Open Computing Infrastructure for Sharing Data Analytics to Support Building Energy Simulations

Omer T. Karaguzel[1]; Mohammed Elshambakey[2]; Yimin Zhu, A.M.ASCE[3]; Tianzhen Hong[4]; William J. Tolone[5]; Sreyasee Das Bhattacharjee[6]; Isaac Cho[7]; Wenwen Dou[8]; Haopeng Wang[9]; Siliang Lu[10]; Mohamed Khalefa[11]; and Yong Tao[12]

Abstract: Building energy simulation plays an increasingly important role in building design and operation. This paper presents an open computing infrastructure, Virtual Information Fabric Infrastructure (VIFI), that allows building designers and engineers to enhance their simulations by combining empirical data with diagnostic or prognostic models. Based on the idea of dynamic data-driven application systems (DDDAS), the VIFI infrastructure complements conventional data-centric sharing strategies and addresses key data-sharing concerns such as the privacy of building occupants. To demonstrate the potential of the VIFI infrastructure, an empirically derived lighting schedule in the US Department of Energy's small office building reference model is simulated. The case-study simulation is used to explore (1) the possibility and potential of integrating data-centric and analytic-centric sharing strategies; (2) the method of combining empirical data with simulations; (3) the creation, sharing, and execution of analytics using VIFI; and (4) the impact of incorporating empirical data on energy simulations. Although the case study reveals clear advantages of the VIFI data infrastructure, research questions remain surrounding the motivation and benefits for sharing data, the metadata that are required to support the composition of analytics, and the performance metrics that could be used in assessing the applications of VIFI. **DOI:** 10.1061/(ASCE)CP.1943-5487.0000857.

Author keywords: Data sharing; Building energy; Simulation; Data standards; EnergyPlus; Operational schedules; Energy simulations.

Introduction

Buildings account for more than one-third of the primary energy consumption in the world; therefore, reducing energy use and greenhouse-gas emissions in the building sector is a key strategy for achieving global energy and environmental goals. Building performance simulation has played an increasingly important role in designing, operating, and retrofitting buildings to predict energy efficiency and utility costs (Roth 2017). As more energy policies and regulations push new buildings or building retrofits toward achieving low- or zero-net-energy (ZNE) goals, building energy simulations can be used as tools for evaluating and comparing design alternatives that employ integrated building technologies and control strategies. ZNE buildings also require balancing demand-side building energy efficiency and on-site generation with the supply-side renewable power, which requires an optimization based on various performance metrics, including energy use, energy cost, return on investment (ROI), or greenhouse-gas (GHG) emissions. Moreover, energy models developed in the building design phase can be adapted for use later in the building operation phase to support performance diagnostics and operation improvements, along with guiding building retrofit strategies.

Building performance simulations rely on diagnostic and/or prognostic models to understand and predict building performance. Often, these models require significant empirical data inputs on building operations, physical and environmental conditions, or occupants and their behaviors (D'Oca and Hong 2015; Kwak et al. 2015). In addition, model calibration with empirical data is critical

[1]Assistant Professor, School of Architecture, Carnegie Mellon Univ., 5000 Forbes Ave., MMCH #415, Pittsburgh, PA 15213. Email: okaraguz@andrew.cmu.edu

[2]Postdoctoral Research Fellow, College of Computing and Informatics, Univ. of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223; Researcher, Knowledge-Based and AI Systems, SRTA-City, Universities and Research District, New Borg Elarab, Alexandria 21934, Egypt. ORCID: https://orcid.org/0000-0001-6059-5032. Email: melshamb@uncc.edu; mshambakey@srtacity.sci.eg

[3]Professor, Dept. of Construction Management, Louisiana State Univ., Baton Rouge, LA 70803 (corresponding author). Email: yiminzhu@lsu.edu

[4]Staff Scientist, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720. Email: thong@lbl.gov

[5]Professor, College of Computing and Informatics, Univ. of North Carolina at Charlotte, Charlotte, NC 28223. Email: william.tolone@uncc.edu

Note. This manuscript was submitted on July 6, 2018; approved on March 29, 2019; published online on July 19, 2019. Discussion period open until December 19, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Computing in Civil Engineering***, © ASCE, ISSN 0887-3801.

[6]Postdoctoral Research Associate, College of Computing and Informatics, Univ. of North Carolina at Charlotte, Charlotte, NC 28223. Email: sdasbhat@uncc.edu

[7]Research Assistant Professor, College of Computing and Informatics, Univ. of North Carolina at Charlotte, Charlotte, NC 28223. Email: icho1@uncc.edu

[8]Assistant Professor, College of Computing and Informatics, Univ. of North Carolina at Charlotte, Charlotte, NC 28223. Email: wdou1@uncc.edu

[9]Ph.D. Student, School of Architecture, Carnegie Mellon Univ., 5000 Forbes Ave., MMCH #415, Pittsburgh, PA 15213. Email: wanghp18@gmail.com

[10]Ph.D. Student, School of Architecture, Carnegie Mellon Univ., 5000 Forbes Ave., MMCH #415, Pittsburgh, PA 15213. ORCID: https://orcid.org/0000-0002-4550-4152. Email: siliang1@andrew.cmu.edu

[11]Assistant Professor, J.B. Speed School of Engineering, Univ. of Louisville, Louisville, KY 40208. Email: mohamed.khalefa@louisville.edu

[12]Professor, Dept. of Engineering, Nova Southeastern Univ., 3301 College Ave., Ft. Lauderdale, FL 33314. Email: ytao@nova.edu

to the quality of building performance simulations (Coakley et al. 2014). However, gaining access to empirical data, particularly monthly or time-interval energy data, can be challenging due to the data privacy concerns of building owners. Thus, on one hand, the incorporation of empirical data into building performance simulations improves simulation quality; on the other hand, it presents a data privacy challenge.

