From in-situ measurement to regression and time series models: An overview of trends and prospects for building performance modelling

Massimiliano Manfren, and Benedetto Nastasi

ARTICLES YOU MAY BE INTERESTED IN

Synthesis of Y\textsubscript{3}Ba\textsubscript{5}Cu\textsubscript{8}O\textsubscript{18} superconductor by auto-combustion reaction
AIP Conference Proceedings 2123, 020085 (2019); https://doi.org/10.1063/1.5117012

Investigation of high-voltage silicon solar modules
AIP Conference Proceedings 2123, 020103 (2019); https://doi.org/10.1063/1.5117030

Improvement of electrical features of SnO\textsubscript{2} based varistor doped with Al\textsubscript{2}O\textsubscript{3}
AIP Conference Proceedings 2123, 020097 (2019); https://doi.org/10.1063/1.5117024
From in-situ measurement to regression and time series models: An overview of trends and prospects for building performance modelling

Massimiliano Manfren 1, a), Benedetto Nastasi 2

1 Faculty of Engineering and Physical Sciences, University of Southampton, United Kingdom
2 Department of Architecture Engineering & Technology (AE+T), TU Delft University of Technology, Netherlands

a) Corresponding author: M.Manfren@soton.ac.uk

Abstract. Data analysis methodologies are crucial to learn insights from data and to create more trust in the assumptions used for energy performance assessment. Indeed, continuous performance monitoring should become a more diffuse practice in order to improve our design and operation strategies for the future. This is an essential step to reduce incrementally the gap between simulated and measured performance. In fact, assumptions in simulation represent a significant source of uncertainty when estimating the energy performance of buildings. This uncertainty affects decision-making processes in multiple ways, from design of new and refurbished buildings to policy making. The research presented aims to highlight potential links between experimental approaches for test-facilities and methods and tools used for continuous performance monitoring, at the state of the art. In particular, we start by exploring the relation between in-situ measurement of thermal transmittance (U) and regression-based monitoring approaches, such as co-heating test and energy signature, for heat load coefficient (HLC) and solar aperture (gA) estimation. After that, we highlight some recent developments in simplified dynamic energy modelling using lumped parameter models. In particular, we want to underline the scalability of these techniques, considering relevant issues in current integrated engineer design perspective. These issues include, among others, the necessity of limiting the number of a sensors to be installed in buildings, the possibility of employing both experimental and real operation data (and compare them with design data as well) and, finally, the possibility to automate performance monitoring at multiple scales, from single components, to individual buildings, to building stock and cities.

Keywords: building performance monitoring; in situ measurements; co-heating test; energy signature; regression models; time series models.

INTRODUCTION

Simple and robust data analysis methodologies are crucial to learn insights from measured data and to reduce the performance gap in building stock [1], considering also the relevant impact of human behaviour [2]. In general the assumptions in building performance simulation are characterized by a significant level of uncertainty [3]. This uncertainty affects estimates of energy performance of buildings and, consequently, decision making processes in multiple ways, from design of new and refurbished buildings [4] to energy planning and policy making. The effect of uncertainty is huge on sizing of building technologies and technical systems [5], as well as on indoor environmental quality assessment [6] (and critical in cases where long-term preservation of appropriate internal microclimate conditions is essential [7], considering also acoustic performance [8] as well as other properties of building).

The possibility of linking transparently design phase performance estimates and measured data in operation is a major challenge today [9], requiring a careful consideration of the methods and tools available at the state of the art, essentially to avoid duplication of efforts and redundancy. The gap between simulated and measured conditions can be huge as acknowledged by numerous studies [1]. The control of building performance needs to be oriented towards
continuous monitoring across its life span, requiring models and experiments to check the status of different technologies, from construction technologies up to technical systems, including performance indicators for heating systems [10], electricity and interaction with the grid [11] and primary energy consumption [12]. In order to verify the robustness of solutions in terms of performance, an increasing research effort has been devoted to the simulation of more realistic operating conditions already in design phase [9], using parametric and/or probabilistic approaches. In any case, continuous performance monitoring should become a more diffuse practice in order to improve our design and operation strategies for the future [13] and to handle the energy transition in existing buildings [14].

