Energy Proceedings
Vol 31, 2023
ISSN 2004-2965

Benchmarking HVAC controller performance with a digital twin

Fazel Khayatian[1*], Hanmin Cai [1], Aaron Bojarski[1], Philipp Heer[1], Andrew Bollinger[1]
[1] Urban Energy Systems Laboratory, Swiss Federal Laboratories for Materials Science and Technology-Empa, Uberlandstrasse 129, Dubendorf, Switzerland (*Corresponding Author)

ABSTRACT

This paper presents a pipeline for creating digital twins of building energy systems, which is shared as an open test-environment for controller benchmarking. The digital twin is calibrated based on an extensive dataset including a wide variety of data with fine temporal resolutions. The comprehensive list of controllable variables, the temporal resolution of measurements, and the real-time capabilities of the digital twin distinguish this work from the existing test environments. A case study is also provided to exemplify the use of this open environment for benchmarking the performance of building automation and control systems.

Keywords: Digital twin, HVAC control, Benchmarking, Building energy performance

NONMENCLATURE

Abbreviations	
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BACS	Building Automation and Control Systems
BEM	Building Energy Model
DHW	Domestic Hot Water
EV	Electric Vehicle
FMU	Functional Mockup Unit
HP	Heat pump
HVAC	Heating, Ventilation, and Air-Conditioning
KPI	Key Performance Indicator
IEQ	Indoor Environmental Quality
MPC	Model Predictive Controller
PV	Photovoltaics
RBC	Rule-Based Controller

1. INTRODUCTION

More than 80% of homeowners personally program their room thermostats, a practice which often leads to energy waste [1]. A solution is to automate the process of adjusting indoor environmental conditions with Building Automation and Control Systems (BACS). BACS are equipment, software, and services that allow monitoring and operation of indoor environments. One application of BACS is the optimal control of Heating, Ventilation, and Air-Conditioning (HVAC) systems while satisfying Indoor Environmental Quality (IEQ). Given that HVAC systems are the dominant consumers of energy in buildings [2], the energy saving potentials of BACS are well-documented. BACS have displayed promising results in load management, where dynamic tariffs are present [3]. Furthermore, incorporating model predictive control into BACS has shown the potential to save energy while increasing comfort for occupants [4]. Aside from energy performance optimization, BACS are suitable for performing system diagnostics and anomaly detection during operation [5].

Depending on the end-use, BACS have various objectives that span from economic (e.g. operation cost) to environmental (e.g. CO_2 emissions) performances, as well as restrictions based on social constraints (e.g. thermal comfort). Consequently, numerous evaluation metrics are described in the literature [2], which report a wide range of improvements on buildings' economic and environmental performances. As of 2022, lack of adequate benchmarking datasets and platforms continues to be a major challenge for evaluating the performance of BACS [6]. However, valuable contributions have greatly facilitated the reproducibility of experiments and evaluations. Energym [6], BEOPTEST [7], and CityLearn [8], provide reference buildings, open-source environments, and Key Performance Indicators (KPIs) for benchmarking BACS' performance.

Despite the major advancements in benchmarking BACS' performance, test environments are still scarce and existing models lack the full range of variables that affect HVAC performance. For example, the existing test environments do not offer the option to override the shading system or allow for window airing. In addition, current test-environments suffice to a ten-minutes sampling rate, and often only cover portions of a full year. Yet, the temporal resolution of a dataset directly affects the assessment of system's performance. Previous studies have argued that temporal aggregation of measurements can distort the load shape and lead to biased assumptions about BACS' performance [9]. Studies have also shown that using a deterministic operation scenario is insufficient for benchmarking controller performance, as it does not encompass the full range of uncertainties that are associated with occupants' behavior [10]. Hence, it is imperative to enrich the existing list of open test environments for more rigorous benchmarking of BACS.