This paper introduces an open computing infrastructure, or Virtual Information Fabric Infrastructure (VIFI), that integrates predictive models with empirical data to enhance building energy simulations. VIFI shares analytics on raw data, complementing conventional data-sharing strategies that require exposing the raw data directly. The framework is designed based on the concept of dynamic data-driven application systems (DDDAS) (Darema 2005). In essence, DDDAS include two parts: computation and measurement. Computation refers to predictive models or simulations for a particular purpose, such as building energy simulations, and measurement refers to instrumentation for data collection, such as building sensors or data archives. The computation piece takes advantage of measurement data to dynamically calibrate simulations or predictions and provides feedback to the measurement piece to inform future data-collection efforts. The feedback loop between computation and measurement allows continuous improvement in simulations and predictions.

Two aspects of VIFI are particularly significant. First, the VIFI computing infrastructure offers a new data-sharing alternative in cases where a conventional data-centric sharing strategy is not applicable. For example, advancements in building sensor and control technologies have significantly improved the ability of building owners and managers to collect building performance data, potentially offering an unprecedented opportunity to improve building energy simulations as they apply to the design of new buildings, building retrofits, or operational strategies. Yet, access to these big data by building designers and engineers is highly constrained due to data privacy issues. In particular, building owners or facility managers may not feel comfortable sharing the data with others because such raw data often contain sensitive occupant information. In addition, the raw data may not be in the correct form for others to use. Nevertheless, data owners may support the general idea of collaborating with others to improve the quality of building energy simulations. In such scenarios, the VIFI data-sharing infrastructure can alleviate the data owner's concerns about privacy and support their ability to collaborate with others without exposing their raw data.

Consequently, the VIFI data-sharing infrastructure potentially allows building designers and engineers to access a larger pool of building data than would be possible using a conventional direct-sharing strategy alone. This potential leads to another significant contribution of the infrastructure: VIFI allows building designers and engineers to enhance their simulations by combining high-quality empirical data with diagnostic and/or prognostic models for simulations. Energy simulations applied to building design often have many input parameters that need to have strong relevance to buildings under design. Such relevance significantly affects the quality of simulations. Improved relevance can be achieved by offering building designers and engineers access to a large set of relevant building data, which the new strategy can offer. Therefore, the infrastructure has the potential to change how building performance simulations are done in the future.

The following sections discuss existing studies that are relevant to building energy data integration and sharing, outline the proposed VIFI for building energy simulations, present a case study to demonstrate the implementation and use of the infrastructure, and summarize key conclusions and areas for future research.

Related Studies

The need to incorporate empirical data in building energy simulations has been extensively explored in the existing literature. Building energy simulations typically require several input variables, many of which are related to human behavior, such as an occupancy or lighting schedule (Crawley et al. 2001). Traditionally, building energy simulations are performed in a closed manner with all inputs fixed or assumed to be constant during a simulation (e.g., Zhu 2006; Pan et al. 2007). To improve the accuracy of simulations, dynamic approaches that incorporate real-time variable inputs have recently gained attention among researchers. For example, Pang et al. (2016) presented a method that integrates real-time data such as weather, plug load power, lighting power, occupancy, and room temperature settings with whole-building energy simulation. Similarly, Kwak et al. (2015) discussed real-time simulation through cosimulation to optimize building operations. However, most such applications have a focus on building operations, not building design.

Although simulations using real-time data during the design stage are rare, the idea of enhancing design-stage simulations through cosimulation with other programs has been previously explored. For example, Feng et al. (2015) and later Chen et al. (2018) developed an occupancy simulation software tool to work with building energy simulation through cosimulation. Software platforms have also been developed for deploying cosimulations, such as the tight-coupling approach using the building controls virtual test bed (BCVTB) (Wetter 2011) or the loose-coupling approach based on the occupant behavior XML (obXML) (Hong et al. 2016). Additionally, OpenStudio SDK provides an ecosystem of software tools that allows a comprehensive suite of modeled variables such as lighting and occupancy to be analyzed in whole-building energy simulations through an open application programming interface (Guglielmetti et al. 2011). However, the deployment of the OpenStudio SDK requires energy models and associated data to be stored on an OpenStudio server for model execution and calibration. Thus, the potential of the proposed VIFI strategy, i.e., moving analytics to data and thus alleviating data privacy concerns, has not already been explored in the OpenStudio SDK. Clearly, whole-building energy simulation trends have demonstrated the need for simulations that address finer component levels. However, even cosimulations cannot fully address the issue of a simulation's relevance to a particular design because the parameters of cosimulations may be tuned to data from dissimilar buildings. Thus, connections must be established between buildings with high relevance to a particular design and design-stage simulations in order to further enhance the robustness of such simulations.

DDDAS offer a method to establish the aforementioned connections between buildings and simulations. Since their inception, DDDAS have been applied in scenarios as diverse as emergency and disaster management (Madey et al. 2007), supply-chain systems (Celik et al. 2010), and threat management for urban water distribution systems (Wang et al. 2014). Potential DDDAS applications to building performance simulations have been identified as well. For example, Spitler (2006) discussed the use of model-based control systems as an application of DDDAS and the need to explore methods of data-gathering processes under a DDDAS framework. Bouktif and Ahmed (2015) discussed the application of DDDAS to the energy consumption of a residential system, where the application includes real-time metering and sensor data and incorporates human decisions into energy analytics to enhance predictions. In current building performance applications, DDDAS fit well for applications such as real-time model-based simulations for optimizing building operations (Pang et al. 2016).

