Characterisation of thermal energy dynamics of residential buildings with scarce data

Palmer Real, Jaume; Rasmussen, Christoffer; Li, Rongling; Leerbeck, Kenneth; Jensen, Ole Michael; Wittchen, Kim B.; Madsen, Henrik

Published in:
Energy and Buildings

DOI (link to publication from Publisher):
10.1016/j.enbuild.2020.110530

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Energy & Buildings 230 (2021) 110530

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Characterisation of thermal energy dynamics of residential buildings with scarce data

Jaume Palmer Real [a,*], Christoffer Rasmussen [a], Rongling Li [b], Kenneth Leerbeck [a], Ole Michael Jensen [c], Kim B. Wittchen [c], Henrik Madsen [a]

[a] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
[b] Department of Civil Engineering, Technical University of Denmark, Denmark
[c] Danish Building Research Institute (SBi), Aalborg University, Denmark

ARTICLE INFO

Article history:
Received 29 May 2020
Revised 22 September 2020
Accepted 30 September 2020
Available online 11 October 2020

Keywords:
Thermal building characterisation
Data analysis
Modelling
Energy flexibility
Demand response

ABSTRACT

Buildings account for a large portion of the total energy consumption and they might serve as a significant thermal storage capacity that can be advantageous for the future energy grid. To utilise this capacity, it is necessary to characterise the thermal dynamics in buildings using methods that are general enough to be applicable to a significant share of the building stock. This work proposes a data-driven method to characterise thermal dynamics of thermostatically controlled buildings with *night setback*. The method includes 1) using Hidden Markov Models to systematically select data periods when the indoor temperature decays steadily during night; 2) model reduction of a Stochastic Differential Equations model of heat transfer to a discrete linear model which is fitted by utilising the selected night-time data; and 3) computing one short time constant and one long time constant, which allows to categorise buildings according to their thermal response. This method is applied to 39 different Danish residential buildings and the results reveal that this simplified model captures the main processes governing the heat transfer: the one-step predictions for the indoor air temperature return 95% of the residuals $\in [-0.05\,^{\circ}\mathrm{C}, 0.05\,^{\circ}\mathrm{C}]$. For all buildings, the short time constants are lower than an hour, and the long time constants range from 20 h to 100 h. Finally, this method is used in simulated data to validate that the time constants provide insight about the energy flexibility potential of a building. The results show that dynamic thermal response of buildings can be discovered using limited data.

1. Introduction

In order to successfully transition from fossil fuel to renewable energy based supply at a large scale, it is necessary to understand buildings' thermal dynamics as they will be a key asset in the future flexible energy systems [1].

A main contributor to this green transition, will be the deployment of *Demand-Response* (DR) strategies [2], i.e., altering the demand-side of the buildings energy load to match the requirements of the grid. These requirements range from, balancing frequency, maximising the renewable energy share, or reducing energy peaks [3].

Space heating takes a major part of the energy load in buildings, specially in colder regions. Then, in order to use DR strategies, buildings need to be prepared to store and release heat when needed, without affecting severely the comfort of their users. Residential buildings in particular, can be used as *thermal storage units* (TES), and heavy buildings are specially suited for that task without risking indoor comfort [4]. In addition, the DR potential of residential buildings can be increased through adequate retrofitting [5]. Thus, there is a need for tools that can assess how buildings are able to store and release heat. Moreover, these tools need to be general, considering the existing variety of residential buildings in place [6].

1.1. Literature review

There are several approaches for estimating the main indicators of the building energy performance: the *heat loss coefficient* (or its inverse, the thermal *resistance*) and the thermal capacity [7]. The approach will depend on the chosen model for the energy flow in the building. In their review, Foucquier et al., separated the main modelling methods for buildings in three categories:

* Corresponding author at: Anker Engelunds vej 1, Building 101A, 2800 Kongens Lyngby, Denmark.
E-mail address: jpre@dtu.dk (J. Palmer Real).

https://doi.org/10.1016/j.enbuild.2020.110530
0378-7788/

Nomenclature

Variables

T_i	Indoor air temperature, K
T_m	Thermal mass temperature, K
T_a	Outdoor air temperature, K
Φ_h	Space heating, kW
I_g	Solar irradiation, kW/m^2
$v_t^{(i)}$	white noise term $\forall i,j \in [1,2]$, –

Parameters

R_i	Thermal resistance: indoor air ↔ thermal mass, K/kW
R_a	Thermal resistance: indoor air ↔ outdoor air, K/kW
C_i	Thermal capacity of the indoor air, kWh/K
C_m	Thermal capacity of the thermal mass, kWh/K
A_w	Effective window area, m^2
σ_1	Incremental variance of T_i, K
σ_2	Incremental variance of T_m, K
θ_i	Auto-regressive parameter of order $i \forall i \in [1,2]$, –
Γ_{ij}	Parameter of the discrete inputs matrix $\forall i,j \in [1,2]$, –
ϕ_{ij}	Parameter of the discrete design matrix $\forall i,j \in [1,2]$, –

Acronyms

SDE	Stochastic Differential Equation
DR	Demand Response
FI	Flexibility Index
FF	Flexibility Function
AR	Auto-Regressive
ARX	Auto-Regressive with eXogenous inputs
HMM	Hidden Markov Model
OLS	Ordinary Least Squares
WLS	Weighted Least Squares
TES	Thermal Energy Storage

physics-based methods (*white box*), purely statistical methods (*black-box*) and hybrid methods (*grey-box*) [8].

In the first category, it is possible to find numerical tools, such as the finite volume methods (FVM) [9]. This approach is computationally expensive, and relies either on simulated data or complex experimental assemble. Often, these methods require going into detail to component level to gain insight about the building's thermal characteristics, such as the outer wall structure and its composition [10,11]. A different approach, was presented as the *co-heating* method [12]. This method is based on simplified heating dynamic equations describing the heat transfer inside a building. The co-heating method showed accurate results assessing the energy performance of buildings; however, it needs a meticulous experimental set up to control and measure the temperature, and the experiment can take multiple days. In addition, the measurements are bound to the external conditions during the experimental period. Finally, due to the necessary infrastructure for the assessment, it can only be performed in empty houses. Alternatively, Alzetto et al. introduced a new method that could evaluate the building thermal response in a two days experiment [13]. This represents an improvement from the traditional co-heating method. Nevertheless, it still requires an extensive experimental set up. In addition, there often exists a performance gap between the model prediction using thermal parameters estimated with the above methods and the real operational energy performance of the building [14,15].

Black-box methods have been commonly applied for predicting energy consumption of buildings. For example, D'Amico et al. used *Multi Criteria Decision-Making* to compare three different methods when forecasting energy demand [16], and Finck et al. trained *Artificial Neural Networks* models in their simulation and demonstration work [17,18]. Nevertheless, these methods are not used for characterisation purposes, as their outputs are difficult to interpret physically. Thus, they have been omitted from this review.

For existing buildings, the grey-box model approach is often used, as it takes into account building physics in data-driven modelling. Bacher and Madsen studied multiple variations of the lumped resistance-capacity (RC) model are presented [19]. In their work, it is shown that such models make it possible to estimate the heat resistance and capacity for different components of the house. This method allows the use of in situ measures of operative houses. However, this approach still needs a considerable amount of data to decouple the different processes that are part of the energy flow. Also, the complexity of these models can easily grow, burdening the computation. For this reason, and specially when working with large data sets, it is important to use strategies to reduce the order of the models, as Goyal and Barooah suggested [20]. Even with high quality data, it could be difficult to gain insights into the building dynamics. For instance, during the periods where the building is thermostatically controlled, the impact of external variables over the indoor air temperature might be masked by the effects of the controller, which makes it impossible to estimate the thermal capacities.

