26].

The fact that the building's own heating and ventilation system are used, and that the occupants do not have to vacate their dwelling, would make this a less resource-intensive and more practical approach. However, these same occupants pose a great challenge to the characterisation because of the hard-to-trace disturbances they induce, such as varying internal gains and the opening and

closing of doors and windows. An additional limitation of on-board monitoring as compared to dedicated experiments is the limited freedom in excitation signal that can be used due to comfort constraints.

The key to tackling these barriers may already lie in the design of the experiment itself. However, to date, the optimal set-up of such on-board monitoring for building energy performance characterisation remains unclear in terms of variables, resolution and analysis methods. This discussion should moreover be put in a perspective of desired accuracy of the outcome and maximum cost of the procedure. Therefore, this paper investigates the potential of the emerging concept of 'in-use characterisation', and more specifically its applicability to assess the Heat Loss Coefficient of a building envelope. The impact of methodological choices on the characterisation outcome is demonstrated by means of a case study and supported by a building physical framework. Such methodological choices include the decisions made during the design of experiment, the methods used to pre-process some of the data and the selected properties of the ARX models. Since the actual HLC of the studied dwelling is unknown, the research cannot declare the accuracy of the examined methods. Instead, this work aims to advance the understanding of the sensitivity of the characterisation procedure and the pitfalls one could encounter on the path from measurement set-up to characterisation results.

In the next section of the paper, the building and monitoring data set used in the case study analysis will be introduced. The description of the research methodology in section 3 has been divided into two parts. First, the physical phenomena relevant to the HLC are identified through a review of the underlying building physics principles, and mapped against the available measurement data. Secondly, the data subsets that are constructed, and the statistical techniques that are applied to infer the HLC estimates are discussed. Section 4 presents and discusses the characterisation results. In a concluding section 5, the main findings are summarized.

## 2. Case Study (data set)

### 2.1. The Building

The case study is a residential building that is part of a development of four social houses in Gainsborough, UK (53.4° N, 0.77° W). All four houses were certified to the Code for Sustainable Homes Level 5 [27] and monitored from October 2012 (three months after completion of the construction works) until November 2015. The houses and the results of the in-use monitoring campaign are extensively described in [28].

This paper focuses on the south-facing end-terrace house (Figure 1). The house is a two-storey dwelling with a total internal gross floor area of 67.2 m<sup>2</sup> and an internal volume of 192.7 m<sup>3</sup>. Floor plans are shown in Figure 2.



Figure 1: South and east elevations of the house in Gainsborough, UK.

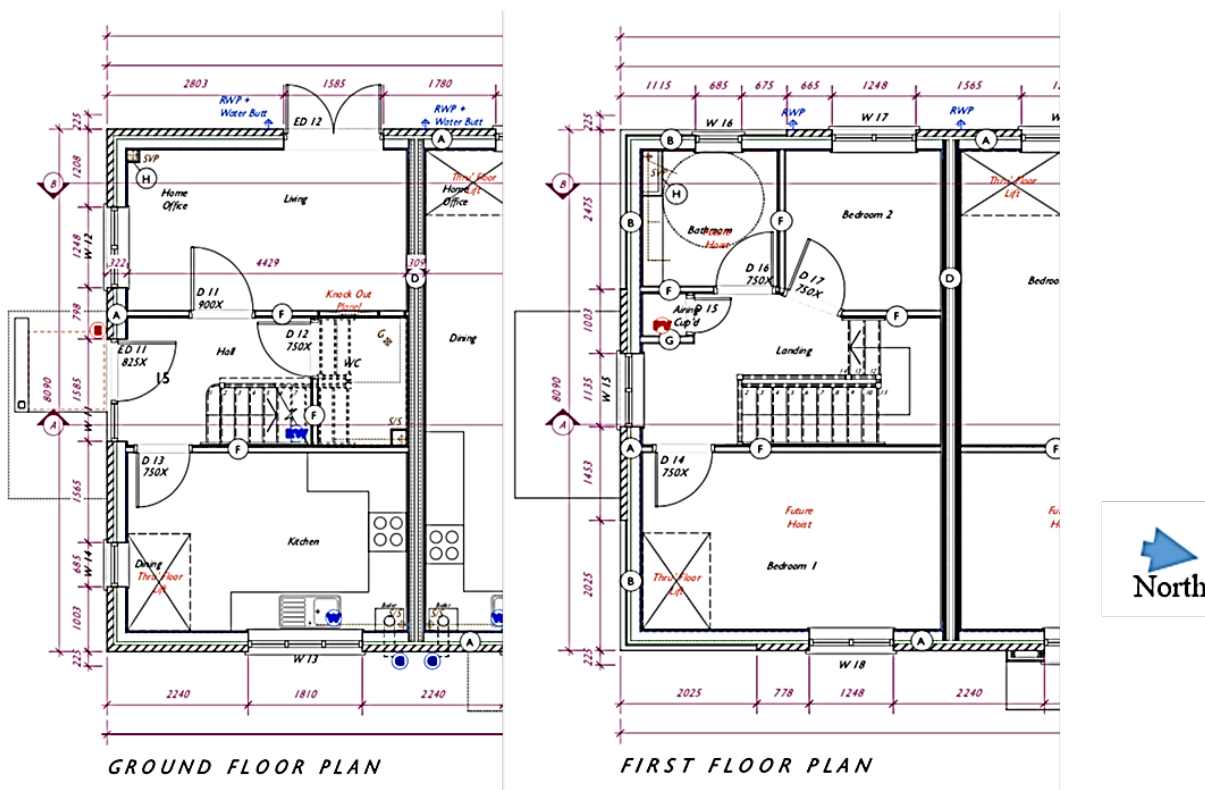


Figure 2: Floor plans of the house (units in mm).

### 2.1.1. Building Envelope

Table 1 summarizes the surface areas and design U-values of the building envelope components. Notably, the ground floor slab is assumed to be in contact with the ambient since it concerns a suspended concrete structure with a ventilated underfloor void.

The design U-value of the external wall was verified on site by means of heat flux measurements performed on the north façade of the opposite end-terrace house of the development, which has an identical construction design. The measurements, that were carried out and analysed in accordance with the ISO 9869-1 standard [29], show a slightly lower U-value (0.12 W/K) compared to the design value of 0.14 W/K.

A thermographic survey, conducted in accordance with ISO 6781 [30], showed that the thermal resistance of the building envelope might furthermore be negatively influenced by the occurrence of thermal bridging at window openings, floor junctions and the interface of different wall finishes.

Table 1: Composition, area and U-value of the building components in contact with the ambient, as detailed in the construction specifications and predicted by SAP calculations [28].

	Composition	Area [m <sup>2</sup> ]	Design U-value [W/m <sup>2</sup> K]
Ground Floor Slab	Proprietary suspended concrete beam and block with 20 mm of insulation	33.62	0.12
External Walls	142 mm Structural Insulated Panels (SIPs) finished in brick or render clad	76.15	0.14
Roof	Single ply roofing membrane fixed to 142 mm Structural Insulated Panels (SIPs) and 50 mm rigid insulation	33.62	0.12
Front Door	Munster EcoClad timber board effect with triple glazed side screen	1.89	1.20
Windows	Munster EcoClad triple glazed windows	15.66 (of which 6.24 south-oriented)	1.15

The house is designed to achieve an air permeability at 50 Pa ( $AP_{50}$ ) of maximum 3.0 m<sup>3</sup>/h.m<sup>2</sup>. Three air leakage tests, conducted following the procedures for whole building pressurization tests detailed in Technical Standard 1 of ATTMA [31], resulted in values for  $AP_{50}$  ranging between 1.9 and 3.7 m<sup>3</sup>/hr.m<sup>2</sup> with an average value of 2.9 m<sup>3</sup>/hr.m<sup>2</sup>.

The limited on-site tests could build a certain degree of confidence in the construction quality and suggest that no large discrepancies (order of magnitude) between the design and real performance of the façade should be expected. Nevertheless, the limited scope of the U-value measurements and the unknown quality of workmanship cast doubt on the actual performance.

### 2.1.2. Technical Building Systems

Space heating (SH) and domestic hot water (DHW) production are provided by a Potterton Promax combination gas boiler with a manufacturer's quoted SEDBUK (Seasonal Efficiency of Domestic Boilers in the UK [32]) efficiency of 91 %. SH is accomplished by hydronic radiators and controlled using two room thermostats located in respectively the hall on the ground floor and master bedroom at the first floor.

Ventilation is provided by a 'Lo-Carbon Astra' mechanical ventilation system with heat recovery (MVHR) from Vent-Axia. According to the manufacturer's specifications, the heat recovery system has a thermal efficiency of up to 91 % for a specific fan power of 0.73 W/l/s.

Part of the electricity demand is generated on site by a 3 kW<sub>peak</sub> PV system, composed of Hengji PV-Tech Mono-crystalline Photovoltaic Panels. The panels are installed on the roof of the dwelling and face south, with an inclination angle of about 35 degrees.

## 2.2. The Monitoring Campaign

Table 2 gives an overview of the data that was collected during the 3-year monitoring campaign. Only those measurements that are used for this study are listed; for the full monitoring set-up the reader is referred to [28].

The interior temperature was monitored in two rooms in both the house ( $\theta_i$ ) and its adjacent dwelling ( $\theta_a$ ). The aggregate gas consumption was monitored with a volumetric gas meter ( $Gas_{mains}$ ). No sub-meters were installed to differentiate between SH and DHW production. Furthermore, the

amount of electricity drawn from the national grid ( $Elec_{mains}$ ) and the PV production ( $Elec_{PV}$ ) were registered. However, a lack of sub-metering made it impossible to distinguish between the percentage of the electrical energy generated by the PV system that was consumed on site and the percentage that was exported to the grid. None of the electrical appliances, except for the MVHR, was fitted with a sub-meter to track individual consumption.

Since some climate variables, e.g. solar radiation, could not be measured on site, the original data set was supplemented with meteorological data recorded at an RAF weather station located in Waddington (53.18° N, 0.55° W).

Throughout the monitoring campaign the property was inhabited by one adult and two children. Occupant behaviour, however, was not monitored.

Table 2: Overview of the variables registered during the earlier performed monitoring campaign and the main properties of the instrumentation used.