These previous applications of DDDAS in buildings mainly focused on daily operations; applications of DDDAS to building designs have not yet been reported in the existing literature. The application of DDDAS during the design of new buildings or building retrofits is different from applications to operations, in particular because the input data for design simulations often cannot be collected in situ. In addition, simulations during the design stage often cannot be calibrated using empirical data from buildings that are exact proxies for the building under design (Lomas et al. 1997). Thus, a better strategy is to offer maximum access to building data sets, such that building designers and engineers are more likely to find relevant data from existing buildings for model input or calibration. As discussed, VIFI provides such a possibility by offering a complementary approach to the existing data-centric sharing strategy, increasing access to shared building performance data.

In summary, existing studies concerning building energy simulations call for component-level cosimulations to better address the need for different types of simulations. However, these component-level simulations still rely on data that may not be representative of a particular building under design. Consequently, better access to building performance data is needed. The increasing availability of big building operation data sets coupled with the DDDAS framework offers the possibility of a new data-sharing approach that allows building designers and engineers to seek those data that best match a given design or operation optimization plan and integrate these data with associated building simulations.

Open Computing Infrastructure for Sharing Analytics

This section presents a high-level conceptual architecture for the VIFI open computing infrastructure. Because the topics of simulation models and data collection from existing buildings have already been discussed extensively and are not the major focus of this work, the application of VIFI to the integration of simulations with empirical data is the focus here.

High-Level Conceptual Architecture

Fig. 1 demonstrates the high-level conceptual architecture of VIFI. There are three key components in the infrastructure: predictive models (such as simulations), existing buildings, and the VIFI platform. The predictive model component represents any predictive

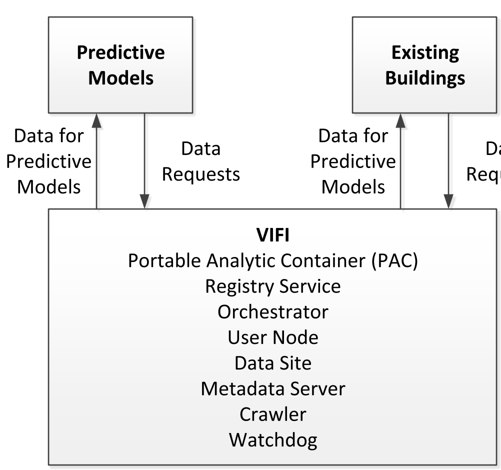


Fig. 1. High-level conceptual architecture of the open data-sharing environment.

model that may be applied during building design process, such as a whole-building energy simulation. The existing building component represents sources of empirical data. Data may be drawn from real-time sensor readings or from other stored logs of building operations, occupant behaviors, indoor or outdoor environmental conditions, and other variables relevant to building research or energy applications. Finally, the VIFI platform represents a computing platform that coordinates the feedback loop between the first two components.

This computing infrastructure extends a conventional application of DDDAS by inserting a generic computing component, the VIFI platform, between computation and instrumentation, the two key components of DDDAS. This extension potentially allows a wider adoption of DDDAS and better support of data sharing. In the following section, the key components of this VIFI platform are detailed.

VIFI Platform

VIFI is an open-source platform that can run on any operating system. Although the final VIFI version will be conveyed on an open-source vault, the Amazon Elastic Compute Cloud (Amazon EC2) was used for the initial implementation, testing, and assessment of the platform. In essence, VIFI brings analytics to locations that contain a large amount of data. VIFI allows users to access data already claimed by others; thus, users of VIFI have the capacity to perform a variety of analyses, such as integrating data from multiple locations, for example. VIFI permits virtual sharing of various sorts of test beds, such as labs that create trial data, without requiring raw data sets that may be too large or sensitive to share outside the control of the primary data owners to be moved. The current implementation of VIFI consists of the following main components (Figs. 2 and 3):

- Portable analytic container (PAC): a PAC is a lightweight virtual machine, called a container, that hosts software (e.g., EnergyPlus), libraries [e.g., Eppy Python module to communicate with EnergyPlus (Eppy 0.5.44)], and any other tools or operating systems required by end-users to analyze data. PACs can receive and execute end-user analysis programs if the required programs are not already contained in the PAC. Leveraging container technology [e.g., Docker (Miell and Sayers 2016)], a PAC is portable and can migrate and execute on heterogeneous host platforms. PACs facilitate reusability by hosting and utilizing different analytical libraries and programs pulled from shared repositories (e.g., DockerHub). Container technology enables the movement of analytics rather than the movement of data, thus alleviating problems related to the transfer of big data, e.g., download times and/or security requirements. Each PAC is a lightweight virtual machine implemented as a Docker image (DockerHub) and containing the required libraries and support programs that are required to execute the desired analysis. PACs offer a number of affordances for distributed analytics: (1) they can be easily transmitted over the network due to their limited size; and (2) they simplify analytics development for inexperienced clients. The VIFI infrastructure, as explained by Talukder et al. (2017) and Elshambakey et al. (2017), is also scalable, i.e., enabling the integration of various VIFI nodes [e.g., a Building Information Model (BIM) server]. The ability for VIFI workflows to access fixed VIFI nodes allows VIFI to cooperate with non-open-source resources, assuming that a VIFI user has the proper credentials.
- Registry service: distinctive PACs are stored, searched, utilized, and shared through a registry service. In the current usage, a docker hub (DockerHub) is utilized to implement the registry service. Future VIFI upgrades will incorporate expansions of