The research presented aims to highlight potential links between experimental approaches for test-facilities and methods and tools used for continuous performance monitoring, at the state of the art. Further extensions are also possible with advanced computing tools, dealing with simulation of interacting systems [15]. Data analysis techniques can offer deeper insights of building performance and their effects on occupants but they have to be simple, robust and scalable, so that they can be easily incorporated in innovative technologies such as BEMS, which are necessary to defined optimal operational strategies of heat pumps [16] or more complex hybrid systems [5], managing the integration of renewable energy in terms of direct use [17] or when interacting with storages [18], to reduce the cost of energy [19]. In all these cases, the ability to define robust performance estimates is crucial. While black-box approaches are possible [20], we consider more appropriate the choice of grey-box, i.e. physical-statistical, approaches, to learn useful insights from data and to link with current state of the art approach for component scale [3] and building scale analysis [21].

**METHODOLOGY**

Many sophisticated simulators are available today for building performance modelling. However, a gap between simulated and measured performance is generally observed [1], due to the relevant uncertainties of the assumptions introduced in building performance assessment. In order to become aware of the potential variability of building performance outcomes, it is necessary to introduce parametric or probabilistic simulation approaches. Physical-statistical lumped parameter models can help reducing computational time in simulation (forward modelling) and enable also inverse parameter estimation (inverse modelling) [22]. In the recently introduced ISO 52016-1 standard [23] (which supersedes the consolidated ISO 13790 [24]) a lumped modelling approach for walls is proposed, with a classification based on walls’ stratigraphy. This approach is part of the strategy of ISO 52000 framework [25] that retains and updates other standards, for example ISO 13789 [26] for building fabric performance, ISO 6946 [27] for construction component stationary thermal performance, and ISO 13786 [28] for dynamic construction components thermal performance. The approach proposed at the normative level for building performance simulation (ISO 52016-1) is substantially similar, in principles, to research focused on lumped parameter modelling using resistance-capacitance (RC) analogy [29] and analytical calculations [30, 31]. The conversion of RC models in state-space form and then in time series models is described in detail in recent literature [32, 33]. More in general, the use of reduced order models for building performance simulation is an active research field at present, with multiple possible applications [34]. In this sense, the substantial compatibility among the formulations proposed in recent technical standards (for building performance simulation) and the more generally applicable formulations that can be found in literature is opening new perspectives for the future research in the area.

Moving from dynamic simulation models to models aimed at identifying stationary properties, we consider particularly interesting the relation between ISO 9869 [35] method for in-situ measurement of thermal transmittance (U), as well as its extensions [36], and regression-based monitoring approaches, such as co-heating test [37] and energy signature [38, 39], for heat load coefficient (HLC) and solar aperture (gA) estimation [40, 41]. All these approaches are already consolidated at the state of the art but further efforts are necessary for their seamless integration, in particular with respect to the comparability of measurements with design phase simulations and the robustness of the quantities identified (estimates of physical parameters) by means of regression on operation phase data.

The importance of creating standardized procedures for large scale statistical analysis of building data has been stressed, for example, by institutions such as NIST (National Institute of Standards and Technology) in the US [42]. In fact, in the next few years, the possibility of collecting and processing data at large scale will be crucial for informing future policies for built environment [43].
OVERVIEW OF TRENDS AND PROSPECTS FOR ENERGY PERFORMANCE MONITORING IN BUILDINGS

Start from the considerations reported in the previous Section we report an overview of approaches for energy performance monitoring and modelling in buildings, basically following a bottom-up logic, from individual construction components (U value), to building fabric assembly (HLC), up to the meter level (energy signature). In the following subsections we will describe first the analogies among methods from U value estimation and co-heating tests, up to energy signature. After that, we will present some recent advances in simplified dynamic modelling of building components and thermal zones (aggregate behavior of building fabric). The general goal of this overview on modelling research developments is highlighting the simplicity and scalability of techniques and the need for harmonization of experimental procedures to increase robustness of estimates (which dependents critically on the amount and quality of data collected) and reduce cost.

From in situ measurement to regression-based approaches (static)

In this section we describe briefly the steps necessary to link U value estimation, HLC estimation and energy signature. We can start from the stationary thermal energy transfer for a multi-layered construction component that can be calculated, using ISO 6946 [27], as follows:

\[ U = \frac{1}{R} = \frac{1}{R_{si} + \sum \frac{s_i}{\lambda_i} + R_{se}} = \frac{1}{\left( \frac{1}{h_{si}} + \sum \frac{s_i}{\lambda_i} + \frac{1}{h_{se}} \right)} \]  

\[ q = UA\Delta T = UA(T_i - T_e) \]  