Measurement		Instrumentation	
Description	Abbreviation	Resolution	Accuracy
<i>1. Measured on site, 5-min sampling interval</i>			
Exterior temperature	$\theta_{e1}$	0.05°C	+/- 0.5°C
Interior temperature in the living room and bedroom 1 (master bedroom) of the house	$\theta_{i\_liv}, \theta_{i\_bed}$	0.05°C	+/- 0.25°C
Interior temperature in the living room and master bedroom of the adjacent dwelling	$\theta_{a\_liv}, \theta_{a\_bed}$	0.05°C	+/- 0.25°C
Mains gas consumption of the house	$Gas_{mains}$	0.001 m <sup>3</sup>	+/-2%
Mains water consumption of the house	$Water_{mains}$	1 litre	+/-2%
Mains electricity consumption of the house	$Elec_{mains}$	1Wh	+/-2%
Electricity generated by PV system of the house	$Elec_{PV}$	1Wh	+/-2%
The room air supply temperature of the MVHR of the house	$\theta_{supply}$	0.05°C	+/- 0.25°C
The return temperature of the MVHR of the house (the average temperature of the extracted air, before the HR unit)	$\theta_{return}$	0.05°C	+/- 0.25°C
<i>2. Available through 'nearby' weather station<sup>1</sup>, 1-hour sampling interval</i>			
Exterior air temperature	$\theta_{e2}$	0.1°C	< +/- 0.2°C
Global horizontal solar irradiance	GHR	< +/- 1W/m <sup>2</sup>	+/- 3%

### 3. Research Methodology

Incorrect assumptions about the relationship between the actual physical processes related to a building envelope, the statistical model used to describe these processes and the available input data for the model undermine the validity of the HLC estimates. Hence, in the following two sections, a detailed analysis of the 3 components (physical case, statistical model, measurement data) and their interactions will be made. In order to adequately define the 'Heat Loss Coefficient' and its influence factors, section 3.1 will systematically deduce the equations describing a building's dynamic thermal behaviour and capture them in a 'building physical framework'. Next, the available OBM data will be mapped to the theoretical variables of the heat balance. In section 3.2, 'Statistical Framework', the heat balance equation will then be translated into an ARX model on the basis of which the HLC can be assessed.

<sup>1</sup> The climate data registered at the weather station in Waddington were obtained from Met Office (UK).

### 3.1. Building Physical Framework (BPF)

The parameter of interest of this paper, the HLC, is a property of the overall building envelope separating the interior from the exterior environment. Its assessment hence calls for a single-zone approach, although the actual case study concerns a two-storey building with six rooms and unknown internal partitions and air flows. The simplified, dynamic, single-zone heat balance that will be taken as a starting point, Eq.1, states that over time the interior temperature  $\theta_i$  [K] of a zone with effective heat capacity  $C_i$  [J/K] is affected by heat transfer through transmission and ventilation ( $\varphi_{tr}$  and  $\varphi_{vent}$  resp., both expressed in [W]), the net heating power supplied by the heating system ( $\varphi_H$ ) and internal, solar and latent gains ( $\varphi_{int}$ ,  $\varphi_{sol}$  and  $\varphi_{LG}$  resp.).

$$C_i \cdot \frac{d\theta_i}{dt} = \varphi_{tr;t} + \varphi_{vent;t} + \varphi_{H;t} + \varphi_{int;t} + \varphi_{sol;t} + \varphi_{LG;t} \quad (1)$$

The effective heat capacity of the zone  $C_i$  is here simplified as a lumped capacity comprising the thermal mass of the building envelope, internal partitions and furniture, and zone air [33]. In what follows, the interior temperature term and the driving forces behind each of the heat flow rate terms will be further detailed. The equations will be based on ISO 52016-1 [34] and Hens [35]. Next, the identified explanatory variables will be compared to the collected monitoring data. This will enable pointing out characteristics of the case study building and monitoring campaign that might cause difficulties or even errors in the characterisation process. Often, the theoretically required input data cannot be measured directly and other measured variables will have to be put forward as an alternative, or assumptions will have to be made. The process is summarized in (the left half of) Table 3 at the beginning of the next subsection. The first column of this table lists the physically relevant variables that are identified in the heat balance equations. The second column sums up the monitoring data that could be used to represent them.

#### 3.1.1. Interior temperature

The simplified heat balance equation, Eq. 1, represents the case study building with its six rooms as a single thermal zone. The positioning of doors, windows and ventilation and heating elements, the occupant behaviour and thermal stack effect make one uniform interior temperature throughout the house, as represented by  $\theta_i$  in Eq.1, highly unlikely. In an attempt to capture this non-uniformity, a sample of the interior temperature was taken at two points across the dwelling:  $\theta_{i,liv}$  and  $\theta_{i,bed}$  (see Table 2). The question arises whether and how the two registered temperatures should be related to obtain one single signal for the dwelling as a whole. Furthermore, it should be noted that the physical representativeness of the sensor data itself is uncertain, since the sensors do not register the *equivalent* temperature as required by the transmission term in the heat balance equation (see further, Eq.2) or the pure *air* temperature as required by the ventilation term (see further, Eq.5), but an unspecified combination of the *air* and *radiative* temperature of their surroundings. The second column of Table 3 presents four different approaches in which the collected interior temperature data can be used as input for the variable  $\theta_i$  in the heat balance:  $\theta_{i,liv}$  or  $\theta_{i,bed}$  can be used individually, or an arithmetic mean or (internal) volume weighted average of both temperatures can be considered. In the latter case, both temperatures are assumed to represent the interior temperature of their entire storey. Another option would be to consider a heat loss area weighted average temperature. However, the limited data available on the building geometry does not allow the calculation of this time series.

#### 3.1.2. Transmission heat flow rate

Theoretically, the transmission heat flow rate of a zone,  $\varphi_{tr}$ , is given by Eq.2 and depends on equivalent indoor and outdoor temperatures ( $\theta_{eq}$ ). Equivalent temperatures differ from air temperatures ( $\theta_{air}$ ) in encompassing longwave and shortwave radiation exchange and thus being

building element specific [35]. To obtain one interior-exterior temperature difference term as in Eq.2, the disparate temperature terms of the individual building elements hence need to be averaged (subscript 'avg') as also discussed in [36].

The overall heat transfer coefficient by transmission  $H_{tr}$  [W/ K] in Eq.2 represents the sum of the direct transmission heat transfer coefficient between the conditioned zone and the exterior ( $H_{tr,e}$ ) and the transmission heat transfer coefficients through the ground ( $H_{tr,g}$ ), unconditioned spaces ( $H_{tr,u}$ ) and adjacent buildings ( $H_{tr,a}$ ) (Eq.3).  $H_{tr,e}$  can in turn be split into the heat transfer through the elements  $j$ , linear thermal bridges  $k$  and point thermal bridges  $m$  of the building envelope in contact with the ambient (Eq.4).  $A_j$  and  $U_j$  hereby respectively represent the building elements' surface area [m<sup>2</sup>] and thermal transmittance [W/m<sup>2</sup>K],  $\psi_k$  is the linear and  $\chi_m$  the point thermal transmittance of the linear [W/mK] and point thermal bridges [W/m<sup>2</sup>K], respectively. The factor  $b_{tr}$  ensures that the appropriate temperature difference (compared to the interior-exterior temperature difference) is taken into account. Hence  $b_{tr,e}$  equals 1.

$H_{tr,a}$ ,  $H_{tr,g}$  and  $H_{tr,u}$  follow analogous equations but are adjusted by a factor  $b_{tr}$  different from 1.

$$\varphi_{tr;t} = H_{tr;t} (\theta_{eq;i,avg;t} - \theta_{eq,e,avg;t}) \quad (2)$$

$$H_{tr;t} = H_{tr,e;t} + H_{tr,g;t} + H_{tr,u;t} + H_{tr,a;t} \quad (3)$$

$$H_{tr,e;t} = \left( \sum_j (A_j \cdot U_{j;t}) + \sum_k (l_k \cdot \psi_{k;t}) + \sum_m \chi_{m;t} \right) \cdot b_{tr,e;t} \quad (4)$$

Since the presented case study concerns a semi-detached dwelling with a suspended ground floor construction, only the transmission heat losses to the ambient and the adjacent dwelling are retained. According to the above-developed formulas, these depend on  $\theta_{eq,e}$ ,  $\theta_{eq,i}$  and the equivalent temperature of the adjacent dwelling  $\theta_{eq,a}$ . Comparison of these variables with those listed in Table 2 shows that a one-to-one matching is impossible: assumptions for plausible substitutes for  $\theta_{eq,a}$  and  $\theta_{eq,e}$  are discussed below.

Just like the case study dwelling, the adjacent dwelling was fitted with two sensors, registering a temperature  $\theta_{a,liv}$  and  $\theta_{a,bed}$  (Table 2). Hence, similar sensor choices and approaches for data aggregation could be evaluated for  $\theta_a$ . In this work however, only the variant with the arithmetic mean of the two available signals will be illustrated (Table 3).

Two time series describe the exterior temperature observed during the monitoring campaign: the first one, referred to as  $\theta_{e1}$ , was collected on the site of the dwelling; the second one,  $\theta_{e2}$ , at a national weather station about 30 km from the site (see Table 2). On the one hand, the closer proximity to the research object and the lower sampling interval are advantages of the sensor registering  $\theta_{e1}$ , on the other hand, the  $\theta_{e2}$  data series have been acquired with a greater accuracy regarding the measurement equipment and potentially also its positioning. The consequences of a choice for one of both will be analysed further on. The reservations concerning the *equivalent* temperature expressed in the paragraph about  $\theta_i$  apply here as well.

### 3.1.3. Ventilation heat flow rate

The ventilation heat flow rate term  $\varphi_{vent}$  from Eq.1 can be further elaborated as described in Eq.5, with  $H_{ve}$  the overall heat transfer coefficient by ventilation and  $\theta_{air,i}$  and  $\theta_{air,e}$  the zone and exterior air temperature respectively.

As can be seen from Eq.6,  $H_{ve}$  [W/K] takes the sum over the air flow rates  $G_v$  [m<sup>3</sup>/s] of all relevant air flow elements, such as air infiltration and mechanical or intended natural ventilation. The unitless factor  $b_{ve}$  adjusts the temperature difference whenever the supply temperature of the air flow element does not equal the temperature of the external environment, e.g. due to the presence of a heat recovery unit.  $\rho_a$  is the density of air [kg/m<sup>3</sup>] and  $c_a$  its specific heat capacity [J/(kgK)].

$H_{ve}$  can thus also be written as the sum of  $H_{ve,wtt}$ , which comprises the heat transfer through the air



flow elements for which the factor  $b_{ve}$  equals 1, implying that the external air flow enters the zone without prior thermal treatment (subscript 'wtt'), and  $H_{ve,tt}$  which takes the sum over the remaining air flow rates undergoing thermal treatment (subscript 'tt') for example in a heat recovery unit.

$$\varphi_{vent;t} = H_{ve;t} \cdot (\theta_{air,i;t} - \theta_{air,e;t}) \quad (5)$$

$$H_{ve;t} = \rho_a \cdot c_a \cdot \sum_{p=1}^r (b_{ve;p;t} \cdot G_{V;p;t}) = H_{ve,wtt;t} + H_{ve,tt;t} \quad (6)$$

As for the *equivalent* temperatures, no monitoring data of the pure *air* temperatures  $\theta_{air,i}$  and  $\theta_{air,e}$  are available (see §3.1.1). The infiltration losses will therefore be accounted for in the same way as the transmission losses to the external environment; using the candidate substitutes for  $\theta_i$  and  $\theta_e$  listed in Table 3. The disregard of the distinction between the air and equivalent temperatures will be further discussed in §3.1.8.

The actual air change rate of the Mechanical Ventilation Heat Recovery (MVHR) system is unknown, but sensors were installed inside the unit to monitor both the supply and return temperature (Table 2). However, due to technical problems, no observations are available for  $\theta_{return}$  during the measurement period selected to train the model (see further in §3.2.1). At times when the return temperature is indeed adequately tracked, the difference between  $\theta_{supply}$  and  $\theta_{return}$  is small (RMSE=0.6 °C), which suggests that the actual energy efficiency of the heat recovery unit is rather high.

