Experiment strategy for evaluating advanced building energy management system

Hanmin Cai, Fazel Khayatian, Philipp Heer

Urban Energy Systems Laboratory, Swiss Federal Laboratories for Material Science and Technology (Empa), Dübendorf, Switzerland

E-mail: hanmin.cai@empa.ch, fazel.khayatian@empa.ch, philipp.heer@empa.ch

Abstract. Buildings are envisioned to play an active role in future low-carbon energy systems. The complexity of building energy management systems increases as they interface more and more subsystems and domains. As an important step to achieve a higher technology readiness level, these energy management systems need to be systematically tested in real-life conditions. Currently, there are no standard testing and experiment strategies in buildings to handle the mentioned complexity. Additionally, the levels of details reported in the existing experimental studies are heterogeneous. This paper summarizes an application of a holistic testing method to a flexible fully-equipped prosumer with the goal of facilitating test preparation, execution, replication, and comparison. Several empirical suggestions are provided, and a hybrid quantification strategy with digital twins is presented.

1. Introduction
Advanced building energy management systems have received much attention in recent years due to increasing digitalization, potentials to contribute to carbon neutrality targets [1] and capability of supporting future low-carbon energy systems [2]. Additionally, they are envisioned to be interactive in future smart energy systems and will operate in dynamic conditions. Hence, testing these energy management systems is increasingly complex as they are interfacing multiple subsystems and domains. However, the levels of details reported in the existing experimental studies are heterogeneous. This may lead to difficulties in replication and comparison. Reference [3] presents an empirical holistic testing method for complex cyber-physical systems. The method provides cross-disciplinary abstraction to facilitate test preparation, execution, replication and comparison. This paper adapts the method for building applications and complements it with additional empirical insights and a hybrid quantification strategy using digital twins. The remainder of the paper is organized as follows: Section 2 adapts the original method for building applications and a case study is presented in Section 3. Section 4 discusses an empirical strategy to reduce bias due to imperfect digital twins. Finally, Section 5 gives a brief summary and areas for future research are outlined.
2. Methodology

The main methodological framework follows the holistic testing method proposed in [3]. This section briefly summarizes the existing method and complements it with empirical suggestions for building applications. Additionally, a quantification strategy with digital twins is motivated and described.

2.1. Testing method

The holistic testing method includes the formulation of a test case, quantification strategy, test specifications and experiment realization plans. The method systematically verifies subsystems’ response according to predefined verification, validation and characterization metrics. According to the technology readiness level definition, testing the system with typical operating conditions is crucial. In the rest of this part, insufficient study of building control under typical operating conditions in the existing literature is reasoned. More specifically, operating conditions here refer to different control objectives and Figure 1 presents an example with room heating control. Figure 1(a) shows the temperature response when minimizing total energy consumption and it is expected that temperature is kept close to the lower limit. This is commonly used in experimental studies to validate the sufficiency of data-driven predictive control [4, 5]. The verification criteria are accumulated constraint violations, especially during critical times such as the grey period \( \circ \) in Figure 1(a). On the other hand, room temperature has higher variances and tends to exploit the comfort zone when reacting to exogenous dynamic signals. Room temperature may be driven closer to the upper limit as shown in Figure 1(b). Comparing with Figure 1(a), erroneous controller decisions such as overheating that could occur at the critical time such as the grey period \( \circ \) in Figure 1(b) cannot be directly compensated. In brief, controller verification experiments presented in many existing literature only examine half of the operating conditions that future building energy management system may encounter.

![Figure 1](image-url)

Figure 1: Temperature response when (a) minimizing energy consumption and (b) reacting to exogenous dynamic signals [6].

2.2. Quantification strategy

Quantifying the controller’s contribution to energy- and cost-savings in real-life conditions can be challenging due to uncertainties and the fact that a perfect control group cannot be established. Reference [7] quantifies controller performance improvement in simulation and further validates the control strategy in an experiment. However, the benefits of the controller identified in simulations are the upper bounds of achievable benefits because controller performance can be significantly affected by uncertainties in real-life conditions. Several experimental strategies are commonly used in the existing literature. A degree-day approach is used in [8] to evaluate energy-saving potential. However, the approach cannot retain the daily variation of exogenous dynamic conditions.
signals such as electricity carbon intensity [9]. Multiple energy schedules with distinct daily carbon footprints may have the same daily energy consumption. Although the method has its merits in benchmarking energy usage, it is not suitable for more complex tasks. References [10] and [4] adopt a “placebo-controlled approach”. Reference [4] designs two-stage experiments and compares a data predictive controller with a hysteresis controller in two identical rooms. By switching the control strategy implemented in each room, the authors assume a sufficient control group can be established. Similarly, reference [10] splits the buildings into two groups and utilizes half of the buildings as the control group. Both studies assume redundancy in the experimental infrastructure and sufficiency of the control group. The comparison also requires both control strategies to be implemented at the same time to preserve the same boundary conditions. Alternatively, this study proposes a hybrid strategy using a digital twin, which refers to an extensive model of the physical system set up in a programming environment, connected to various streams of data, and accurately calibrated to mimic the system’s dynamic behavior through simulation. The notion of a digital twin that is adopted in this study is closer to the descriptions in [11], which defines a digital twin as a high fidelity computational model of a product that improves its performance by leveraging operational data. With calibrated digital twins, the requirement of redundancy in the infrastructure can be relaxed.

