A data driven method for optimal sensor placement in multi-zone buildings

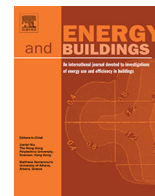Suryanarayana, Gowri; Arroyo, Javier; Helsen, Lieve; Lago, Jesus

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A data driven method for optimal sensor placement in multi-zone buildings

Gowri Suryanarayana [a,b], Javier Arroyo [a,b,c], Lieve Helsen [b,c], Jesus Lago [a,b,d,*]

[a] Algorithms, Modeling, and Optimization, VITO NV, Boerentang 200, Mol, Belgium
[b] EnergyVille, Thor Park, Genk, Belgium
[c] Department of Mechanical Engineering, KU Leuven, Leuven, Belgium
[d] Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

ABSTRACT

In this paper, we propose a data-driven methodology to identify the optimal placement of sensors in a multi-zone building. The proposed methodology is based on statistical tests that study the (in) dependence of measurements from various available sensors. The tests advice on a set of most dissimilar sensors to be retained, as they would convey the maximum information. The method starts with an initial setup that can provide measurements of every building zone to carry out this study; any of these sensors can be removed eventually to decrease costs in normal operation. The method has the advantages of being purely data driven and computationally efficient, as against several methods proposed in the scientific literature, that operate under the premise that detailed building models are available, to evaluate the number/position of the required sensors. This property makes the method scale to different buildings, in an expert free manner. The methodology can help towards better characterization of a building for optimal control and monitoring applications. It is validated against a widely used method – Kalman filtering with Grey-box models, using two different case studies. In both cases, the proposed approach agrees with the results using grey box models, suggesting that the method is reliable, while being quick and efficient.

1. Introduction

Buildings account for over one third of the final energy usage worldwide [1]. Optimal control of their energy use is hence crucial in advancing towards a sustainable future. To accommodate this, advanced building control is being intensively researched in the recent years to achieve significant energy savings [2]. Moreover, the European Energy Performance in Building Directive (2018) sets direct requirements for Building Automation System (BAS) [3]. However, these advanced controllers usually need a predictive model that requires a set of sensors to estimate the thermal state of the building [4]. The thermal characterization of the building state is hence a critical step to ensure optimal performance of these controllers. While there is ample literature and ongoing research in data driven building models, there is still a lack of methodologies for the selection of the number and location of sensors that provide these data. For instance, the Energy in Buildings and Communities Programme of the International Energy Agency carried out an extensive study in Annex 58 to gather an inventory of full scale test facilities for evaluation of building energy performance [5] with an overview of methods to analyse dynamic data [6]. This project was oriented towards the thermal characterization of building components [7] using the data, and not on the design of where the data comes from, i.e. number and location of sensors. Particularly, the sensors are currently installed in an ad hoc manner without considering operational costs and control performance. In this context, the number and placement of sensors are two important factors. In general, the more the sensors, the more the information that can be retrieved directly from the system. However, the setup and maintenance of multiple sensors is likely to be expensive and can lead to redundant information and hence, a trade off will be required here.

To address the challenge for optimal placement of sensors, two types of methods exist: model based methods and data driven methods. The majority of the research that exists focuses on model based methods. For instance, [8] uses a Kalman filter approach to identify the number and placement of sensors. In [9], a building energy simulation is used for the same purpose. A model based

* Corresponding author at: Algorithms, Modeling, and Optimization, VITO NV, Boerentang 200, Mol, Belgium.

E-mail address: gowri.suryanarayana@vito.be (G. Suryanarayana).

approach is also suggested in [10] for optimal placement of sensors for fault detection and diagnosis in smart buildings. The field of optimal experimental design is a closely related topic, where experiments are designed to get the best results for system identification [11,12]. However, these are also model based methods, and require an *a priori information about the application and models.*

In another related field, research on optimal placement of sensors for monitoring the structural strength/health of buildings focus on the use of information entropy based approaches, see for e.g., [13,14] and the references there in. These methods also need a clear definition of the model to determine the entropy or uncertainty in the model parameters and differ from the purely data-driven technique proposed in this manuscript in that sense.

Although model based approaches have the advantage that physical characteristics of the system are taken into account, their disadvantage is that they heavily rely on a modeling procedure. Specifically, the sensor placement decisions will depend on the model training procedure, and can further vary based on the model that is chosen. Moreover, such a procedure often needs a good understanding of the system, and substantial human expert intervention would be necessary to get the method working for each individual building.

Data driven models overcome these limitations by depending purely on measurements, and can be used across buildings in a human expert free manner. Despite their advantages, to the best of our knowledge, research on purely data driven methods that rely on advanced metrics is limited. In [15], a purely data driven technique based on *k*-means is suggested for ranking the sensors. However, the information loss here is biased towards sensors with larger amplitudes. Moreover, the *k*-means is an incomplete (low-dimensional) metric when analyzing the complex dependencies between sensors, and the method is not validated against other methods.

1.1. Contributions

In this work we aim at filling the above mentioned gap. We propose a purely data driven methodology, which uses advanced metrics, to evaluate and identify the particular zones of the building where zone air temperature sensors should remain. The main motivation behind this methodology is that it is model agnostic, which means that the results are generalizable to any type of model or control strategy. As an additional advantage, the methodology has small computational times.

The proposed methodology is based on statistical tests to identify the probabilistic dependency between different pairs of sensors within a building. We start with the premise that the initial setup provides measurements from multiple sensors already in place, and that any of them can be removed afterwards to decrease costs in normal operation. In detail, the methodology assumes that the dependency between any pair of sensors comes from proximity and the ability of zones to influence one another. Therefore, a pair of sensors that are highly dependent would imply that one of the sensors could be redundant and could be removed. In this context, the simplest criterion to test dependency between variables would be to check the covariance or the Pearson correlation coefficient. However, the problem with using this metric is that it captures only linear relationships, i.e., it falls short in giving information about more complex non-linear relationships, and is also sensitive to outliers. For this reason, the proposed methodology uses the Hilbert Schmidt Independence Criterion (HSIC), which is based on kernel independence measures using reproducing kernel Hilbert spaces (RKHSs) [16,17], to determine linear and non-linear dependencies. This metric is more robust than the Pearson correlation coefficient as it is less sensitive to outliers and can identify non-linear relationships. Yet, it can be interpreted in a similar way:

the larger the metric, the larger the dependence. Although we refer to [17] for a detailed explanation on how HSIC can capture nonlinear relations, it is important to note here that the method is based on expressing functions in infinite dimensional spaces. In this context, the function spaces impose some *norm* and *smoothness* constraints to avoid overfitting the data, e.g. to avoid having all sensors related to each other (in the context of sensor placement). A second property to be noted is that the results are independent of the kernel used as long as the kernels have the so called "universal" property [17] and, in practice, there are a wide range of kernels that satisfy this property (see [17] for details).

To study and demonstrate the use of the proposed methodology, we consider two case studies. First, we apply the proposed methodology to both case studies to perform the feature selection. Then, we implement a standard model-based approach to validate the results of the novel methodology. Based on the empirical results we show that the proposed methodology is a reliable method to select the important sensors in a building.