Assuming that the heat recovery efficiency of the MVHR, which equals  $(1 - b_{ve})$  (Eq.6), is constant, the temperature difference driving the mechanical ventilation heat losses is a multiple of the one driving the air infiltration losses. Therefore, the former losses will be modelled using the same  $\theta_i$  and  $\theta_e$  signals as the latter.

#### 3.1.4. Net heating power

The net heating power supplied by the heating system  $\varphi_H$  (Eq.1) is linked to the energy use of this system  $\varphi_{H,sys}$  [W] by the overall system efficiency  $\eta_{H,sys}$ , which accounts for unrecoverable generation, storage, distribution and emission losses:

$$\varphi_{H;t} = \eta_{H,sys;t} \cdot \varphi_{H,sys;t} \quad (7)$$

During the monitoring campaign, the mains gas consumption [m<sup>3</sup>] was tracked. The conversion factor between this variable and  $\varphi_{H,sys}$  is the volumetric calorific value (CV) of the gas (1). As postulated in Eq.7,  $\varphi_{H,sys}$  should be multiplied with the system efficiency  $\eta_{H,sys}$  (2) to obtain the sought-after net heating power.

The volumetric CV of natural gas can be calculated as the product of (a) the volume correction factor, a national fixed factor equal to 1.02264, and (b) the gross calorific value. Daily averaged, charging area specific values of the latter may be taken from the website of National Grid UK [37]. For Gainsborough, the mean daily averaged gross calorific value for the 2014-2015 heating season (October 1 till March 31), for instance, equals 39.3 MJ/m<sup>3</sup>. Lander [38] demonstrates how employing these conversion factors ensures that the bias on the volume and energy conversion is limited to less than 1%, notwithstanding the fact that the precise composition of the delivered gas, and its pressure and temperature over the meter may vary over time.

The overall efficiency  $\eta_{sys}$  of the heating system installed in the case study dwelling is unknown, except for one of its components: the production efficiency of the combi boiler, which reaches 91% according to the manufacturer's specifications. However, this SEDBUK rating is a measure of the annual average energy performance achieved in typical domestic conditions. Full and 30 % part load efficiency are averaged out and in addition, no differentiation is made between the production efficiency of respectively Space Heating (SH) and Domestic Hot Water (DHW) [32].

Furthermore, the monitoring only covers the total fuel consumption of the combi boiler and does not differentiate the consumption for SH from the consumption for DHW production (gas is not used

for cooking). Although some of the gas consumption for DHW production can be expected to result in internal gains (see further) a significant fraction of the DHW leaves the house without having contributed to a raise of  $\theta_i$ . Hence, the energy use breakdown into these two end-uses has to be estimated. In Senave et al. [39] different options to accomplish this were compared using the same monitoring data set as the present study. The approach resulting in the gas consumption for SH profile with which the best scores were obtained on some simulation cross-validation tests (following the same principles and on the same data periods as used here) will be adopted for this investigation. It is a deterministic method that requires monitoring data of the mains water consumption and draws its disentanglement on the following assumptions: (1) the gas boiler cannot operate in SH and DHW mode at the same time; and (2) hot water is produced instantaneously (hence, no storage tank). For a more detailed explanation, the reader is referred to [39]. For now, it is important to notice that the split-up of total gas consumption in SH and DHW fractions is a first approximation. The validity of the approach and its applicability for this specific case study could not be verified.

In conclusion, based on the available measurement data,  $\varphi_H$  can be approximated by  $Gas_{mains}$ , a variant of  $Gas_{mains}$  altered by the known boiler efficiency, and an estimate of the gas consumption for SH (Table 3).

### 3.1.5. Internal heat gains

$\varphi_{int}$  in Eq.1 includes the heat flow rate from occupants  $\varphi_{int,Oc}$ , appliances  $\varphi_{int,Ap}$  and lighting  $\varphi_{int,Li}$ , the heat dissipated from or absorbed by hot and mains water and sewage  $\varphi_{int,Wa}$ , and the recoverable losses from or to heating, cooling and ventilation systems  $\varphi_{int,HVAC}$ :

$$\varphi_{int;t} = \varphi_{int,Oc;t} + \varphi_{int,Ap;t} + \varphi_{int,Li;t} + \varphi_{int,Wa;t} + \varphi_{int,HVAC;t} \quad (8)$$

The electricity imported from the grid ( $Elec_{mains}$ ) and the energy yield of the PV installation ( $Elec_{PV}$ ) were being registered during the monitoring campaign. However, there was no meter installed to specifically measure the amount of energy exported to the grid. The mains meter also does not work in reverse. The total electricity consumption of the dwelling  $Elec_{total}$  will therefore be calculated as the sum of  $Elec_{mains}$  and a cover factor for self-consumption  $\gamma_s$  times  $Elec_{PV}$  (see Table 3). In a first approximation, it will be assumed that 50 % of the generated electricity has been exported and that the remaining half is used on site ( $\gamma_s=0.50$ ), an approach adopted from BSRIA (BSRIA, personal communication, January 4, 2018). Alternatively, a cover factor of 0.31 will be used as suggested by Reynders et al. [40] for a system sized to annually produce as much as the house consumes and where no heat pump is installed.

To a large extent, this total electrical energy will be converted into internal gains by appliances and lighting ( $\varphi_{int,Ap}$  and  $\varphi_{int,Li}$ , resp.). Notable exceptions are for example the energy used by a washing machine to elevate the water temperature and the heat generated by an electric cooking stove that is extracted by a range hood.

The internal gains  $\varphi_{int,Wa}$  cannot be measured directly and no accurate estimate can be made. Similarly, uncertainty exists about  $\varphi_{int,Oc}$  and  $\varphi_{int,HVAC}$  (Eq.8). Ultimately, since no direct useful alternative inputs are available,  $\varphi_{int}$  will crudely be represented by  $Elec_{total}$  alone (Table 3).

### 3.1.6. Solar heat gains

A fifth driving force of  $\theta_i$  in Eq.1,  $\varphi_{sol,i}$  are the solar gains through each of the  $k$  transparent envelope elements (Eq.9).  $I_{sol;k}$  represents the solar irradiance per square metre of collecting area of surface  $k$  [ $W/m^2$ ], and encompasses both direct and diffuse solar radiation.  $F_{sh,ob;k}$  is a time-dependent ( $t$ ) shading reduction factor for external obstacles such as trees and roof overhangs.  $A_{sol;k}$  the effective collecting area of surface  $k$ , corresponds to the overall projected area of the glazed element altered by the frame area fraction, the total solar energy transmittance of the transparent part of the

element, and a shading reduction factor for movable shading provisions. Both  $A_{sol;k}$  and  $I_{sol;k}$  are applicable for a given orientation and tilt area and thus moment in time  $t$ .

$$\varphi_{sol;t} = \sum_k \varphi_{sol;k;t} = \sum_k (F_{sh;ob;k;t} \cdot A_{sol;k;t} \cdot I_{sol;k;t}) \quad (9)$$

Ideally, the direct and diffuse radiation for each orientation would have been registered on site during the OBM. Since this was practically unfeasible, the available GHR data, measured 30 km off site will be used instead. Limitations associated with this measurement variable are for example that it does not account for the site-specific shading and that it might suffer from time delays regarding for instance clouds.

Alternatively, the registered  $Ele_{cpv}$  (Table 2) can be used to represent the solar irradiance  $I_{sol}$ . Although this variable might better represent the actual incident radiation on the site, the raw PV production data might lack a correction for its area and electricity generation. Furthermore, the GHR can be converted into the incident radiation on the vertical surfaces, by means of TRNSYS Type 99 [41]. Yet, to ensure practical identifiability in the ARX model, only the incident radiation on the surface with the most dominant solar orientation, the south surface, will be used.

### 3.1.7. Latent heat gains

Over time, the building elements surrounding the zone load and unload hygroscopically, for newly built dwellings mainly by evaporating encapsulated moisture. These latent heat flows  $\varphi_{LG}$  can be mathematically expressed as the product of (1) the latent heat of evaporation  $h_L$  [J/(kg)] and (2) the moisture exchange rates  $G$  [kg/s] between the zone air on the one hand and the interior surfaces of the building elements separating the considered zone from the exterior environment ( $G_{si,e}$ ), the ground ( $G_{si,g}$ ), unconditioned spaces ( $G_{si,u}$ ) and adjacent buildings ( $G_{si,a}$ ) on the other hand:

$$\varphi_{LG;t} = h_L \cdot (G_{si,e;t} + G_{si,g;t} + G_{si,u;t} + G_{si,a;t}) \quad (10)$$

In both the case study house and the adjacent dwelling, two sensors were installed to register the relative humidity ( $RH$ ). The derived level of vapour pressure could give an indication of the latent heat gains, as well as of the presence of occupants and thus the internal gains. However, an in-depth study is needed to examine this relationship. Since  $\varphi_{LG}$  can be assumed to represent only a small fraction in the overall heat balance (up to 5 % of the heating power, [12]), it will further be neglected.

### 3.1.8. Definition of Heat Loss Coefficient

In this case, the difference between the air and equivalent temperature is small, as it concerns a well-insulated, airtight dwelling. In practice, it is even hard to measure on site with low-cost equipment. Therefore, the temperature difference terms in  $\varphi_{tr}$  (Eq.2) and  $\varphi_{vent}$  (Eq.5) can be approximated by one general  $(\theta_i - \theta_e)$  term. It is then possible to rewrite the transmission and ventilation losses as demonstrated in Eq.11. The sum of  $H_{tr}$  and  $H_{ve}$ , the heat transfer coefficient by transmission and ventilation respectively, is commonly (subscript 'c') defined as the '(overall) Heat Loss Coefficient' or  $HLC$ .

$$\varphi_{tr;t} + \varphi_{vent;t} = (H_{tr;t} + H_{ve;t}) \cdot (\theta_{i;t} - \theta_{e;t}) = HLC_c \cdot (\theta_{i;t} - \theta_{e;t}) \quad (11)$$

In this paper, however, a narrower definition is adopted, which excludes the transmission heat transfer coefficient through adjacent buildings ( $H_{tr,a}$ ), the subterm of  $H_{tr}$  in Eq.3 in which a temperature difference adjustment factor  $b_{tr}$  different from 1 is embedded ( $H_{tr,g}$  and  $H_{tr,u}$  are not present in this case study). The considered HLC thus forms a subterm of  $HLC_c$ , only taking account of the transmission heat loss to the ambient and the ventilation losses, as demonstrated in Eq.12a-b.