The thermal characteristics of buildings, can also be computed statically, by using daily averages of the consumption and outdoor temperature. Nielsen et al. and Rasmussen et al. studied the impact of weather conditions in a set of danish residential buildings [21,22]. These approach relies on having consumption data. In addition, by using daily values, they omit the dynamic nature of using buildings thermal mass as TES.

1.2. Work description

In this study, we suggest a simple and efficient method to scan residential buildings and retrieve their thermal characteristics using only indoor and outdoor temperature measurements. Specifically, we compute the *time constants* of the two main dynamic processes that govern the heat loss in buildings.

Most of the characterisation work presented in the literature review rely on complex model configurations that try to fit in many building components taking part in the heat transfer of buildings. Such studies are based on data sets with a large number of variables, e.g. temperatures of all components, which are not commonly measured in existing buildings. In this work, in an effort to pursue generality, we take the opposite direction to build simple models that can capture energy dynamics of buildings using scarce data that is easy to measure. Moreover, this method does not require a particular experimental set-up as it is developed based on the data that is available in many existing dwellings in Denmark.

We focus the study to data periods that are following a *night setback* schedule; i.e., periods with no heat input, where the indoor temperature decays steadily. This schedule is a commonly used strategy to decrease the energy consumption during night with a lower temperature setpoint than during daytime. During night, it can be assumed that there is no significant influence on the indoor temperature from the users as they most likely are asleep. There is also no solar irradiation affecting the indoor temperature. Thus, the

pattern of decreasing temperature can be understood as a fingerprint that explains the energy storage performance of the building.

There are multiple methods to identify decaying behaviours. One example is the use of statistical change point detection of the signal of interest [23]. In the present work, a Hidden Markov Model (HMM) of the indoor temperatures to identify the decaying periods. The same method was used by Wolf et al. to detect various human activities using CO_2 concentration data [24]. The relevant data is selected and it is used to fit an auto-regressive model. This model is derived from a Stochastic Differential Equation system describing the thermal dynamics in buildings.

Our method is general and can be applied to a significant portion of the current building stocks. Moreover, the method presented in this work is easily scalable and can be used to identify groups of buildings with similar energy response. Grouping buildings into clusters where the energy response is known, can be specially relevant since it could reduce the uncertainty of DR policies [25].

This study is divided into three sections, followed by a conclusion. In Section 2, the mathematical background of the applied method is explained; the model structure and its physical interpretation are discussed, and the method for data selection and the concepts used in the flexibility analysis are described. Section 3 describes the data used in the study. Section 4 shows the modelling results, an interpretation of the system dynamics and the estimated values of time constant for a number of buildings. In Section 5, a simulation framework is used to show how the building intrinsic parameters affect the time constants, and how the time constants provide information about the flexibility potential of the building.

2. Method

In this section, it is shown how an auto-regressive model is derived from a stochastic RC model. The purpose of this approach is to offer a physical interpretation of the parameters. Then, it is explained how the time constants are computed using the *transfer function* form of the auto-regressive model.

Our proposed method only works for specific periods of time; here, the method to select the relevant data periods is presented.

Finally, the concepts of *flexibility index* (FI) and *flexibility function* (FF) that we use in Section 5 are presented.

2.1. Building as a second order dynamical system

Eq. (1) describes a general continuous time model for heat dynamics inside of a building. It tracks the temporal evolution of two main variables inside the building: the indoor air temperature, T_i, and the thermal mass temperature, T_m. The model is represented as a second order linear stochastic differential equation (*SDE*) system. This system has three main external inputs: the outdoor temperature, T_a, the global solar irradiation, I_g, and the space heating input, Φ_h. This model has five parameters $\{R_i, R_a, C_i, C_m, A_w\}$ that are described in the nomenclature section. The uncertainty in the system is captured by the stochastic term, $d\omega_i \forall i \in [1,2]$. This term represents a Wiener process with incremental variances $\sigma_i^2 \forall i \in [1,2]$.

The external inputs affect only the indoor air temperature; in turn, there is a heat transfer between the indoor air and the thermal mass. Madsen described this model in detail [26].

$$\begin{cases} dT_i = \frac{1}{C_i}\left(\frac{1}{R_i}(T_i - T_m) - \frac{1}{R_a}(T_i - T_a) + \Phi_h + I_g A_w\right)dt + \sigma_1 d\omega_1 \\ dT_m = \frac{1}{R_i C_m}(T_m - T_i)dt + \sigma_2 d\omega_2 \end{cases} \tag{1}$$

Since the system in Eq. (1) is linear, it can be re-written using the following matrix form,

$$\begin{pmatrix} \frac{dT_i}{dt} \\ \frac{dT_m}{dt} \end{pmatrix} = \begin{pmatrix} \frac{1}{R_i C_i} & -\left(\frac{1}{R_a C_i} + \frac{1}{R_i C_i}\right) \\ \frac{-1}{R_i C_m} & \frac{1}{R_i C_m} \end{pmatrix} \begin{pmatrix} T_i \\ T_m \end{pmatrix} + \begin{pmatrix} \frac{1}{R_a C_i} & \frac{1}{C_i} & \frac{A_w}{C_m} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} T_a \\ \Phi_h \\ I_g \end{pmatrix} + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \frac{d\omega_1}{dt} \\ \frac{d\omega_2}{dt} \end{pmatrix}. \tag{2}$$

Now, the model variable is a vector: $\boldsymbol{T}(t) = (T_i(t), T_m(t))^\top$, and $\boldsymbol{U}(t) = (T_a(t), \Phi_h(t), I_g(t))^\top$ is the vector of external inputs. We can write Eq. (2) in a compact form as

$$\frac{d}{dt}\boldsymbol{T}(t) = \boldsymbol{A}\boldsymbol{T}(t) + \boldsymbol{B}\boldsymbol{U}(t) + \boldsymbol{\Sigma}\frac{d}{dt}\boldsymbol{\omega}(t), \tag{3}$$

where $\boldsymbol{A}$ is the design matrix describing the dynamic characteristics of the building, and $\boldsymbol{B}$ describes how the input variables enter the system. Finally, $\boldsymbol{\omega}(t) = (\omega_1(t), \omega_2(t))^\top$ is the stochastic term, and $\boldsymbol{\Sigma}$ is the matrix of incremental variances.

In this work, only the indoor air temperature, T_i is observed. It is important to notice that the system described in Eq. (1) has no measurement equation. Thus, we assume that the error measurements for T_i are small enough to be disregarded.

2.2. From SDEs to auto-regressive

The system in Eq. (3) can be discretized by integrating over a sample interval, $[t, t+s)$, where s is the *sampling time* of the system. Then, assuming that the input $(U(t))$ is constant in the sampling interval, the system can be re-written as

$$\boldsymbol{T}(t+s) = \boldsymbol{\Phi}(s)\boldsymbol{T}(t) + \boldsymbol{\Gamma}(s)\boldsymbol{U}(t) + \boldsymbol{v}(t,s). \tag{4}$$

If the sampling time is small enough, the discrete time model structure will capture the relevant dynamics described in the continuous case. We can fix the sampling time to an arbitrary time unit, $s = 1$, and find he explicit expression for the elements describing Eq. (4):

$$\boldsymbol{\Phi}(s=1) = \exp(\boldsymbol{A}\cdot 1) = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \tag{5}$$

$$\boldsymbol{\Gamma}(s=1) = \int_0^{s=1} \exp(\boldsymbol{A}\cdot r)\boldsymbol{B}dr = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} & \Gamma_{23} \end{pmatrix} \tag{6}$$

$$\boldsymbol{v}(t, s=1) = \begin{pmatrix} v_t^{(1)} \\ v_t^{(2)} \end{pmatrix} \text{ with } v_t^{(i)} = \mathcal{N}(0, \varsigma_i^2)\ \forall i \in [1,2]. \tag{7}$$

In this case, $\boldsymbol{\Phi}$ describes the discrete dynamics of the system and $\boldsymbol{\Gamma}$ the effect from the external inputs. The vector $\boldsymbol{v}$ is normally distributed white noise with zero mean and variance $\varsigma_i^2 \forall i \in [1,2]$, and it accounts for the stochastic part of the system.