3. Case study
This section summarizes a case study of a flexible building energy management system reported in [12]. The studied system is first described, which is followed by application of the digital twin to examine room temperature control results. The energy management system, as shown in Figure 2, interacts with the appliances within the building and exports quantified flexibility to support collaborative smart grid [2]. Due to page limits, we briefly describe the overall system setup and refer interested readers to [13] for detailed test forms, in which several experiments are designed with incremental complexity such that erroneous sub-functions can be located properly. The overall system configuration is shown in Figure 2(a), in which multiple systems, stakeholders and domains are illustrated. All the controlled appliances and their connections are shown in a single-line diagram in Figure 2(b). Both Figure 2(a) and Figure 2(b) illustrate the complexity of the studied system that interfaces ICT, electrical and thermal domains.

Of all controlled subsystems illustrated in Figure 2(b), space heating accounts for a notable portion of total energy usage. It is considered as a highly nonlinear process affected by various environmental variables and occupants. The rest of the part present the application of a digital twin to establish a control group. In the digital twin setup, the physical twin is the UMAR unit of NEST demonstrator at Empa [14]. The digital counterpart is a detailed representation of the unit as an EnergyPlus model. The digital copy is set up based on the unit’s geometry and construction characteristics, as well as the heating and cooling distribution system, as shown in Figure 3. Data streams are logged at 1-minute intervals and include measurements of the indoor air temperature, setpoint temperature, and hydronic valve status in every zone, as well as the hydronic flow rate and temperature, electricity energy consumption, and climate characteristics for the entire unit. The digital twin model is calibrated at 1-minute temporal resolution with heating, cooling, domestic hot water, electricity, and indoor air temperature as described in [15].

To enable a component level control of the digital twin through co-simulation, the EnergyPlus model is wrapped into a Functional Mockup Unit (FMU). The FMU export package provided by the Lawrence Berkeley National Laboratory enables communication with EnergyPlus simulation engine through the functional mockup interface standard version 2.0 [16]. The FMU is imported into a Simulink environment and a co-simulation is executed at 1-minute intervals. The FMU receives two sets of data streams as inputs for co-simulation: (1) non-controllable inputs and
controllable inputs. The former group consists of on-site weather data, electricity loads, and measurements of hydronic flow and temperature from the district heating system at the unit’s inlet. The latter group consists of setpoint temperature, hydronic valves of conditioned zones, and the unit’s heat-exchanger status. Given that in this study, only setpoint temperature is controlled, the status of other components (i.e., room valves and heat-exchanger) is determined by the intended setpoint temperature and the simulated indoor temperature at every time step. Although a perfect calibration of the digital twin is impossible due to uncertainties associated with occupant behavior, the digital twin guarantees a maximum temperature divergence of 1 °C from the physical system during the initialization of experiments in all rooms. Application of the digital twin, which is benchmarked with the full-year data of 2019, to quantify controller’s impacts in improving indoor comfort level is shown in Figure 4 [6]. Figure 4(a) shows the simulated room temperature responses if a hysteresis controller were implemented, whereas Figure 4(b) shows the real measurements of an experiment in which Model Predictive Control (MPC) was implemented. The digital twin calibrated with full-year data and simulated with measured weather conditions can provide a reasonable virtual testbed to establish a control group.
4. Discussion
In this section, an empirical difference-of-difference strategy to reduce biases due to imperfect digital twins is discussed. The strategy is illustrated in Figure 5. Denote the physical system as $\mathcal{M}$ and its digital twin as $\hat{\mathcal{M}}$. The impact of a control policy $\pi_t$ is denoted as $c := \sum_{t \in \mathcal{T}} \mathcal{M}(\pi_t)$ for the physical experiment and $\hat{c} := \sum_{t \in \mathcal{T}} \hat{\mathcal{M}}(\pi_t)$ for the virtual experiment with the digital twin, where $t$ is the time index and $\mathcal{T}$ denotes the experiment period. The strategy is composed of two experiment periods $\mathcal{T}$ and $\mathcal{T}'$ of equal duration. During the first experiment period $\mathcal{T}$, the impacts of two policies $\pi_{1,t}$ and $\pi_{2,t}$ implemented on the physical system and the digital twin are denoted as $c_1 := \sum_{t \in \mathcal{T}} \mathcal{M}(\pi_{1,t})$ and $\hat{c}_2 := \sum_{t \in \mathcal{T}} \hat{\mathcal{M}}(\pi_{2,t})$ respectively. The real impact of $\pi_{1,t}$ over $\pi_{2,t}$, referred to as $\text{diff 1}$ in Figure 5, is $\Delta c = c_1 - \hat{c}_2 - \epsilon$ where $\epsilon$ denotes the bias. When the digital twin is a perfect replicate of the physical system, there is $\epsilon = 0$. However, this is not possible because of unmeasured variables and noises. A persistent bias $\epsilon \in \mathbb{R}_+$ is therefore assumed. During the second experiment period $\mathcal{T}'$, the implementation platforms of the two policies are switched with $\hat{c}_1 := \sum_{t \in \mathcal{T}} \mathcal{M}(\pi_{1,t})$ and $c_2 := \sum_{t \in \mathcal{T}} \hat{\mathcal{M}}(\pi_{2,t})$. The real impact, referred to as $\text{diff 2}$ in Figure 5, is $\Delta c' = c_2 - \hat{c}_1 - \epsilon'$. The total impacts of both experiments, referred to as counterfactual impacts in Figure 5, are $\Delta c - \Delta c'$ if $\epsilon$ converges to $\epsilon'$ over time. Additionally, with the hybrid strategy, it is not restricted to implement controllers at the same time to compare their performance. Hence, it can promote more efficient usage of the infrastructure because any historical experiment can also be reused to compare with a new control strategy.