$\theta_a$  represents the temperature of the adjacent building and  $\varphi_{tr,a}$  the transmission heat flow rate to the adjacent building.

$$\varphi_{tr;t} + \varphi_{vent;t} = (H_{tr,e;t} + H_{ve;t}) \cdot (\theta_{i;t} - \theta_{e;t}) + (H_{tr,a;t}/b_{tr,a;t}) \cdot (\theta_{i;t} - \theta_{a;t}) \quad (12a)$$

$$= HLC \cdot (\theta_{i;t} - \theta_{e;t}) + \varphi_{tr,a;t} \quad (12b)$$

Finally, since it was decided to neglect  $\Phi_{LG}$  in this case study, Eq.1 may be reformulated as:

$$C_i \cdot \frac{d\theta_i}{dt} = HLC \cdot (\theta_{i;t} - \theta_{e;t}) + \varphi_{H;t} + \varphi_{int;t} + \varphi_{sol;t} + \varphi_{tr,a;t} \quad (13)$$

The outcome of the building physical framework is thus a definition of the HLC and its variables of influence, with some of the OBM data mapped to it (Table 3). But, it must be stressed that none of the 'monitoring variables' is a direct, one-to-one match with the 'theoretical variables' from the heat balance.

### 3.2. Statistical Framework

This work uses a transfer function modelling approach to characterise the HLC. More precisely ARX models will be identified on the available dataset using a combination of the variables listed in Table 3 as inputs and the interior temperature as output. The following paragraphs describe the steps needed to prepare the data set, estimate and validate the model coefficients and map the identified model coefficients to the physical parameter HLC.

#### *3.2.1. Pre-processing of the data*

Pre-processing of the collected monitoring data was carried out to adjust for Daylight Saving Time, discriminate and convert the *G<sub>as</sub>mains* data (as discussed in §3.1.4) and calculate the incident radiation on the south facade based on the *GHR* data (§3.1.6).

To ensure a relatively large interior-exterior temperature difference and limit the influence of the solar radiation on the overall heat balance, the heating season of 2014-2015 (October 1 till March 31) was selected as training period for the characterisation of the HLC. The monitoring data of October and November 2013 and January, February and March 2014 are used for cross-validation. The resulting data set was resampled to hourly values. Consequently, the percentage of missing data points per variable and per month of the considered periods ranges between 0.0 % and 6.7 % (disregarding the completely missing  $\theta_{return}$  time series).

To validate the choice of resampling time, variants with resampling times of 4, 6 and 24 hours were explored as well.

#### *3.2.2. Development of Data Subsets*

To gain insight into the relative influence of the above-discussed heat flows on the HLC estimate, 5 subsets are taken of the original dataset. The data subsets, described in columns 3 to 7 of Table 3, are numbered according to increasing extent (regarding the number of sensors installed) and thus cost of the measurement campaign.

To be able to evaluate all the candidate substitutes for the variables of the heat balance equation, it was furthermore decided to consider both a basic ('B') and some alternative ('A') versions of data subsets 1, 3 and 4.

Table 3: LEFT SIDE: Mapping of the available measurement data to the variables present in the heat balance. RIGHT SIDE: Overview of the data included in the 5 established subsets, with '1' the most rudimentary data set and '5' the most extensive one. 'B' marks elements present in the Base case, 'A' denotes investigated Alternatives with regard to the aspect included in the name tag.

Heat Balance	Available from collected data or assumptions	Data subset 1	Data subset 2	Data subset 3	Data subset 4	Data subset 5
$\theta_i$	• $\theta_{i\_liv}$ [°C]	1_A_ $\theta_{i\_liv}$				
	• $\theta_{i\_bed}$ [°C]	1_A_ $\theta_{i\_bed}$				
	• $\theta_{i\_mean}$ (arithmetic mean of $\theta_{i\_liv}$ and $\theta_{i\_bed}$ ) [°C]	1_B	2_B	3_B	4_B	5_B
	• $\theta_{i\_vol}$ (volume weighted average of $\theta_{i\_liv}$ and $\theta_{i\_bed}$ ) [°C]	1_A_ $\theta_{i\_vol}$				
$\theta_e$	• $\theta_{e1}$ [°C]	1_B	2_B	3_B	4_B	5_B
	• $\theta_{e2}$ [°C]				4_A_ $\theta_{e2}$	
$\theta_a$	• $\theta_{a\_mean}$ (arithmetic mean of $\theta_{a\_liv}$ and $\theta_{a\_bed}$ ) [°C]					5_B
$\varphi_H$	• $Gas_{mains}$ [W]	1_B				
	• $Gas_{mains}$ * boiler efficiency [W]	1_A_Efficiency				
	• (Estimated) gas consumption for SH [W]		2_B	3_B	4_B	5_B
$\varphi_{int}$	• $Elec_{total} = Elec_{mains} + 0.50 * Elec_{PV}$ [W]			3_B	4_B	5_B
	• $Elec_{total} = Elec_{mains} + 0.31 * Elec_{PV}$ [W]			3_A_Int		
$I_{sol}$	• GHR [W/m <sup>2</sup> ]				4_B	5_B
	• (Modelled) Incident radiation on the south surface [W/m <sup>2</sup> ]				4_A_South	
	• $Elec_{PV}$ [W]				4_A_PV	

- In this manner, the *base* version of data subset 1 ('1\_B') is limited to three time series:  $\theta_{i\_mean}$ ,  $\theta_{e1}$  and  $Gas_{mains}$ . To analyse the sensitivity of the HLC estimate to the sensor set-up the *alternative* candidates for  $\theta_i$ , listed in column 2 of Table 3, are subsequently used in '1\_A\_ $\theta_{i\_liv}$ ', '1\_A\_ $\theta_{i\_bed}$ ' and '1\_A\_ $\theta_{i\_vol}$ '. Likewise, the characterisation exercise is executed with and without taking account of the theoretical boiler efficiency in respectively cases '1\_A\_Efficiency' and '1\_B'.
- Data subset 2 (DS2) entails a larger measurement set-up with an extra sensor to register  $Water_{mains}$ . This way it is possible to estimate the gas consumption for space heating and hence assess the impact on the HLC of the exclusion of the energy consumption for DHW production from  $\varphi_H$ .
- An even more elaborate measurement campaign also tracks  $Elec_{mains}$  and  $Elec_{PV}$ . At its most extreme, all electricity use ( $Elec_{total}$ ) can be attributed to internal gains. Hence, DS3 examines the expected increase of the estimated HLC when including these gains.
- In the case of DS4 we furthermore consider the accessibility to the data of a 'local' weather station. This not only makes it possible to investigate the difference between an HLC estimation based on exterior temperatures measured on site (basic version '4\_B') or 30 km off site (alternative version '4\_A\_ $\theta_{e2}$ '), but also to include the solar radiation explicitly in the model in the form of the GHR ('4\_B') or its alternatives.
- Finally, DS5 additionally encompasses detailed information on the temperature of the neighbouring dwelling  $\theta_{a\_mean}$ . This is in contrast to the other cases for which the heat losses to the adjacent building were not modelled explicitly (see further).

### 3.2.3. Model Parametrization and Selection

The model structure of the ARX, as it was defined by Madsen et al [42], is being tailored to this application. Since the case study concerns a real-life occupied building in which the temperature fluctuates, it is opted to analyse the data subsets using a model with  $\theta_i$  as output variable. The previously-discerned physical phenomena influencing  $\theta_i$  (Table 3) determine the input variables. Hence, the translation of the considered heat balance (Eq.13) into an ARX model formula is

presented in Eqs. 14 and 15.  $\varphi_x(B)$  and  $\omega_y(B)$  are, respectively, output and input polynomials of order  $p_x$  and  $p_y$  in the backshift operator  $B$ .  $Int$  is a constant intercept term and  $\varepsilon_t$  the residual (error) [42].

Eq.14 was applied on DS1 and DS2. Next, the model structure was gradually extended with more variables to analyse the more elaborate data sets (Eq.15). An approach which sums up the gas use for space heating and the total electricity use (crudely representing estimable internal gains by appliances and lighting) is adopted since both phenomena are assumed to mainly use convection as heat transfer mode.

$$\varphi_i(B) \cdot \theta_{i;t} = \omega_H(B) \cdot \varphi_{H;t} + \omega_e(B) \cdot \theta_{e;t} + Int + \varepsilon_t \quad (14)$$

$$\underbrace{\varphi_i(B) \cdot \theta_{i;t}}_{\text{Data set 3}} = \underbrace{\omega_H(B) \cdot (\varphi_{H;t} + \varphi_{int;t})}_{\text{Data set 4}} + \underbrace{\omega_e(B) \cdot \theta_{e;t} + Int + \varepsilon_t + \omega_{sol}(B) \cdot I_{sol;t} + \omega_a(B) \cdot \theta_{a;t}}_{\text{Data set 5}}$$

(15)

To decide on the model order, a backward elimination procedure is followed. In a first step, a model including 24 lags for each of the considered polynomials ( $p_x=p_y=24$ ) is fitted on the data using the `lm` function in R [43]. It was chosen to involve 24 data points from the past since this corresponds, for a 1-hour resampling time, to a daily cycle<sup>2</sup>. Based on the results of the fitting, in subsequent runs all model coefficients of the highest order present (initially: 24) which prove insignificant according to a t-test (threshold of 0.05) are removed. Hence, following this iterative selection procedure, a model with the lowest possible order for all polynomials and significant model coefficients is retained.

#### 3.2.4. Model Validation and Comparison

A number of performance tests are carried out to verify the validity of the developed models. First, the fitted model coefficients are evaluated by means of a marginal t-test. Since the model order selection process already uses the *coefficients' significance* as a criterion, all final models should pass this first test.

Next, the prediction accuracy of the model is assessed.

The autocorrelation function (ACF) and cumulated periodogram (CP) of the model residuals are inspected to verify that they show no significant *autocorrelation*. If any autocorrelation is still existing, the *cross-correlation* function of the residuals with the input variables and the time series plots of the residuals and input variables are examined for any clear patterns, indicating model deficiencies. To ensure that the developed models are not *over-fitted*, the normalized root mean square error nRMSE [%] between the observed interior temperature and 1-step ahead predictions is furthermore determined for both a two-month period embedded in the training period (November-December 2014) and a two-month period that has not been used for the parameter estimation process, a so-called 'cross-validation period' (November-December 2013). The nRMSE is hereby defined as Eq.16, with  $\theta_i(t)$  and  $\hat{\theta}_i(t)$  the observed and predicted signals, respectively, and  $n$  the number of samples considered. The result of this cross-validation test should be that there is no significant difference between the model's goodness-of-fit for both periods.

<sup>2</sup> For the model variants fitted on data resampled to 4-hour, 6-hour and 24-hour values, starting orders of respectively  $p=6$ ,  $p=4$  and  $p=1$  were used.

$$nRMSE = \sqrt{\frac{\sum(\theta_i(t) - \hat{\theta}_i(t))^2}{(n-1)}} \cdot \frac{100}{(\theta_{i,max} - \theta_{i,min})} \quad (16)$$

Taking this *cross-validation* even a step further, the developed models are challenged to *simulate* the model output  $\theta_i$  for 3 different validation periods, selected based on the condition that no observations of the input variables or their lags are missing. The validation periods for simulation hence are (1) January 7 till January 16, 2014; (2) January 27 till February 5, 2014 and (3) February 20 till March 2, 2014. The nRMSE between the simulated ( $\hat{\theta}_i$ ) and measured ( $\theta_i$ ) time series gives insight into the ability of the model to predict the dynamic behaviour of the real system. If the developed model has trouble or even fails to explain the observed variances, its internal structure and thus the inferred HLC can be questioned.