This study is focused on the dynamics of buildings with a night heating setback. Thus, if there is no solar irradiation and the space heating is turned off, the input term can be simplified to $\boldsymbol{U} = (T_a, 0, 0)^\top$. It is then possible to write discrete explicit expressions for the indoor air temperature and the thermal mass:

$$\begin{cases} T_{t+1}^i &= \phi_{11}T_t^i + \phi_{12}T_t^m + \Gamma_{11}T_t^a + v_t^{(1)} \\ T_{t+1}^m &= \phi_{21}T_t^i + \phi_{22}T_t^m + \Gamma_{21}T_t^a + v_t^{(2)}, \end{cases} \tag{8}$$

where the notation has been changed to highlight the discrete nature of the expression. Notice that the terms from the matrices described in Eqs. (5)–(7), appear now in the transformed difference equation system (8).

Now, we want to remove the thermal mass variable from the model, since we do not have access to its associated temperature measurements. In order to do this, the equations from Eq. (8) have been merged, adjusting the index t,

$$T^i_{t+1} = \phi_{11} T^i_t + \phi_{12} \times \overbrace{\left(\phi_{21} T^i_{t-1} + \phi_{22} T^m_{t-1} + \Gamma_{21} T^a_{t-1} + v^{(2)}_{t-1}\right)}^{T^m_t} + \Gamma_{11} T^a_t + v^{(1)}_t. \quad (9)$$

For houses with regular heating schedules that have not been ventilated recently, the temperature of the thermal mass should be very similar to the temperature of the indoor air [27]. Assuming that $T^i_{t-1} \approx T^m_{t-1}$, it is possible to reduce the expression to Eq. (10), where all variables are observed.

$$\begin{aligned} T^i_{t+1}, &= \phi_{11} T^i_t + (\phi_{12}\phi_{21} + \phi_{12}\phi_{22}) T^i_{t-1} + \Gamma_{11} T^a_t + \phi_{12}\Gamma_{21} T^a_{t-1}, \\ &+ v^{(1)}_t + \phi_{12} v^{(2)}_{t-1}. \end{aligned} \quad (10)$$

In Denmark, the cold weather conditions of winter, present slow outdoor temperature variations during night, as shown in Fig. 1. For a small enough time step, there is little change from one measurement to the next one, i.e. $T^a_t \approx T^a_{t-1} \forall t$. This simplification allows us to reduce the system complexity:

$$T^i_{t+1} = \overbrace{\phi_{11}}^{\theta_1} T^i_t + \overbrace{(\phi_{12}\phi_{21} + \phi_{12}\phi_{22})}^{\theta_2} T^i_{t-1} + \overbrace{(\Gamma_{11} + \phi_{12}\Gamma_{21})}^{\omega} T^a_t + v^{(1)}_t + \phi_{12} v^{(2)}_{t-1}, \quad (11)$$

where the final model parameters, $\{\theta_1, \theta_2, \omega\}$, have been introduced.

Finally, a new stochastic variable is defined, $v_t = v^{(1)}_{t-1} + \phi_{12} v^{(2)}_{t-2}$. Since $v^{(1)}_t$ and $v^{(2)}_t$ are independent and normally distributed with zero mean $\forall t$; the new variable, v_t, is also normally distributed with zero mean. Thus, it is possible to write the final model as the following *auto-regressive* model,

$$T^i_t = \theta_1 T^i_{t-1} + \theta_2 T^i_{t-2} + \omega T^a_t + v_t, \quad (12)$$

where the time index, t, has been adjusted for clarity. Fig. 2 summarizes the process leading from Eq. (1) to Eq. (12), highlighting the necessary assumptions.

2.3. A system with two time constants

In order to study the heat loss between the indoor air and the outdoor air, we focus on the interaction between T^i_t and T^a_t from model (12) using the *transfer function* as shown in Eq. (13),

$$T^i_t = \underbrace{\frac{\omega}{1 - \theta_1 B - \theta_2 B^2}}_{H(B)} T^a_t, \quad (13)$$

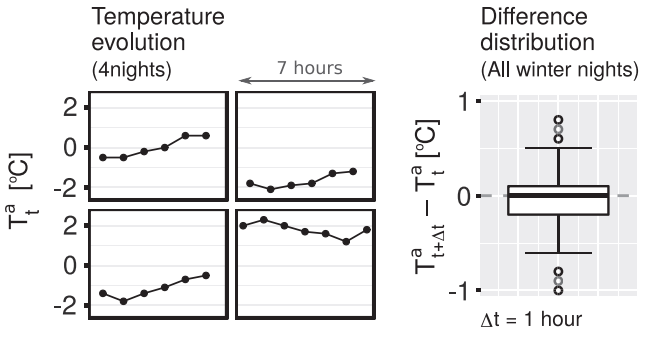

Fig. 1. The lattice on the left shows the temporal evolution of the outdoor temperature for 4 randomly selected nights. On the right, the temperature hourly differences for all winter nights is shown.

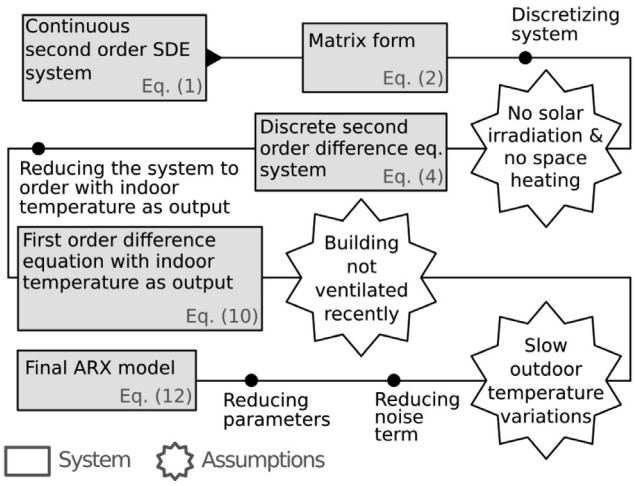

Fig. 2. Flowchart describing the method to transform the initial SDE system into a reduced discrete linear model.

where B is the *backshift operator* defined as $B^j X_t = X_{t-j}$ for an arbitrary dynamic random variable $\{X_t\}$, and $H(B)$ is the transfer function of the system (12). As described in [28], in order to compute the time constant of the system (12) it is necessary to find the roots of the denominator of $H(B)$, i.e., the *poles* of the system. In this case, there are two poles, q_1 and q_2, as the polynomial in the denominator is order two. Finally, each pole has an associated time constant which can be computed as:

$$\tau_j = \frac{s}{\ln |q_j|} \quad \forall j \in [1, 2]. \quad (14)$$

Hence, in this case there are two different time constants that characterize the heat flow between indoor air and outdoor air. Each of these time constants has time units and they reveal information about the two processes described by the initial system (2) as illustrated in Fig. 3. In general, when the heating is shut down, there is a quick heat transfer between the indoor air and the thermal mass due to the low thermal capacity of the air. The initial heat transfer is captured by the parameter τ_1. The heat transfer between indoor air and outdoor air is the second process, which is characterized by τ_2, is slower and will dominate the dynamics as indoor temperature keeps decreasing [26].