J. Comput. Civ. Eng., 2019, 33(6): 04019037

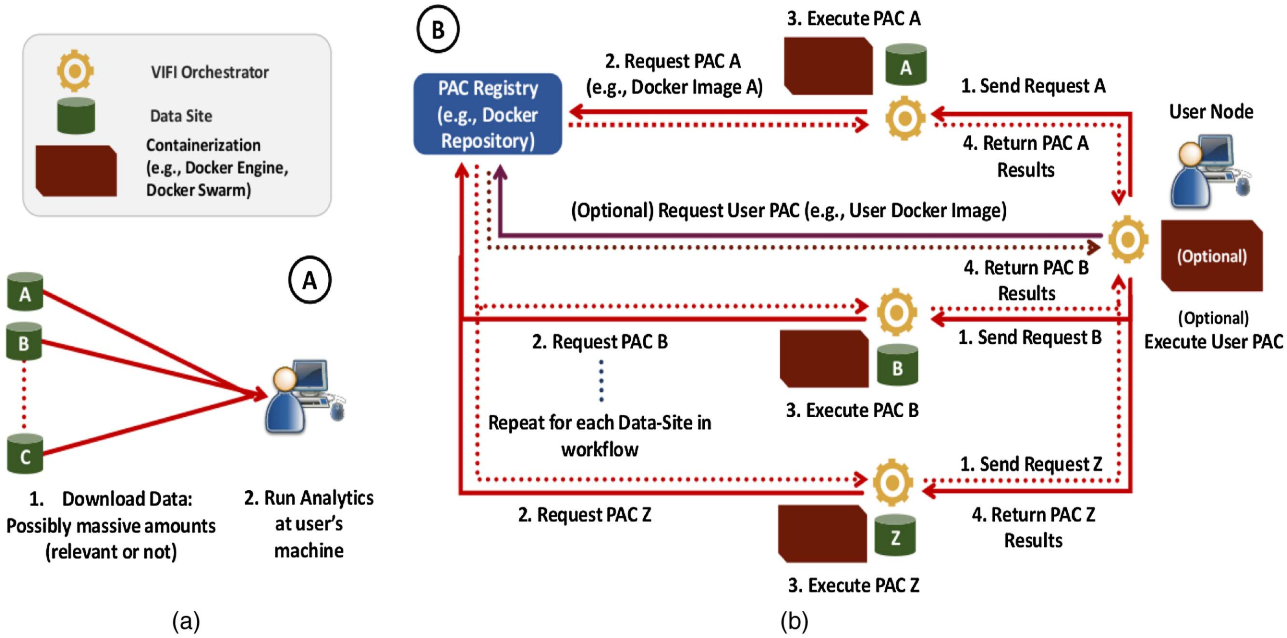


Fig. 2. VIFI system compared with a traditional data analysis system: (a) traditional information fabrics; and (b) VIFI information fabric.

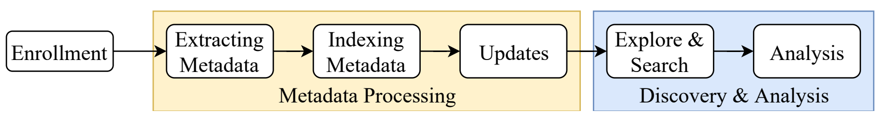


Fig. 3. Data life cycle in VIFI.

appropriated registry repositories to advance download and transfer time of PACs.

- Orchestrator: the orchestrator provides automatic correspondence between distinctive workflow segments and the coordination of infrastructure at various locales. Every segment can be as basic as a solitary procedure running on a host, or the host itself. In the current VIFI implementation, NiFi version 1.1.1 functions as the orchestrator, coordinating the execution of workflow segments across the VIFI infrastructure. NiFi is an open-source tool for automating and managing the flow of information between systems. Workflow segments can dwell on the same host or distributed hosts through NiFi site-to-site communication The VIFI orchestrator includes loosely coupled NiFi interaction and coordination between destinations that empower fault-tolerance and scalability with an increasing number of data stores as well as end-users.
- User node: the user node is the passageway for a client to interface with the VIFI framework. The user node provides a user interface, communication, and basic computation capacities. The current VIFI implementation uses NiFi at every user node to empower correspondence with other workflow parts, as well as workflow inception. Clients can plan, alter, and submit systematic contents as well as other required sources of info to VIFI through a NiFi web interface.
- Data site: data sites are locations in the VIFI infrastructure that have varying types of heterogeneous information from different sources and models. Each VIFI-empowered information site uses NiFi and Docker Swarm (Vohra 2017). NiFi empowers appropriated coordination and correspondence between every data site and other parts of the workflow, and Docker Swarm executes clients' contents utilizing determined PACs as Docker orchestrators. Consequently, Docker Swarm conveys a group of services at every data site for parallel investigation execution without expecting clients to have any previous knowledge of the framework, stages, or situations at every datum site.
- Metadata server: the metadata server stores and lists gathered data and comments about data sets and records, which can be created by utilizing a set of enlisted extractors or physically included by clients and information proprietors. For movability, extractors are containerized as PACs. A metadata server supports data set search and discovery across different sites, as shown in Fig. 3.
- Crawler: the crawler provides data-crawling through extractors that send extracted data to the metadata server.
- Watchdog: The watchdog updates the metadata server for any modifications of the registered distributed data sets.