### 3.2.5. Determination of 'HLC'

The identified model coefficients of Eqs.14-15 are physically interpreted as outlined in Eq.17: by taking the quotient of the exterior air and heating power polynomials, the parameter HLC is obtained. This parameter represents the sum of  $H_{tr,e}$  and  $H_{ve}$  as stated in Eqs.12, if there is no cross-correlation between unmodelled phenomena and  $\theta_e$ . The backshift operator B is set equal to 1 in order to obtain the steady-state gain, as desired in the case of the HLC.

$$\frac{\omega_e(1)}{\omega_H(1)} = HLC \quad (17)$$

## 4. Results and Discussion

As can be deduced from Eqs. 6 and 12, the heat loss coefficient of the case study dwelling comprises three components:  $H_{tr,e}$ ,  $H_{ve,wtt}$  and  $H_{ve,tt}$ . The information available on the building allows to make an initial estimate for each of them:

- Based on the building components' surface areas and design U-values (Table 1) a design  $H_{tr,e}$  of 39.0 W/K is calculated (Eq.4).
- Considering that there are no natural ventilation systems and that a target threshold was set for the infiltration rate,  $H_{ve,wtt}$  should not exceed 8.2 W/K (Eq.18, with  $A_E$  the building envelope area).

$$H_{ve,wtt} = 0.34 \cdot (0.05 \cdot (AP_{50} \cdot A_E)) \quad (18)$$

- Given the missing measurement data and the limited information on the mechanical ventilation system, an assumption will have to be made on the value of  $H_{ve,tt}$  (Eq.6). An air change rate of 0.4 1/h, which implies a  $G_v$  of 0.02 m<sup>3</sup>/s, and an efficiency of the heat recovery unit of 91.0 % ( $b_{ve}=0.09$ ), result in a value of 2.4 W/K for  $H_{ve,tt}$ . By setting  $b_{ve}$  equal to 0.09, the heat recovery unit is assumed to work at its maximum efficiency, as the limited measurement data of  $\theta_{supply}$  and  $\theta_{return}$  seem to suggest. Meanwhile, with a lower efficiency of 80 % or 60 %,  $H_{ve,tt}$  could grow as large as 5.2 W/K or 10.5 W/K respectively.

Adding up these three heat transfer coefficients, a design HLC ranging between 49.6 W/K and 57.7 W/K is obtained (Eqs. 6 and 12).

On-site measurements of the U-value of the external walls and the  $AP_{50}$ , as well as infrared thermography, already indicated that the actual as-built HLC might deviate from this theoretical value, which for example excludes any thermal bridging effects (§2.1.1). Moreover, the actual air change rate and energy efficiency of the MVHR are unknown and no information is available on the opening of windows. Therefore, this design value 'range' will merely be considered as an indicative reference, not as a strict criterion to evaluate the accuracy of the estimation outcomes.

In what follows, an overview will be given of the characterisation results for the 5 data subsets. This will provide insight into the sensitivity of the HLC estimate to the extent of the monitoring set-up and the assumptions made in case of missing input data.

However, first, the impact of the duration and timing of the monitoring campaign and the applied data analysis techniques will be demonstrated, since this will underpin some of the choices made in setting up the experiment. Hereto the most limited and most extensive data subset will be used (DS1 and DS5 respectively).

It is important to emphasise that the case study building has a highly insulated and airtight envelope. The results should be looked at in this context.

#### 4.1. Sensitivity of Estimate to Duration and Timing of the Monitoring Campaign

For data subset 1\_B, which only includes monitoring data of the total gas consumption and an assessment of the temperature difference over the building envelope, an estimate of the HLC of 69.6 W/K (range 95% confidence interval (CI)=21.5 W/K) is obtained. This result was deduced from monitoring data for a full heating season (Oct 1-Apr 1, 26 weeks). Nonetheless, it can be seen from Figure 3 that a similar result could have been found for a monitoring campaign of at least 16 weeks. From this point on, the mean estimates for HLC vary consistently around approximately 70 W/K and within 13.3 % of their mean. The range of the 95% CI too drops significantly until a duration of 16 weeks, but remains almost constant thereupon. It is hence essential to consider a sufficiently long monitoring period (at least 16 weeks in this case) and to verify whether the estimates converge. The limited convergence could for this first data subset be attributed to the crude input data, since for the more extensive data set 5, the results seem to stabilize earlier in time and from the 16week-mark on, the results only vary within 11.8 % of their mean and the range of the 95 % CI is limited to 17.5 W/K.



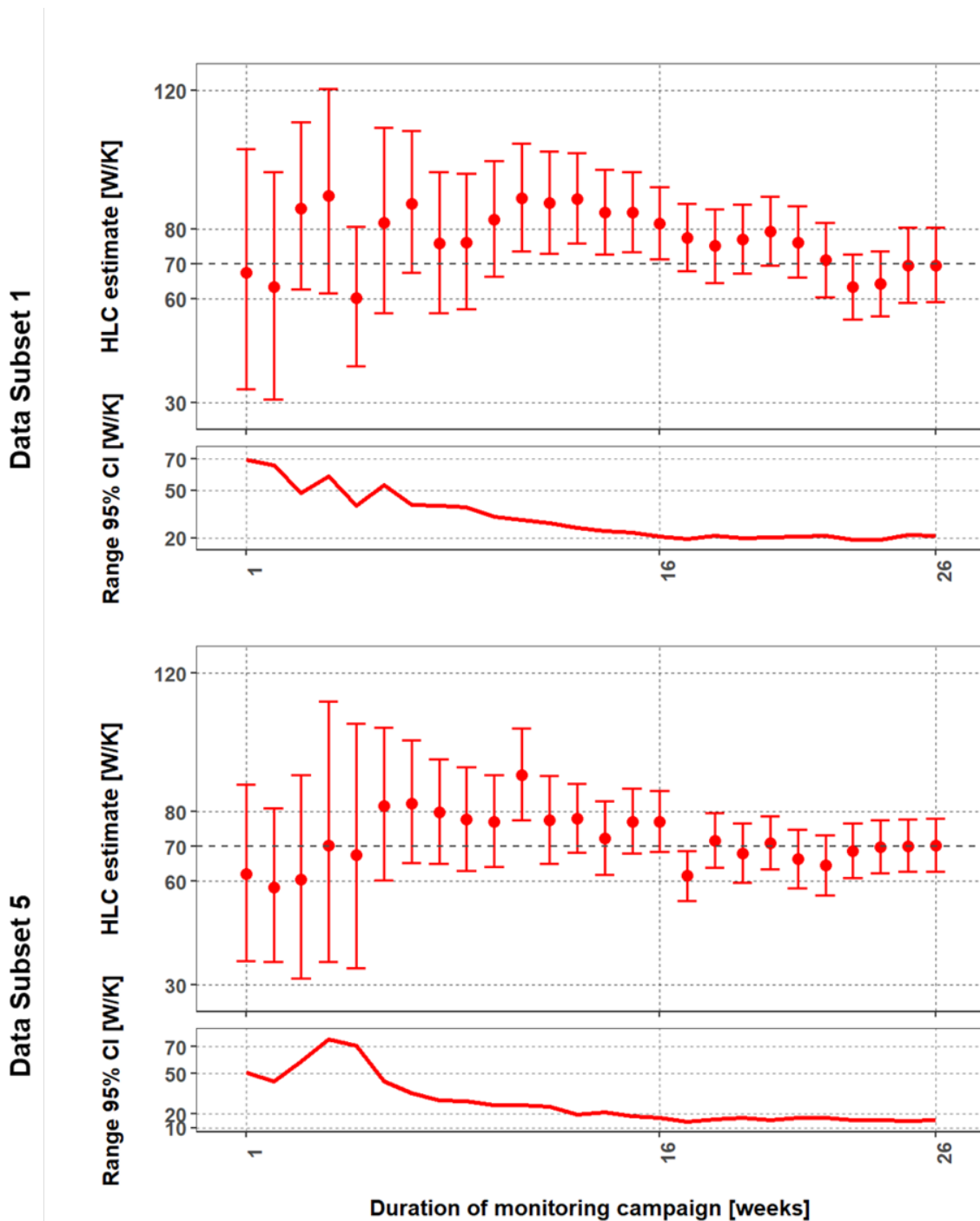


Figure 3: From a monitoring period of 16 weeks onwards, about the same range of 95 % confidence interval ('CI', indicated by the whiskers) is obtained for Data Subset 1 and the mean estimates (dots) appear to stabilize. Full convergence, however, cannot be expected from this plain model; the mean estimates continue to fluctuate between approximately 60 W/K and 80 W/K.

For the more extensive Data Subset 5, the results seem to stabilize 3 weeks earlier: the range of the 95 % CI remains quasi-constant from week 13 onwards.

Naturally, both the characterisation outcome and convergence rate depend on the quality of the training data. This is demonstrated in Figure 4, which shows the base case estimates for DS1 and DS5 on the far left, and the results of five investigated variants on the right. A variant of case 1\_B with a 1 month shifted monitoring data set, spanning from November 1 to May 1, resulted in a 21.6 % higher mean estimate (Figure 4, variant 'Starting point'). However, the estimate for the shifted data set is

considered less reliable than that for case 1\_B: the model residuals show a higher autocorrelation and in addition, the estimate did not converge towards the result obtained for the 26-week data set. Performing the same period-shifting exercise on Data Subset 5, results in a less deviant estimate. Hence, the seemingly limited robustness of estimate 1\_B might among other things be attributed to the lack of modelling of the solar gains, which have more importance in spring time. Nonetheless, repeating the characterisation exercise on a different measurement period, as is also done in the cross-validation tests, is a good approach to assess the robustness of the outcome.