It is important to note that the proposed method is the first that avoids using a model (existing literature relies on model-based approaches) for selecting the sensors in buildings. This is important and beneficial because it removes/reduces the need for model identification, a process that in the case of buildings is nonconvex and without converge guarantees. In addition, although we focus on the paper on temperature sensors, the method can be applied on a heterogeneous set of sensors, and finds the dependencies in the underlying probability distributions.

Second, it is also important to remark that the method assumes that the initial setup can provide measurements of every zone to carry out this study, but that any of them can be removed afterwards to decrease costs in normal operation. It is specifically useful, when temporary reusable sensors may be located in the building to generate the data required for the novel sensor placement approach. The final number of sensors to be retained then depends on the trade off between the costs and the application requirements.

1.2. Organization of the paper

The rest of the paper is organized as follows. Section 2 explains the methodology in detail. Section 3 describes the procedure to validate the proposed approach with a model-based solution. Section 4 introduces the two case studies and Section 5 presents the results. Finally we make some concluding remarks and suggestions in Section 6.

2. Studying sensor dependence with HSIC

The most widely used metric while studying the dependence of two variables is the Pearson correlation coefficient. Given pairs of samples $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, it is defined as

$$r(x, y) := \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{1}$$

The problem with using the above metric is that it captures only linear relationships. It fails to give information about complex nonlinear relationships and is also sensitive to outliers. These limitations are illustrated with several data sets. See Fig. 1, where there is a clear dependence of variable y on variable x; however the Pearson correlation coefficient between the variables is only 0.02, indicating negligible dependence. Also for instance, the data sets shown in Figs. 2 and 3 are vastly different on visual inspection. The variables in Data set 2 seem not to be correlated. In contrast, the variables in Data set 3 seem to be very correlated. However the Pearson correlation coefficient of both these data sets is 0.812.
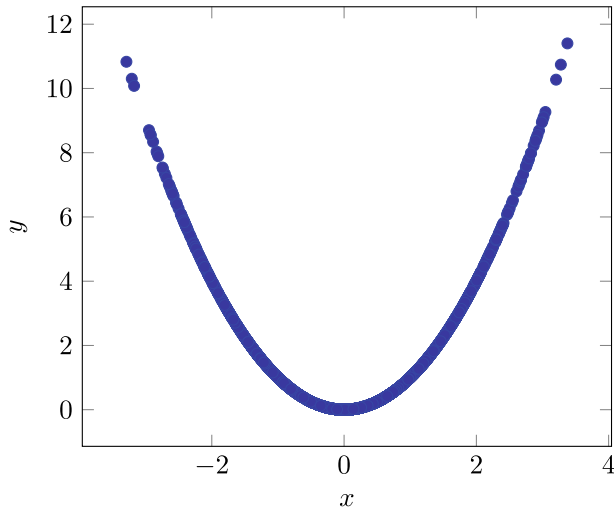
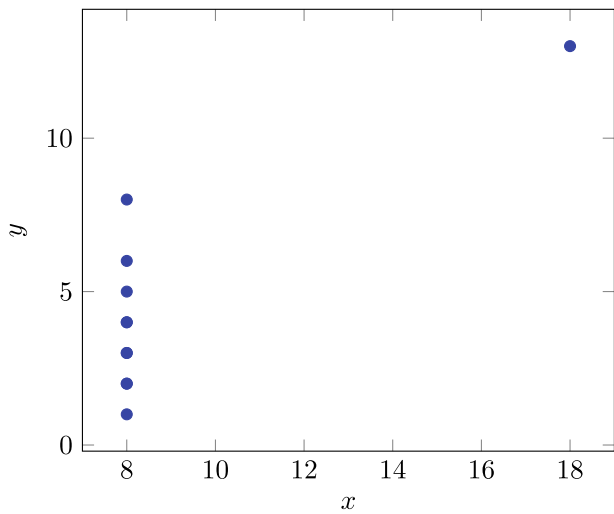**Fig. 1.** Data set 1: correlation coefficient = 0.02, *p*-value of HSIC based statistical test = 0.0.

Fig. 2. Data set 2: correlation coefficient = 0.812, *p*-value of HSIC based statistical test = 0.9.
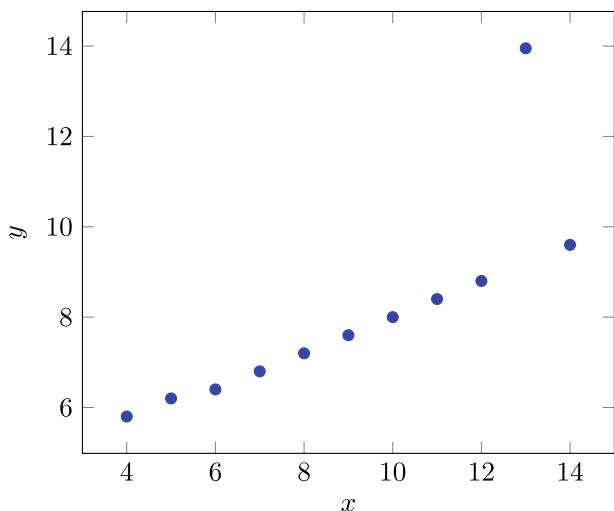

Fig. 3. Data set 3: correlation coefficient = 0.812, *p*-value of HSIC based statistical test = 0.0.

Hence, the Pearson correlation coefficient is not a reliable metric in trying to understand the dependence between various variables in a data set.

To overcome the shortcoming of the Pearson correlation coefficient we use another metric called the HSIC (The Hilbert–Schmidt Independence Criterion). The HSIC is a measure for the (in) dependence of variables, i.e., the higher the score, the more dependent the variables are on each other. The metric itself and the approach where it is used in are explained in the following two subsections.

2.1. HSIC

The explanation of HSIC presented in this section is based on [16,17]. Let us define $P(X, Y)$ as a Borel probability measure [18] defined on the domain $\mathscr{X} \times \mathscr{Y}$, and X and Y as the associated random variables. The problem that we are addressing here is to find out whether $P(X, Y)$ factorises as $P(X)P(Y)$ where $P(X)$ and $P(Y)$ are the (marginal) distributions defined on $\mathscr{X}$ and $\mathscr{Y}$ respectively, i.e. whether the random variables X and Y are independent. To answer that question, we use a *kernel based approach* for finding dependence between variables. Particularly, we use the HSIC, which is an empirical estimate of the *Hilbert–Schmidt norm of the cross-covariance operator*. The HSIC has a zero expected value if and only if the corresponding random variables are independent.