5. Conclusion
Advanced building energy management systems that interface multiple subsystems, domains and stakeholders are complex cyber-physical systems that need to be systematically tested. An existing holistic testing method is investigated with additional empirical suggestions. More specifically, this paper presents a case study in which a collaborative and flexible building energy management system is the focus. Additionally, an empirical method is discussed to further
reduce the biases introduced by the imperfect digital twins. However, the method relies on strong assumption on the biases. In future work, we will examine the empirical method in Monte-Carlo simulations and seek to relax the assumption to extend its applicability.

References

[1] Global Alliance for Buildings and Construction, International Energy Agency, and United Nations Environment Programme, “2020 global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector,” 2020.
[2] J. Gasser, H. Cai, S. Karagiannopoulos, P. Heer, and G. Hug, “Predictive energy management of residential buildings while self-reporting flexibility envelope,” Applied Energy, vol. 288, p. 116653, 2021.
[3] K. Heussen, C. Steinbrink, I. F. Abdulhadi, V. H. Nguyen, M. Z. Degefa, J. Merino, T. V. Jensen, H. Guo, O. Gehrke, D. E. M. Bondy, and Others, “Erigrid holistic test description for validating cyber-physical energy systems,” Energies, vol. 12, no. 14, p. 2722, 2019.
[4] F. Bünning, B. Huber, P. Heer, A. Aboudonia, and J. Lygeros, “Experimental demonstration of data predictive control for energy optimization and thermal comfort in buildings,” Energy and Buildings, vol. 211, p. 109792, 2020.
[5] Y. Lian, J. Shi, M. P. Koch, and C. N. Jones, “Adaptive robust data-driven building control via bi-level reformulation: an experimental result,” arXiv preprint arXiv:2106.05740, 2021.
[6] Swiss Federal Office of Energy, “alumid - versorgung neu: Feldtest 1.1.2019 – 30.6.2020.” https://www.aramis.admin.ch/Beteiligte/?ProjectID=43459. accessed: 19.04.2021.
[7] D. Sturzenegger, D. Gyalistras, M. Morari, and R. S. Smith, “Model predictive climate control of a swiss office building: Implementation results, and cost–benefit analysis,” IEEE Transactions on Control Systems Technology, vol. 24, no. 1, pp. 1–12, 2015.
[8] L. Lundström and F. Wallin, “Heat demand profiles of energy conservation measures in buildings and their impact on a district heating system.” Applied energy, vol. 161, pp. 290–299, 2016.
[9] D. Beloin-Saint-Pierre, “Ecodynbat - ecobilan dynamique des bâtiments.” https://www.aramis.admin.ch/Grunddaten/?ProjectID=41804, 2019. accessed: 09.04.2020.
[10] M. H. Christensen, R. Li, and P. Pinson, “Demand side management of heat in smart homes: Living-lab experiments,” Energy, vol. 195, p. 116993, 2020.
[11] G. P. Lydon, S. Caranovic, I. Hischier, and A. Schlueter, “Coupled simulation of thermally active building systems to support a digital twin,” Energy and Buildings, vol. 202, p. 109298, 2019.
[12] H. Cai and P. Heer, “Experimental implementation of an emission-aware prosumer with online flexibility quantification and provision,” Applied Energy, submitted.
[13] H. Cai, “Test specifications and experiment plans.” https://info.nestcollaboration.ch/wikipedia/public/research/currentprojects/alumid/, 2020. accessed: 17.06.2021.
[14] Empa. “Nest wiki.” https://info.nestcollaboration.ch/wikipedia/public/. accessed: 12.12.2020.
[15] F. Khayatian, A. Bollinger, and P. Heer, “Temporal resolution of measurements and the effects on calibrating building energy models,” arXiv preprint arXiv:2011.08974, 2020.
[16] M. Wetter, “Co-simulation of building energy and control systems with the building controls virtual test bed,” Journal of Building Performance Simulation, vol. 4, no. 3, pp. 185–203, 2011.