

# Estimation of thermophysical properties from in-situ measurements in all seasons: Quantifying and reducing errors using dynamic grey-box methods



Virginia Gori\*, Clifford A. Elwell

*Physical Characterisation of Buildings Group, UCL Energy Institute, 14 Upper Woburn Place, London WC1H 0NN, UK*

## ARTICLE INFO

### Article history:

Received 23 October 2017

Revised 27 January 2018

Accepted 22 February 2018

Available online 3 March 2018

### Keywords:

U-value  
Heat transfer  
In-situ measurements  
Bayesian statistics  
Error quantification  
Grey-box methods  
Uncertainty analysis  
Inverse modelling

## ABSTRACT

Robust characterisation of the thermal performance of buildings from in-situ measurements requires error analysis to evaluate the certainty of estimates. A method for the quantification of systematic errors on the thermophysical properties of buildings obtained using dynamic grey-box methods is presented, and compared to error estimates from the average method. Different error propagation methods (accounting for equipment uncertainties) were introduced to reflect the different mathematical description of heat transfer in the static and dynamic approaches.

Thermophysical properties and their associated errors were investigated using two case studies monitored long term. The analysis showed that the dynamic method (and in particular a three thermal resistance and two thermal mass model) reduced the systematic error compared to the static method, even for periods of low internal-to-external average temperature difference. It was also shown that the use of a uniform error as suggested in the ISO 9869-1:2014 Standard would generally be misrepresentative. The study highlighted that dynamic methods for the analysis of in-situ measurements may provide robust characterisation of the thermophysical behaviour of buildings and extend their application beyond the winter season in temperate climates (e.g., for quality assurance and informed decision making purposes) in support of closing the performance gap.

© 2018 The Authors. Published by Elsevier B.V.  
This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

It is essential to understand the energy performance of the building stock to significantly reduce its associated energy demand and meet climate-change mitigation and carbon-emission targets [1]. In this regard, a number of software tools and calculation methods [2,3] have been developed to forecast and evaluate the energy consumption of buildings both at stock (e.g., for policy-making applications) and building scales (e.g., to issue energy performance certificates or evaluate retrofitting interventions). However, several studies [4–8] have shown a performance gap between simulation outputs and as-built thermal behaviour of buildings, identifying the thermophysical properties of the building envelope as one of the most important causes of uncertainties in energy performance models [9,10].

To address the current limited understanding of the as-built energy performance across the whole construction sector, the need

for novel (or improved) standardised diagnostic tests and robust analysis methods has been identified as a priority to ensure consistency and repeatability of outcomes [11–14]. Within this context, a wider use of in-situ measurements for the estimation of the thermophysical performance of the building fabric has been identified as key to tackle the problem [11]. However, stationary methods (particularly the average method [15]) are currently still the most common approaches adopted for the characterisation of the as-built thermophysical properties of building elements [16], effectively limiting an extensive use of in-situ measurements for large-scale applications. The steady-state assumptions underlying stationary methods imply time-independent thermophysical properties of building materials, constant boundary conditions, and neglect heat storage effects [17]. Since these conditions are very unlikely to be achieved on site, stationary methods require long monitoring periods and large temperature differences (preferably above 10 °C [18]) between the two sides of the element surveyed to minimise the error introduced by neglecting dynamic effects and provide useful estimates [15,19]. Dynamic methods (e.g., [20–24]) may be able to overcome some of these limitations by accounting for

\* Corresponding author.

E-mail address: [virginia.gori.12@ucl.ac.uk](mailto:virginia.gori.12@ucl.ac.uk) (V. Gori).

## Nomenclature

$\delta_{\varepsilon}^a$	Total absolute uncertainty for each data stream $\varepsilon$
$\delta_{\varepsilon}^r$	Total relative uncertainty for each data stream $\varepsilon$
$\Delta T$	Temperature difference between the internal and external surface [°C]
$\sigma_{\varepsilon}$	Systematic measurement error on each data stream
$C_n$	$n$ -th lumped thermal mass (starting from the internal side) [ $Jm^{-2}K^{-1}$ ]
$E_{m,\varepsilon}^p$	Measured observation for data stream $\varepsilon$ at time step $p$
$Q_{m,in}, Q_{m,out}$	Measured heat flux into and out of the internal and external surfaces [ $Wm^{-2}$ ]
$R_n$	$n$ -th lumped thermal resistance (starting from the internal side) [ $m^2KW^{-1}$ ]
$T_{C_n}^0$	Initial temperature of the $n$ -th lumped thermal mass [°C]
$T_{int}, T_{ext}$	Measured internal and external surface temperature [°C]
$U$	U-value [ $Wm^{-2}K^{-1}$ ]
$P(\theta H)$	Prior probability distribution of the parameters ( $\theta$ ) of model $H$
$P(\theta y, H)$	Posterior probability distribution of the parameters, given the observations ( $y$ ) and the model
$P(y \theta, H)$	Likelihood function
$P(y H)$	Evidence
$\theta$	Vector of the unknown parameters
1TM	Single thermal mass model
2TM	Two thermal mass model
MAP	Maximum a posteriori estimation

the dynamic fluctuations of the system instead of neglecting them [16], in addition to providing insights into the thermal mass of the element. Dynamic methods may enable an extension of the monitoring period to all times of the year and potentially a reduction of the monitoring length, depending on the specific conditions encountered and method used [25].

Robust characterisation of the thermophysical properties of building elements cannot disregard the quantification of the associated error, to provide necessary context and facilitate the interpretation of results, for example to understand the significance of differences between estimates, or the range of potential pay-back times for interventions. Error estimates themselves may vary according to the context in which they are applied and interpreted. For example, the purpose of a measurement may be to compare the U-values at different points on a building element at a specific time or to compare the estimated U-values to a “true” value. In the former case, seasonal variations in U-values may be neglected, whilst the second case should address changes in parameters such as the moisture content of the element and variation in the radiative and convective heat transfer in surface resistances; both should consider error associated with the equipment (e.g., calibration and its impact on heat flow) and consider uncertainty in the ability of the applied model to represent the concept of a U-value. The causes and enumeration of error and uncertainties in the estimation of the thermophysical properties of buildings is discussed further in Section 2.

Mathematically, identified errors in the model inputs (e.g., on in-situ measurements in this case) must be propagated to its outputs (e.g., the estimated R-values and U-value). Although error propagation is straightforward for static methods, this may be more complex for dynamic methods where the parameters are

often estimated by means of optimisation techniques. This paper presents a method for the quantification of systematic errors (i.e. errors caused by biases in the system that affect all observations in a consistent manner, following a constant or fixed pattern) affecting the estimates of grey-box dynamic methods. The error propagation method is then contextualised to the dynamic grey-box method described in [22,25,26], and tested on two in-situ walls using measurements collected at different times of the year. The average method is also applied to the same data to compare and contrast the performance of the two approaches across the seasons.

## 2. Uncertainties in the estimation of thermophysical properties from in-situ measurements

The thermophysical properties of in-situ building elements are generally estimated from non-destructive measurements of the heat flowing through the structure and of the temperature (either of the surface or the air) on either side of it [27,28]. Depending on the purpose of the investigation, sensors are mounted at representative locations ideally using thermal imaging to inform the process [18]. Good thermal contact between the equipment (i.e. heat flux meters and temperature sensors) and the structure must be ensured, avoiding the formation of air pockets. The case study and experimental methods are discussed in detail in Section 4, whilst the methods of data analysis are reviewed in Section 3.1.

Once the data are collected, their analysis (whether using a steady-state or a dynamic method) should not disregard error analysis to identify all the quantifiable uncertainties affecting the observations (i.e. the model's input) and estimate how these combine and propagate to the model estimates, as discussed above. An overview of the main sources of uncertainties is provided below according to the definitions of the international vocabulary of metrology [29], contextualised to in-situ measurements and estimates of the thermophysical properties of building elements. This is followed by a description of the error analysis scheme suggested in the ISO 9869-1:2014 Standard (standardised in the UK in [15]) and a discussion of the need for different error propagation approaches depending on the mathematical model adopted.

### 2.1. Monitoring and modelling uncertainties

Several uncertainties affect estimates of the as-built thermophysical performance of buildings and potentially contribute to the observed discrepancies between thermophysical parameter estimation from published material properties and in-situ measurements. The international vocabulary of metrology defines *measurement uncertainty* as the “dispersion of the quantity values being attributed to a measurand,<sup>1</sup> based on the information used” [29, Section 2.26] and usually includes *definitional uncertainties* and *measurement errors*. Definitional uncertainties are a consequence of the finite amount of detail in the definition of the quantity of interest [29, Section 2.27], and practically define the minimum uncertainty achievable even in case no measurement errors have been introduced by the monitoring process. As a result, there is no unique true value of the measurand but rather a range of true values that are equally consistent with its definition [30, p.87]. For example, in the context of thermophysical characterisation of buildings, definitional uncertainties occur when the measurand is defined as the U-value of a building element without specifying the exact location within the element.