2.4. Identification of night setback temperature curves

Hidden Markov Models consist of two components: an observed sequence of states and a corresponding hidden state sequence. The current state only depends on the state of the previous observation. The states change according to a fitted transition matrix, which is a matrix providing the probability of a state change. For each state, the observations are Gaussian, and the mean and variance depends on the state.

All studied houses showed a temperature decay of the indoor air during night hours. The goal, was to identify the state where the temperature is decaying constantly. In order to identify it, a new variable was created to be used as an input for the model:

$$d_t = T^i_t - T^i_{t-1}. \quad (15)$$

When temperature is continuously dropping, the data points of d_t are distributed differently from the rest, as can be seen in Fig. 4. The Viterbi algorithm [29] was used to find the most likely sequence of states given a sequence of observed d_t.

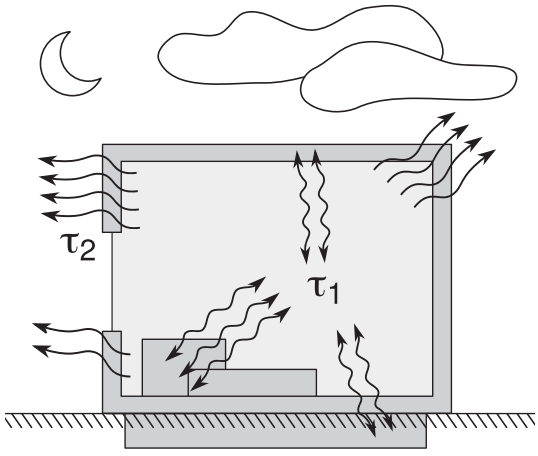

Fig. 3. This figure shows a schematic of the two main processes governing the heat flow between indoor air and outdoor air. First, the heat transfers from the indoor air to the thermal mass, characterized by τ_1. Meanwhile, τ_2 characterises the heat transferred from the indoor air to the outdoor air.

For some houses, the selected periods are not only the long night decaying periods, but also shorter decays during the day time due to the dead-band of the controller. The long night decays are selected with a threshold. This threshold was computed by clustering all d_t data points into two based on the length of the period. In general, it is concluded that this process is sufficient to reveal the natural split between short daytime decays (unwanted) and long night decays, although the time spans vary from house to house because of their different time constants.

2.5. Flexibility Index (FI) and Flexiblity Function (FF)

In order to gain insights into the relationship between the time constant and the flexibility of a building, the concepts of the Flexibility Index (FI) [30] and the Flexibility Function (FF) [31] are used. The FI quantifies the savings caused by allocating energy consumption in a flexible way. This is done by comparing the cost of the consumption adapted to a flexible control signal (*flexible cost*) and the cost of the consumption if the system was unaware of the price signal (*ignorant cost*). The idea behind the FI can be seen in Eq. (16). FI = 1 characterizes a building with an extreme flexibility potential, and FI = 0 the opposite.

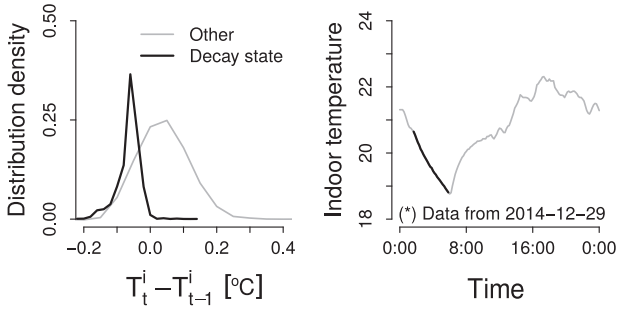

Fig. 4. An example of the distribution of decay points for one arbitrary building with night setback. The figure on the left shows that the decay states distribute different from other states. The figure on the right shows the indoor temperature of a day with the thicker curve indicating the decay period identified using the HMM method.

$$\mathrm{FI} = 1 - \frac{\text{Flexible cost}}{\text{Ignorant cost}} \tag{16}$$

The *flexibility function* (FF) describes the energy that is available at a particular moment, or *state of charge*; and the resources it can allocate and how it can allocate them before reaching the system limits. This function provides information about how an energy system would adapt to changes in the control signal or changes in its state of charge. Moreover, an energy system has limited resources that can be turned on/off in case of need, and the rate at which it is able to move those resources depends on the dynamic characteristics of each system. The FI and FF used in this work are based on the work introduced by Junker et al. [30,31].

When calculating the FI and FF in this work, the energy systems are buildings that have their indoor air temperature controlled with an MPC controller. Then, the state of charge translates into the room temperature: the building is completely charged when the indoor temperature reached its daily maximum inside the comfort boundaries and vice versa. The heating input is controlled by a price signal which is built using the wind speed data. This price signal can also be based on other data and be used for other purposes such as peak shaving, or lowering CO_2 emissions.

3. Data description

This study is based on measurements from a set of 39 Danish single family houses in the Middelfart region in Denmark. None of the studied buildings was built after 1980s, and their construction years vary as described in Fig. 5. They also vary in size and plant blueprint: the built area ranges from 80 to 226 m^2. Most of the houses are detached family houses, although there are also, semi-detached houses, town–houses and a few farms. In this work, these details have been omitted as the focus is set on characterising the buildings solely based on how they lose heat.

There are two data sets. The first data set contains the indoor temperature measurements from December 2014 to December 2015. We used only the winter data between December 2014 and March 2015 and with a 10 min resolution. The second data set consists of hourly outdoor temperature values. This data set has been selected to match the dates of the indoor temperature data. In order to compensate the lower resolution, the hourly values are interpolated using linear interpolation. This interpolation are carried under the assumption that, during winter nights the fluctuations in outdoor temperature are slow.

For better presentation and visualisation of the work, the rest of this work focuses on three buildings with characteristic decay patterns, as shown in Fig. 6. The three buildings have been chosen due to they present qualitative differences in their continuous heat loss pattern. It can be noticed that building B shows a sharp decay at the beginning and then a slower decaying trend; building A has a fast temperature drop; and building C has a shorter decay curve,

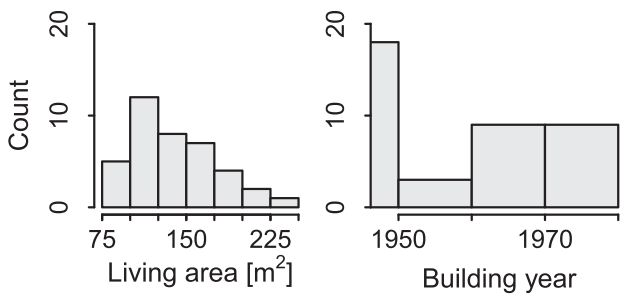

Fig. 5. Summary description of the studied set of 39 buildings. Almost half of the buildings were built before 1950.

but still steeper than the curve in B. For all the three example buildings, the decaying patterns are consistent during different nights. For instance, it can be noticed that in building C, the decaying trend is similar in every case even though the initial indoor temperature varies.