where \( U \) (W/m²K) is thermal transmittance, \( R \) (m²K/W) is total thermal resistance, \( \lambda_i \) (W/m²K) is thermal conductivity of layer \( i \), \( s_i \) (m) is depth of material layer of layer \( i \), \( R_{si} \) (m²K/W) is thermal resistance on internal side, \( R_{se} \) (m²K/W) is thermal resistance on external side, \( h_i \) (W/m²K) is thermal heat transfer coefficient on internal side, accounting for convection and radiation, \( h_e \) (W/m²K) is thermal heat transfer coefficient on external side, accounting for convection and radiation, \( q \) (W/m²) is heat flux per unit of surface in stationary conditions, \( \Delta T = T_i - T_e \) is internal air temperature, \( T_e \) (°C) is external air temperature.

In order to estimate U value from in situ measurement (with an experimental setup such as the one reported later in Figure 1) following ISO 9896 [35] methodology, we consider the subsequent averaging method:

\[ U = \frac{\sum q_{in}}{n} \]  

where \( U \) (W/m²K) is thermal transmittance experimentally determined, \( q_{in} \) (W/m²) is the heat flux entering in the wall, \( n \) is the number of data points (at least 72h of monitoring data are necessary according to the standard).

If we consider then the daily average heat flux (instead of instantaneous measurements for a single component) to maintain a thermal zone in constant internal temperature conditions, we can formulate a simplified representation of the energy balance of the zone as shown in Equation 4. This equation corresponds to simplification used in co-heating test method [37]:

\[ q_h = HLC\Delta T - gA_{sol}l_{sol} - q_{int} \]  

where \( q_h \) (kW) is average daily heat flux, \( HLC \) (W/K) is heat loss coefficient, \( gA_{sol} \) (m²) is solar aperture, \( l_{sol} \) is the average daily solar irradiation (on horizontal or south plane), \( q_{int} \) (kW) is average daily heat flux due to internal gains (people, light, appliances). \( gA_{sol} \) parameter is a very approximated estimation of the actual solar geometry of the building, as it is highly dependent on the orientation chosen for \( l_{sol} \) (horizontal or south).

Assuming the possibility of performing measurement without internal gain (\( q_{int}=0 \)), we can reduce Equation 4 to Equation 5, which may be fit with a linear bivariate regression with intercept equal to 0.

\[ q_h = HLC\Delta T - gA_{sol}l_{sol} \]
If we then divide both sides of Equation 5 by $\Delta T$ we obtain Equation 6, which represents an alternative formulation of co-heating test. The left hand side of this equation is similar in principles to Equation 2. This property has been exploited in recent research [40, 41] to estimate the value of HLC for thermal zones with a dynamic averaging method conceptually similar to the one proposed by ISO 9869 (in situ measurement of U value) for individual construction components.

$$q_h = HLC - gA_{sol}I_{sol} \Delta T$$  \hspace{1cm} (6)

Finally, we can use the same measurements to obtain an energy signature model [38], represented in Equation 7:

$$q_h = a(T - T_{bp})$$  \hspace{1cm} (7)

where $a$ is the slope of the linear regression (negative for heating), $T_e$ represent daily average external air temperature and $T_{bp}$ represent balance-point temperature ($q_h = 0$).

Energy signature uses regression to derive parameters, without assuming internal air temperature as an input, and can be used for inexpensive long term monitoring (e.g. using utility bills collected on a monthly base). For this approach also an approximated physical interpretation of the coefficients can be found in literature [21]. Therefore, energy signature can essentially complement (for long-term monitoring) co-heating test, which needs measures of indoor air temperatures as an input and short term measurements under controlled conditions. Further, extensions of this methodology can be achieved by means of integration with variable-base degree-days method, as shown in recent literature [44], potentially extending the applicability for large scale energy system planning [43].

Finally, the approach proposed can be improved in particular in terms of identification of daily dynamic components of energy balance in regression formulations [45] and extension to hourly regression models. The final goal is linking this approach transparently with surrogate modelling techniques to be used already in the design phase [9] to fit multiple possible building conditions.

**Simplified dynamic modelling and time series**

The development of simplified time series model for simulation and parameter identification of walls is described in recent literature [29, 33], and is based substantially on linear algebra formulations. A schematic representation of a simplified model is reported in Figure 1.