Measurement errors describe the differences between the magnitude of the signal monitored by a sensor and a hypothetical reference quantity value of the phenomenon being surveyed [29, Sec-

<sup>1</sup> The measurand is the “quantity intended to be measured” [29, Section 2.3].

tion 2.16]. These are introduced for example by the presence and intrusiveness of the equipment compared to the undisturbed configuration, accuracy (including calibration and precision errors) and state of repair of the sensors deployed, installation and placement strategy. Some measurement errors can be mitigated by adopting good-practice precautions. For example, performing preliminary surveys of the case study to assess the placement of sensors, or selecting appropriate monitoring set up and sampling intervals in relation to the time period over which the physical process investigated is expected to occur and to the response time of the equipment used [27,28]. This is crucial to avoid the introduction of artefacts in the recorded signal (e.g., aliasing) or failing to capture useful information [31]. Although the selection of the sampling interval depends on several case-specific aspects (including the expected environmental conditions, the time constant of the structure, the specifications of the equipment, the method later adopted for data analysis), sampling intervals between five and sixty minutes are generally considered appropriate [28,31].

Additional uncertainties are introduced when the measurand is obtained indirectly from other measured quantities through a *measurement model* [29, Section 2.48]. These uncertainties arise both from the process undertaken to abstract the reality into its model representation required for parameter inference, and from the incremental approximations and assumptions the modeller may introduce during the modelling process (e.g., when representing the reality by means of mathematical descriptions, or when discretising continuous processes on digital computers) [32,33]. Modelling simplifications are generally introduced by neglecting in the model features that are difficult or practically impossible to foresee and/or to account for (e.g., structural and situational inhomogeneities, defects), and by compromising between computational efficiency and adequacy of the complexity of the model in relation to the purpose of describing the physical phenomenon.

Errors and uncertainties can further be categorised as *random* or *systematic*. Random errors [29, Section 2.19] represent fluctuations of a quantity of interest around the value that would be obtained by averaging the outcome of an infinite number of repeated observations collected under consistent conditions [34]. Random errors are inversely related to the *precision* [29, Section 2.15] of the measurement process. In case random errors can be considered independent of each other (*i.e.* the occurrence of a specific error does not imply an increase or decrease in others), the precision of the combined result is improved by increasing the number of observations. Conversely, systematic errors are introduced by bias in the system that affects repeated observations in a consistent manner, following a fixed or predictable pattern [29, Section 2.17]. Unlike random errors, systematic errors are not minimised by increasing the number of observations and are inversely related to measurement *trueness* [29, Section 2.14], which can be improved by an attentive experimental procedure (e.g., use of calibrated sensors, or care with placement and fixation method) or an improvement of the model specification (e.g., modelling of additional physical effects, or use of additional data streams). Measurement errors are a combination of random and systematic errors, and are inversely related to the measurement *accuracy* [29, Section 2.13].

Within the context of the estimation of the thermophysical characteristics of building elements from in-situ measurements, four types of errors can be defined from the combination of those listed above. Specifically, systematic measurement errors are generated by offsets in the monitoring equipment (e.g., due to sensor drift, or erroneous experimental set up); random measurement errors are generally introduced by noise in the system (e.g., grounding and shielding noise in the equipment); systematic modelling errors occur when adopting an unrepresentative or incomplete model (e.g., due to the use of a limited number of parameters or data streams, or to assumptions made during model specifica-

tion such as the one-dimensional or the steady-state assumption); finally, random modelling errors (often referred to as “statistical errors” [26,34, Ch.4.2]) may be introduced by the use of digital computers and can be generated by rounding effects, discretisation of continuous time series, tolerance in the termination criteria of iterative algorithms.

## 2.2. Error analysis in ISO 9869-1:2014 Standard

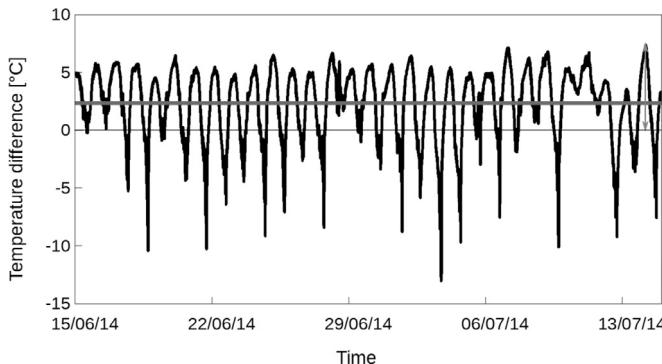
Besides methods for U-value estimation from in-situ measurements, the ISO 9869-1:2014 Standard [15, p.12] provides a guideline for the calculation of the systematic error affecting estimates. Specifically, it lists the main uncertainties affecting the measurements and quantifies their proportional effect on the U-value. The uncertainties cover: the accuracy of heat flux plates (HFPs), temperature sensors, and data logging system(s); variations due to uneven thermal contact between the sensors and the surface; an operational error of the HFPs caused by their presence; variations over time of temperatures and heat flow; temperature variation within the space and between radiant and air temperature. Although these are useful references, the Standard does not detail or refer to how these percentages (generally reported as a fixed value) were evaluated, nor the methods for the quantification of more appropriate values in specific circumstances. As an example, it is stated that “the accuracy of the measurement depends on errors caused by the variations over time of the temperatures and heat flow” and that “such errors can be very large but, if the criteria described in 7.1 and 7.2 or Annex C are fulfilled, they can be reduced to less than  $\pm 10\%$  of the measured value” [15, p.13]. However, a procedure to reduce the error when the criteria are met (which should represent the majority of cases if the Standard is applied correctly) is not provided nor is addressed the variation in error according to the magnitude of the temperature difference and heat flow (Section 2.3).

Combining the uncertainties stated and provided that the conditions listed in the guideline are met, the Standard suggests that the total error on the U-value is expected between 14% (combining the uncertainties in quadrature sum, *i.e.* considering them independent) and 28% (arithmetic sum of uncertainties, representing correlations between errors). However, adopting the percentages suggested may result in an underestimation of the systematic error affecting the estimates. Specifically, the standard lists the “accuracy of the calibration of the HFP and the temperature sensors” among the uncertainties suggesting that “the error is about 5% if these instruments are well calibrated” [15, p.12]. However, it does not provide guidelines to determine whether the 5% uncertainty suggested is representative of the case study nor a method for its quantification based on sensor specifications.<sup>2</sup> Similarly, for the data logging system the Standard only provides a procedure for the selection of an appropriate device based on the ratio between the measurement minimum output value and the density of heat flow rate [15, Annex E], but it does not warn the user on the need to account for the accuracy of the logger based on manufacturer specifications.

## 2.3. Systematic errors for static and dynamic methods in conditions with low average temperature difference

Since the average and the dynamic method adopted in this paper for the estimation of the thermophysical properties of the

<sup>2</sup> For example, assuming the HFP has an accuracy of 5% (in line with commercial HFP sensors, e.g., [35]) and temperature sensors of 0.1 °C, the relative systematic error on the U-value for average temperature differences between 5 °C and 10 °C is in the range [5.1, 5.4]%. However, if the accuracy of the temperature sensors is 0.5 °C (in line with commercial products often used for in-situ measurements in buildings) and leaving all other parameters unchanged, the relative systematic error increases to the range [7, 11]%.



**Fig. 1.** Measured temperature difference at each sampling interval (black line) and average temperature difference over the monitoring period (grey line). The arrow shows the amplitude of the temperature difference at a given sampling interval.

building fabric are based on a different mathematical model to describe the heat transfer, the error estimates are different in the two cases. In particular, the mathematical formulation of the average method (AM) introduces a fundamental limitation to the applicability of this method when the average temperature difference (and consequently the heat flux) is close to zero (e.g., when the heat flux reverses over the monitoring period) since the temperature difference appears at the denominator of the U-value equation (see Eq. (1), Section 3.1.1). This limitation affects both the U-value estimate and its associated systematic error. Conversely, this is not the case for the dynamic method where the parameters are estimated from the comparison of the predicted and measured time series (e.g., the heat flux in this work) at each sampling interval. This issue is illustrated in Fig. 1, where the average internal-to-external temperature difference is low (so would return a high error using static methods) but the temperature difference at each sampling interval is considerably higher, returning a lower systematic error. Given the limitation imposed on the AM by small temperature differences, the application of this method restricts the conditions in which measurements may be analysed yet return moderate error compared to the dynamic method, as the average temperature difference over the whole monitoring period is generally lower than the temperature difference at each time step.

Although the propagation of systematic errors is easily quantifiable for the AM based on the U-value definition (Section 3.2.1), this is not the case for dynamic methods (like the grey-box method adopted in this paper) where the thermophysical parameters are estimated by means of optimisation techniques. A method for the estimation of the systematic errors on the estimates of dynamic methods is presented in Section 3.2.2.

### 3. Thermophysical properties and systematic measurement error estimation from in-situ measurements

The average method [15] and the grey-box dynamic method described in [22,25,26] were used to evaluate the thermophysical properties of in-situ building elements (Section 3.1). The systematic measurement error associated with the parameters estimated with the two frameworks were also evaluated (Section 3.2). Different approaches were used for the propagation of systematic uncertainties to reflect the different mathematical formulation of the two methods, as discussed in Section 3.2; the two methods of U-value estimation are summarised in Section 3.1.