Fig. 2 outlines the current VIFI implementation [Fig. 2(a)] and compares this implementation against conventional data-centric sharing approaches [Fig. 2(b)]. Under conventional data-centric sharing approaches [Fig. 2(a)], users must first download data locally and then run required analytics. Under the VIFI implementation [Fig. 2(b)], users submit required analytics (including a Docker image and scripts) to data locations and then retrieve post-processed results. The example in Fig. 2 shows the execution of multiple requests simultaneously in both cases.

This paper focuses on the data privacy benefits of VIFI more than the benefits of VIFI in handling large data sets; nevertheless, VIFI was previously evaluated for an earth science application, where observational data (≅11 GB) and model data (≅21 GB) were analyzed and compared to evaluate the fitness of the model

Table 1. Comparison of transmission time between the conventional approach and proposed VIFI approach for the earth science use case

Data type	Data size (MB)	Source	Destination	Time (s)
Conventional approach				
Model data	21,296	Data site 3	Data site 1	246.088
Observation data	11,373.308	Data site 2	Data site 1	126.934
VIFI approach				
Docker image (one-time transfer)	578.13	AWS ECR	Data site 1	0.156
Docker image (one-time transfer)	578.13	AWS ECR	Data site 2	0.128
Docker image (one-time transfer)	578.13	AWS ECR	Data site 3	0.135
User script and configuration file	0.008	Data site 1	Data site 2	0.183
User script and configuration file	0.008	Data site 1	Data site 3	0.145
Result files	0.156	Data site 2	Data site 1	0.186
Result files	0.156	Data site 3	Data site 1	0.246

(Talukder et al. 2017; Elshambakey et al. 2017). In this setting, it was observed that VIFI consumed far less processing time than a conventional data-centric approach. This is because under VIFI, Docker images need be downloaded only once and then again as updates are available. The data update rate at each data site under the conventional approach is expected to be much higher than the update rate for Docker images; accordingly, users of the conventional approach must download updated data far more often than they would need to update Docker images under the VIFI approach. Table 1 illustrates the execution time saving by VIFI, as found by Talukder et al. (2017).

As indicated in Table 1, the data-transfer time in the case of the conventional approach is 246.088 s where model and observation data are transferred simultaneously from Data sites 2 and 3 to the user node (Data site 1). Due to the workflow implementation in the case of VIFI, the maximum transfer time is calculated

$$\mathrm{Max}(T1, T2) + T3$$

where $T1$ = transfer time of user scripts and configuration files from Data site 1 to Data site 2 + Docker image transfer time from Amazon Web Services (AWS) Elastic Container Registry (ECR) to Data site 2 + transfer time of results from Data site 2 to Data site 1; $T2$ = transfer time of user scripts and configuration files from Data site 1 to Data site 3 + Docker image transfer time from AWS ECR to Data site 3 + transfer time of results from Data Site 3 to Data site 1; and $T3$ = Docker image transfer time from AWS ECR to Data site 1.

Thus, the maximum transfer time under VIFI is 0.682 s compared with 246.088 s in the conventional approach.

Case Study

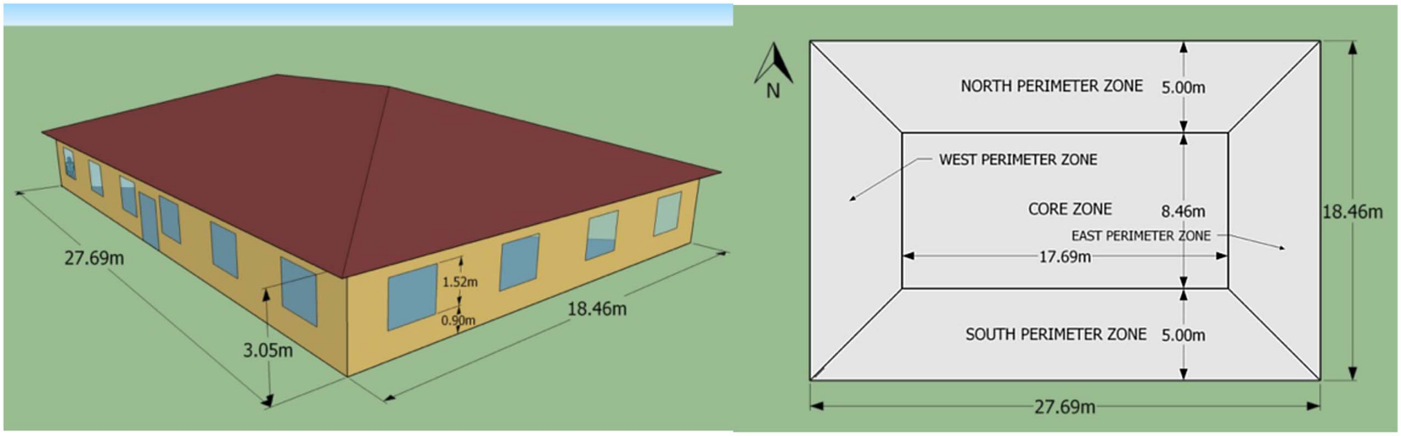
A case study is used to demonstrate the application of VIFI for the integration of empirical data with whole-building energy simulations in the early building design phase. The case study generates dynamic lighting schedules using data processing algorithms within a PAC. Algorithms extract the necessary information from empirical building data collected from the Intelligent Workspace at Carnegie Mellon University. The Intelligent Workplace at Carnegie Mellon University is a living and lived-in research laboratory dedicated to prototyping and testing of integrative high-performance building technologies such as shading integrated solar photovoltaics, automated operable windows, automated adjustable external shading devices, and daylight-controlled dimmable electric lights. Building data (including occupancy presence) are collected from these systems through an extensive wired and wireless sensor network and data-acquisition systems.