The first step of the method is to define the HSIC. For that, we start with two reproducing kernel Hilbert spaces (RKHSs) $\mathscr{F}$ and $\mathscr{G}$ (see [19]) on the compact domains $\mathscr{X}, \mathscr{Y}$, with orthogonal feature sets $\phi(x) \in \mathscr{F}$ ($\forall \, x \in \mathscr{X}$) and $\psi(y) \in \mathscr{G}$ ($\forall \, y \in \mathscr{Y}$). Then, the HISC can be defined as:

$$
\begin{aligned}
HSIC(P(X,Y),\mathscr{F},\mathscr{G}) \quad &= \mathbb{E}_{x,x',y,y'}[k(x,x')\ell(y,y')] \\
&+ \mathbb{E}_{x,x'}[k(x,x')]\mathbb{E}_{y,y'}[\ell(y,y')] \\
&- 2\mathbb{E}_{x,y}[\mathbb{E}_{x'}[k(x,x')]\mathbb{E}_{y'}[\ell(y,y')]]
\end{aligned}
\tag{2}
$$

where $k(x, x')$ and $l(y, y')$ are the kernels of the RKHSs $\mathscr{F}$ and $\mathscr{G}$ respectively. Note that, for some chosen kernels, $HSIC(P(X,Y),\mathscr{F},\mathscr{G})$ is zero (see [16]) if and only if $X \perp Y$.

The second step of the method is to estimate the HSIC based on the data samples that are available. Particularly, given independent and identically distributed samples $\{(x_i, y_i)\}_{i=1}^n$, a biased estimator of the HSIC that converges at a rate of $\mathcal{O}(n^{-1/2})$ is given by:

$$
HSIC(Z) = \frac{1}{n^2} trace(KHLH)
\tag{3}
$$

where $K, L, H \in \mathbb{R}^{n \times n}$, and where $K_{ij} = k(x_i, x_j), L_{ij} = k(y_i, y_j)$, and $H = I - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n$, with $\mathbf{1}_n$ being a vector of n ones. The cost of computing the above estimate is $\mathcal{O}(n^2)$. Note that the kernel that we used for the computations in this paper was the Gaussian kernel, that satisfies all the conditions for the results to hold.

As a final step, a statistical test is built on top of the metric defined above to determine the statistical significance of dependence/independence inferred from using the HSIC. For more details on the method, see [16,17].

2.2. Statistical testing for variable dependence

In order to state that two sets of (temperature) measurements $\boldsymbol{x}$ and $\boldsymbol{y}$ are dependent or independent, we build a hypothesis test set up as shown in Eq. (4) and Eq. (5).

$$
\mathscr{H}_0 : \mathbf{P}_{xy} = \mathbf{P}_x\mathbf{P}_y
\tag{4}
$$

$$
\mathscr{H}_1 : \mathbf{P}_{xy} \neq \mathbf{P}_x\mathbf{P}_y
\tag{5}
$$

where Eq. (4) is the null hypothesis stating that both variables are independent and Eq. (5) is the alternate hypothesis that indicates that both variables are dependent.

To evaluate the hypothesis test, we compute the HSIC between $\boldsymbol{x}$ and $\boldsymbol{y}$, and we compute the HSIC under the null distribution, i.e., assuming that they are indeed independent. To compute the distribution of the HSIC under the null hypothesis, we randomly permute one of the two series and compute the HSIC. As the permutation breaks the pair-wise interaction between $\boldsymbol{x}$ and $\boldsymbol{y}$, it provides a correct representation of the HSIC when $\boldsymbol{x}$ and $\boldsymbol{y}$ are independent. Repeating this step for multiple permutations, the empirical distribution of the HSIC under the null distribution is obtained. Finally, we use this distribution to evaluate the real HSIC. If the probability of the real HSIC between $\boldsymbol{x}$ and $\boldsymbol{y}$ under the null assumption is below a threshold, we reject the null hypothesis, i.e., we consider the variables as dependent. This threshold is usually 5 %, and the probability of the HSIC is the p-value of the statistical test given by

$$p = \frac{\sum_{\psi \in S} \mathbf{1}_{h_{xy} < h_{x\psi(y)}}}{|S|}. \tag{6}$$

In Eq. (6), ψ indicates the permutation operator, S is the set of all permutations and h_{xy} is the HSIC value of the measurements $\boldsymbol{x}$ and $\boldsymbol{y}$. If $p < 0.05, H_0$ is rejected and the variables are dependent with 95% confidence.

To highlight the benefits of this test in comparison to that of Pearson correlation coefficient, we compare the results of applying this test for the data sets in Figs. 1–3, where the correlation coefficient had suffered from non-linearities and outliers. For the data set in Fig. 1, the p-value is 0, which suggests that the null hypothesis can be rejected and hence the variables are dependent. The non-linear dependence in the data set is thus well captured by the statistical based on HSIC, in contrast to the Pearson correlation coefficient. For the data set in Fig. 2, the p-value is 0.905, which implies that the null hypothesis cannot be rejected, suggesting this test is resistant to outliers that give false positives for dependence. For the data set in Fig. 3, the p-value is 0, which suggests that the null hypothesis can be rejected and hence the variables are dependent. This is also a test for robustness against outliers, as the p-value is not affected by the single outlier.

2.3. A greedy algorithm

As mentioned previously, we assume that the initial setup can provide measurements of every zone to carry out this study, but that any of them can be removed afterwards to decrease costs in normal operation. Due to heat flows and proximity, the sensors with in a building can (very) often be dependent on one another. In this case, the p-value alone is not enough for the sensor placement problem, and a more closer look at the HSIC itself is necessary. Given a set of dependent sensors, the following algorithm can be further used to choose which sensor among them can be removed. Let $h(s_1, s_2)$ denote the value of HSIC between sensors s_1 and s_2. Further, let S be a set of sensors, and define the aggregated HSIC of a sensor $s \notin S$ w.r.t the set S as

$$r(s, S) = \sum_{s_i \in S} h(s, s_i).$$

Clearly, the larger the value of $r(s, S)$, the larger the dependence between the given sensor s with those from the set S. Conversely, the smaller its value, the further away is the dependence between s and the sensors in S. Next we define by n the number of sensors to be retained, and by U the set of all sensors. Given these definitions, the proposed methodology is defined by Algorithm 1.

Algorithm 1: Algorithm for ranking sensors.

Result: Sensor set S with the sensors to be retained
$S = \{\arg\max_{s \in U} r(s, U \setminus \{s\})\}$.
$U = U \setminus S$
while $|S| < n$ **do**
$\quad | \quad S = S + \arg\min_{s \in U} r(s, S)$
$\quad | \quad U = U \setminus S$
end

The proposed methodology is based on selecting one sensor at a time. For that, it considers two sets: a set $\mathscr{U}$ containing the sensors that have not yet been selected and a set $\mathscr{S}$ containing the sensors already selected.

The first step of the methodology is to select the sensor with the largest aggregated HSIC w.r.t. the other sensors, i.e., largest dependency with the other sensors. This first step guarantees that, if a single sensor is needed, the selected sensor maximizes the amount of thermal information. After this first step, $\mathscr{S}$ contains one sensor and $\mathscr{U}$ all sensors but the selected one.

If a second sensor is required, the methodology selects the sensor in $\mathscr{U}$ with the smaller aggregated HSIC w.r.t. the other sensors in $\mathscr{S}$. As the first sensor was already highly correlated, this step guarantees that the second sensor is the most independent from the sensor in $\mathscr{S}$, i.e., it guarantees that the second sensor includes information from the area where information is less redundant. Any sensor added to $\mathscr{S}$ should be appropriately deleted from $\mathscr{U}$ (denoted by $U = U \setminus S$ in the algorithm).