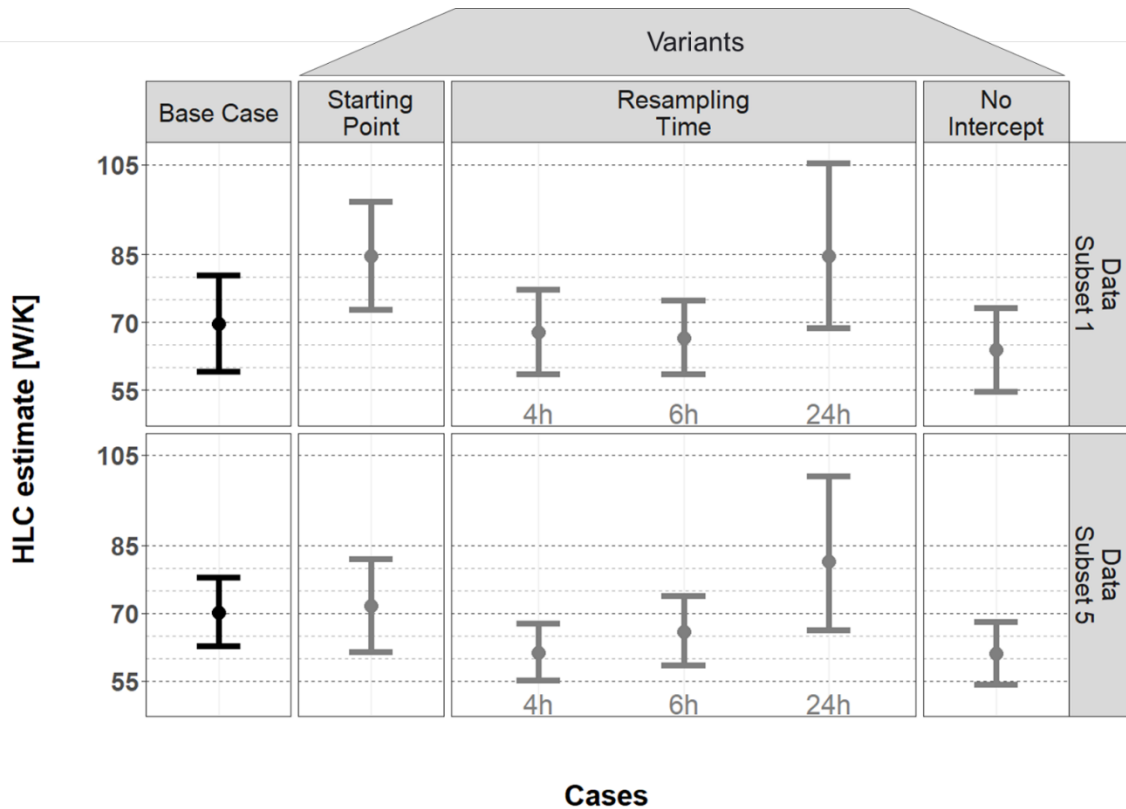


Figure 4: An investigation of some variants of the base cases 1\_B (upper row) and 5\_B (lower row) learns that shifting the monitoring campaign in time or adopting an altered data analysis approach can result in deviations of the HLC estimate up to 21.6%. The figure presents the results obtained from 26 weeks of hourly data

#### 4.2. Sensitivity of Estimate to Data Analysis Technique

As stated earlier (§3.2.1) it was decided to fit the ARX models on data with the highest available frequency: hourly values. In Figure 4, however, the sensitivity of the characterisation result to the resampling time (RT) of the data is explored. For 4-hourly and 6-hourly averaged data, slightly lower mean estimates are obtained (2.5 % and 4.5 % resp. for DS1, 12.7 % and 6.3 % resp. for DS5). The uncertainty on the outcome reduces even more (12.8 % and 23.6 % resp. for DS1, 17.1 % and 0.1 % resp. for DS5). Analysis of diurnal data, on the other hand, results in higher mean HLCs and larger 95 % CIs. For both the base model and its variants, all model coefficients are significant and the residuals pass the ACF and CCF tests. Although the hourly resampling time does not provide the smallest CI, the cross-validation tests via simulation argue in its favour, by pointing out that the models 1\_B and 5\_B are better at predicting the dwelling's behaviour (indoor temperature profile) than their variants, as expressed by lower nRMSE-values (Table 4). At larger data resampling intervals, the faster dynamics are averaged out to a certain extent. This appears to make the considered ARX model less appropriate to analyse the data.

Table 4: The nRMSE between  $\theta_i$  and  $\hat{\theta}_i$  for some prediction tests performed on both training data and validation data. The results of the one-step-ahead predictions signal no significant problems of over-fitting, the nRMSEs of the cross-validation tests via simulation are used for model comparison. RT= resampling time.

Cases	nRMSE [%]				
	One-step-ahead		Simulation		
	Nov-Dec '14 (Training)	Nov-Dec '13 (Validation)	Jan '14 (Validation)	Jan-Feb '14 (Validation)	Feb-Mar '14 (Validation)
<b>1_B</b>	1.7	2.7	31.0	21.3	28.7
<b>1_B, RT=4u</b>	5.4	8.1	31.9	23.0	29.9
<b>1_B, RT=6u</b>	6.0	11.1	39.3	25.9	39.2
<b>1_B, RT=24u</b>	9.8	17.9	97.4	26.8	39.1
<b>1_A_Efficiency</b>	1.7	2.7	31.0	21.3	28.7
<b>2_B</b>	1.7	2.7	14.5	10.7	12.8
<b>3_B</b>	1.5	2.6	25.5	17.7	21.6
<b>3_A_Int</b>	1.6	2.6	27.0	18.0	20.0
<b>4_B</b>	1.5	2.6	22.7	16.4	23.5
<b>4_A_θ<sub>e2</sub></b>	1.5	2.6	27.0	17.0	25.5
<b>4_A_PV</b>	1.5	2.6	22.7	16.6	23.1
<b>4_A_South</b>	1.5	2.6	25.1	17.5	21.4
<b>5_B</b>	1.5	2.6	22.1	16.0	24.0
<b>5_B, RT=4u</b>	4.7	7.9	30.3	22.7	28.2
<b>5_B, RT=6u</b>	5.1	10.7	29.9	20.1	28.4
<b>5_B, RT=24u</b>	8.8	19.5	41.1	24.4	41.4

Forcing the model structure through the origin by setting the intercept term in Eqs. 14-15 equal to zero, might come across as an attractive option, as it gives realistic results with smaller confidence intervals (13.6 % and 8.9 % smaller ranges compared to the 95 % CIs of 1\_B and 5\_B respectively, as demonstrated by the rightmost estimates in Figure 4). Nevertheless, deliberately excluding a statistically significant constant term seems incorrect as the term might represent the unmodelled steady-state phenomena. The type of physical phenomena that could be lumped into the constant term are also discussed by Bauwens [18].

Hence, to ensure comparability, a training data set spanning over 26 weeks, hourly resampling time and non-zero intercept term will be assumed for all subsequent analyses.

#### 4.3. Sensitivity of Estimate to Extensiveness of the Monitoring Set-up and Assumptions Made

Figure 5 displays an overview of the estimated HLCs for all the base case data sets and the investigated alternatives ('B' and 'A', respectively) listed in Table 3.

Overall and ignoring some outliers, the characterisation results for HLC show remarkable agreement, considering the rather distinct underlying data sets and assumptions. The majority of the estimated HLC values are to be found within the range of 60.7 to 71.6 W/K, and their confidence intervals show considerable overlap. The uncertainty on the estimates reduces with 30.7 % from the base case

model of DS1 ('1\_B') to case 3\_B, but does not decrease further for cases 4\_B and 5\_B.

A more detailed description of the statistical and physical interpretation of the different cases will be given below. The base case and alternative cases will be discussed subsequently, starting from DS1.

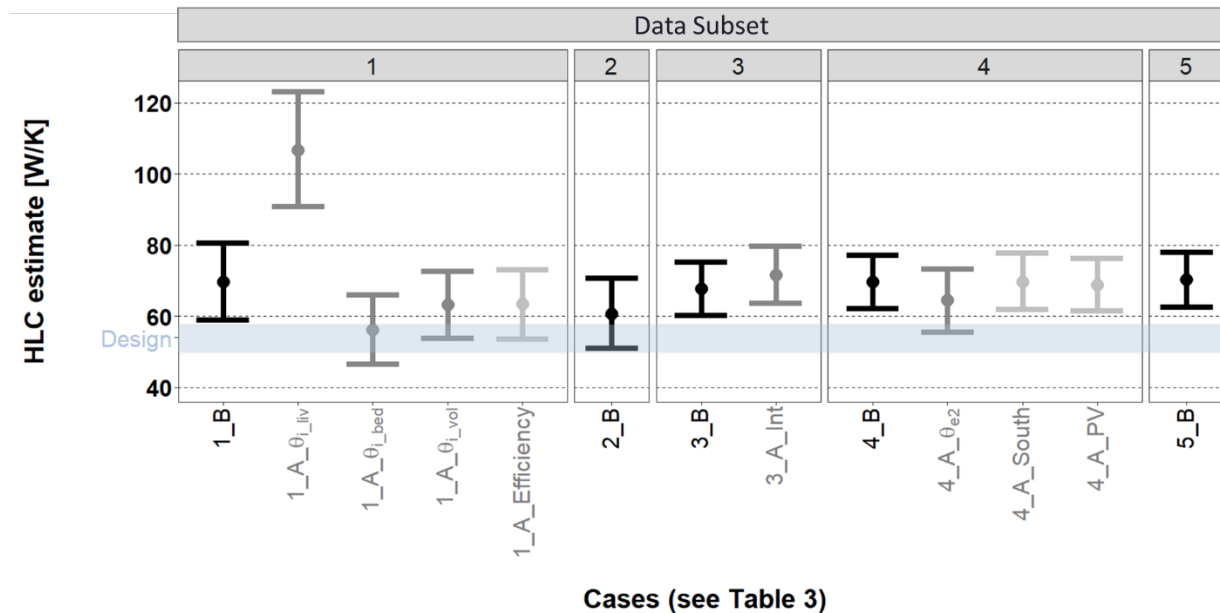


Figure 5: The deviation between the HLC estimates for the different cases exhibits the characterisation outcome's sensitivity to decisions made during the monitoring campaign and data analysis.

#### 4.3.1. Data subset 1

The base case 1\_B uses  $\theta_{i,mean}$  as interior temperature. The three alternatives for  $\theta_i$ , outlined in Table 3, were evaluated as well. An overview of a few statistical features of all four time series is provided in Table 5. Figure 5 displays the sensitivity of the HLC estimate to the choice for  $\theta_i$ . Strikingly, the interior temperature signals cannot be used interchangeably for the characterisation. Although their mean values and variance closely align and they are highly correlated, as can be expected from a well-insulated, airtight building, the HLC estimates deviate considerably (up to 89.5 %). Especially the use of  $\theta_{i,liv}$  leads to a deviant, highly uncertain estimate.

The ARX model used for the characterisation is a black-box model; it thus entirely depends on the provided data to create its internal structure. Since the parameter which we ultimately aim to deduce from the model coefficients, the HLC, depends on the interior temperature of a homogeneously assumed zone (see BPF §3.1.1), it seems appropriate to initiate the model with a combination of  $\theta_{i,liv}$  and  $\theta_{i,bed}$  as model input. Volume weighted averaging requires prior knowledge of the building geometry which will not always be readily available for remote assessments.

Therefore, preference will be given to the use of  $\theta_{i,mean}$  in the remainder of the paper.

The reservations concerning the suitability of the available time series to represent  $\theta_i$  (see §3.1.1), and the above-demonstrated sensitivity of the estimate to  $\theta_i$ , should be kept in mind when interpreting the obtained HLC estimates. Moreover, the results shown in Figure 5 stress the need for further research into the optimal set-up of interior temperature sensors, as they evidently have strong implications on the HLC estimate, even in a well-insulated, relatively small building.

Table 5: At first sight, a statistical description of the four candidate time series for ' $\theta_i$ ' (taking the integral heating season into consideration) does not reveal any major differences that could give rise to divergent HLC estimates.