Building on previous research, this study resorts to long-term measurements with fine temporal resolutions for developing a high-fidelity building energy model (BEM) labelled as a "Digital Twin". Here, the term digital twin refers to a virtual replica of the physical system, which accurately mimics its components and operation. The objective of using a digital twin is to facilitate benchmarking the performance of controllers while gathering insights into system performance. As an original contribution, we propose a pipeline for creating integrated BEM-control workflows and describe the iterative process of calibrating BEMs with minimum computation cost. The study also introduces BAC-Bench, an open test environment that forms the core of the digital twin with options to run assessments offline or in real-time. BAC-Bench allows researchers to benchmark the performance of new controllers against that of a baseline controller. BAC-Bench also enables researchers to assess the reliability of their controllers amid uncertainties in climate or building operation. A real-world application of BAC-Bench is demonstrated through a case study of benchmarking three controllers with dissimilar strategies and objectives.

2. MATERIAL AND METHODS

This section first presents the workflow for creating a high-fidelity BEM, hereafter referred to as the digital twin. The section also describes the physical setup, on which the BAC-Bench test environment is based. Then the setup of a field experiment is briefly summarized, and the digital twin's application for investigating the impacts of alternative controllers is discussed in detail.

Fig. 1. The case-study unit (UMAR) marked in white (top), and its floor plan (bottom) © Zooey Braun, Stuttgart.

2.1 *Model setup and calibration*

In its core, BAC-Bench [11] is an EnergyPlus model of the Urban Mining and Recycling (UMAR) unit at NEST in Empa [12]. The EnergyPlus model is wrapped into a functional mockup unit (FMU) to enable co-simulation. The UMAR unit is a residential apartment hosting two occupants. The apartment consists of a living room, two bedrooms, two bathrooms, as well as a storage cell (Fig. 1). Air conditioning at the UMAR unit is though ceiling radiant panels, which are supplied by heat exchangers that are connected to district heating and cooling networks. Measurements at UMAR are conducted at various nodes and registered in one-minute intervals.

Benchmarking the performance of BACS requires high-fidelity BEMs. Thus, model calibration and modularity play a vital role in the flexibility of the BEM as well as the reliability of the benchmarking process [13]. Calibration of BEMs often requires hundreds to thousands of simulations depending on the number of variables. Furthermore, there is no guarantee of optimality in many cases, particularly when numerous variables are calibrated simultaneously.

In this study, we adopt the calibration strategy presented in [14] with small modifications as described in the following. Instead of using Genetic Algorithm, we opt for Subset Simulation to reduce the computation time for optimizing the inputs and calibrating the BEM [15]. Subset Simulation has proven highly effective for multivariate calibration, while minimizing the compute through efficient sampling from posterior distributions. Given the wide range of variables that simultaneously affect Energy consumption and IEQ, the proposed pipeline executes calibration at four consecutive stages (Fig. 2).

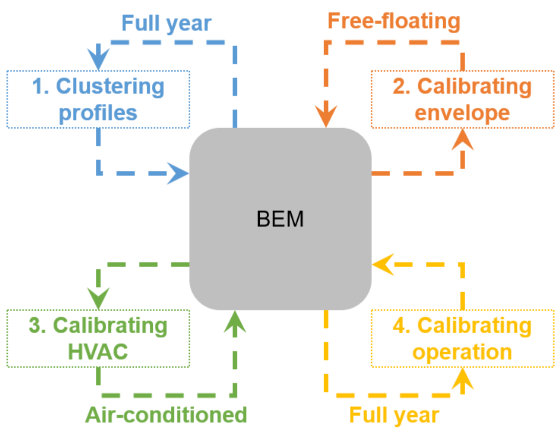

Fig. 2. The workflow for calibrating the BEM.

First, typical daily occupancy profiles are created for the entire year through clustering electricity and water consumption measurements. Further details on the clusters and representative profiles are available in [9]. Second, free floating instances are extracted from the dataset and utilized for calibrating model inputs based on IEQ measurements. At this stage, only construction characteristics of the envelope (e.g., thermal conductivity, thermal mass, g-value, infiltration, etc.) are calibrated. Third, air-conditioned periods are extracted from the dataset and utilized for calibrating model inputs based on energy consumption and IEQ measurements. At the third stage, only the HVAC system properties (e.g., thermal conductivity of radiant ceiling panels, length of fluid pipes, system's throttling range, etc.) are calibrated. Forth, model inputs for the full year are calibrated based on energy consumption and IEQ measurements. The final stage of the process focuses on the occupancy profiles by executing the calibration process for each day of the year.