Measured building operation data from the Intelligent Workspace are utilized remotely by a building energy simulation tool, EnergyPlus, via the VIFI platform. The VIFI platform facilitates the operation of an analysis procedure that converts the collected empirical data into hourly building lighting schedules. Without moving any raw empirical data from its original location, VIFI executes analytics that transform the raw data collected by sensors in the Intelligent Workspace every 5 min into an hourly lighting schedule in a specific format that is readable by the remote EnergyPlus simulation engine. VIFI supports the integration of this empirically derived lighting schedule with EnergyPlus as well as the execution of annual energy performance simulations that reflect the schedule.

Description of the Building Energy Model

The small-sized commercial reference office building provided by the US Department of Energy Building Technologies Office (Deru et al. 2011) was used for the case-study simulations. This model has been developed in the format of EnergyPlus version 8.7, which has been extensively utilized for computational evaluations of integrated building energy performance, mechanical and electrical systems, and thermal interactions of building spaces with increased spatial and temporal granularity (Field et al. 2010). The specific small-sized office model is selected from a comprehensive database containing 16 different hypothetical reference building definitions that represent about 70% of the new commercial building stock in the United States (DOE-EERE 2018).

The DOE small office reference model is a single-story commercial office with a total conditioned and usable floor area of 511 m^2. The building has a rectangular shape and an aspect ratio of 1.5, with its long axis on a north–south orientation (Fig. 4). Punched windows are used to represent vertical fenestration configurations, which are uniformly distributed to each cardinal orientation with a window-to-wall ratio (WWR) of 21.2% (total window area is 59.6 m^2). Thermal zone layouts are formed using the perimeter-core zoning approach, resulting in five thermal zones and an unconditioned and unoccupied attic space. The building envelope is composed of thermally massive and insulated construction assemblies that comply with the minimum requirements imposed by the American National Standards Institute/American Society of Heating, Refrigerating and Air-Conditioning Engineers/Illuminating Engineering Society of North America (ANSI/ASHRAE/IESNA) Energy Standard 90.1 for nonresidential building types (ASHRAE 2004). The thermophysical properties of the main building envelope assemblies are listed in Table 2. The model is equipped with a double-pane glazing system defined in a simplified manner with U-values, the solar heat gain coefficient (SHGC), and V_T inputs of 3.241, 0.385, 0.305, respectively (including the thermal and optical effects of frames and dividers). Internal loading conditions pertaining to artificial lights, electrical appliances, and occupancy are given in Table 3 together with outdoor ventilation rates, infiltration rates, and thermostatic control limits for indoors.



(a) (b)

Fig. 4. Overall geometry of the reference building model: (a) three-dimensional exterior view; and (b) plan layout of a typical floor.

Table 2. Envelope thermophysical properties

Assembly	U-factor (W/m^2 K)	Thickness (m)
Walls	0.698	0.298
Attic soffit	0.193	0.264
Roof	5.082	0.011
Ground slab	2.193	0.101

Table 3. Internal loading conditions

Model parameter	Value
Lighting power density	10.76 W/m^2
Equipment power density	10.76 W/m^2
Occupancy density	18.58 m^2/person
Outdoor air ventilation rate	10 L/s/person
Envelope infiltration rate	0.000302 m^3/(s · m^2)
Heating set-point/set-back temperature	21°C/15.6°C
Cooling set-point/set-back temperature	24°C/26.7°C

The HVAC system type modeled for the building is a group of packaged single-zone air conditioners (PSZ-AC) with single-speed direct expansion (DX) coils and compact gas furnaces for heating defined for each of the five thermal zones. There is not a central air conditioning unit (AHU) or an economizer in this configuration. System fans are of the constant-volume type without variable-speed drivers (Table 4). The reference office model is simulated under the environmental boundary conditions for the location of Pittsburgh, Pennsylvania. A detailed statistical climate summary is provided in Table 5.

Table 4. HVAC system properties

System parameter	Definition/value
Cooling system	Single-speed DX cooling coils COP = 3.66W/W
Heating system	Gas burner nominal efficiency = 0.80
Fans (only supply)	Constant-volume fans with efficiency of 0.54 at maximum 622 Pa
Minimum air flow rate fraction (turndown ratio)	Autosized to meet zone ventilation requirements (from occupancy)
OA controller	No economizers
Design cooling supply air temperature	14.0°C
Design heating supply air temperature	40.0°C

Note: COP = coefficient of performance; DX = direct expansion; and OA = outside air.

Table 5. Building location and climate characteristics

Parameter	Pittsburgh, Pennsylvania ASHRAE climate 5A
Weather station number	TMY3—725200
Coordinates	40°30′ N; 80°13′ W
Elevation (m)	350
Global solar radiation (Wh/m^2)	3,804.5
Direct solar radiation (Wh/m^2)	3,112.5
Diffuse solar radiation (Wh/m^2)	1,956.3
Wind speed (m/s)	3.94
Dry-bulb temperature (°C)	10.50
Relative humidity (%)	66.41
Heating degree days (18.3°C)	3,124
Cooling degree days (18.3°C)	417

VIFI Implementation of the Case Study

Fig. 5 shows the VIFI implementation of the case study. Although the case study was implemented on the Amazon EC2, the authors emulated a distributed environment as in actual applications, including a building information model (BIM) server, a user node, a data server, and an energy server. The BIM server contains the BIM of the case building. The user node represents the location where building designers and engineers work. The data server represents the access point to empirical data, where raw lighting schedule data are stored in this case. The energy server is where EnergyPlus is located.