$\theta_i$ candidate $\rightarrow$	$\theta_{i,liv}$	$\theta_{i,bed}$	$\theta_{i,mean}$	$\theta_{i,vol}$
Average( $\theta_i$ ) [°C]	21.1	21.6	21.3	21.4
Standard deviation( $\theta_i$ ) [°C]	1.7	1.5	1.4	1.4
Correlation( $\theta_i, \theta_{el}$ )	0.5	0.1	0.4	0.4
Correlation( $\theta_i, \theta_{i,liv}$ )	1	0.5	0.9	0.9

Another variant of data subset 1 assesses the consequence of using the gross energy input ( $Gas_{mains}$ ), instead of an energy input that takes account of a constant boiler efficiency of 91%. As might be expected, Figure 5 shows a semi-net HLC ('1\_A\_Efficiency') that is 9.0 % lower than its gross variant ('1\_B'), and has an accordingly reduced CI. Since the boiler efficiency only represents a part of the overall system efficiency, the actual overestimation associated with an HLC solely based on gas consumption data will be even larger. This raises the question whether it is possible to perform accurate estimations of the heat loss coefficient based on smart meter gas data - certainly without prior knowledge of the building systems.

#### 4.3.2. Data subset 2

In a first attempt to align the outcome of the characterisation more closely with the actual physical HLC, the  $\varphi_H$  variable was modified towards DS2 to only represent the gas use for space heating (see §3.1.4). This lowers the mean HLC estimate of the base case with 12.8 %, from 69.6 W/K to 60.7 W/K (cases 1\_B and 2\_B, Figure 5). The (range of the) associated 95 % CI decreases as well, with 8.5 %. The rather large reduction of the nRMSE (Table 4) between cases 1\_B and 2\_B indicates a significant improvement of the model. The disentanglement of the gas consumption for SH and the production of DHW hence is an important adaptation, certainly in this case study of a well-insulated house. Nonetheless, given the fact that the applied decomposition method has not yet been validated, a significant degree of uncertainty remains with regard to the assessed values for both  $\varphi_H$  and the HLC [39].

#### 4.3.3. Data subset 3

For DS3, which additionally includes the measurements  $Elec_{mains}$  and  $Elec_{PV}$ , the assumption is made that all electricity consumption results in internal gains that, together with the net heating input, cover the heating demand in winter (Eq.15). A cross-correlation plot of the residuals of the ARX model of case 2\_B with the  $Elec_{total}$  time series of 3\_B indeed corroborates the view that this variable is a useful extension of the model description (Figure 6).

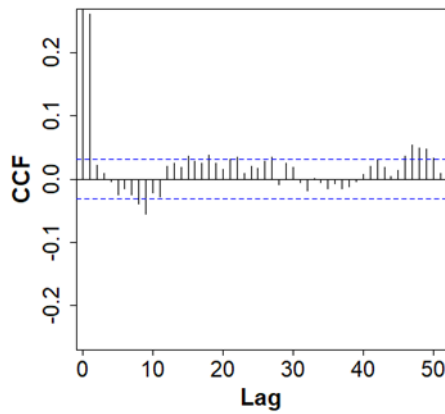


Figure 6: The cross-correlation function (CCF) between the residuals of case 2\_B and the estimated total electricity consumption of case 3\_B shows significant correlation (especially at the first lag, but also at some further lags), which emphasises the potential of  $EleC_{total}$  to improve the model fit.

Since the actual total electricity consumption is unknown, two estimates were made using different cover factors  $\gamma_s$  for the gains from local photovoltaic energy generation (see Table 3). Figure 5 shows that the base and alternative case are characterised by a respectively 11.6 % and 18.0 % higher HLC than case 2\_B. The fact that the higher supply cover factor does not result in the higher estimate may be attributed to the ARX model's black-box structure that considers the correlation between the input and output variables, which does not necessarily imply conservation of energy.

Both models score well on the aforementioned validation tests (ACF, CCF), and a comparison of the nRMSEs of the one-step-ahead predictions indicates that the models are not over-fitted (Table 4). However, the cross-validation tests via simulation mark the models as less suitable candidates to represent the dwelling system than model 2\_B. The relatively higher nRMSEs indicate that arbitrarily adding the whole mains electricity use plus either 50 % or 30 % of the PV-production in each time step might be an overestimation of the actual internal heat gains. Notably, however, this finding might emerge from a time shift between the electricity consumption and thermal gains or from a mismatch between the position of the temperature sensors and the electrical appliances. In addition, it is important to emphasize that the heat flow rate from occupants, which forms part of the total internal heat gain, is not explicitly incorporated in either of the models.

Since case 3\_B approximates the dwelling's behaviour slightly closer than 3\_A\_Int, the 50 % cover factor will be applied in DS4 and DS5.

#### 4.3.4. Data subset 4

The dynamic heat balance also includes the incident solar radiation as an explanatory variable for the variation of  $\theta_i$  in time. Not monitoring it, and hence not being able to include it in the ARX model, may therefore lead to a model structure that is missing out on the fast dynamics and result in an incorrect HLC estimate. However, including it improperly, might be even more harmful while giving a false sense of accuracy. Data subset 4 is intended to examine the impact on the HLC estimate of, firstly, the mere inclusion of an explicit solar radiation term and, secondly, the form in which it is included. Three different variables, feasible to monitor in practice, were investigated (see Table 3). It must be noted that none of these three time series corresponds to the solar radiation as requested by Eq.9, since the actual incident radiation, per orientation and tilt angle and accounting for external shading, is not available. This may have repercussions for the interpretability of the estimated HLC values.

Cross-validation shows that all three models (4\_B, 4\_A\_PV and 4\_A\_South) are better at explaining the variance in the interior temperature than their equivalent without an explicit  $I_{sol}$  term (case 3\_B).

Still, the addition of the solar radiation enhances the model only moderately. Possible reasons are not hard to find. Firstly, the three discussed models include the solar radiation in a very crude way: the solar radiation was averaged to hourly values on a horizontal or south facing surface and the solar aperture is modelled as a single term (instead of for example as a function of time and orientation). Moreover, it should be noted that the window area of the case study dwelling is rather small (Table 1). Finally, the solar registrations at a weather station 30 km off site and the PV yields potentially do not fully capture the actual solar heat gains.

Regardless of whether or not the model has sufficiently improved, Figure 5 shows that the estimates obtained with the base case model and its two solar radiation related alternatives are nearly equal (largest difference between their means is 0.9 W/K), which suggests that in this case inquiring off-site GHR data and modelling the solar radiation in TRNSYS were unnecessary.

Another variable related to the local climatic conditions is the outdoor air temperature. Substituting the exterior temperature measured on site ( $\theta_{e1}$ ) in model 4\_B with the off-site registrations ( $\theta_{e2}$ , see Table 3), results in a 7.3 % lower mean HLC estimate with a larger CI (17.0 % higher range). The RMSE between both exterior temperature time series is 1.2 °C with a maximum absolute difference of 5.3 °C. The cross-correlation of  $\theta_{e1}$  and  $\theta_{e2}$  with  $I_{sol}$  is almost the same (maximum cross-correlation over 100 lags of 0.29 and 0.26 respectively). A detailed investigation of the positioning of the sensors registering  $\theta_{e1}$  and  $\theta_{e2}$  with respect to local shading and wind flow patterns should reveal which time series is most representative. For the next data subset preference will be given to  $\theta_{e1}$  since model 4\_B scores better at the cross-validation tests.

#### 4.3.5. Data subset 5

The assumption for the 5<sup>th</sup> and final data subset is that the temperature of the neighbouring dwelling is monitored as well. The model structure of the ARX (Eq.15) can now be extended with this variable, as a result of which the heat loss through the party wall is modelled explicitly. This model extension should not only enable the explanation of a larger amount of the dynamics and variance available in the observed interior temperature. It should also give the means for purging other model coefficients, whose variables are somewhat correlated with  $\theta_a$ , such as possibly  $\omega_e$ , from disturbances and uncertainty.

Despite these expectations, it appears in practice impossible to distinguish between  $\theta_a$  and the constant intercept term. Therefore, the  $\omega_a$  polynomial had to be eliminated from the model description for being 'insignificant'.

## 5. Conclusion

Currently, the question is raised whether and how on-board monitoring could be used to characterise the building energy performance of occupied buildings on site. This paper presents a first step in answering this question by assessing the heat loss coefficient (HLC) of a well-insulated, semi-detached, occupied house in the UK based on on-board monitoring and data analysis using ARX models.

The HLC estimates that are labelled 'more trustworthy' on the basis of a validation process consistently range between 60.7 W/K and 71.6 W/K (this compared to a design-value of 49.6 -57.7 W/K). With a range of about 17.2 W/K the 95% confidence intervals are rather high, but they overlap.

However, the paper also shows that estimates as far as 89.5 % apart can be obtained due to decisions made by the user throughout the characterisation process. With the aid of a carefully designed system of 'data subsets', 'variants' and 'alternative variables', insight was offered into the origin of these deviations.

Firstly, the sensitivity of the outcome to the timing and duration of the monitoring campaign was addressed. It was shown how, for a data subset only including the exterior and interior temperatures and total gas consumption, the monitoring should span over at least 16 weeks. For a more extensive data set with more sensors, a 3-week shorter monitoring period appeared to suffice. Reapplying the characterisation on a different part of the monitoring data set proved to be a good approach to evaluate the robustness of the outcome.

Secondly, the influence of the extensiveness of the monitoring set-up was evaluated. The number of interior temperature sensors, and their position, proved to have an important impact on the HLC estimate, even in a well-insulated dwelling. The estimate based on a single temperature signal (registered in the living room) was for instance 68.9 % higher than the one based on a volume weighted average of two sensors installed on different floors (in the living and bedroom). The study furthermore demonstrated the uncertainty and the overestimation associated with HLC characterisations based solely on smart meter gas consumption data. Not accounting for the boiler efficiency can already cause an overestimation of 9.9 %. However, accounting for solar gains and the temperature of the neighbouring dwelling did not appear to significantly alter the assessed HLC for this case. The inability to explicitly model the ventilation losses on the other hand causes a lingering uncertainty about the accuracy of the fitted HLC.

Furthermore, the analyses provided a deeper understanding of the assumptions surrounding the handling of unmeasured phenomena. For this specific case, (1) GHR data registered in a weather station 30 km off site, (2) incident radiation on the south surface calculated based on these data, and (3) the electricity generated by the on-site PV-system could serve interchangeably as a placeholder for the incident solar radiation, with a difference of the HLC estimate smaller than 1.3 %. However, larger differences (7.3 %) were observed when substituting the exterior temperature measurements of a sensor installed in a professional weather station by on-site registrations.

Regarding the internal gains, too, assumptions have to be made as to which share of the total electricity consumption can be classified as internal heat gains. In the case where PV generation is present but not submetered, additional assumptions will have to be made concerning the supply cover factor. It was shown how these can introduce deviations of up to 5.7 %. In the same vein, different approaches to disentangle the non-submetered gas consumption for space heating and domestic hot water can lead to significant deviations in the HLC estimate.

A fourth element that was analysed is the applied data analysis technique. It was found that it is necessary to add an intercept term to the ARX model structure as long as it proves to be statistically significant. Furthermore, ARX models proved to be better able to describe the dynamical behaviour of the system, and hence can be assumed to yield more accurate HLC estimates, when the faster dynamics are not averaged out of their training data.