#### 3.1. Estimation of thermophysical properties

##### 3.1.1. Average method

Among static approaches, the average method is one of the most commonly adopted to analyse in-situ measurements and

evaluate the thermophysical properties of building elements [16]. According to the AM, the thermal transmittance (U-value) of a building element is defined as the ratio of the mean integral heat flow rate density and the mean integral temperature difference collected over a sufficiently long period of time [15, p.6]:

$$U = \frac{\frac{\tau}{n} \sum_{p=1}^n Q_m^p}{\frac{\tau}{n} \sum_{p=1}^n (T_{int}^p - T_{ext}^p)} = \frac{\sum_{p=1}^n Q_m^p}{\sum_{p=1}^n (T_{int}^p - T_{ext}^p)} \quad (1)$$

where  $Q_m$  is the measured heat flow rate density (usually taken on the interior side of the element investigated, i.e.  $Q_m = Q_{m,in}$  in Fig. 2) at each time step  $p$ ;  $\tau$  is the duration of the time step between successive observations (i.e. the recording interval for the measured quantities);  $n$  is the number of observations; and  $T_{int}^p, T_{ext}^p$  are the internal and external temperatures at each time step.

#### 3.1.2. Grey-box dynamic method

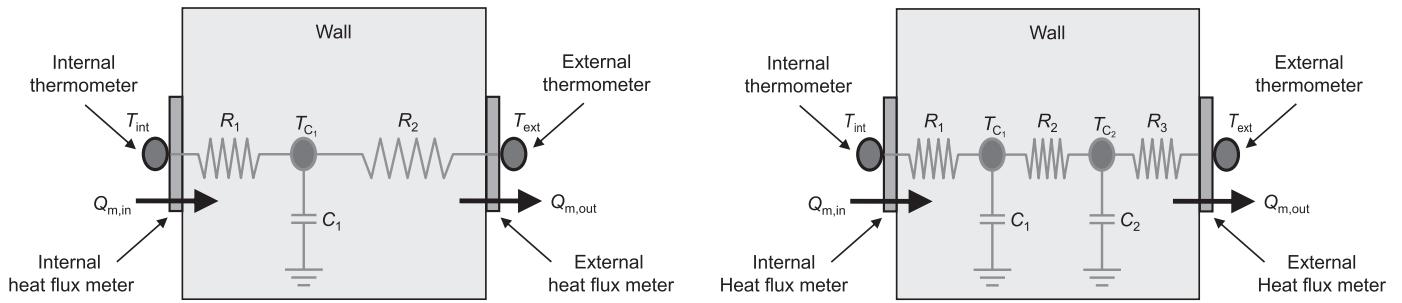
A grey-box dynamic method based on Bayesian statistics was adopted to estimate the thermophysical properties of the building element investigated (i.e. walls in this paper). The method (introduced and described in detail in [22,25,26]) combines lumped-thermal-mass models to simulate the heat transfer through a building element, and a Bayesian framework to estimate the set of parameters that best reproduce the monitored data (i.e. the heat flux into and out of the wall in this research).

The two simplest equivalent electrical circuits of one-dimensional heat flow incorporating thermal mass effects (Fig. 2) are applied here, consisting of one lumped thermal mass with two lumped thermal resistances (one thermal mass model, 1TM) and two lumped thermal masses with three lumped thermal resistances (two thermal mass model, 2TM) [26]. More complex models of heat flow through the element are possible, such a three and four thermal mass models [25]; the simpler models are presented here to focus the discussion on error analysis. For the 1TM model, the parameters were estimated both optimising the heat flux measured on the internal side of the wall only (1 HF), and the internal and external heat fluxes simultaneously (2 HF) [26]. The latter configuration was also used to optimise the parameters of the 2TM model.

Applying a Bayesian approach, the best-fit parameters ( $\theta_{MAP}$ ) were estimated using the maximum a posteriori (MAP) approach. Specifically, these were estimated by maximising their posterior probability distribution (i.e. the probability of the vector of parameters,  $\theta$ , given the measured data,  $y$ , and a model,  $H$ , of the underlying physical process of interest):

$$\theta_{MAP} = \arg \max_{\theta} P(\theta|y, H) = \arg \max_{\theta} \frac{P(y|\theta, H)P(\theta|H)}{P(y|H)}, \quad (2)$$

where  $P(y|\theta, H)$  is the likelihood function;  $P(\theta|H)$  is the prior probability distribution of the parameters;  $P(y|H)$  is the evidence (or marginal likelihood). The likelihood describes the ability of the model to explain the measurements, the prior represents the initial estimated probability distribution of each parameter based on expert knowledge before observing any data, and the evidence is a normalisation factor. The method used in this paper follows that in [26], but with an improved formulation for the likelihood function as in [25]. Owing to the Bayesian framework, the widely used assumption of independent and identically distributed (i.i.d.) residuals (i.e. the difference between the measured and modelled time series) previously made in [26] was obviated by introducing a prior distribution for the residuals that accounts for their potential auto-correlation [25]. This prior was chosen such that its scale parameter coincided with the variance of the noise term computed from all the known and quantifiable systematic uncertainties affecting the data streams optimised (i.e. the heat flow rate density in this paper). Consequently, no tests on the residuals (e.g., analysis of the



**Fig. 2.** Schematic of the one (left) and two thermal mass (right) models showing the equivalent electrical circuit for heat transfer simulation. Parameters of the models are the thermal resistances ( $R_1, R_2, R_3$ ), the effective thermal masses ( $C_1, C_2$ ), and their initial temperatures ( $T_{C_1}^0, T_{C_2}^0$ ). The measured quantities are the internal ( $T_{int}$ ) and external ( $T_{ext}$ ) temperatures, and the heat flow rate density entering the internal ( $Q_{m,in}$ ) and leaving the external ( $Q_{m,out}$ ) surfaces.

autocorrelation function or the cumulated periodogram [36]) are required.

Bayesian model comparison (described in [26]) was undertaken to select the model among the several devised that is most likely to describe the underlying physical process. To be a fair comparison of models, the same input data is used: the 1TM (2 HF) and 2TM models are therefore compared in this work.

### 3.2. Estimation of systematic measurement error on thermophysical properties

Different approaches for the propagation of the systematic measurement error on the thermophysical estimates were implemented for the two methods to reflect their different mathematical modelling. Although the propagation of systematic errors is easily quantifiable for the AM based on the U-value definition (Section 3.2.1), this is not the case for dynamic methods (like the grey-box method adopted here) where the thermophysical parameters are estimated by means of optimisation techniques. A method for the quantification of the systematic errors on the estimates of dynamic methods is presented below (Section 3.2.2).

#### 3.2.1. Error propagation for the average method

The systematic error affecting the U-value estimates obtained with the AM was quantified using a linear error propagation by means of a first-order Taylor expansion [34, Ch.4.3] of the U-value definition (Eq. (1)):

$$\begin{aligned} dU &= d\left(\frac{Q_m}{\Delta T}\right) = \frac{\partial U}{\partial Q_m} dQ_m + \frac{\partial U}{\partial \Delta T} d\Delta T \\ &= \frac{1}{\Delta T} dQ_m - \frac{Q_m}{\Delta T^2} d\Delta T \end{aligned} \quad (3)$$

where  $Q_m$  and  $\Delta T$  are respectively the measured heat flux (usually the heat flux into the internal surface) and the difference between the internal and external temperatures;  $dQ_m$  and  $d\Delta T$  are their differentials. Applying the propagation of error formulas [37] to Eq. (3) and assuming that the systematic measurement errors on the heat flux observations ( $\sigma_{Q_m}$ ) and the temperature data streams ( $\sigma_{T,\varepsilon}$ ) are independent, the relative systematic error on the U-value can be computed as:

$$\frac{\sigma_U}{U} = \sqrt{\frac{\sigma_{Q_m}^2}{Q_m^2} + \frac{\sigma_{T,\varepsilon}^2}{\Delta T^2}} = \sqrt{\frac{\sigma_{Q_m}^2}{Q_m^2} + \frac{\sigma_{T,\varepsilon}^2}{(T_{int} - T_{ext})^2}}. \quad (4)$$

#### 3.2.2. Error propagation for the dynamic method

Propagation of error formulas [37] (like those adopted in Section 3.2.1) cannot be directly applied for methods estimating the parameters of interest by means of optimisation techniques, as in this case the parameters are estimated by minimising a given

cost function (e.g., Eq. (2)) instead of being explicitly calculated from a formula. The systematic measurement error affecting the thermophysical parameters estimated with optimisation methods can be quantified from the analysis of the global optimum of the unnormalised posterior probability distribution [25]. Defining  $\ell(y, \theta)$  as the log-posterior of the parameters given all observations (i.e. both heat flux(es) and temperature data streams), its gradient with respect to the parameters has to be zero at the maximum of the function (i.e. the MAP):

$$\frac{d\ell(y, \theta)}{d\theta} \Big|_{y, \theta_{MAP}(y)} = 0. \quad (5)$$

Since  $\theta_{MAP}$  depends on the observations and Eq. (5) describes a stationary point, using the chain rule<sup>3</sup> the derivative of Eq. (5) with respect to the data is:

$$\frac{\partial^2 \ell(y, \theta)}{\partial \theta \partial y} \Big|_{y, \theta_{MAP}(y)} + \frac{\partial^2 \ell(y, \theta)}{\partial \theta^2} \Big|_{y, \theta_{MAP}(y)} \frac{d\theta_{MAP}}{dy} = 0. \quad (6)$$