4. Results and discussions

This section presents the results from the modelling process. These results are also discussed to understand how the thermal dynamics are captured. Afterwards, this section offers a classification of the buildings based on their time constants.

4.1. Model validation

The parameters from the model in Eq. (12) were estimated using *Ordinary Least Squares* (OLS) method in the first place, since the model is linear and the noise is supposed to be normally distributed. This revealed that the term related to the outdoor temperature, ω, was not significant in most of the tested buildings. However, as presented in Section 2, the outdoor air temperature affects the heat loss dynamics. This impact is more noticeable at the end of the decaying curve, when the heat exchange with the outdoor air is more significant than the heat exchange with the thermal mass. In addition, at the beginning of the decay trend, the assumption that $T_t^i \approx T_t^m$ is less accurate, since in reality both temperatures converge over time. Lastly, at the beginning of the decaying curve there is a higher chance than the users might be awake. All these reasons suggest that the noise over time, $\{v_t\}$, might not be independent. In fact, the system noise is expected to be at its maximum at the beginning of the decay and then continuously decrease over time. In order to cope with this, the *weighted least squares* (WLS) method was used, where the models is fitted giving a specific weight to each observation. These weights are proportional to the time since the temperature decay started, given that the observations become more reliable as time moves forward. The following equation was used for the weighting process,

$$w(k) = 1 - \frac{1}{\alpha + (1-\alpha)k} \quad \text{with } 0 \leqslant \alpha \leqslant 1, \tag{17}$$

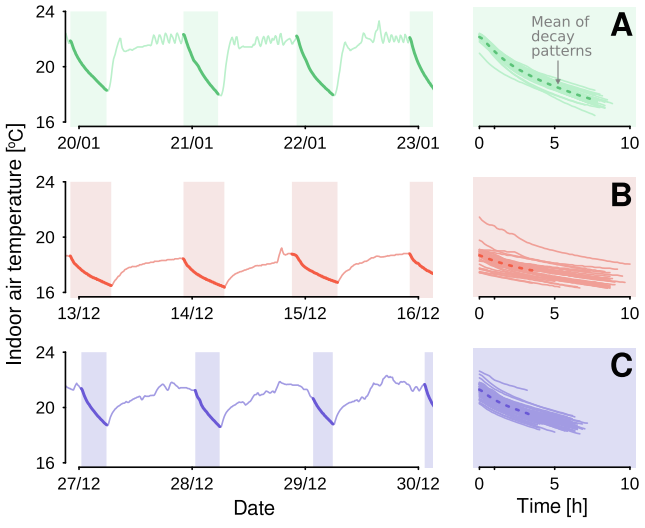

Fig. 6. On the left, 4 day period for the three example households (A, B, C). On the right, the lines corresponding to all decay periods present in the data are shown. Each building has its clear decay pattern.

where k is an integer counting the number of measurements since the beginning of the decay period and α is a tuning parameter. Notice that $w \in [0, 1]$, it is minimum right at the beginning of the decay period, $w(1) = 0$, and then grows monotonically. The parameter α defines how fast is this growth; i.e. how the weights are distributed along the decay period. Notice that this parameter depends on the time constants of the building and it could be fine-tuned in a recursive method in order to fit a particular building. Thus, it is important to remark that this weighting function is not unique and it could be adapted to each case.

Table 1 compares the results after fitting the model with OLS and WLS. Notice how, after using WLS, the p-values of the parameter estimates are below 0.1, confirming that the parameters are highly significant. It can also be noticed that the estimate of the contribution of the outdoor temperature, $\hat{\omega}$, increased significantly using the WLS, especially for houses A and C, which confirms the influence of the external conditions.

Fig. 7 reveals that the ordinary residuals, after using the WLS, behave like white noise regardless of the outdoor conditions. It also can be seen that most of the residuals are in the $[-0.1, 0.1]$ range, which show the accuracy of the one-step prediction. The values that fall outside of this range are from the beginning of the decaying periods when it is expected to be noisy, as explained previously. The distribution of residuals can be seen in Fig. 8, which shows that the errors are small and centered around zero.

In Figs. 7 and 8 it is not possible to see if the residuals are biased on a daily basis. For this reason, the evolution of residuals are plotted for 3 arbitrary days picked at random for each example building in Fig. 9. It can be seen from this figure that in all cases the prediction follows the same pattern. At the beginning of the decay period, the model is over-predicting so the residuals are negative and larger. This bias comes from the simplification $T_t^i \approx T_t^m$, since it forces the stochastic part to account for the changes in T_t^m. However, shortly after the start the residuals converge to white noise, since $T_t^i \to T_t^m$ quickly [26]. These transient periods are the cause of the larger residuals seen in Fig. 7.

4.2. Time constant of 39 houses

The time constants of the three example buildings are shown in Table 2; it is important to notice that τ_1 is expressed in minutes, whereas τ_2 is expressed in hours. These results can be compared with the qualitative behaviour observed in Fig. 6. In comparison with the other two, building B shows a flatter decay trend at the end of each night. This translates into having $\tau_2^{(B)} > \tau_2^{(C)} > \tau_2^{(A)}$. In addition, the numerical results in Table 2 reveal that building B loses energy swiftly at the beginning of the decay, which is different from house A where the heat transfer between the indoor air and thermal mass also has a significant contribution to the decay period.

Moreover, the results of house A were compared with the results computed using a different method presented in [32], where a more complex auto-regressive model (including heating data) was fitted using a time span of 11 days. The difference between results were smaller than an 8%. In addition, it is important to remark that it is possible to further reduce the difference between the results by fine-tuning the weighting function (17), specifically for house A.

The simplicity and generality of this method has allowed us to use it in the total pool of 39 buildings and to categorize their parameters τ_1 and τ_2. Fig. 10 shows the distribution of the time constants for all 39 buildings. It can be observed that: i) the long time constant, τ_2, has the same order of magnitude as the time constant one would expect from Danish family houses [33,34]; ii) for the short time constant, τ_1, all values are shorter than one hour,

Table 1
Table comparing the parameter estimates for each example building. The table compares the results using WLS (weighted with the function in (17)) and using OLS.

		Estimate		Std. error		p-value	
		OLS	WLS	OLS	WLS	OLS	WLS
A	$\hat{\theta}_1$	1.768	1.779	0.016	0.017	< 1E−16	< 1E−16
	$\hat{\theta}_2$	−0.769	−0.781	0.016	0.017	< 1E−16	< 1E−16
	$\hat{\omega}$	4.3E−4	3.3E−4	2.5E−4	2.9E−4	0.09	0.26
B	$\hat{\theta}_1$	1.111	1.273	0.024	0.023	< 1E−16	< 1E−16
	$\hat{\theta}_2$	−0.113	−0.275	0.024	0.023	< 1E−16	< 1E−16
	$\hat{\omega}$	1.8E−3	1.5E−3	2.4E−4	2.6E−4	< 1E−16	< 1E−16
C	$\hat{\theta}_1$	1.624	1.707	0.015	0.015	< 1E−16	< 1E−16
	$\hat{\theta}_2$	−0.625	−0.708	0.015	0.015	< 1E−16	< 1E−16
	$\hat{\omega}$	2.8E−4	1.3E−4	1.4E−4	1.7E−4	0.05	0.45

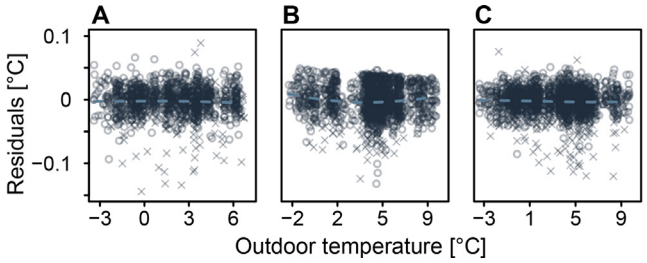

Fig. 7. The figure compares the outdoor temperature and the residuals using the WLS model. The residuals from the first hour of each decay are marked with an x. The trend of all residuals is plotted as a dashed line. All residuals are centered around 0 and evenly spread across all temperature range.