2.2 *Experiment setup*

The calibrated digital twin is then utilized to evaluate a model predictive controller in a field experiment, during which all common electrical loads were controlled to minimize equivalent carbon emission due to electricity imported from the power grid. The controller was implemented on top of a standard BACS [16]. Optimal decisions were obtained combining models of the physical system and forecasts. Specific to the HVAC, continuous controller decisions were executed by modulating the duty cycle of valve opening in individual rooms. Note that the thermal power into individual rooms were not measured but calculated by dividing the total power of the entire unit by the size of inlet pipe. The interested reader is referred to [17] for a detailed description of the experimental setup. Given that the controller's sampling time differs from the simulation timestep, timely communication between the physical system (UMAR), digital twin (BAC-Bench), controller, and miscellaneous systems (EV, PV, batteries, etc.) is imperative. Figure 3 shows how the MOSAIK tool manages communication between different FMUs and synchronizes the physical and digital systems, a feature that is necessary for hardware-in-the-loop simulations. Online deployment of BAC-Bench is currently at the testing and validation phase.

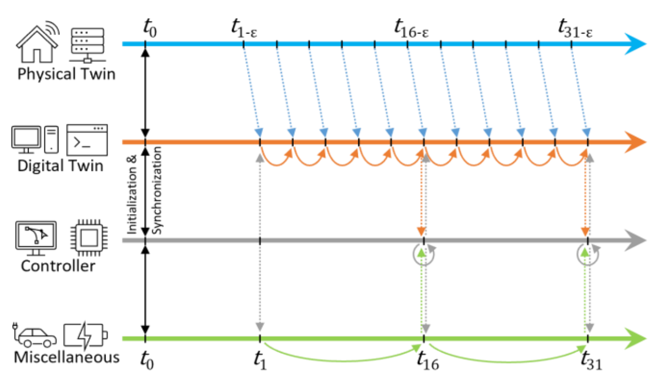

Fig. 3. Synchronization mechanism between BAC-Bench and the other components for real-time operation.

3. RESULTS

This section is composed of two parts. First, the accuracy of the digital twin is assessed and compared with the existing recommendations. Second, comparisons with alternative controllers of the field experiment are reported.

3.1 *Calibration results*

The BEM is calibrated on one-minute measurements from the entire year of 2020, covering the indoor air temperature in Rooms 1, 2, and 3, as well as the heating and cooling energy consumptions for the entire unit. The accuracy of the calibrated model is reported in Table 1. Since there are no standards and thresholds for evaluating the calibration of one-minute measurements, the performance values reported in Table 1 are based on one-minute data loggings that are resampled to hourly temporal resolution.

Table 1. Calibration error of the BEM.

	Room1 temp.	Room2 temp.	Room3 temp.	Heating energy	Cooling energy
nMBE	-0.002	0.003	-0.003	-0.03	0.09
CVRMSE	0.019	0.035	0.019	2.4	3.7
R^2	97.51	96.87	98.42	92.1	87.4

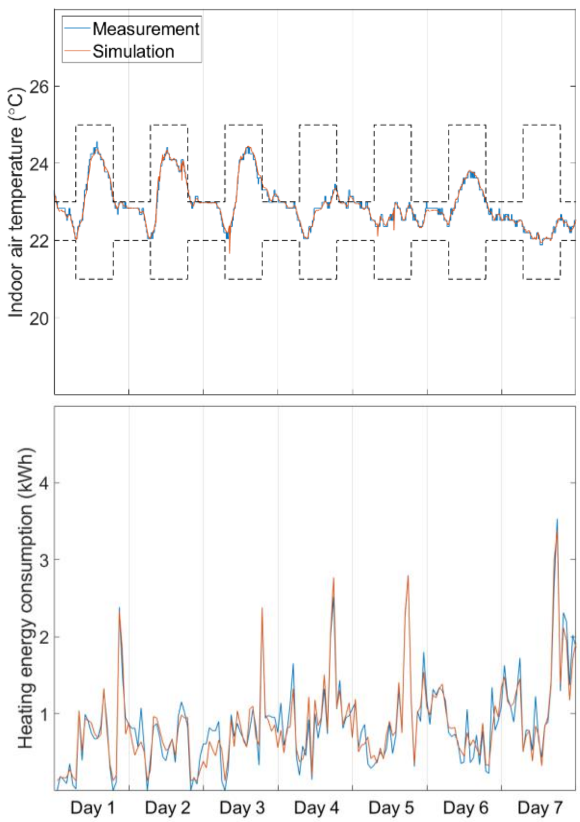

Fig. 4. Contrasting simulated data against measurements for the experiment period.