![FIGURE 1. Example of experimental setup for dynamic wall behavior simulation using lumped parameter models](image)

This model formulation is compatible with the ongoing normative evolution in building performance assessment (ISO 52016-1 [23]), considering, in this example, a wall with insulation layer on the outer side. Model is defined as follows:

$$U = \frac{1}{\frac{1}{U_1} + \frac{1}{U_2}} = \frac{1}{R_1 + R_2} = \frac{1}{R}$$  \hspace{1cm} (8)
\[ C_m \frac{dT_m}{dt} = q_{in} - q_{out} \]  
(9)

\[ q_{in} = U_1(T_i - T_m) \]  
(10)

\[ q_{out} = U_2(T_m - T_e) \]  
(11)

\[ C_m \frac{dT_m}{dt} = U_1(T_i - T_m) + U_2(T_e - T_m) \]  
(12)

\[ T_{m,i+1} = T_{m,i} \left( 1 - \left( \frac{U_1 \Delta t}{C_m} \right) - \left( \frac{U_2 \Delta t}{C_m} \right) \right) + \left( \frac{U_1 \Delta t}{C_m} \right) T_{i,i} + \left( \frac{U_2 \Delta t}{C_m} \right) T_{e,i} \]  
(13)

where \( U \) (W/m²K) is thermal transmittance experimentally determined using ISO 9869, \( q_{in} \) (W/m²) is the heat flux entering in the lumped capacity of wall, \( q_{out} \) (W/m²) is the heat flux exiting from the lumped capacity of wall, \( n \) is the number of data points, \( U_1 \) and \( U_2 \) (W/m²K) are conductances on the internal and external side, \( R \) (m²K/W) is the total resistance, \( R_1 \) and \( R_2 \) (m²K/W) are thermal resistances on the internal and external side, \( T_i \) (°C) is internal air temperature, \( T_e \) (°C) is external air temperature, \( \Delta T = T_i - T_e \), \( C_m \) (J/m²K) is the lumped thermal capacity per unit of wall area, \( T_m \) is the temperature of the lumped thermal mass, \( t \) (s) is time, \( i \) is a time index in the time series.

Equation 13, the time series model used for calculation, corresponds to the explicit discretization of Equation 12, the model is in the form ARX (Autoregressive with Exogenous Input). Assumptions for the calculation of \( U_i \) can be made considering the thermal mass lumped on the internal side. Further, the lumped thermal capacity \( C_m \) assumed in the simulation is very near to the value of internal areal heat capacity \( k_1 \) that can be calculated using ISO 13786 [28]. A detailed explanation on this assumption can be found in [33].

**Future prospects for modelling research development**

There are many possible future developments for the research on building performance monitoring, following the bottom-up logic depicted in the previous section, we report here a selection of relevant topics. First of all, dynamic physical-statistical models can be used for Model Predictive Control [46], considering in particular the thermal storage capabilities of building fabric [47], and also the possibility to predict free-running operation [48].

With respect to regression models (including energy signatures), it is possible to work on the automated model selection [49], which is particularly important to ease the process of continuous monitoring and commissioning, and on the use of energy signatures specifically for the analysis of heat pump performance [16]. Further, the formulation of psychrometric equations for air-handling processes using linear algebra [50] offers interesting perspectives for optimization of air-handling systems in HVAC with surrogate models [51]. Finally, extensions of the variable base degree days method [44] could enable large scale energy performance analytics, ideally linking it with energy system planning and policy studies [42, 43] together with R&D strategies [52].

**CONCLUSIONS**

In this research we presented an overview of building performance monitoring methodologies, following a bottom-up perspective, from individual components up to the system level perspective. We highlighted a potential continuity among different methodological approaches acting on different building elements and scales, indicating the conceptual simplicity and scalability of the underlying numerical techniques. Further, comparability and ease of use also are important factors to improve the ability to learn from data on continuous base. However, further research efforts should be devoted to harmonization of computational techniques and experimental analysis procedures as done in other disciplines[53], as all these approaches depend critically on the quantity and quality of data. In this paper we discussed also about the compatibility between reduced order models used in research and model simplifications adopted for building performance simulation at the normative level. These dynamic models are also using stationary properties that can be determined by means of regression-based approaches. The possible advantages related to future advances in this direction are multiple, from the improvement of performance of individual technologies, up to energy systems planning and operation at large scale. This can lead to an integrated strategy for continuous improvement based on a constant feed-back from data, taking place at multiple levels with the different values chains of built environment products and services. Clearly, in order to exploit these potential advantages, it is necessary to automate data
acquisition and processing procedures (in terms of computing tools and rules) at multiple scales, from single components, to individual buildings, to building stock and cities.

REFERENCES