After this, the second step is repeated to include as many sensors in $\mathscr{S}$ as needed, each time picking the most independent sensor from the sensors in $\mathscr{S}$.

2.4. Properties of the method

The proposed method has several properties that make it a great alternative to model-based approaches; some of these properties improve upon the drawback of model-based approaches, while others are useful features in real-life applications. In this section, we describe and explain these properties.

2.4.1. Computational cost

A first advantage of the proposed method is its computational cost. In particular, given n sensors, the computational cost of the method is a deterministic value that can be computed as follows:

$$c_{\text{comp}} = \binom{n}{2} c_{\text{HSIC}} = \frac{n(n-1)}{2}, \tag{7}$$

where c_{HSIC} is the cost of computing the HSIC metric. Thus, the complexity of the algorithm is $\mathcal{O}(n^2)$, i.e. it has polynomial running time.

In contrast with the cost of the proposed method, model-based approaches have to rely on non-convex optimization problems to identify the building parameters. Such approaches are NP-hard [20] and have no convergence guarantees.

Besides the algorithm complexity, the proposed approach is fully parallelizable. In particular, as it computes the HSIC metrics for each pair of sensors, such a computation can be parallelized. Thus, assuming that we have available $\frac{n(n-1)}{2}$ cores/threads, the cost of the proposed approach is the cost of a single HSIC computation (which is typically a matter of seconds).

As a result, the proposed method improves upon the existing literature by providing a fast and reliable method with convergence guarantees. To validate this claims, in this article we will compare a model-based approach and the proposed method; for details, we refer to Section 5.2.

2.4.2. Initial sensor placement Independence

A key property of the method is that it is independent of the initial sensor placement. In particular, the initial sensor placement is determined by the distribution of zones. Thus, if we have a building with 5 thermal zones, e.g. rooms, we would initially have a sensor per zone. That is fixed and, given a building, has no room for change.

While it is true that within a zone a sensor could be placed in more than one location, the paper does not study the distribution of sensors within a zone as it is a science by itself where sensor placement must satisfy multiple requirements, e.g. it has to be located in an internal wall and/or out of solar radiation.

2.4.3. Multiple sensors per zone

Although the case of multiple sensors per zone is not specifically studied in this paper, the proposed method is still applicable for the same. In such a scenario, the method is model and zone agnostic and simply provides the optimal selection of sensors that contains the most information. Thus, even for the case of large areas with multiple sensors, the method will evaluate nonlinear relations from the sensors and select those sensors that maximize the system information.

Unlike the proposed method, standard model-based approaches are more dependant on sensor location as they require mapping sensors to zones/model state. Hence, in the case of having multiple sensors per zone, model-based approaches will struggle to select the optimal sensors.

2.4.4. Limited data requirements

To ensure that the proposed method works, the data-set must contain enough data to accurately represent the underlying probability distribution. Based on experimental results, this translates to having 100–200 well spaced and distributed data-points. In the case of buildings, that means that 1 week of data at hourly intervals would suffice: 1 week ensures that the variability in the data is enough; one hour intervals guarantees that there are 150+ data-points available.

The data requirements for the methods is also a nice property in comparison with model-based approaches. The latter, as they rely on a non-convex parameter identification procedure to estimate an approximated model, often require a larger number of data-points to reliably identify the dynamical model.

2.5. Applications of the method

The proposed method has multiple applications where it is expected to perform better than existing model-based approaches. A first application is the characterization of the building independently of the final usage of the sensors. Usually, sensor data has multiple applications in the context of building: verifying simulations, system identification, model predictive control, etc. For each of these applications, a different model is usually considered, i.e. simulating the building would normally consider a higher fidelity model than a controller. Thus, if a model-based approach is used to optimally select the sensors, it is clear that we would have as many model-based approaches (and in turn sensor selections) as final applications for sensor data. As only one sensor selection is possible, we need a method that can provide the optimal subset of sensors independently of the final application, i.e. model. The proposed method can satisfy this requirement and accurately select a subset of sensors independently of any underlying model.

In addition to that application, the proposed method is also very useful in the construction phase of the building. There, sensors are placed within the walls of the building and a method is needed to select the optimal location. At this stage, the building is not yet characterized and accurate models are often lacking. In this situation, the proposed approach can provide an accurate selection of the locations even when models are not available. Not only that, but the method can even be used to select the best location for each sensor within a given room as no model assumptions are needed.

2.6. Limitations of the method

Despite its numerous advantages, the proposed approach might have some limitations w.r.t certain applications. In this section, we discuss the potential limitations, we explain how these limitations compare to model-based approaches, and describe how these limitations can be overcome.

2.6.1. Data gathering

A potential drawback of the proposed method is that it requires the evaluation of sensors in multiple places. This requirement can be infeasible in several real-life scenarios as the cost of gathering temperature data in multiple locations can be difficult and expensive. To tackle this, one could use high fidelity models to gather reliable data from simulations. Initiatives such as the IDEAS [21] Modelica library or the BOPTEST project [22] may be of great help for this task. These libraries facilitate the modeling of not only the heat transfer between the different building components, but also of the air exchanged among the building zones.

2.6.2. Optimality of the solutions

A second potential limitation of the proposed method is that the obtained solutions are not optimal in a strict sense. However, it is important to note that, since model-based approaches rely on a parameter identification process that is non-convex, the obtained solutions of traditional model based approaches cannot guarantee global optimal minima either. However, unlike model-based approaches, the proposed method is faster and avoids the lengthy and costly parameter identification step.

2.6.3. Cost evaluation

Finally, the proposed model does not provide a direct link to the cost vs. gains of adding a new sensor. In particular, with a model-based approach, we could compare the gains due to the increased accuracy of adding one extra sensor. However, performing the cost benefit analysis w.r.t improved accuracy is nevertheless not possible even with model based methods.

3. Validation based on grey-box models with Kalman filter

For validating the proposed methodology, we implement a conventional model based approach. The model used here is based on a grey-box building model structure. Grey-box modeling is a typical approach for optimal control in buildings because of its balanced trade-off between accuracy, robustness and complexity [23,4,24].

3.1. Grey-box model

First, an *RC* network is built by connecting thermal resistances and capacitances as lumped parameters. An initial guess of the parameters is derived from actual physical properties of the building system. Then, monitoring/measurement data is used to train the model to identify the parameters of the model. For the purpose of validation, a *centralized multi-zone* grey-box model is identified for each of the test cases envisaged in this study. A multi-zone model is required because the model should output the temperature estimates of each individual zone in the building. Eq. (8)

shows the dynamics driving the variations of each of the zonal temperatures as described by a grey-box model:

$$C_z \frac{dT_z}{dt} = \dot{Q}_{hz} + \sum_{a \in A} \frac{T_a - T_z}{R_{az}} + \frac{T_e - T_z}{R_w} + g\dot{Q}_{Sun}, \qquad (8)$$

where T_z is the temperature of zone z with an air thermal capacitance of C_z and $\dot{Q}_{hz}$ is the heat released to that zone for building acclimatization. The second term at the right side of the equation represents the heat exchanged between zone z and an adjacent zone a through the thermal resistor R_{az}, where A refers to the set of all adjacent zones connected to z. Notice that doors remain either opened or closed during the experiments, and therefore a single thermal resistor is considered enough to model the heat transmitted between the zones. T_e denotes the exterior temperature, and the corresponding term accounts for the heat exchange with the zone through a wall thermal resistor R_w. The last term indicates the thermal power added to the zone due to solar irradiation, $\dot{Q}_{Sun}$, where g is the parameter representing the solar admittance the wall glazing-structure of that zone. Other thermal states may be included in a single zone model as well, such as an internal wall state, or an extra internal state representing furniture and indoor walls. For the sake of simplicity, we limit the description of the grey-box models to this extent.