The current capability of VIFI can be extended by incorporating fixed VIFI nodes (e.g., BIM server and EnergyPlus server) where analytics do not have to run inside containers (Talukder et al. 2017; Elshambakey et al. 2017). The ability of VIFI workflows to access fixed VIFI nodes allows VIFI to cooperate with non-open-source resources given that a VIFI user has the proper credentials. The services provided by the BIM server, the data server(s), and the

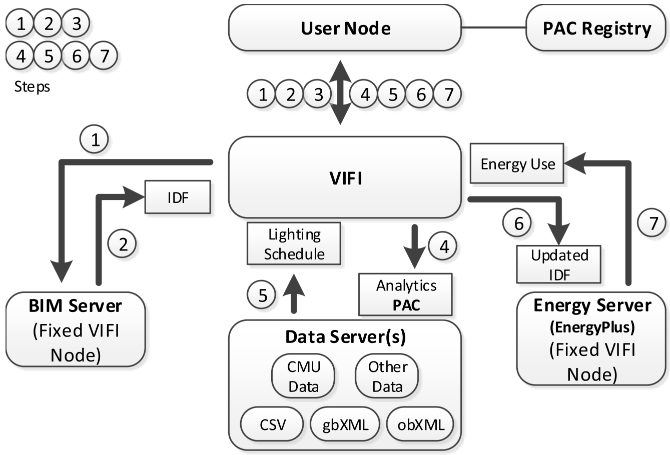


Fig. 5. VIFI application case study for building energy simulations.

energy server are registered with VIFI so that a user can determine what operations or analytics are available for their applications. In the case study, the following operations and analytics are registered with VIFI first in the PAC registry (Fig. 5):

- TheBIM server (fixed VIFI node) accepts requests to create an input data file (IDF) based on the BIM model for energy simulations and returns an IDF.
- The data server creates and returns hourly lighting schedules based on the raw data.
- The energy server (fixed VIFI node) accepts an IDF and returns energy simulation results.

The following is the workflow of the VIFI implementation of the case study:

In Step 1, a building designer or engineer contacts the BIM server to get an EnergyPlus Input Data File (IDF) (Step 2), which contains default schedule settings. At this point, some or all parts of the IDF need to be updated based on proper data analysis (e.g., lighting schedule or occupancy schedule). In this case study, the BIM server was accessed directly to simplify the implementation. In future implementations, the BIM server can be implemented as a fixed VIFI node that does not require a PAC to access. Sharing a BIM directly with all partners is a conventional data-centric strategy; therefore, the building designer or engineer can access BIM on a shared network. The integration of BIM with VIFI represents an extension to the conventional data-sharing strategy. The new strategy (i.e., sharing analytics) and the conventional strategy (i.e., sharing data directly) are complementary to each other and can work hand-in-hand to solve practical problems.

Steps 3, 4, and 5 create an IDF that better reflects design or operation conditions. In Step 3, the building designer or engineer searches for lighting schedules that best match the expected design or operation conditions. The metadata server of VIFI (Fig. 3) helps identify relevant data sources that are accessible through VIFI. In the case study, the authors assume that the collected data source from the Intelligent Workspace represents the best match.

Once the data source or sources are identified, the building designer or engineer sends a request to the data source(s) to generate the required schedule in Step 4. In the case of energy simulations, the operations needed to develop better schedules are implemented as standard analytics in VIFI. The analytics are small-sized applications that run inside Docker containers (Miell and Sayers 2016) using Docker Swarm (Vohra 2017). Docker containers are produced using Docker images hosted at different VIFI registries (e.g., DockerHub). Docker images contain all required libraries/dependencies to run the user's analytics. Thus, end users such as a building designer or engineer do not have to worry about configuration of the required analytics.

The analytics are executed on data at the data source(s). In the case study, the raw data were collected by sensors every 5 min continuously over multiple years at the Intelligent Workspace. The raw data cannot be used by others directly for different reasons: (1) the occupancy or lighting schedule requires hourly data but the raw data are collected every 5 min; (2) the data may be in different formats and schemas [e.g., csv, gbXML (GreenBuildingXML), or obXML (Hong et al. 2015a, b)] from those needed for creating the IDF; (3) the raw data may contain data that are irrelevant to the analysis; and (4) the data are too sensitive to share in their raw format because they may compromise occupant privacy. Thus, instead of downloading data from the Intelligent Workspace to the building designer's or engineer's local site, the data owner [Carnegie Mellon University (CMU) in this case] can publish the analytics that are allowed to operate on its data sources to VIFI, and the building designer or engineer can find the analytics at the PAC registry. The building designer or engineer then uses the analytics to access data and obtain the desired schedule updates. Importantly, this approach does not require a user to have significant knowledge about programming, so the person can stay focused on the main simulation tasks.

Once the building designer or engineer receives the updated hourly lighting schedule, the original IDF is updated with this new information in Step 5.

Finally, the updated IDF file is used to run energy simulations using the EnergyPlus server (Cao et al. 2011) in Step 6. In Step 7, the building designer or engineer receives simulation results.

Creating, Sharing, and Executing Analytics

Currently VIFI supports two mechanisms for creating, sharing, and executing analyses. The first mechanism uses Docker containers to encapsulate the analytics. This mechanism supports the reusability of different analysis applications and building new analytics from existing ones. The second mechanism uses a fixed-point VIFI node to host a specific software such as a BIM. Fixed-point VIFI nodes are useful in cases where VIFI must interact with software that is difficult to containerize (e.g., proprietary software without available Docker images).