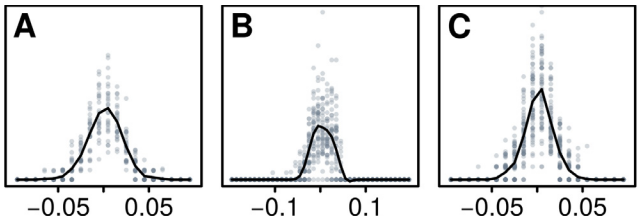

Fig. 8. Distribution of residuals after using the WLS model. The figures include all the decays in the time span of interest ignoring the first hour of each decay. For each building a smooth curve was plotted for better visibility.

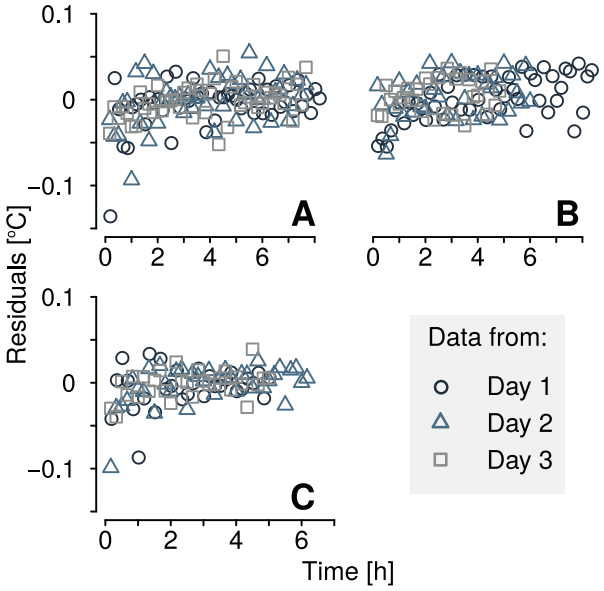

Fig. 9. Temporal evolution of the residuals during each decay period. Only 3 decays are plotted for better visibility. It can be seen that the ordinary residuals are larger at the beginning of the decay period, but shortly after there is no observable trend.

Table 2
Results for the example houses.

	A	B	C
τ_1 [min]	38.8	4.6	21.5
τ_2 [hour]	36.7	65.6	49.7

which highlights the importance of a small time step to capture this part of the dynamics. From Fig. 10 it can be seen that the 39 buildings are clustered in groups. K-means clustering method [35] was used and as the result three clusters are marked in the figure. Note that each of the three selected buildings falls in different categories, confirming the qualitative differences in their heat loss dynamics spotted at the beginning of this work.

In Fig. 10, one can get a clear picture at the available classes of buildings in the studied set. On the short time constant axis, τ_1, the time values are mostly scattered, contrarily to the τ_2 where most of the values lie around the bottom half of the plot. Furthermore, it can be seen how the main driver for clustering comes from the long time constant τ_2, i.e., the three main regions reveal three different steps along the y-axis. The difference among three clusters could be due to the difference in insulation level, house size, the amount of thermal mass, etc. However, to investigate this level of detail would require more information about these houses which we do not have, thus it was left out of the current study.

5. Flexibility assessment

This section uses simulated data to show that the computed time constants of a building reveal information about its energy potential.

The simulations are based on the model described in (2) and carried out using different parameter values to gain an overview of their impact. Specifically, the parameters of the indoor air heat resistance, R_i, and the capacity of the indoor air, C_i, were fixed; while the heat resistance between the indoor and outdoor air, R_a, and the capacity of the thermal mass, C_m, were changed in each simulation. The first two parameters depend on intrinsic magnitudes of the air; meanwhile, the last two parameters are easier

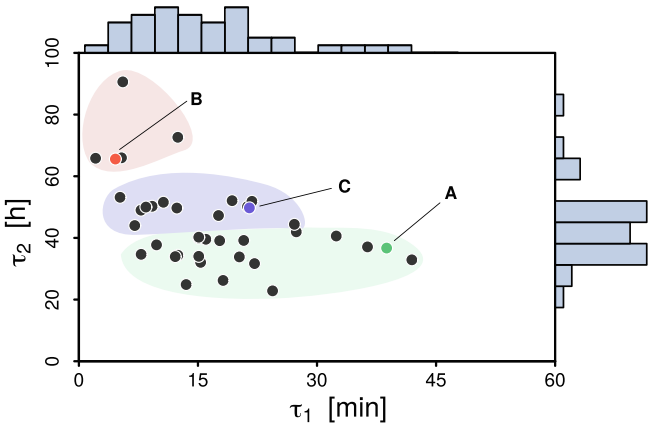

Fig. 10. Scatter plot of the two time constants for 39 buildings. As the result of a K-means clustering, three regions are marked in the figure. In each cluster, the example house is highlighted.

to interpret and they characterize magnitudes from a building that are easier to correct through renovation. Finally, the only external output of the model, the outdoor temperature, follows an arbitrary pattern that matches the order of magnitude of danish winter time.

For each combination of parameters, a time series of four days was simulated with a night-setback schedule. Using those values, the time constants were computed, and the results can be seen in Fig. 11. It can be seen that both parameters are directly proportional to the value of the time constants, as expected. Three cases have been highlighted (H1, H2 and H3) to represent three buildings with a different parameter combination. These have been chosen to further assess the effects of the parameters on the building thermal performance.

For each combination of parameters shown in Fig. 11, we simulated four days using two different control strategies: one ignores the price signal and only tries to keep the indoor temperature inside a defined comfort region. In the other simulation, the heating system is controlled using *Model Predictive Control* (MPC), where the control signal is the price of energy. The MPC controller tries to minimize the operation cost using the aforementioned price signal, while also keeping the indoor temperature within the comfort region. These two strategies represent the *Ignorant Cost* and *Flexible Cost* as described in Section 2.

For the simulations, we created a price signal that depends on the wind speed during an arbitrary period of time, to simulate a system powered by wind energy. The energy price decreases as the wind speed increases, assuming that energy supply is always sufficient.

The results of the controlled simulation for the three highlighted cases (H1, H2, H3) are presented in Fig. 12. It can be seen that when the price is low, the heating is switched on and when the price increases the heating is turned off until the temperature approaches the lower boundary. It can be noticed that the heating in H3 could be turned off for a longer time due to the building's higher time constants. The indoor air temperature in building H2 never reached the upper boundary of the comfort region due to higher heat losses.

It can be seen in Fig. 13 that the resistance R_a is the key parameter to increase savings using the flexibility of the building.

Lastly, the results of the flexibility function of the three houses can be seen in Fig. 14. This figure summarizes how the three different houses react to a change in the two main drivers of the energy consumption: the room temperature and the energy price. This reaction is presented as the deviation from the ignorant consumption; i.e. the consumption of the system ignoring the flexible price signal. The three houses display a similar response to the state of charge: a flat section, where the system ignores the variations of the indoor temperature, and two steep curves at the ends of the temperature range. When the room temperature reached the low boundary, the system was forced to increase consumption to maintain comfort; similarly, the system decreases consumption when the room temperature reached the high boundary. It can be noted that H3 is able to stay on the flat region for a wider temperature range than the other two. This suggests that H3 is more resilient to changes in the room temperature than the H1 and H2.