We also contrast the measurements of indoor air temperature and energy consumption against those of the simulation for the experiment period. Fig. 4 compares the measured and simulated indoor air temperature as well as the heating energy consumption. For brevity, we only plot the calibration results of indoor air temperature for Room 3, which will be further discussed in the next section. The simulated energy consumption in one-minute temporal resolution occasionally shows large deviations from the measurements. This is because measurements are collected at a heat exchanger that is placed between the unit and the backbone, thus logging the energy consumption with some time lag. However, the observed deviations at one-minute temporal resolution does not invalidate the accuracy of the model. Evaluating the hourly data (resampled from one-minute measurements) as shown in Table 1 reaffirms that the calibration error is considerably lower than ASHRAE recommendations [18].

3.2 Benchmarking the experiment

Among the benefits of a digital twin is to validate the suitability of a technology by preserving an experiment's settings while manipulating one or more features. Given that this study focuses on controllers' performance, the experiment is replicated with identical boundary conditions and dissimilar control strategies. Table 2 details the characteristics of three controllers (one rule-based controller and two predictive controllers) with dissimilar strategies, objectives, and system configurations. Controllers 1 and 3 are borrowed from a previous study [17] that highlights the potential of BACS for reducing carbon emissions. Controller 2 is designed to share some characteristics with the other two controllers. Controllers 1 and 2 have identical system components, while controllers 2 and 3 both opt for a predictive control strategy. All three controllers are constrained by thermal comfort bounds.

Table 2. Characteristics of the controllers tested in this study

ID	Control strategy	Objective	Constraints	System components
1	RBC	N/A	Thermal comfort	HVAC
2	MPC	Minimize energy consumption	Thermal comfort	HVAC
3	MPC	Minimize carbon emissions	Thermal comfort	HVAC, DHW, EV, PV, Battery

In this study, we leverage the potential of the digital twin and dig deeper into the performance of BACS by analyzing the controllers at room level. The digital twin allows us to break down the energy consumption of the whole unit by each room. Such spatial granularity of measurements is not available from the energy meter