3.2. Kalman filter definition

We use the Kalman filter [25] along with the grey-box model explained above, to track the hidden states, i.e., states where sensor measurements are (considered) absent. For this purpose, the continuous time state space model described by Eq. (8) is first discretized in time using *zero order hold* on the controllable heat inputs to the zones $\dot{Q}_{hz}$ to obtain the following set of equations:

$$\begin{aligned} \boldsymbol{x}_k &= A_d \boldsymbol{x}_{k-1} + B_d \boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1} \\ \boldsymbol{y}_k &= \boldsymbol{C_d} \boldsymbol{x}_k + \boldsymbol{v_k}, \end{aligned} \qquad (9)$$

where A_d, B_d, C_d and D_d are the matrices of the discretized dynamics. Further, $\boldsymbol{x}_k$ is the vector of the states at time step k, i.e., the temperature of each zone, $\boldsymbol{u}_k$ represents the input (heat input, solar irradiation, and external temperature) to each zone at time step k, $\boldsymbol{w}$ represents the model error and $\boldsymbol{v}$ is the measurement noise. Note that the matrix C_d encodes the number of states for which measurements are available. For the derivation of the state space matrices, we refer to [26]. Next, the Kalman filter iterations are performed as follows:

- First *a-priori covariance is calculated as* $P_k^- = A_d P_{k-1}^+ A_d^T + Q_d$, *where Q_d is the covariance matrix of the model error.*
- Next, the Kalman gain is calculated as $K_k = P_k^- C_d^T (C_d P_k^- C_d^T + R_d)^{-1}$, *where R_d is the covariance matrix of the measurement noise.*
- The a priori state estimate is then $\hat{\boldsymbol{x}}_k^- = A_d \hat{\boldsymbol{x}}_{k-1}^+ + B_d \boldsymbol{u}_{k-1}$.
- The a posteriori state estimate is given by $\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + K_k(\boldsymbol{y_k} - \boldsymbol{C_d}\hat{\boldsymbol{x_k}}^-)$
- Finally, the a posteriori covariance is calculated with $P_k^+ = (I - K_k C_d) P_k^-$

These steps are repeated for every time step. Note that P_0 (or P_0^+), $\hat{x}_0, Q_d$ and R_d need to be initialized properly, based on measurements and model errors.

3.3. Implementation

We implement the Kalman filter as follows:

- We start from a case with temperature readings from the sensors in all zones. Then, we simulate the values of the indoor temperatures using all measurement updates in every iteration.
- The error from the above procedure is assumed to be the minimal achievable error as the state is estimated with the maximum available information. This error is used as the reference for comparison.
- Then, for each case of sensor combinations, the vector $\boldsymbol{v}$ and matrix C_d is updated to have measurements only from the *retained* sensors. This means that the model is updated only with the measurements available from the included sensors, and thus, the estimates at the other zones need to be done in the absence of sensor measurements.
- The error is calculated for all estimates (including those for which sensor measurements were not taken into account), and is compared against the reference error.
- Based on the errors calculated above, ranks are assigned. For instance, if from a set of 5 sensors, 4 have to be retained, there are 5 ways of doing so, and each of the 5 combinations is given a rank using the errors calculated above.
- These ranks are then used to validate the combination picked by the statistical method.

4. Case studies

The novel sensor placement methodology has been implemented in two test cases. The first one consists of real data coming from the Twin Houses experiment [27]. The added value of using real data is that the building is prone to actual disturbances and measurement noise leading to a realistic scenario to run the sensor placement methodology. The second one uses the Building Optimization Testing framework (BOPTEST) framework [22] to interact with a virtual building that allows the generation of different evaluation data sets.

4.1. Twin house data

The twin houses experiment is a well-known experiment in the building modeling community that has been used by different authors e.g., [28–30], to prove their modeling techniques. It provides detailed measurement data of an experimental dwelling in Fraunhofer IBP (Holzkirchen, Germany) with not only the temperature of every zone, but also the heat input released to each. Two data-sets are available for about 40 days each: the first one between August and September 2013, the second one between April and May 2014. We decide to use the second one because the lower outdoor temperatures are more representative of a heating season and because the blinds were open during such experiment, leading to a more realistic scenario accounting for the impact of irradiance.

The house consists of a cellar, a ground floor and an attic. The layout of the house is shown in Fig. 4. The adjacent zones (cellar and attic) are held at a constant temperature of 22 °C. Also the three rooms at the north face of the dwelling are sealed from the others and kept at 22 °C. On the contrary, the rest of the rooms at the ground floor maintain their doors opened, substantially facilitating heat exchange between them.
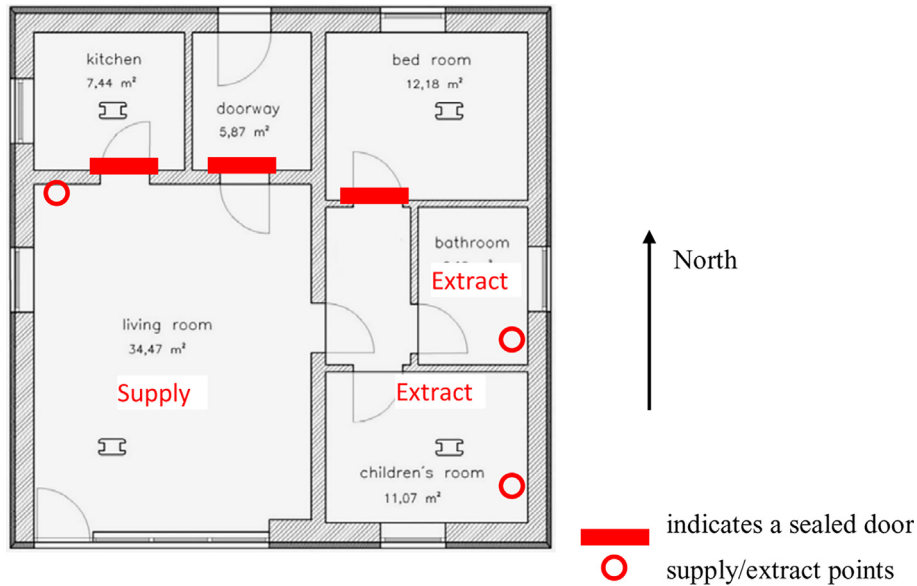
Fig. 4. Layout of the experimental dwelling (called Twin house) ground floor.