The research did not only address the pitfalls when characterising the as-built HLC, but also provided some handles for appraising the validity of the characterisation results. First, the model description should be carefully compared to the underlying building physical framework, to verify if any relevant physical phenomena are not (accurately) included or are counterintuitively classified as insignificant. Further, the statistical validity of the model should be checked, especially by investigating the model residuals. Finally, it was shown how cross-validation via simulation (the nRMSE) can be used to test the model structure's ability to explain the dynamics present in the measurement observations.

## 6. Acknowledgements

The authors gratefully acknowledge the Research Foundation Flanders (FWO) and the Flemish Institute for Technology (VITO) for funding this research (application number 1131918N).



## 7. References

- [1] Eurostat - Statistical Office of the European Union, Energy, transport and environment indicators, 2017. doi:10.2785/964100 KS-DK-17-001-EN-N.
- [2] EU, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast), Off. J. Eur. Union. L153 (2010) 13–35. doi:10.3000/17252555.L\_2010.153.eng.
- [3] H. Hens, Energy efficient retrofit of an end of the row house: Confronting predictions with long-term measurements, *Energy Build.* 42 (2010) 1939–1947. doi:10.1016/j.enbuild.2010.05.030.
- [4] O.G. Santin, Actual energy consumption in dwellings. The effect of energy performance regulations and occupant behaviour., PhD Thesis TUDelft, The Netherlands, 2010. doi:10.3233/978-1-60750-651-5-i.
- [5] D. Majcen, L.C.M. Itard, H. Visscher, Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications, *Energy Policy.* 54 (2013) 125–136. doi:10.1016/j.enpol.2012.11.008.
- [6] Zero Carbon Hub, Closing the gap between design & as-built performance: Evidence Review Report, London, 2014. Available from: [www.zerocarbonhub.org/full-lib](http://www.zerocarbonhub.org/full-lib) (accessed May 2, 2019).
- [7] D. Cali, T. Osterhage, R. Streblow, D. Müller, Energy performance gap in refurbished German dwellings: Lesson learned from a field test, *Energy Build.* 127 (2016) 1146–1158. doi:10.1016/J.ENBUILD.2016.05.020.
- [8] G. Dall’O’, L. Sarto, A. Galante, G. Pasetti, Comparison between predicted and actual energy performance for winter heating in high-performance residential buildings in the Lombardy region (Italy), *Energy Build.* 47 (2012) 247–253. doi:10.1016/J.ENBUILD.2011.11.046.
- [9] H. Hens, A. Janssens, W. Depraetere, J. Carmeliet, J. Lecompte, Brick cavity walls: A performance analysis based on measurements and simulations, *J. Build. Phys.* 31 (2007) 95–124. doi:10.1177/1744259107082685.
- [10] F.G.N. Li, A.Z.P. Smith, P. Biddulph, I.G. Hamilton, R. Lowe, A. Mavrogianni, E. Oikonomou, R. Raslan, S. Stamp, A. Stone, A.J. Summerfield, D. Veitch, V. Gori, T. Oreszczyn, Solid-wall U - values: Heat flux measurements compared with standard assumptions, *Build. Res. Inf.* 43 (2015) 238–252. doi:10.1080/09613218.2014.967977.
- [11] R. Gupta, D. Dantsiou, Understanding the Gap between ‘as Designed’ and ‘as Built’ Performance of a New Low Carbon Housing Development in UK, in: Springer; Berlin, Heidelberg, 2013: pp. 567–580. doi:10.1007/978-3-642-36645-1\_53.
- [12] D. Johnston, D. Farmer, M. Brooke, M. Brooke-Peat, D. Miles-Shenton, Bridging the domestic building fabric performance gap, *Build. Res. Inf.* (2014). doi:10.1080/09613218.2014.979093.
- [13] A. Marshall, R. Fitton, W. Swan, D. Farmer, D. Johnston, M. Benjaber, Y. Ji, Domestic building fabric performance: Closing the gap between the in situ measured and modelled performance, *Energy Build.* 150 (2017) 307–317. doi:10.1016/J.ENBUILD.2017.06.028.
- [14] J. Wingfield, M. Bell, D. Miles-Shenton, T. South, B. Lowe, Evaluating the impact of an enhanced energy performance standard on load-bearing masonry domestic construction - Understanding the gap between designed and real performance: lessons from Stamford Brook, London, 2011.

- [15] J. Wingfield, M. Bell, D. Miles-Shenton, J. Seavers, Elm Tree Mews Field Trial – Evaluation and Monitory of Dwellings Performance, Final Technical Report, 2011. Available from: [http://www.leedsmet.ac.uk/as/cebe/projects/elmtree/elmtree\\_finalreport.pdf](http://www.leedsmet.ac.uk/as/cebe/projects/elmtree/elmtree_finalreport.pdf) (accessed May 10, 2019).
- [16] S. Roels, Reliable building energy performance characterisation based on full scale dynamic measurements (Annex 58). Project Summary Report, 2017.
- [17] D. Johnston, D. Miles-Shenton, J. Wingfield, D. Farmer, M. Bell, Whole House Heat Loss Test Method (Coheating), 2013. Available from: [https://scholar.google.com/scholar\\_lookup?title=Whole House Heat Loss Test Method \(Coheating\)&author=D. Johnston&publication\\_year=2013](https://scholar.google.com/scholar_lookup?title=Whole House Heat Loss Test Method (Coheating)&author=D. Johnston&publication_year=2013) (accessed May 10, 2019).
- [18] G. Bauwens, In situ testing of a building's overall heat loss coefficient, PhD thesis KU Leuven, Belgium, 2015.
- [19] L. Ljung, System Identification: Theory for the User (2nd edition), PTR Prentice-Hall; Upper Saddle River, USA, 1999.
- [20] G. Pandraud, R. Fitton, QUB: Validation of a Rapid Energy Diagnosis Method for Buildings, in: 4th Expert Meet. IEA EBC Annex 58, 2013.
- [21] P. Boisson, R. Bouchié, ISABELE method : In situ assessment of the building envelope performances, in: Proc. Ninth Int. Conf. Syst. Simul. Build., 2014: pp. 302–320.
- [22] R. Bouchié, F. Alzetto, B. Adrien, B. Pierre, T. Simon, Short Methodologies for in-Situ Assessment of the Intrinsic Thermal Performance of the Building Envelope, Sustain. Places. (2014).
- [23] K. Subbarao, J.D. Burch, C.E. Hancock, A. Lekov, J.D. Balcomb, Short-Term Energy Monitoring (STEM ): Application of the PSTAR Method to a Residence in Fredericksburg, Virginia, 1988.
- [24] EU, Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency, Off. J. Eur. Union. L156 (2018) 75–91.
- [25] European Commission, Clean energy for all - The new Energy efficiency measures, Technical memo, 2016. Available from: [https://ec.europa.eu/energy/sites/ener/files/documents/technical\\_memo\\_energyefficiency.pdf](https://ec.europa.eu/energy/sites/ener/files/documents/technical_memo_energyefficiency.pdf) (accessed May 10, 2019).
- [26] D. Saelens, G. Reynders, IEA EBC Annex 58: Report of Subtask 4b: Towards a characterisation of buildings based on in situ testing and smart meter readings and potential for applications in smart grids., KU Leuven; Leuven, 2016.
- [27] Communities and Local Government, Code for Sustainable Homes: Technical Guide., 2010. Available from: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/5976/code\\_for\\_sustainable\\_homes\\_techguide.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/5976/code_for_sustainable_homes_techguide.pdf) (accessed June 1, 2018).
- [28] B. Sodagar, D. Starkey, The monitored performance of four social houses certified to the Code for Sustainable Homes Level 5, Energy Build. 110 (2016) 245–256. doi:10.1016/j.enbuild.2015.11.016.
- [29] ISO (International Organization for Standardization), ISO 9869-1: Thermal insulation: Building elements: In-situ measurement of thermal resistance and thermal transmittance: Part 1: Heat flow meter method, International Organization for Standardization; Geneva, Switzerland,

- 2014.
- [30] ISO (International Organization for Standardization), ISO 6781: Thermal insulation: Qualitative detection of thermal irregularities in building envelopes: Infrared method, International Organization for Standardization; Geneva, Switzerland, 1983.
- [31] The Air-tightness Testing and Measurement Association [ATTMA], TECHNICAL STANDARD 1. MEASURING AIR PERMEABILITY OF BUILDING ENVELOPES, 2007. Available from: [https://www.attma.org/wp-content/uploads/2013/10/ATTMA\\_TS1\\_Issue2\\_July07.pdf](https://www.attma.org/wp-content/uploads/2013/10/ATTMA_TS1_Issue2_July07.pdf) (accessed June 1, 2018).
- [32] B. Young, J. Hayton, Technical Papers supporting SAP 2009: Changes to the treatment of boilers (gas and oil) in SAP 2009, 2009.
- [33] K.A. Antonopoulos, E. Koronaki, Apparent and Effective Thermal Capacitance of Buildings, 23 (1998) 183–192. doi:10.1016/S0360-5442(97)00088-1.
- [34] ISO (International Organization for Standardization), ISO 52016-1: Energy performance of buildings: Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads: Part 1: Calculation procedures, International Organization for Standardization; Geneva, Switzerland, 2017.
- [35] H. Hens, Bouwfysica: warmte- en massatransport. Gebouwenteknik - deel 1, Acco; Leuven, 2013.
- [36] M. Senave, G. Reynders, S. Verbeke, D. Saelens, A simulation exercise to improve building energy performance characterization via on-board monitoring, Energy Procedia. 132 (2017) 969–974. doi:10.1016/j.egypro.2017.09.687.
- [37] National Grid UK, Calorific Value (CV): CV Data: Data Item Explorer, (n.d.). Available from: <http://mip-prod-web.azurewebsites.net/DataItemExplorer/Index> (accessed June 1, 2018).
- [38] D.F. Lander, Accuracy of CV Determination Systems for Calculation of FWACV. Appendix 5 of EMIB Review Report DLC/0030, 2012.
- [39] M. Senave, G. Reynders, B. Sodagar, D. Saelens, Uncertainty in Building Energy Performance Characterization : Impact of Gas Consumption Decomposition on Estimated Heat Loss Coefficient, in: Proc. 7th Int. Build. Phys. Conf., IBPC 2018, 2018: pp. 1491–1496.
- [40] G. Reynders, T. Nuytten, D. Saelens, Potential of structural thermal mass for demand-side management in dwellings, Build. Environ. 64 (2013) 187–199. doi:10.1016/j.buildenv.2013.03.010.
- [41] S.A. Klein, W.A. Beckman, J.W. Mitchell, J.A. et al. Duffie, TRNSYS 17: A Transient System Simulation Program: Volume 4: Mathematical Reference, Solar Energy Laboratory, University of Wisconsin; Madison, USA, 2014.
- [42] H. Madsen, P. Bacher, G. Bauwens, A.H. Deconinck, G. Reynders, S. Roels, E. Himpe, G. Lethé, IEA EBC Annex 58: Report of Subtask 3, part 2: Thermal performance characterisation using time series data – statistical guidelines, KU Leuven; Leuven, 2016.
- [43] R-core, lm | Fitting Linear Models, R Doc. (n.d.). Available from: <https://www.rdocumentation.org/packages/stats/versions/3.4.3/topics/lm> (accessed May 28, 2018).