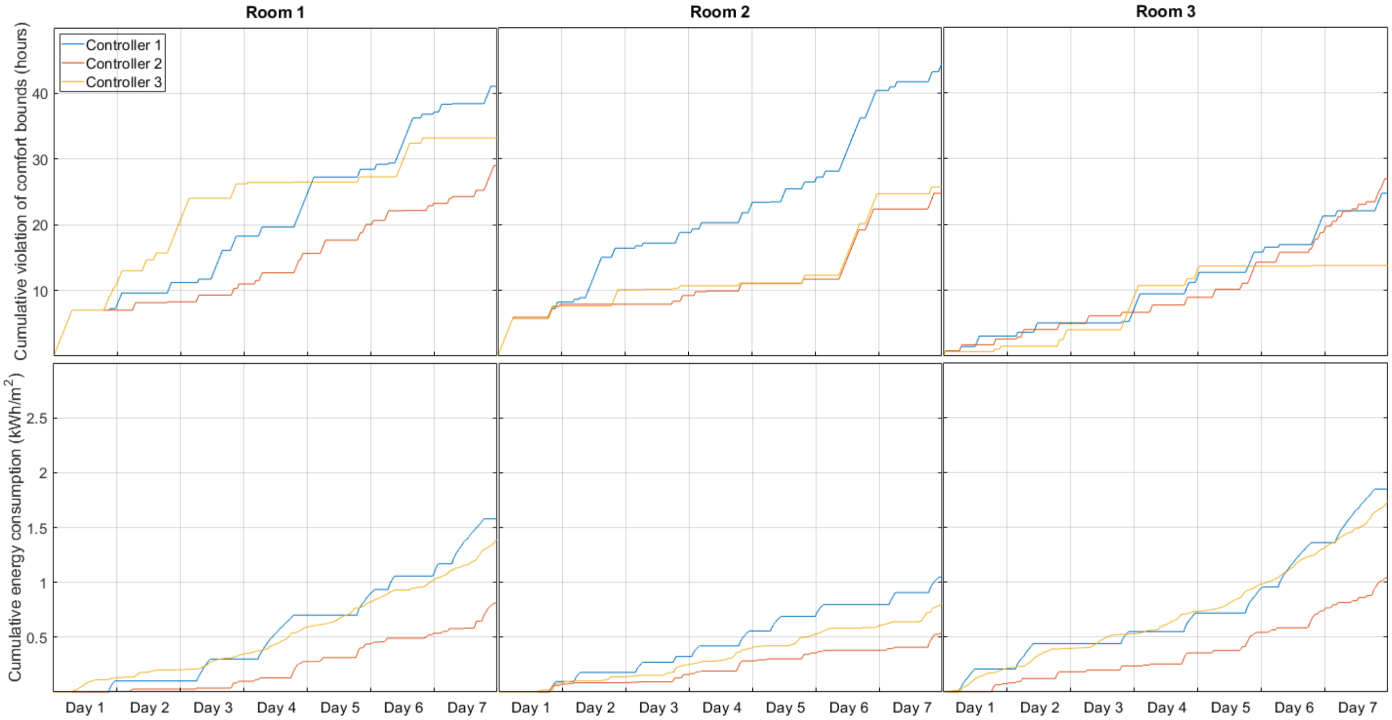

Fig. 5. Comparison of the three controllers listed in Table 1 In terms of: violation of thermal comfort (top) and cumulative energy consumption (bottom).

that is installed at UMAR and provides new insights into controllers' response to building operation.

It is observed that the performance of controllers vary significantly among the three rooms (Fig. 5), which can be attributed to the room operation. Room 3 was unoccupied during the experiment with the shades deployed, thus blocking the solar gains. While suppressing the solar gains through a deployed shading results in less violations of the upper comfort bound (Fig. 5, top-left), it comes at the cost of higher heating energy consumption (Fig. 5, bottom-left) when compared to Room 1. We also observe an unusual phenomenon when studying Room 3, i.e. controller 2 violates the comfort bounds more frequently when compared to the other two controllers. This observation contradicts the performance of controller 2 in Rooms 1 and 2, where it outperforms the other two controllers. Plotting the indoor air temperature of Room 3 shows that the absence of solar gain in the room allows controller 2 to keep the indoor air temperature extremely close to the lower comfort bound (Fig 6). As a result, controller 2 occasionally crosses the comfort limits, particularly when transitioning between day and night comfort bounds. The other two controllers, however, adopt a completely different strategy. Controller 1 does not have any incentives to minimize the energy consumption, and thus, is less prone to cross the lower comfort bound. Controller 3 weighs emissions from other components such as the PV and battery, and thus, does not keep the indoor air temperature close to the lower bound; consequently reducing the risks of crossing beyond the thermal comfort limits.

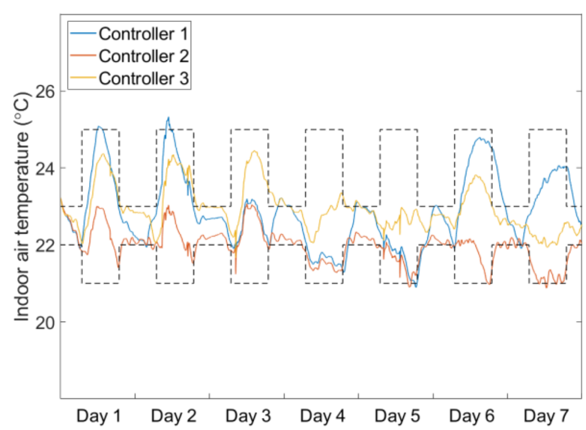

Fig. 6. Comparing thermal comfort violations in Room 3.

4. CONCLUSIONS

The high-fidelity model in this study labelled as a "Digital Twin" allows us to benchmark the performance of an experiment with identical boundary conditions.

This freedom is particularly important, as re-executing an experiment on the physical system with a different controller and identical climate and operation conditions would be impossible. Furthermore, the digital twin allows us to explore what-if scenarios, in which the controller is exposed to out-of-sample disturbances and uncertainties. Digital twin's capability of swapping climate and operation conditions on the fly enables us to study the robustness of a controller to unseen events. Thus, there is a great potential to leverage the interoperability of digital twins for designing robust controllers.