4.2. BOPTEST variable air volume building

For the second case study, a simulation model is chosen to generate and test the methodology with different sets of data using the BOPTEST framework. The Building Optimization Testing framework (BOPTEST) is an open source initiative offering an ensemble of detailed building models to be used as a common benchmark to evaluate different advanced control algorithms in buildings. The data sets used here are generated by implementing different heating input profiles in one of the BOPTEST building test cases, while keeping the same disturbances, i.e., weather, occupancy and internal gains. This allows us to fairly evaluate the impact of the data.

We use the BOPTEST commercial multi-zone air-based test case that comprises the variable air volume building model from the Buildings Modelica library [31]. This freely accessible model is fully documented and has been accommodated for its use in BOPTEST. However, it is not yet considered a final test-case version thus slight changes may apply for future BOPTEST users.

A schematic of the building energy system is shown in Fig. 5. The building consists of five thermal zones: core, south, east, north and west, an air based HVAC system, and air flow models of the building leakage and between the zones.

The first set of data (called *baseline*) is generated from the simulation of the model using the default baseline controller. Two weeks are simulated in the winter period and with the weather conditions coming from a TMY3 file of the location of Chicago, Illinois. Different variations of the baseline controller are realized in order to obtain richer dynamic data. These data variations are gathered from the same time period and ambient conditions, but with different controllable inputs.

In the first variation (called *1 day*) the lower bound temperature set-point of the comfort range is excited to its upper bound for one zone at a time. The excitation at each zone persists for one day. Then it stops and a new zone starts being excited. The goal is to increase the temperature of one room at a time to its maximum allowed comfort bound. This may increase the energy use, but should not significantly affect comfort since the set-points remain within the actual allowed comfort range. As a result, we stimulate the thermal interaction between zones to unlock any possible dependencies that could be hidden in the baseline scenario due to its intrinsic steady-state functioning.

The second variation (called *constant*) supposes only a slight modification to the first one. An excitation period remains where every zone is excited one day at a time. However, an extra day with the same constant temperature in all zones is included as well.
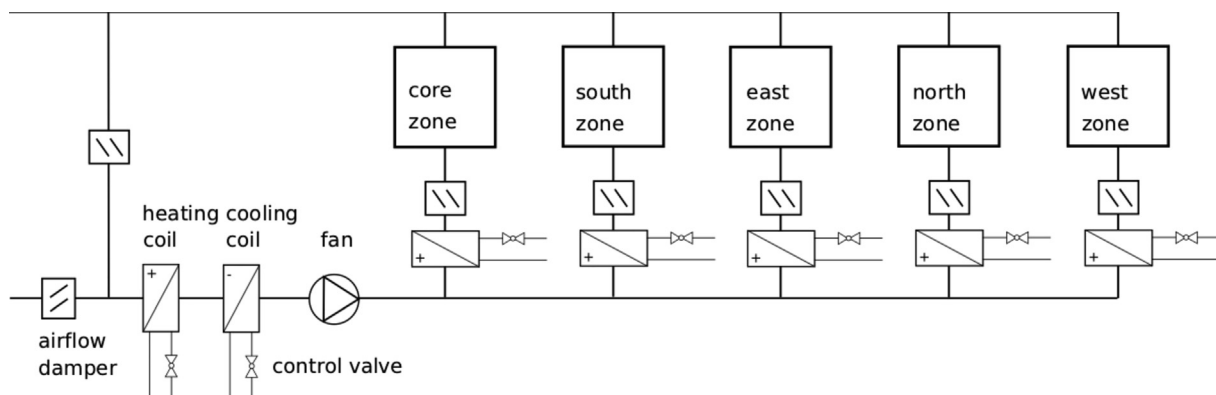

Fig. 5. Schematic of the BOPTEST variable air volume building system used to test the methodology. The model consists of a five-zone building envelope and an air-based HVAC system with reheating. Figure obtained from the documentation of the Modelica Buildings Library [31].

During the latter period, all zone temperatures are kept at the middle of the comfort range. This practice is found to be favorable for the grey-box parameter estimation process.

Finally, in the last variation (called *random*) from the baseline controller, a more variable set of input signals is applied. In this case, the comfort range is collapsed into a unique temperature set-point per zone. These set-points are randomly drawn from a uniform distribution ensuring that there is no dependency between the heat delivered to the rooms (pair-wise).

4.3. Limitations of the case study

It is important to note that the case studies only consider a limited number of sensors. While this might seem a limitation, it is actually a design choice to ensure that the proposed method can be properly validated: if we were to consider large-scale complex buildings, while the proposed method would not suffer from scalability issues, the traditional methods required to compare the proposed method will. In detail, as explained in Section 2.4.1, the proposed method is fully parallelizable and has a complexity of $\mathcal{O}(n^2)$, where n is the number of sensors. In contrast, standard model-based approaches require solving NP-hard problems and have no guarantee of convergence. As a result, in this study, we have limited the analysis to small-scale buildings where the standard model-based approach is tractable.

A second potential limitation of the case study is the fact that we only consider multi-zone models but not open areas with multiple sensors per zone. The limitation in the case study is a design choice to compare the proposed approach with a state-of-the-art model-based approach. Particularly, using a model based approach for large-scale open areas with multiple sensors is very difficult as it requires mapping multiple sensors to a single room/model state. By contrast, the proposed method is model agnostic and provides the optimal selection of sensors independently of the underlying model. Thus, even for the case of large-scale open areas with multiple sensors, the method will evaluate nonlinear relations from the sensors and select those sensors that maximize the system information.

Finally, a third potential limitation of the case study is that the study only provides the optimal sensors on average. In particular, although buildings employ different working regimes, e.g. day and night, and each mode might require different number of sensors, the case study only considers the optimal number of sensors when all modes are considered. That being said, this is not a limitation of the model but a limitation of the case study. In fact, the

advantage of the method is that, in a data driven way, without prior knowledge of the occupancy and heating schemes, the method can learn these patterns from probability distributions - it will be able to group the sensors that vary more during the day from the sensors that vary more during the night.

5. Results

In this section, we will present the results of applying our methodology to the two case studies. Since the twin–house setup has only two isolated thermal zones, we first use the statistical test method with this data to verify whether it can indeed detect the two thermal zones. Subsequently, we apply the sensor selection procedure on the BOPTEST building.

5.1. Twin-house

Applying the statistical testing method to the twin-house data reveals the following observations (see Fig. 6)

- We cannot claim dependence between the heat inputs of the south zones and the temperatures of the north zones (*p*-value > 0.8), i.e., we observe that the null hypothesis (4) cannot be rejected. These pairs of sensors are hence *likely* to be independent. However the *p*-value is 0 for the heat inputs and the temperatures in the south zones.
- For the temperature sensors, the *p*-values are 0, suggesting that we cannot reject the null hypothesis that any of the sensor pairs are independent. However, inspecting Fig. 6 reveals that the HSIC is higher for the sensors pairwise with in north or with in south zones. For a pair with one sensor from north and the other from the south, the HSIC is considerably smaller.