Therefore, the dependency of the MAP from the observations can be calculated from Eq. (6) as:

$$\frac{d\theta_{MAP}}{dy} = \left( -\frac{\partial^2 \ell(y, \theta)}{\partial \theta^2} \Big|_{y, \theta_{MAP}(y)} \right)^{-1} \frac{\partial^2 \ell(y, \theta)}{\partial \theta \partial y} \Big|_{y, \theta_{MAP}(y)} \quad (7)$$

where  $-\frac{\partial^2 \ell(y, \theta)}{\partial \theta^2} \Big|_{y, \theta_{MAP}(y)}$  is the Hessian of the minus logarithm of the posterior probability distribution and its inverse coincides with the covariance matrix under the Laplace approximation;  $\frac{\partial^2 \ell(y, \theta)}{\partial \theta \partial y}$  is a matrix whose elements  $i, j$  contain the derivative of the log-posterior with respect to the  $i$ -th parameter and to perturbations of the  $j$ -th data stream (this term can be easily computed using finite differences).

Since the variations of the total R-value ( $R_{tot}$ ) are simply the sum of the variations of the R parameters contributing to it, this can be formally expressed as:

$$\frac{dR_{tot,MAP}}{dy} = \nabla_f^T \frac{d\theta_{MAP}}{dy}. \quad (8)$$

As the U-value is the inverse of the total R-value, its variations can be calculated according to the formulas for the error propagation of a ratio [34, Ch.4.3]:

$$\frac{dU_{MAP}}{dy} = -\frac{1}{R_{tot,MAP}^2} \frac{dR_{tot,MAP}}{dy} \quad (9)$$

<sup>3</sup> The chain rule for a composition of functions  $f(x, g(x))$  states that:  $\frac{df(x,g(x))}{dx} = \frac{\partial f(x,w)}{\partial x} \Big|_{x,g(x)} + \frac{\partial f(x,w)}{\partial w} g'(x)$ .

The absolute systematic error on the U-value is represented by the quadrature sum of the uncertainties on each data stream, assuming that these are independent:

$$\sigma_U = \sqrt{\sum_{\varepsilon} \left( \frac{dU_{MAP}}{dy_{\varepsilon}} \sigma_{\varepsilon} \right)^2} \quad (10)$$

where  $\sigma_{\varepsilon}$  is the systematic measurement error on each data stream, which comprises both the error on the heat flux ( $\sigma_{Q,\varepsilon}$ ) and temperature ( $\sigma_{T,\varepsilon}$ ) measurements.

The random modelling (or statistical [26,34]) error on the parameter estimates is quantified from the covariance matrix obtained during the Bayesian inference [26]. A first-order Taylor series expansion is applied to calculate the modelling error on the U-value given the thermal resistances contributing to it.

## 4. Experimental method and analysis

### 4.1. Case studies

Two in-situ walls of different construction (i.e. one solid and one full-fill cavity wall) were monitored long term and used as case studies to investigate the effects of seasonal and temperature variations on the estimation of thermophysical properties and the associated systematic errors. A description of the monitoring campaigns is provided below.

#### 4.1.1. Solid wall in an office building (OWall)

The OWall case study was a traditional north-west-facing solid wall located on the first floor above ground of an occupied office building in London (UK). From the outside, the wall is made of a layer of exposed brick  $350 \pm 5$  mm and a layer of plaster  $20 \pm 5$  mm expected to be lime, for a total thickness of  $370 \pm 7$  mm. The wall was instrumented with a pair of heat flux plates [35] and type-T thermocouples, placed in-line with each other on opposite sides of the wall [26]. The internal HFP was secured to the wall using a layer of low-tack tape on the wall-facing side of the sensor followed by a layer of double-sided tape [38], while for the external HFP a thin layer of water-resistant elastomeric polymer was used on the edges (only) and a layer of heat compound on the remaining area. The thermocouples were taped on the guard ring of each HFP, using thermal paste on the hot junction to ensure good thermal contact with the wall [38]. Measurements were sampled every 5 s and averaged over 5-min intervals using a Campbell Scientific CR1000 [39] data logger. The OWall was monitored for a full year, from the 2nd of November 2013 to the 1st of December 2014 [40].

#### 4.1.2. Cavity wall in an unoccupied residential building (UHWall)

The UHWall case study was a 1970s north-facing filled cavity wall located at the ground floor of an unoccupied residential building in Cambridgeshire (UK). The wall is  $275 \pm 10$  mm thick and consists of four layers. From the exterior,  $100 \pm 5$  mm of exposed bricks are followed by a  $65 \pm 5$  mm cavity likely to be filled with urea formaldehyde foam,  $100 \pm 5$  mm aerated concrete blocks, and  $10 \pm 5$  mm plaster. According to visual inspection (also on a neighbouring property of same structure and period of construction) and to literature on the performance of urea formaldehyde foam [41], the insulation layer is expected to have shrunk inside the wall cavity and the thermal resistance of the wall is expected to have decreased accordingly.

A pair of HFPs [35] and thermistors was placed in-line with each other on opposite sides of the wall. A thin layer of silicon-free heat compound was used under each sensor to ensure good thermal contact with the wall. Indoor sensors were secured using masking tape (only on the guard ring for the HFP), while a thin

layer of silicon sealant was applied on the edges of the external equipment. Data were sampled every 5 s and averaged over 5-min intervals using Eltek 451/L and 851/L data loggers [42]. The UHWall was monitored from the 12th of March 2015 to the 30th of August 2015 [43].

### 4.2. Experimental analysis

A number of quantities have to be pre-computed to initialise the Bayesian analysis, including the quantification of the systematic measurement error on each data stream and the prior probability distribution on the parameters of the dynamic model. The data analysis framework is also described below.

#### 4.2.1. Systematic measurement error on the data streams

To estimate the systematic error affecting the thermophysical properties of the element under study (Sections 3.2.1 and 3.2.2), all the known and quantifiable systematic uncertainties affecting both the heat flux and temperature measurements have to be considered. Depending on the nature of the uncertainties, these can be classified as relative (if proportional to the magnitude of the observation) or absolute. Assuming that the uncertainties affecting the observations can be considered independent, the systematic measurement error on each data stream can be calculated as the quadrature sum of the individual relative and absolute uncertainties:

$$\sigma_{\varepsilon} = \sqrt{(\delta_{\varepsilon}^a)^2 + \left( \frac{\sum_{p=1}^n |E_{m,\varepsilon}^p|}{n} \delta_{\varepsilon}^r \right)^2} \quad (11)$$

where  $\delta_{\varepsilon}^a$  is the total absolute uncertainty for each data stream;  $\delta_{\varepsilon}^r$  is the total relative uncertainty for each data stream;  $E_{m,\varepsilon}^p$  are the measured observations for each data stream at each time step;  $n$  is the number of observations analysed.

In this paper, the systematic measurement error on the temperature data streams was calculated combining the accuracy of the temperature sensors and the data logging system in quadrature sum according to Eq. (11). Similarly, the following uncertainties were considered to calculate the systematic measurement error on each heat flux data stream: (a) the accuracy of the equipment (i.e. HFP and data logging system(s) involved in the analysis, according to manufacturers' specifications); (b) the effect of random variations caused by imperfect thermal contact between the sensor and the wall (5% according to [15, p.13]); (c) an uncertainty due to the modification of the isotherms caused by the presence of the HFP (3% according to [15, p.13]). Owing to the use of surface temperatures, the calculation omitted the 5% uncertainty in [15, p.13] to account for differences between air and radiant temperature, and temperature variations within the space [26]. For the AM, an extra 10% was added in quadrature sum to account for errors caused by the variations over time of the temperatures and heat flow, as suggested in [15, p.13].

#### 4.2.2. Priors on the parameters of the model

The use of prior probability distributions enables the coupling of information extracted from measurements with tabulated values; utilising this expert information enhances the robustness of the estimates and potentially reduces the monitoring time and costs. In general, either uniform or non-uniform priors (e.g., log-normal distributions as in [25]) can be defined depending on the amount of information available on the parameters of the problem. Uniform priors were used in this paper since the distributions of the thermophysical properties for some of the materials constituting the cavity wall (e.g., aerated concrete blocks) were not readily available in the literature. It is hoped that stochastic data bases such as the one developed by Zhao and colleagues [32] will be readily available for all building materials in the future as these would

be a useful resource for the development of probabilistic methods for the assessment of the hygrothermal behaviour of buildings and building components.