ACKNOWLEDGEMENT

This research has been carried out within the context of the Empa-funded project Algorithmic Regulation and Control. The work of H. Cai and P. Heer was supported by the Swiss Federal Office of Energy through the Sustainable Demand-Side Management for the Operation of Buildings (S-DSM) project under the contract number SI/502165-01.

REFERENCE

[1] Peffer T, Pritoni M, Meier A, Aragon C, Perry D. How people use thermostats in homes: A review. Building and Environment. 2011 Dec 1;46(12):2529-41.
[2] Aste N, Manfren M, Marenzi G. Building Automation and Control Systems and performance optimization: A framework for analysis. Renewable and Sustainable Energy Reviews. 2017 Aug 1;75:313-30.
[3] Oldewurtel F, Ulbig A, Morari M, Andersson G. Building control and storage management with dynamic tariffs for shaping demand response. 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies 2011 Dec 5 (pp. 1-8). IEEE.
[4] Bünning F, Huber B, Heer P, Aboudonia A, Lygeros J. Experimental demonstration of data predictive control for energy optimization and thermal comfort in buildings. Energy and Buildings. 2020 Mar 15;211:109792.
[5] Lu Q, Xie X, Parlikad AK, Schooling JM. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. Automation in Construction. 2020 Oct 1;118:103277.
[6] Scharnhorst P, Schubnel B, Fernández Bandera C, Salom J, Taddeo P, Boegli M, Gorecki T, Stauffer Y, Peppas A, Politi C. Energym: A building model library for controller benchmarking. Applied Sciences. 2021 Jan;11(8):3518.
[7] Blum D, Jorissen F, Huang S, Chen Y, Arroyo J, Benne K, Li Y, Gavan V, Rivalin L, Helsen L, Vrabie D. Prototyping the BOPTEST framework for simulation-based testing of advanced control strategies in buildings. 2019.
[8] Vázquez-Canteli JR, Kämpf J, Henze G, Nagy Z. CityLearn v1. 0: An OpenAI gym environment for demand response with deep reinforcement learning. InProceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation 2019 Nov 13 (pp. 356-357).
[9] Khayatian F, Bollinger A, Heer P. Temporal resolution of measurements and the effects on calibrating building energy models. arXiv preprint arXiv:2011.08974. 2020 Nov 4.
[10] Khayatian F, Nagy Z, Bollinger A. Using generative adversarial networks to evaluate robustness of reinforcement learning agents against uncertainties. Energy and Buildings. 2021 Nov 15;251:111334.
[11] Khayatian F, Cai H, Bojarski A. BAC-Bench: An open-source environment for benchmarking building automation and control systems [Computer software]. https://github.com/Khayatian/DTpy
[12] Richner P, Heer P, Largo R, Marchesi E, Zimmermann M. NEST-A platform for the acceleration of innovation in buildings. Informes de la Construccion. 2017;69(548):e222.
[13] Aird G, Coakley D, Kerrigan R. Application of an optimisation approach for the calibration of high-fidelity building energy models to support model predictive control (mpc) of HVAC systems. BSO 2016 proceedings.
[14] Ruiz GR, Bandera CF, Temes TG, Gutierrez AS. Genetic algorithm for building envelope calibration. Applied energy. 2016 Apr 15;168:691-705.
[15] Meshkinkiya M, Loonen RC, Paolini R, Hensen JL. Calibrating Perez model coefficients using subset simulation. IOP Conference Series: Materials Science and Engineering 2019 Jul 1 (Vol. 556, No. 1, p. 012017). IOP Publishing.
[16] Cai H, Khayatian F, Heer P. Experiment strategy for evaluating advanced building energy management system. Journal of Physics: Conference Series 2021 Nov 1 (Vol. 2042, No. 1, p. 012030). IOP Publishing.
[17] Cai H, Heer P. Experimental implementation of an emission-aware prosumer with online flexibility quantification and provision. arXiv preprint arXiv:2110.12831. 2021 Oct 25.
[18] ASHRAE. Guideline 14-2014: Measurement of Energy Demand and Water Savings. ASHRAE Standards Committee: Atlanta, GA, USA. 2014 Dec.