Note that although from the illustration in Fig. 4, the rooms in the North are thermally separated by walls, the sensors are found to be dependent. This is because, as mentioned in Section 4.1, the temperatures of the rooms in the north were maintained at 22 degrees. The methodology has hence captured the correlation in the heating methods used for the three rooms.

The results from the statistical test show that the probabilistic dependence of each sensor on the other is similar, and based on the proximity of the values, they can thus not be ranked in a statistically meaningful way for importance. Hence, with in controlled thermal zone (containing living room, bathroom and children's room) in the south, if only two (out of three) sensors were to be
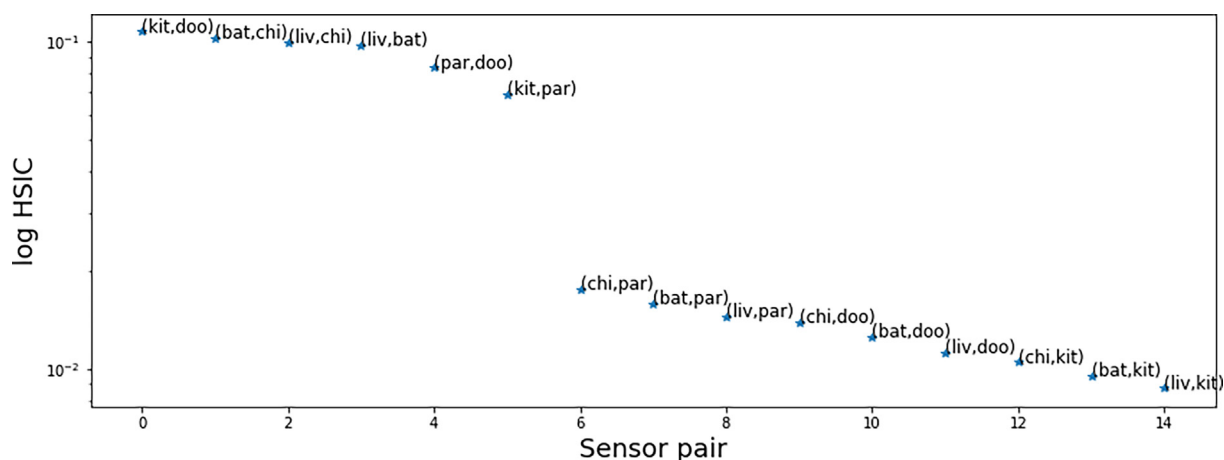

Fig. 6. HSIC readings for the pairs of sensors in the south area of the twin house. The label of each of the markers shows the first three letters of the zone (kitchen, doorway, children's room, living room, bathroom, and bedroom).

Table 1
HSIC and Kalman rank for BOPTEST.

Data set		1 sensor	2 sensors	3 sensors	4 sensors
Baseline	HSIC	East	East South	East South North	East South North West
	Kalman rank	2	1	1	1
1 day	HSIC	West	West South	West South North	West South North East
	Kalman rank	1	1	2	1
Constant	HSIC	West	West South	West South North	West South North East
	Kalman rank	1	1	2	1
Random	HSIC	Core	Core South	Core South North	Core South North East
	Kalman rank	4	7	5	2

retained, the results suggest that any of the three sensors can be removed. It is important to note that for accurate tracking of comfort, at least one sensor should be retained in each of the isolated thermal zones (one in north and one in south in this case).

Fig. 6 shows that the HSIC of the kitchen, doorway and parent's room are much higher than their HSICs with other rooms. Likewise, the HSIC of bathroom, children's room and living room are again much higher than their HSICs with other rooms. We are hence able to clearly identify the thermal zones without having prior knowledge of the building.

Although these results seem straightforward for a simple building like the one of this experiment, they suggest that the proposed methodology might be able to distinguish between areas with different thermal regimes for large-scale buildings. In particular, considering the $\mathcal{O}(n^2)$ computational complexity of the model (with n being the number of sensors), the proposed approach should easily scale to other applications even if they are based on large-scale buildings. Ideally, one would test such an assumption by analyzing the behavior of the algorithm for a large-scale example; as this is out of the scope of this paper, we leave such a study as future research.

5.2. BOPTEST variable air volume building

Here we present the results of applying the novel methodology on the data from the building described in Section 4.2. For all data sets investigated, the statistical tests showed that the sensors are all dependent on each other. This means that all the thermal zones are connected to one-another. We thus look at the HSIC value and apply Algorithm 1.

To evaluate the method, we compare each subset selection S given by the proposed method with the rank obtained by the selected subset S using the grey-box models with Kalman filter, e.g. if the rank of the subset using the Kalman method is 1, the proposed method and the Kalman method agree 100%. To build a rank for the Kalman filter method, we consider all possible subsets of sensors and we rank them by their RMSE value. The results are presented in Table 1. The following analysis can be made.

- For 3 of the 4 data sets (baseline, 1 day, and constant), the method is in perfect agreement with the Kalman method. In particular, the ranking of the Kalman method is 1 in 75% of

the cases and 2 in the remaining 25%. However, the RMSE difference between the ranks 1 and 2 are usually not significant (< 0.01 in RMSE), especially considering that for the 2 and 3 sensor cases there are 12 possible choices.

- While the proposed method disagrees with the Kalman method in the case of the random data set, this behavior is expected as the data set is very unrealistic and does not represent real life conditions in the building. Particularly, unlike the three other data sets that are based on typical HVAC settings, the random data set is driven by a very a-typical data set that breaks the usual behavior of the thermal interaction between the zones. Moreover, due to this unusual data set, the RMSE's of the Kalman method for the random data set are comparatively large and are not reliable in this case to establish meaningful ranks. Nonetheless, the statistical method is still able to give a reasonable estimate of the sensor combination, where except for the first sensor to be retained, the subsequent sensors are chosen in the same order as the other cases.

6. Conclusions

This paper describes, evaluates and validates a data-driven model free approach to rank the importance of sensor locations, and hence an advice system for the placement/eventual maintenance of the sensors. The approach has several advantages: it can be deployed expert free, requires less computational time and is more generalizable than the model-based approaches. Validation was performed in two case studies, and different data sets against a well-known methodology based on grey-box models and Kalman filters. It is shown that the novel approach based on HSIC and statistical tests can reliably identify connected thermal zones, and further rank them in the order of priority for scheduled maintenance.

As future research, firstly, we will validate the proposed approach in large-scale buildings. Such a study was out of the scope of this paper because we did not have benchmark methods for large-scale buildings. Thus, as future research we will propose alternative methods for validation and test the proposed approach for buildings with a larger number of sensors. Additionally, we will extend the method to evaluate the value of adding an extra sensor. While the current method answers the sensor placement question of which sensors can be removed, one could potentially use the method in combination with a state estimation model (that esti-

mates hidden states) or detailed white box models that can generate sensor data, to identify where sensors can be placed next. This can also be potentially supplemented with other absolute metrics for information gain w.r.t adding sensors such as Fischer information matrix, confidence interval and entropy studies to give a better view on the minimum number of sensors to be retained in a building.