For both case studies large priors were defined to encompass all expected values for the thermophysical parameters, with significant safety margin. These ranged between  $[0.01, 4.00] \text{ m}^2\text{KW}^{-1}$  for all thermal resistances, between  $[0.1, 2 \cdot 10^6] \text{ Jm}^{-2}\text{K}^{-1}$  for all effective thermal masses, and in  $[-5, 40] \text{ }^\circ\text{C}$  for their initial temperature.

#### 4.2.3. Hypothetical monitoring campaigns

During experimental analysis, it is good practice to ensure that the length of the time series used for parameter estimation is appropriate, and consequently the estimates obtained are representative of the performance of the building element surveyed. Short monitoring campaigns are preferable for practical reasons (such as minimising the inconvenience to the occupants and costs) to expand the use of in-situ measurements for the characterisation of buildings in practice. Consequently, it is important to determine the minimum number of observations that return a robust estimation of the thermophysical parameters of the building element investigated. This requirement arises from the contrasting need for a time series that is sufficiently long to ensure that the estimates are accurate and have small variability, but that is not too long to ensure that the assumption of a unique model to explain the data over the monitoring period (e.g., constant parameters) holds. At the beginning of the monitored time series, when little data is available, the estimates are noisy and prone to overfitting. Subsequently, with the supplement of new observations, the estimates improve until the addition of new data does not enhance the prediction of the parameters significantly and the values stabilise around a final value. This concept is referred to as "stabilisation" in this paper.

A number of stabilisation criteria are listed in Section 7.1 of [15, p.9] to determine the minimum length of the time series analysed while ensuring that: (a) the steady-state assumption at the basis of the AM holds for the period investigated, and (b) the estimates have converged to an asymptotic value. No standardised criteria are available (to the authors' knowledge) to determine the minimum length of the time series to be analysed with a dynamic method. Therefore, in this work the criteria in [15, p.9] were also imposed for the dynamic analysis, although these may be too conservative in this case (as shown by the evolution of the U-value over time in [25, Ch.5]).

To test and compare the performance and robustness of the average and dynamic method (both in terms of U-value estimations and associated systematic measurement errors) at different times of the year, shorter time series were extracted from the two long-term monitoring campaigns. Each shorter time series was started seven days apart and lasted until the stabilisation criteria [15, p.9] were met using the AM.<sup>4</sup> The time series so obtained were adopted with all data analysis methods (*i.e.* average and dynamic). This approach, referred to as "hypothetical monitoring campaigns" in this paper, effectively synthesises a large number of repeated measurements at different times of the year to test the performance of the methods when the buildings were exposed to different environmental conditions.

## 5. Results and discussion

The OWall and UHWall case studies were used to test and compare the performance of the average and dynamic methods at different times of the year, both in terms of U-value estimates and

associated systematic measurement error. The modelling errors on the U-values are not reported below as these were always substantially smaller than the measurement errors (at least an order of magnitude lower); the total error is dominated by the latter. Initially, the time series were checked and cross-referenced with the metadata to exclude periods where problems in the data collection were identified or data were repeatedly missing. Hypothetical monitoring campaigns that: (a) had not met the stabilisation criteria (described in Section 4.2.3) before one of these periods, or (b) had not met the stabilisation criteria within 30 days were excluded from the analysis. The 30-day threshold on the length of the time series was selected to both reflect practical timescales to complete such measurements and to ensure that the assumption of constant model parameters is reasonable. For a longer monitoring campaign the likelihood that this latter assumption holds reduces since as the length of the monitoring campaign rises, the risk of changes to parameter values increases (e.g., due to changes in the environmental conditions the building element is exposed to during the survey, such as variations in moisture content, moisture penetration depth of wind-driven rain, wind patterns). A plausible range of U-values was also determined on the basis of each wall structure using thermophysical properties in the literature.

### 5.1. Literature U-values

A range of possible U-values was defined for each case study due to the lack of specific information about the thermophysical properties of its materials. The U-value ranges were determined combining the upper and lower values of tabulated thermal conductivity for the materials constituting the layers and their known thickness, and adding constant internal ( $0.13 \text{ m}^2\text{KW}^{-1}$ ) and external ( $0.04 \text{ m}^2\text{KW}^{-1}$ ) air film resistances [44].

Following the procedure above, the literature U-value for the OWall is between 1.11 and  $2.16 \text{ Wm}^{-2}\text{K}^{-1}$ , as the thermal conductivity of solid brick is expected to lie in the range 0.50 to  $1.31 \text{ mKw}^{-1}$  and that of lime plaster between 0.70 and  $0.80 \text{ mKw}^{-1}$  [45]. The brick layer was considered homogeneous in the calculation (*i.e.* mortar joints were not accounted for separately) as the range of thermal conductivity for mortar falls within the values for solid brick. Similarly, the U-value for the UHWall ranged between 0.32 and  $0.40 \text{ Wm}^{-2}\text{K}^{-1}$ , using a thermal conductivity between: 0.22 and  $0.81 \text{ mKw}^{-1}$  for plaster, 0.15 and  $0.24 \text{ mKw}^{-1}$  for aerated concrete blocks, 0.031 and  $0.035 \text{ mKw}^{-1}$  for urea formaldehyde foam, and 0.50 and  $1.31 \text{ mKw}^{-1}$  for the outer-leaf brick work.

### 5.2. In-situ U-value estimates for the solid wall

From the long-term time series collected on the OWall, two weeks of data (between the 16th and the 30th of May 2014) were excluded from the analysis due to repeated missing data for the internal thermocouple. Of the remaining data, fifty-two hypothetical monitoring campaigns were analysed after removing the periods where the acceptance criteria in Section 5 were not met. The length of each hypothetical monitoring campaign (*i.e.* the number of days required for the AM to stabilise, Section 4.2.3) ranged between three and thirty days. Time series of up to ten days were needed in the autumn and winter period, whilst the required length increased in warmer seasons. Note that three days is the minimum length imposed by the ISO Standard [15, p.9], however [25, Ch.5] showed that the evolution of the U-value over time may be stable in a much shorter time span with the dynamic method.

The U-value estimates from the AM and dynamic method (using the 1TM (1 HF), 1TM (2 HF) and the 2TM models) were within the margin of the absolute systematic measurement error for all hypothetical monitoring campaigns. The U-value estimates were

<sup>4</sup> It might be possible that the time series used for consecutive hypothetical monitoring campaigns partly overlap.

**Table 1**

Minimum, maximum, mean and standard deviation of the U-value estimates and the associated relative systematic measurement error for the OWall over the hypothetical monitoring campaigns, estimated with the average and the dynamic method. The range of U-values from tabulated thermophysical properties is also reported.

		Min	Max	Mean	St dev	Units
U-value	Literature	1.11	2.16	–	–	$\text{Wm}^{-2}\text{K}^{-1}$
	AM	1.28	1.92	1.71	0.14	$\text{Wm}^{-2}\text{K}^{-1}$
	1TM (1 HF)	1.34	1.88	1.71	0.13	$\text{Wm}^{-2}\text{K}^{-1}$
	1TM (2 HF)	1.61	2.00	1.77	0.11	$\text{Wm}^{-2}\text{K}^{-1}$
	2TM	1.60	1.85	1.72	0.07	$\text{Wm}^{-2}\text{K}^{-1}$
Rel sys err	AM	14%	50%	22%	8%	–
	1TM (1 HF)	10%	31%	17%	6%	–
	1TM (2 HF)	9%	51%	17%	8%	–
	2TM	9%	27%	16%	5%	–

also within the range of values expected from the literature for a wall similar to the OWall ([Table 1](#)). The minimum, maximum, mean and standard deviation of the U-value estimates for each model are summarised in [Table 1](#). The mean U-value for the AM, 1TM (1 HF) and the 2TM models virtually coincided, while it was slightly higher (about 4% increase) for the 1TM (2 HF) model (although within the standard deviation of the other cases). The 2TM model had the smallest standard deviation (*i.e.* the spread of the U-value estimates from their mean value over the fifty-two periods), while the AM had the largest. The standard deviation of the 2TM model was half that of the AM. The ranges of the relative systematic error on the U-values estimated with the different methods and models over the fifty-two hypothetical monitoring campaigns are also summarised in [Table 1](#). The AM presented the highest minimum and maximum values, while the dynamic method was generally characterised by smaller ranges. In particular, the 2TM model had the smallest range of relative systematic error throughout the year (*i.e.* between 9% and 27%). The mean relative systematic error obtained from the dynamic method was comparable for all models considered.

The relationship between U-value estimates and their associated relative systematic error as a function of the average temperature difference between the internal and external environment<sup>5</sup> was investigated ([Fig. 3](#)). As expected from the mathematical formulation of the AM, the relative systematic error increased as the average temperature difference decreased, reaching a maximum of 50% error for the smallest average temperature difference observed (1.6 °C). Similarly to the AM method, for the dynamic method the magnitude of the relative systematic error increased as the temperature difference decreased, although within a narrower range (generally up to 30%). This result can be ascribed to the different mathematical modelling of the dynamic method, where the temperature difference is calculated at each observation instead of being averaged over the monitoring period (as discussed in [Section 3.2](#)). The 13% uniform relative systematic error obtained combining in quadrature sum the uncertainties listed in the ISO 9869-1 Standard [15] excluding the temperature variations within the space (as discussed in [Section 4.2.1](#)) is also shown in [Fig. 3](#).