The response to energy price follows a decreasing curve for the three houses. For lower prices, the power demand of H1 is below H2 and H3. As the price increases, H3 consumption gets below H1 and H2. This result is in line with the results in Fig. 13, and confirms that the high value of the FI of H3 is the result of avoiding expensive prices. The savings of H1 compared to H2 come mainly from decreasing consumption during cheaper hours.

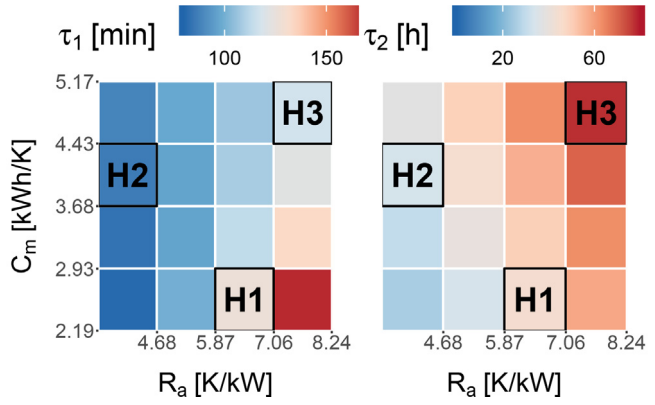

Fig. 11. Resulting time constant values for different values of C_m and R_a. The increase of R_a yields higher values of both time constants. The increase of C_m benefits only τ_2. H1, H2 and H3 represent three different houses with different characteristics.

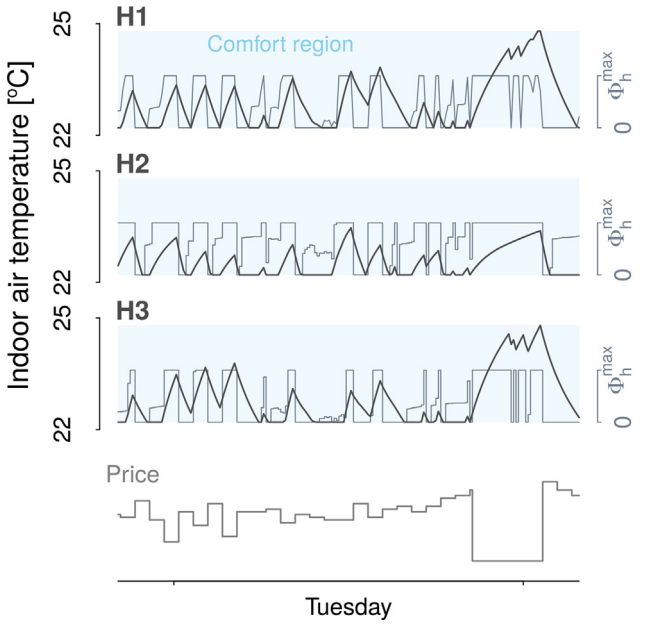

Fig. 12. Simulation of the flexible control scenario for the representative houses. The indoor temperature follows the heat supply, which is controlled by the price signal. It can be seen that H3, is able to keep indoor temperature within the comfort region with the heating system running for a shorter time comparing to H1 and H2.

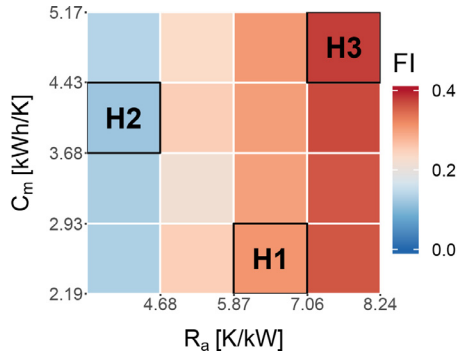

Fig. 13. Resulting Flexibility Index value for different values of C_m and R_a. It can be noticed that the main driver for the FI is the thermal resistance. H1, H2 and H3 represent three different houses with different characteristics.

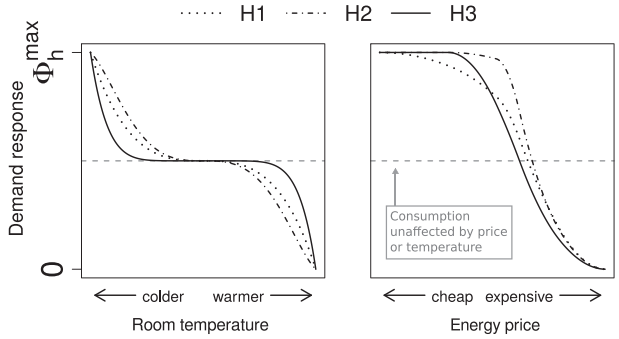

Fig. 14. The figure shows that the energy demand is affected by changes on the state of the system (room temperature) and the variations in the control signal (energy price). The curve above the dashed line means that the consumption is increased, and vice versa. It can be seen that H3 is more resilient to changes in the room temperature, and it is able to consume less during the most expensive hours.

6. Conclusions

This study shows how one can obtain insights from the thermal characteristics of a building with limited data. First of all, Hidden Markov Models were used to select the relevant periods to extract measurements from the periods with a night-setback. By focusing on these long decaying periods, it has been possible to transform a complete physical system to a simple auto-regressive structure. This model structure only uses the indoor air temperature and the outdoor air temperature, which are normally easy to measure.

It is important to use a high resolution sampling to capture the fast dynamics inside the building. In this study, 10-min time interval data was used. This small time step made it possible to reduce the model structure by taking advantage of the slow changes and small variations of the outdoor temperature. Additionally, the resulting time constants highlighted the importance of choosing a small time step. The only external input of our model, the outdoor temperature, was measured hourly and was transformed to 10-min data using linear-interpolation taking advantage of its slow dynamics. The linear interpolation is expected to represent accurately the real outdoor temperature, without affecting the posterior analysis.

It is critical to understand the physical meaning of the model. In order to fit the model, it was necessary to take into account the decreasing trend of the system noise during the night by using the WLS method. Only by doing this, all parameters in the model became significant. This is important because the temperature decay inside of a building could potentially be caused by other factors: such as the air mixing in the same room, or a heat transfer to a much colder contiguous room. The significance of the parameter corresponding to the outdoor air confirms that the indoor air decreases due to a heat loss to the outdoor air, which validates our model assumptions.

This method offers a general and computationally light model that can be scaled to a large portion of the existing building stocks. By visualising the two time constants for all 39 buildings as shown in Fig. 10, three clusters of buildings with similar characteristics could be easily found. In this study, we used a simple clustering method to identify those building clusters.

The usability of each time constant depends on the specific problem. The long time constant is the one that gives a clearer picture of buildings' characteristics for thermal storage and it is also closely related to the classical time constant used in building physics. However, the short time constant could be relevant for studying short term flexibility and indoor comfort.

Finally, it is confirmed that there is a clear connection between the time constants and the flexibility potential of buildings. It is shown that the long time constant dominates the potential usage of a house as an energy storage unit in a flexible energy grid. Moreover, using the FF, it is possible to assess qualitatively the flexible response of the simulated houses. In conclusion, the results show that a house with higher values of τ_1 and τ_2 can implement more flexible strategies.