CRediT authorship contribution statement

Gowri Suryanarayana: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Javier Arroyo:** Methodology, Software, Resources, Investigation, Data curation, Writing - original draft, Writing - review & editing. **Lieve Helsen:** Supervision, Writing - review & editing. **Jesus Lago:** Conceptualization, Methodology, Investigation, Writing - review & editing, Supervision.

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.

Acknowledgement

References

[1] International Energy Agency and International Partnership for Energy Efficiency Cooperation. Building Energy Performance Metrics. Supporting Energy Efficiency Progress in Major Economies. Technical report, OECD/IEA, April 2015..

[2] Abdul Afram, Farrokh Janabi-Sharifi, Theory and applications of HVAC control systems – a review of model predictive control (MPC), Building and Environment 72 (2014) 343–355.

[3] Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 Amending Directive 2010/31/EU on the Energy Performance of Buildings and Directive 2012/27/EU on Energy Efficiency (Text with EEA Relevance). Technical report, Official Journal of the European Union, 2018..

[4] R. De Coninck, L. Helsen, Practical implementation and evaluation of model predictive control for an office building in Brussels, Energy and Buildings 111 (2016) 290–298.

[5] Arnold Janssens. Reliable building energy performance characterisation based on full scale dynamic measurements. Report of Subtask 1a: Inventory of full scale test facilities for evaluation of building energy performances. Technical report, International Energy Agency, EBC Annex 58, 2016..

[6] Arnold Janssens. Reliable building energy performance characterisation based on full scale dynamic measurements. Report of Subtask 1b: Overview of methods to analyse dynamic data. Technical report, International Energy Agency, EBC Annex 58, 2016..

[7] Dirk Saelens, Glenn Reynders. Reliable building energy performance characterisation based on full scale dynamic measurements. Report of Subtask 4b: Towards a characterisation of buildings based on in situ testing and smart meter readings and potential for applications in smart grids. Technical report, International Energy Agency, EBC Annex 58..

[8] V. Tzoumas, A. Jadbabaie, G.J. Pappas, Sensor placement for optimal kalman filtering: Fundamental limits, submodularity, and algorithms, in: 2016 American Control Conference (ACC), July 2016, pp. 191–196..

[9] Germán Ramos Ruiz, Carlos Fernández Bandera, Tomás Gómez-Acebo Temes, Ana Sánchez-Ostiz Gutierrez, Genetic algorithm for building envelope calibration, Applied Energy 168 (2016) 691–705..

[10] Max Emil S. Trothe, Hamid Reza Shaker, Muhyiddine Jradi, Krzysztof Arendt, Fault isolability analysis and optimal sensor placement for fault diagnosis in smart buildings, Energies 12 (9) (2019) 1–12..

[11] FEDOROV Valerii Vadimovich, Theory of Optimal Experiments, Elsevier, 2013..

[12] Cristian R. Rojas, James S. Welsh, Graham C. Goodwin, Arie Feuer, Robust optimal experiment design for system identification, Automatica 43 (6) (2007) 993–1008.

[13] Costas Papadimitriou, James L. Beck, Au Siu-Kui, Entropy-based optimal sensor location for structural model updating, Journal of Vibration and Control 6 (5) (2000) 781–800.

[14] Jie Zhang, Kristof Maes, Guido De Roeck, Edwin Reynders, Costas Papadimitriou, Geert Lombaert, Optimal sensor placement for multi-setup modal analysis of structures, Journal of Sound and Vibration 401 (2017) 214–232.

[15] Duwaraka Yoganathan, Sekhar Kondepudi, Balaji Kalluri, Sumanth Manthapuri, Optimal sensor placement strategy for office buildings using clustering algorithms, Energy and Buildings 158 (2018) 1206–1225.

[16] Arthur Gretton, Olivier Bousquet, Alex Smola, Bernhard Schölkopf, Measuring statistical dependence with Hilbert-Schmidt norms, in: International Conference on Algorithmic Learning Theory, Springer, 2005, pp. 63–77.

[17] Arthur Gretton, Kenji Fukumizu, Choon H Teo, Le Song, Bernhard Schölkopf, Alex J. Smola, A kernel statistical test of independence, in: Advances in Neural Information Processing Systems, 2008, pp. 585–592..

[18] K.R. Parthasarathy, Probability Measures on Metric Spaces, Elsevier, 1967.

[19] Nachman Aronszajn, Theory of reproducing kernels, Transactions of the American Mathematical Society 68 (3) (1950) 337–404.

[20] Prateek Jain, Purushottam Kar, Non-convex optimization for machine learning, Foundations and Trends in Machine Learning 10 (3–4) (2017) 142–336.

[21] Filip Jorissen, Glenn Reynders, Ruben Baetens, Damien Picard, Dirk Saelens, Lieve Helsen, Implementation and verification of the IDEAS building energy simulation library, Journal of Building Performance Simulation 11 (2018) 669–688.

[22] David Blum, Filip Jorissen, Sen Huang, Yan Chen, Javier Arroyo, Kyle Benne, Yanfei Li, Valentin Gavan, Lisa Rivalin, Lieve Helsen, Draguna Vrabie, Michael Wetter, Marina Sofos, Prototyping the BOPTEST framework for simulation-based testing of advanced control strategies in buildings, in: Building Simulation Conference, Rome, 2019..

[23] Samuel Prívara, Jíí Cigler, Zdeněk Váa, Frauke Oldewurtel, Carina Sagerschnig, Eva Žáčeková, Building modeling as a crucial part for building predictive control, Energy & Buildings 56 (January 2013) 8–22.

[24] Javier Arroyo, Bram van der Heijde, Fred Spiessens, Lieve Helsen, A Python-based toolbox for model predictive control applied to buildings, in: Proc. of the Intelligent Building Operations (IBO) Workshop in Purdue, Purdue University, West Lafayette, Indiana, July 2018..

[25] Dan Simon, Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches, John Wiley & Sons, 2006.

[26] C.T. Chen, Linear System Theory and Design, The Oxford Series in Electrical and Computer Engineering. Oxford University Press, Incorporated, 2013.

[27] Paul Strachan, Twin Houses Empirical Validation Dataset: Experiment 2. Technical report, University of Strathclyde (2016).

[28] Roel De Coninck, Fredrik Magnusson, Johan Åkesson, Lieve Helsen, Toolbox for development and validation of grey-box building models for forecasting and control, Journal of Building Performance Simulation 9 (3) (May 2016) 288–303.

[29] Paul Strachan, Katalin Svehla, Ingo Heusler, Matthias Kersken, Whole model empirical validation on a full-scale building, Journal of Building Performance Simulation 9 (4) (2016) 331–350.

[30] F. Monari, P. Strachan, Characterization of an airflow network model by sensitivity analysis: parameter screening, fixing, prioritizing and mapping, Journal of Building Performance Simulation 10 (1) (2017) 17–36.

[31] Michael Wetter, Wangda Zuo, Thierry S. Nouidui, Xiufeng Pang, Modelica buildings library, Journal of Building Performance Simulation 7 (4) (Jul 2014) 253–270.