### 5.3. In-situ U-value estimates for the cavity wall

Twenty-four hypothetical monitoring campaigns were analysed for the UHWall, as all the periods met the acceptance criteria in [Section 5](#). The length of each hypothetical monitoring campaign ranged between three and twenty-eight days. A summary of the minimum, maximum, mean and standard deviation of the U-value and the associated relative systematic error for the UHWall es-

**Table 2**

Minimum, maximum, mean and standard deviation of the U-value estimates and the associated relative systematic measurement error for the UHWall over the hypothetical monitoring campaigns, estimated with the average and the dynamic method. The range of U-values from tabulated thermophysical properties is also reported.

		Min	Max	Mean	St dev	Units
U-value	Literature	0.32	0.40	–	–	$\text{Wm}^{-2}\text{K}^{-1}$
	AM	0.59	1.00	0.71	0.08	$\text{Wm}^{-2}\text{K}^{-1}$
	1TM (1 HF)	0.63	0.79	0.69	0.04	$\text{Wm}^{-2}\text{K}^{-1}$
	1TM (2 HF)	0.47	0.72	0.59	0.08	$\text{Wm}^{-2}\text{K}^{-1}$
	2TM	0.63	0.82	0.70	0.05	$\text{Wm}^{-2}\text{K}^{-1}$
Rel sys err	AM	13%	21%	16%	3%	–
	1TM (1 HF)	8%	18%	12%	3%	–
	1TM (2 HF)	7%	14%	10%	2%	–
	2TM	6%	16%	11%	2%	–

timated with the average and dynamic method (using the 1TM (1 HF), 1TM (2 HF) and 2TM models) is presented in [Table 2](#). For each hypothetical monitoring campaign, the U-value estimates obtained with the average and dynamic method were within the margin of the systematic measurement error. However, the in-situ estimates were higher than the range of values that would be expected from literature calculation making the assumption that the cavity is fully-filled [25] ([Table 2](#)). This result supports the expectation that the foam insulation may have shrunk close to the measurement location, leading to only partial-fill ([Section 4.1.2](#)).

Similarly to the OWall case study, the mean U-value was comparable in all cases except the 1TM (2 HF) model ([Table 2](#)), where the mean U-value was lower. As previously observed for the OWall, the AM presented the highest minimum and maximum value for the relative systematic error on the U-value, although the error ranges were usually smaller for the UHWall than the OWall. The reduction of the systematic error may be ascribed to the use of higher accuracy temperature sensors in this case study and the larger minimum average temperature difference observed (2.4 °C) compared to the OWall (1.6 °C). The thermistors used on the UHWall had an accuracy of 0.1 °C, while the type-T thermocouples installed on the OWall had an accuracy of 0.5 °C.

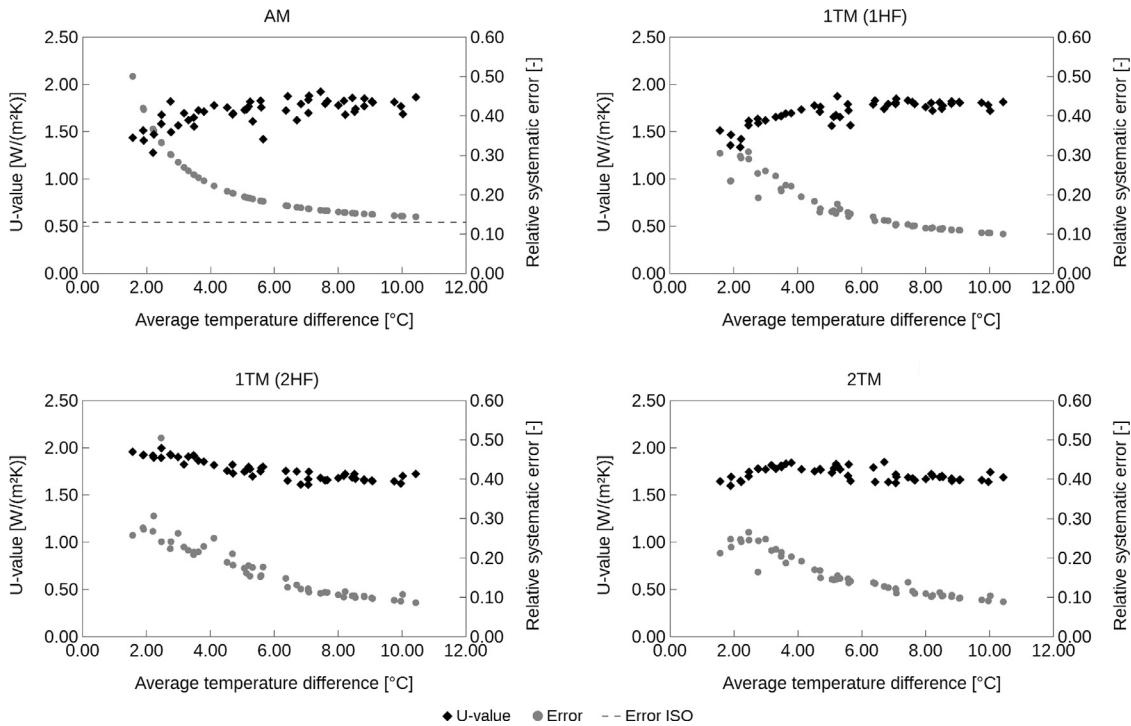
The U-value estimates and the associated relative systematic error as a function of the average temperature difference observed during the corresponding hypothetical monitoring campaign are presented in [Fig. 4](#), as well as the 13% uniform relative systematic error suggested in the ISO 9869-1 Standard [15] (as discussed in [Section 5.2](#)). Similarly to the OWall, the U-value estimates obtained with the 1TM (2 HF) had a different trend than the estimates obtained with the other dynamic models and the AM. Furthermore, the relative systematic error had comparable behaviour to the OWall case study, although the error ranges were generally smaller ([Table 2](#) and [Fig. 4](#)). As previously observed, the relative systematic errors on the U-values tended to increase when the wall was exposed to smaller average temperature differences.

### 5.4. Comparison of the error estimates from static and dynamic methods across the case studies

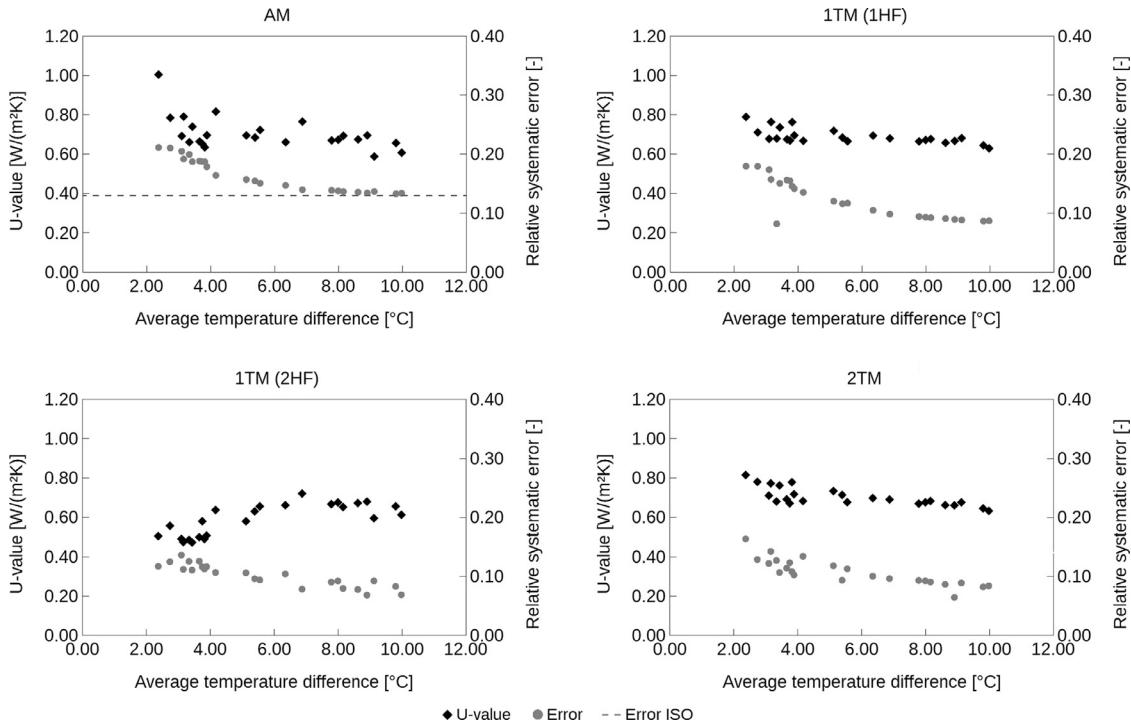
The analysis presented above shows that the insights derived from in-situ measurements may be considerably limited by the associated relative systematic error. This error generally increases as the average temperature difference between the internal and external environment decreases, even when the stabilisation criteria ([Section 4.2.3](#)) are met and the U-value estimates look plausible<sup>6</sup>

<sup>5</sup> Referred to as “average temperature difference” in the following for conciseness.