CRediT authorship contribution statement

Jaume Palmer Real: Conceptualization, Methodology, Software, Visualization, Writing - original draft. **Christoffer Rasmussen:** Methodology, Validation. **Rongling Li:** Writing - review & editing, Supervision. **Kenneth Leerbeck:** Software. **Ole Michael Jensen:** Resources. **Kim B. Wittchen:** Resources. **Henrik Madsen:** Supervision, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is part of the CITIES project (nr. DSF1305-00027B) as well as the *Smart Energi i Hjemmet* project. I would like to thank my colleague Rune G. Junker for many valuable discussions and for sharing his work and vision on the future of flexible energy systems.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.enbuild.2020.110530.

References

[1] K. Foteinaki, R. Li, A. Heller, C. Rode, Heating system energy flexibility of low-energy residential buildings, Energy and Buildings 180 (2018) 95–108.
[2] S. Ø. Jensen, H. Madsen, R.A. Lopes, R.G. Junker, D. Aelenei, R. Li, S. Metzger, K.B. Lindberg, A.J. Marszal, M. Kummert, B. Bayles, E. Mlecnik, R. Lollini, and W. Pasut. Annex 67: Energy Flexible Buildings – Energy Flexibility as a key asset in a smart building future, 2017. Contribution of Annex 67 to the European Smart Building Initiatives..
[3] A.J. Marszal-Pomianowska, H. Johra, A. Knotzer, J. Salom, T. Péan, S.Ø. Jensen, E. Mlecnik, H. Kazmi, R. Pernetti, K. Klein, L. Frison, P. Engelmann, J. Parker, L. Aelenei, R.A. Lopes, D. Aelenei, R.G. Junker, H. Madsen, A.Q. Santos, B.N. Jørgensen, Z. Ma, Principles of Energy Flexible Buildings: Energy in Buildings

and Communities Programme Annex 67 Energy Flexible Buildings, International Energy Agency (2020).
[4] Johan Kensby, Anders Trüschel, Jan-Olof Dalenbäck, Potential of residential buildings as thermal energy storage in district heating systems – results from a pilot test, Applied Energy 137 (2015) 773–781.
[5] T.H. Pedersen, R.E. Hedegaard, S. Petersen, Space heating demand response potential of retrofitted residential apartment blocks, Energy and Buildings 141 (2017) 158–166.
[6] G. Reynders, J. Diriken, D. Saelens, Generic characterization method for energy flexibility: Applied to structural thermal storage in residential buildings, Applied Energy 198 (2017) 192–202.
[7] V. Dimitriou, S.K. Firth, T.M. Hassan, T. Kane, M. Coleman, Data-driven simple thermal models: The importance of the parameter estimates, Energy Procedia 78 (2015) 2614–2619.
[8] A. Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: A review, Renewable and Sustainable Energy Reviews 23 (2013) 272–288.
[9] C. Luo, B. Moghtaderi, H. Sugo, A. Page, A new stable finite volume method for predicting thermal performance of a whole building, Building and Environment 43 (2008) 37–43.
[10] L. Wang, P. Ma, E. Hu, D. Giza-Sisson, G. Mueller, N. Guo, A study of building envelope and thermal mass requirements for achieving thermal autonomy in an office building, Energy and Buildings 78 (2014) 79–88.
[11] L. Evangelisti, C. Guattari, F. Asdrubali, R. de Lieto Vollaro, In situ thermal characterization of existing buildings aiming at nzeb standard: A methodological approach. Developments in the, Built Environment 2 (2020) 100008.
[12] G. Bauwens, S. Roels, Co-heating test: A state-of-the-art, Energy and Buildings 82 (2014) 163–172.
[13] F. Alzetto, G. Pandraud, R. Fitton, I. Heusler, H. Sinnesbichler, Qub: A fast dynamic method for in-situ measurement of the whole building heat loss, Energy and Buildings 225 (2018) 175–182.
[14] S. Roels, P. Bacher, G. Bauwens, S. Castanño, M.J. Jiménez, H. Madsen, On site characterisation of the overall heat loss coefficient: Comparison of different assessment methods by a blind validation exercise on a round robin text box, Energy and Buildings 153 (2017) 179–189.
[15] D. Cali, T. Osterhage, R. Streblow, D. Müller, Energy performance gap in refurbished german dwellings: Lesson learned from a field test, Energy and Buildings 127 (2016) 1146–1158.
[16] A. D'Amico, G. Ciulla, L. Tupenaite, A. Kaklauskas, Multiple criteria assessment of methods for forecasting building thermal energy demand, Energy and Buildings 224 (2020) 110220.
[17] C. Finck, R. Li, W. Zeiler, Optimal control of demand flexibility under real-time pricing for heating systems in buildings: A real-life demonstration, Applied Energy 263 (2020) 114671.
[18] C. Finck, R. Li, R. Kramer, W. Zeiler, Quantifying demand flexibility of power-to-heat and thermal energy storage in the control of building heating systems, Applied Energy 209 (2018) 409–425.
[19] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (2011) 1511–1522.
[20] S. Goyal, P. Barooah, A method for model-reduction of non-linear thermal dynamics of multi-zone buildings, Energy and Buildings 47 (2012) 332–340.
[21] H. Nielsen, S. Mortensen, P. Bacher, H. Madsen, Analysis of energy consumption in single family houses, 01 (2010)..
[22] C. Rasmussen, P. Bacher, D. Calì, H.A. Nielsen, H. Madsen, Method for scalable and automatised thermal building performance documentation and screening, Energies 13 (15) (2020) 3866.
[23] C. Rasmussen, R. Relan, H. Madsen, Identification of occupancy status by statistical change point detection of CO_2 concentration, in: 2018 IEEE Conference on Control Technology and Applications (CCTA), IEEE, 2018, pp. 1761–1766.
[24] S. Wolf, J. Kloppenborg Møller, M.A. Bitsch, J. Krogstie, H. Madsen, A markov-switching model for building occupant activity estimation, Energy and Buildings 183 (2019) 672–681.
[25] A. Wang, R. Li, S. You, Development of a data driven approach to explore the energy flexibility potential of building clusters, Applied Energy 232 (2018) 89–100.
[26] H. Madsen, J. Holst, Estimation of continuous-time models for the heat dynamics of a building, Energy and Buildings 22 (1995) 67–79.
[27] K. Foteinaki, R. Li, A. Heller, M.H. Christensen, C. Rode, Dynamic thermal response of low-energy residential buildings based on in-wall measurements. E3S Web of Conferences, 111 (2019)..
[28] H. Madsen, Time Series Analysis, Chapman & Hall, 2008.
[29] C. Bishop, Pattern Recognition and Machine Learning, chapter 13, Springer, 2006, p. page 629..
[30] R.G. Junker, A. Gashem Azar, R. Amaral Lopes, K. Byskov Lindberg, G. Reynders, R. Relan, H. Madsen, Characterizing the energy flexibility of buildings and districts, Applied Energy 225 (2018) 175–182.
[31] R.G. Junker, C.S. Kallesøe, J. Palmer, B. Howard, R.A. Lopes, H. Madsen, Stochastic nonlinear modelling and application of price-based energy flexibility, Applied Energy 275 (2020).
[32] S. Nordli, Statistical methods for optimizing renovation projects. Master's thesis, DTU Compute (2018).
[33] P. Bacher, H. Madsen, H.A. Nielsen, B. Perers, Short-term heat load forecasting for single family houses, Energy and Buildings 65 (2013) 101–112.
[34] J. Le Dréau, P. Heiselberg, Energy flexibility of residential buildings using short term heat storage in the thermal mass, Energy 111 (2016) 991–1002.
[35] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, chapter 13, Springer, 2009, p. page 460..