<sup>6</sup> The U-value estimates for the OWall were within the literature range, and those for the UHWall were in line with the observed shrinkage of the full-fill insulation layer.



**Fig. 3.** U-value and associated relative systematic measurement error for the OWall as a function of the average temperature difference. Estimates were obtained using the average (AM) and the dynamic method (with the 1TM (1 HF), 1TM (2 HF), 2TM models). The 13% relative systematic error suggested in the ISO 9869-1:2014 Standard (Section 4.2.1) is also shown (dashed line).



**Fig. 4.** U-value and associated relative systematic measurement error for the UHWall as a function of the average temperature difference. Estimates were obtained using the average (AM) and the dynamic method (with the 1TM (1 HF), 1TM (2 HF), 2TM models). The 13% relative systematic error suggested in the ISO 9869-1:2014 Standard (Section 4.2.1) is also shown (dashed line).

compared to literature values (Figs. 3 and 4; Tables 1 and 2). The average temperature difference is particularly relevant for analysis undertaken using the AM, as the measured heat flux and temperature difference constitute the denominator of Eq. (4). Conversely, the use of monitored data at each sampling interval for parameter estimation with the dynamic method was shown to reduce

systematic uncertainties (Figs. 3 and 4) throughout the year. This suggests that the choice of the estimation method for the thermal performance of building elements from in-situ measurements should account for several factors, including: the location of the case study, the time of the year, and the purpose of the survey. For example, the use of a dynamic approach may decrease error

estimates throughout the year for mild climates where the average temperature difference may rarely be higher than 10 °C (as commonly recommended for best-practice in-situ monitoring [18]) even in the winter period [19].

Figs. 3 and 4 show that for the AM the error on the U-value is underestimated assuming independent uncertainties but without accounting for the accuracy of the equipment (*i.e.* applying the 13% uniform relative systematic error obtained combining in quadrature sum the uncertainties in [15], as discussed in Section 5.2). Specifically, for the AM the systematic error including the equipment specification fell within the value in the ISO 9869-1 Standard [15] in only 12.5% of the cases (*i.e.* 3 hypothetical monitoring campaigns out of 24) for the UHWall and never for the OWall; this comparison cannot be directly made for the dynamic method as in this case the 10% uncertainty for errors caused by the variations over time of the temperatures and heat flow does not apply (see Section 4.2.1). In line with previous work [19], this result highlights the role of the accuracy of the monitoring equipment ( $\pm 0.1$  °C for the temperature sensors on the UHWall compared to 0.5 °C for those on the OWall) in the total error estimates from in-situ measurements and the potentially significant impact of propagating errors on the basis of the equipment used rather than applying assumed values.

Applying the dynamic method, smaller relative systematic errors were generally obtained (Figs. 3 and 4; Tables 1 and 2) from the models optimising two heat flux data streams (1TM (2 HF) and 2TM models) compared to that optimising only one (1TM (1 HF) model) heat flux time series (and also the AM). Additionally, the relative systematic errors for the 2TM model were generally smaller than those for the 1TM (2 HF) model, and the U-value estimates were more stable throughout the year. Comparable results were observed for an additional fully-filled cavity wall belonging to a different case-study building monitored long term by the authors and reported in [25].

The Bayesian analysis method enabled the comparison of the probability of different models accurately describing the observed data, using the same input time series for the 1TM (2 HF) and the 2TM models (as introduced in Section 3.1.2). The odds ratio strongly supported the 2TM model for all hypothetical monitoring campaigns in both case studies. The result suggests that although the 1TM (2 HF) model is useful for model comparison purposes, the availability of only one effective thermal mass to simultaneously describe heat flux measurements from the two heat flux data streams is not sufficient. This complements the insights gained from Tables 1 and 2, where the 1TM (2 HF) model presented higher mean U-value, and from Figs. 3 and 4, where the U-value estimates using the 1TM (2 HF) model had a different trend (especially for low average temperature differences) compared to the other methods and models. The lumped thermal mass model consisting of two thermal masses and three thermal resistances was shown to provide a good description of data collected at all times of the year for building elements of different construction (*i.e.* solid and cavity walls), providing robust estimates of their thermophysical properties.

## 6. Conclusions

This research highlights the importance of error analysis to gain robust insights into the actual thermal behaviour of buildings from measurements collected in situ. The propagation of systematic measurement uncertainties on the thermophysical properties of building elements (*e.g.*, R-values and U-value) was investigated in this paper by means of two case studies (*i.e.* a solid and a cavity wall) monitored long-term. U-values were estimated both using the average method and a grey-box dynamic method based on Bayesian statistics. Different approaches were used to quantify the

systematic measurement error on the U-value estimates to reflect the different mathematical description of heat transfer in the two methods adopted. While a linear error propagation from the definition of U-value was applied for the AM, a method for the quantification of the propagation of uncertainties on the parameters estimated using optimisation techniques was proposed.

The results highlight the importance of the quantification of the error based on the equipment used and the conditions experienced. Whilst the required accuracy of a result may vary depending on the purpose of the analysis, quantifying the error estimate facilitates informed decision making. In this work systematic errors estimated through propagation, including equipment accuracy, only fell within the 13% relative systematic error suggested in the ISO 9869-1 Standard [15] for 12.5% of the cases for the cavity wall and never for the solid wall; errors were assumed independent throughout. Besides showing that the application of a uniform 13% error would be misrepresentative, this work highlights the importance of the equipment error and the impact of using lower accuracy equipment on the certainty of results.

The analysis showed that, as expected, total relative error increases as the difference between internal and external average temperatures decreases. The relative systematic errors on the AM estimates were highly sensitive to the average temperature difference observed during the monitoring period (*i.e.* the lower the average temperature difference the higher the relative systematic error). This result emphasises that the estimation of thermophysical properties from in-situ measurements cannot disregard error analysis even when the estimates may look plausible and in line with literature calculation, as the magnitude of the systematic measurement error may effectively nullify the insights gained from the analysis.

Use of the dynamic method significantly reduced systematic error at all times of year compared to the AM. A dynamic method, and derivation of error accounting for the use of optimisation methods, may therefore be applied to provide modest error estimates also in cases that the average internal-to-external temperature difference is considerably lower than 10 °C. The application of a dynamic method of U-value estimation may extend the use of in-situ measurements outside the winter period or to climatic regions where the average temperature difference is generally low.

Results from three different dynamic models associated lower systematic measurement error with the use of more data, in the form of measurements from both an internal and external heat flux plate. The 2TM model presented the smallest relative errors among all models and its U-value estimates were the most stable throughout the year; Bayesian model comparison also favoured this model compared to the 1TM (2 HF) model, which also displayed unexpected trends in the results, most likely associated with challenges of optimising the size of one effective thermal mass using two heat flux data streams.

Quantifying the errors associated with U-value estimates is important to facilitate informed decision making and quality assurance, for example in the investigation into the quality of installation of a building component or to determine the cost-effectiveness of a retrofitting intervention. Whilst the error that may be tolerated in U-value estimates may vary according to the purpose and use of the results, lower error (assuming that it is representative) is generally desirable. This study shows that use of a dynamic method of U-value estimation from in-situ measurements, with an error quantification method that accounts for the optimisation technique used, can result in significantly lower error estimates than those from propagation of error with a static method. The associated decrease in internal-to-external temperature difference required during in-situ measurements that results in U-value estimates with moderate errors extends the applicability of such methods in support of closing the performance gap.

## Acknowledgements

This research was made possible by support from the EPSRC Centre for Doctoral Training in Energy Demand (LoLo), grant numbers EP/L01517X/1 and EP/H009612/1, and the RCUK Centre for Energy Epidemiology (CEE), grant number EP/K011839/1.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.enbuild.2018.02.048](https://doi.org/10.1016/j.enbuild.2018.02.048).

## References

- [1] International Energy Agency, Transition to Sustainable Buildings Strategies and Opportunities to 2050, Technical Report, 2013.
- [2] D.B. Crawley, J.W. Hand, M. Kummert, B.T. Griffith, Contrasting the capabilities of building energy performance simulation programs, *Build Environ.* 43 (4) (2008) 661–673, doi:[10.1016/j.buildenv.2006.10.027](https://doi.org/10.1016/j.buildenv.2006.10.027).
- [3] L.G. Swan, V.I. Ugursal, Modeling of end-use energy consumption in the residential sector: a review of modeling techniques, *Renewable Sustainable Energy Rev.* 13 (8) (2009) 1819–1835, doi:[10.1016/j.rser.2008.09.033](https://doi.org/10.1016/j.rser.2008.09.033).
- [4] L.K. Norford, R.H. Socolow, E.S. Hsieh, G.V. Spadaro, Two-to-one discrepancy between measured and predicted performance of a 'low-energy' office building: insights from a reconciliation based on the DOE-2 model, *Energy Build.* 21 (2) (1994) 121–131, doi:[10.1016/0378-7788\(94\)90005-1](https://doi.org/10.1016/0378-7788(94)90005-1).
- [5] J.R. Stein, A. Meier, Accuracy of home energy rating systems, *Energy* 25 (4) (2000) 339–354, doi:[10.1016/S0360-5442\(99\)00072-9](https://doi.org/10.1016/S0360-5442(99)00072-9).
- [6] P. Baker, Historic Scotland Technical Paper '10 - U-values and Traditional Buildings, Technical Report, Historic Scotland Alba Aosmhor, 2011.
- [7] F.G.N. Li, A. Smith, P. Biddulph, I.G. Hamilton, R. Lowe, A. Mavrogianni, E. Oikonomou, R. Raslan, S. Stamp, A. Stone, A. Summerfield, D. Veitch, V. Gori, T. Oreszczyn, Solid-wall u-values: heat flux measurements compared with standard assumptions, *Build. Res. Inf.* 43 (2) (2014) 238–252, doi:[10.1080/09613218.2014.967977](https://doi.org/10.1080/09613218.2014.967977).
- [8] C. van Dronkelaar, M. Dowson, E. Burman, C. Spataru, D. Mumovic, A review of the regulatory energy performance gap and its underlying causes in non-domestic buildings, *Front. Mech. Eng.* 1 (2016) 17, doi:[10.3389/fmech.2015.00017](https://doi.org/10.3389/fmech.2015.00017).
- [9] D. Feuermann, Measurement of envelope thermal transmittances in multifamily buildings, *Energy Build.* 13 (2) (1989) 139–148, doi:[10.1016/0378-7788\(89\)90005-4](https://doi.org/10.1016/0378-7788(89)90005-4).
- [10] M. Hughes, J. Palmer, V. Cheng, D. Shipworth, Global sensitivity analysis of England's housing energy model, *J. Build. Perform. Simul.* 8 (5) (2015) 283–294, doi:[10.1080/19401493.2014.925505](https://doi.org/10.1080/19401493.2014.925505).
- [11] Zero Carbon Hub, Closing the Gap Between Design & As-Built Performance (end of term report), Technical Report, 2014. [http://www.zerocarbonhub.org/sites/default/files/resources/reports/Design\\_vs\\_As\\_Built\\_Performance\\_Gap\\_End\\_of\\_Term\\_Report\\_0.pdf](http://www.zerocarbonhub.org/sites/default/files/resources/reports/Design_vs_As_Built_Performance_Gap_End_of_Term_Report_0.pdf).
- [12] International Energy Agency, IEA EBC Annex 58: Reliable building energy performance characterisation based on full scale dynamic measurements, 2011.
- [13] International Energy Agency, IEA EBC Annex 70: Building energy epidemiology: analysis of real building energy use at scale, 2016a.
- [14] International Energy Agency, IEA EBC Annex 71: Building energy performance assessment based on in-situ measurements, 2016b.
- [15] BS ISO 9869-1, Thermal insulation – Building elements – In-situ measurement of thermal resistance and thermal transmittance. Part 1: Heat flow meter method, 2014.
- [16] A.-H. Deconinck, S. Roels, Comparison of characterisation methods determining the thermal resistance of building components from onsite measurements, *Energy Build.* 130 (2016) 309–320, doi:[10.1016/j.enbuild.2016.08.061](https://doi.org/10.1016/j.enbuild.2016.08.061).
- [17] H. Hens, Building physics - heat, air and moisture: Fundamentals and engineering methods with examples and exercises, Ernst & Sohn Verlag für Architektur und technische Wissenschaften GmbH & Co. KG, Germany, 2012.
- [18] Energy Saving Trust, CE128/GIR64: Post-construction testing - a professionals guide to testing housing for energy efficiency, 2005.
- [19] G. Desogus, S. Mura, R. Ricciu, Comparing different approaches to in situ measurement of building components thermal resistance, *Energy Build.* 43 (10) (2011) 2613–2620, doi:[10.1016/j.enbuild.2011.05.025](https://doi.org/10.1016/j.enbuild.2011.05.025).
- [20] C. Roulet, J. Gass, I. Marcus, In-situ u-value measurement: reliable results in shorter time by dynamic interpretation of measured data, *Thermal Perf. Exterior Envelopes Build.* III (1987) 777–784.
- [21] O. Gutschker, Parameter identification with the software package LORD, *Build. Environ.* 43 (2) (2008) 163–169, doi:[10.1016/j.buildenv.2006.10.010](https://doi.org/10.1016/j.buildenv.2006.10.010).
- [22] P. Biddulph, V. Gori, C.A. Elwell, C. Scott, C. Rye, R. Lowe, T. Oreszczyn, Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements, *Energy Build.* 78 (2014) 10–16, doi:[10.1016/j.enbuild.2014.04.004](https://doi.org/10.1016/j.enbuild.2014.04.004).
- [23] I. Náveros, C. Ghiaus, D.P. Ruíz, S. Castaño, Physical parameters identification of walls using ARX models obtained by deduction, *Energy Build.* 108 (2015) 317–329, doi:[10.1016/j.enbuild.2015.09.021](https://doi.org/10.1016/j.enbuild.2015.09.021).
- [24] A. Janssens, Report of subtask 1b: Overview of methods to analyse dynamic data, Reliable building energy performance characterisation based on full scale dynamic measurements, International Energy Agency, 2016.
- [25] V. Gori, A Novel Method for the Estimation of Thermophysical Properties of Walls from Short and Seasonally Independent In-Situ Surveys, PhD, University College London, London, UK, 2017. <http://discovery.ucl.ac.uk/1568418>.
- [26] V. Gori, V. Marincioni, P. Biddulph, C. Elwell, Inferring the thermal resistance and effective thermal mass distribution of a wall from in situ measurements to characterise heat transfer at both the interior and exterior surfaces, *Energy Build.* 135 (2017) 398–409, doi:[10.1016/j.enbuild.2016.10.043](https://doi.org/10.1016/j.enbuild.2016.10.043).
- [27] ASTM C1046-95 (2013), Standard Practice for In-Situ Measurement of Heat Flux and Temperature on Building Envelope Components, Technical Report, 2013.
- [28] ASTM C1155-95 (2013), Standard Practice for Determining Thermal Resistance of Building Envelope Components from the In-Situ Data, Technical Report, 2013.
- [29] JCGM 200:2012, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM), Technical Report, Joint Committee for Guides in Metrology (JCGM/WG1), 2012.
- [30] L. Huber, Standardization in measurement: Philosophical, historical and socio-logical issues, Routledge, 2015.
- [31] H. Madsen, P. Bacher, G. Bauwens, A.-H. Deconinck, G. Reynders, S. Roels, E. Himpe, G. Lethé, Report of Subtask 3, Part 2: Thermal Performance Characterisation Using Time Series Data – Statistical Guidelines, Reliable building energy performance characterisation based on full scale dynamic measurements, International Energy Agency, 2015.
- [32] J. Zhao, R. Plagge, N.M.M. Ramos, M. Lurdes Simões, J. Grunewald, Concept for development of stochastic databases for building performance simulation – a material database pilot project, *Build. Environ.* 84 (2015) 189–203, doi:[10.1016/j.buildenv.2014.10.030](https://doi.org/10.1016/j.buildenv.2014.10.030).
- [33] C. Winship, B. Western, Multicollinearity and model misspecification, *Sociol. Sci.* 3 (2016) 627–649, doi:[10.15195/v3.a27](https://doi.org/10.15195/v3.a27).
- [34] G. Bohm, G. Zech, Introduction to Statistics and Data Analysis for Physicists, DESY, 2010.
- [35] Hukseflux, HFP01 Heat flux plate, 2015. <http://www.hukseflux.com/product/hfp01>.
- [36] H. Madsen, Time Series Analysis, CRC Press, 2007.
- [37] H.H. Ku, Notes on the use of propagation of error formulas, *J. Res. Natl. Bureau Standards, Section C* 70C (4) (1966) 263, doi:[10.6028/jres.070C.025](https://doi.org/10.6028/jres.070C.025).
- [38] C. Rye, The SPAB - U-value Report, Technical Report, 2012.
- [39] Campbell Scientific, CR1000 Measurement and Control Datalogger, 2015.
- [40] V. Gori, C.A. Elwell, Long-term in-situ measurements of heat flux and temperature on a solid-brick wall in an office building in the UK, 2018, doi:[10.14324/000.ds.10041899](https://doi.org/10.14324/000.ds.10041899).
- [41] V. Kalthod, H.N. Knickle, The effect of aging of urea-formaldehyde foam on thermal conductance, *J. Build. Phys.* 6 (1) (1982) 14–38, doi:[10.1177/109719638200600102](https://doi.org/10.1177/109719638200600102).
- [42] Eltek, Squirrel 450/850 Series Data Logger, 2015.
- [43] V. Gori, C.A. Elwell, Long-term in-situ measurement of heat flux and temperature on a filled cavity wall in a residential building in the UK, 2018, doi:[10.14324/000.ds.10041900](https://doi.org/10.14324/000.ds.10041900).
- [44] BS EN ISO 6946, Building components and building elements - thermal resistance and thermal transmittance - Calculation method, 2007.
- [45] CIBSE, Environmental Design - Guide A, 2007.