Fig. 6 shows the sample application of the first mechanism to create a lighting schedule for energy simulation. The authors used lighting schedules as an example in the case study to demonstrate creating, sharing, and executing analytics—lighting schedules are one of many input schedules to EnergyPlus. Analytics take the form of a dedicated Python script that generates a lighting schedule in input definition file (IDF) format using the empirical data collected at CMU. The input to this analytics script is a file in comma-separated values (CSV) format with two columns: the first column contains timestamps, and the second column indicates the artificial light status, either on or off, as recorded by sensors. The output is a whole year *Schedule:Compact* IDF object. Between the input and the output, the analytics perform data parsing and extraction using the Python Pandas library. The script is generic and can thus be shared among different application cases to create improved lighting schedules.

The implementation of the second mechanism is straightforward. The Eppy Python module (Eppy 0.5.44) was used to integrate third-party software such as EnergyPlus with VIFI. Despite EnergyPlus being open-source software, EnergyPlus is used in this use case as an example of integrating non-open-source software with VIFI [i.e., VIFI has to interact with the machine hosting EnergyPlus to use it, rather than downloading a Docker image hosting EnergyPlus to the machine(s) executing analytics]. In the case study, the

Input Empirical Field Data	VIFI Docker Image	Output IDF Lighting Schedule
TimeStamp Artificial Light Status 6/22/2016 0:00 off 6/22/2016 0:05 off … … 6/22/2016 7:00 on 6/22/2016 7:05 on … … 6/22/2016 10:00 on 6/22/2016 10:05 on … … 6/25/2016 0:00 off 6/25/2016 0:05 off … … 6/25/2016 7:00 off 6/25/2016 7:05 on … … 6/25/2016 10:00 on 6/25/2016 10:05 on	Analytics: 1. Sort input data into categories: Weekdays, Saturday, SummerDesignDay, WinterDesignDay, and All Other Days based on the TimeStamp, 2. In each category, sort data by hour or a desired time period, 3. For each hour or time period, calculate the frequency and the probability of light-on using the sample data, 4. Use the probability as the fraction value for each hour or time period in the schedule.	Schedule:Compact, BLDG_LIGHT_SCH, !- Name Fraction, !- Schedule Type Limits Name Through: 12/31, !- Field 1 For: Weekdays, !- Field 2 Until: 05:00, !- Field 3 0.05, !- Field 4 … Until: 08:00, !- Field 7 0.3, !- Field 8 Until: 17:00, !- Field 9 0.9, !- Field 10 … Until: 24:00, !- Field 19 0.05, !- Field 20 For: Saturday, !- Field 21 Until: 06:00, !- Field 22 0.05, !- Field 23 Until: 08:00, !- Field 24 0.1, !- Field 25 … Until: 24:00, !- Field 30 0.05, !- Field 31 For: SummerDesignDay, !- Field 32 Until: 24:00, !- Field 33 1.0, !- Field 34 For: WinterDesignDay, !- Field 35 Until: 24:00, !- Field 36 0.0, !- Field 37 For: AllOtherDays, !- Field 38 Until: 24:00, !- Field 39 0.05; !- Field 40

Fig. 6. Sample application of analytics.

Linux-based EnergyPlus Server version 8.7.0 is used. EnergyPlus Server runs the energy simulation based on the updated IDF file and returns the energy use of the sample building.

Simulations and Results

To demonstrate the usefulness of VIFI platform in improving building energy simulations from the point of occupancy representativeness, the authors performed a series of simulation runs using comparable models. A baseline simulation model representing common energy modeling practices uses a standard lighting schedule provided by the Building Technologies Office of the US Department of Energy (BTO 2018). Other simulation models use empirically derived lighting schedules derived based on three occupancy profiles, i.e., a professor, a graduate student, and a postdoctoral researcher. The standard lighting schedule lacks relevance to actual design or operational conditions, whereas empirically derived schedules hold certain relevance to specific design or operation conditions. The main objective of this demonstration and subsequent comparative analyses is to reveal the impact of using empirically derived schedules on the predictive capability of whole-building energy models. In turn, the impact shows the importance of using VIFI to improve building energy simulations.

Fig. 7 clearly indicates that empirically derived electric lighting schedules pertaining to different occupancy types do not match well with standardized schedules (shown in dashed lines), which lack dynamism within a typical weekday. The results also signify the need to use design-specific or building case-specific operational schedules to enhance the predictive capability of building energy models. Meanwhile, empirically derived lighting schedules also show some degree of variations among themselves, which is due to highly varied space use styles for the selected occupant types. Compared with the graduate student (with 1,633 h of occupancy), the user types of professor and postdoctoral researcher spend much more time in their offices, with occupied hours of 3,094 and 2,676 h, respectively. However, all occupancy time is shorter than the assumption of standard schedules, 5,097 h in a typical year.

Given fractional hourly schedules, which are drawn from the outputs from EnergyPlus' annual energy simulation runs, the authors calculated the cumulative power fraction (CPF) for electric lighting systems, defined as the sum of the number of hours in a year when systems are operating in full power, and conducted the comparisons given in Table 6. The CPF metric, together with its derivative of full-power equivalent (FPE) metric where FPE $=$ CPF/8,760, revealed that although occupancy types of professor (Type I) and postdoc (Type III) had very close CPFs compared with the standardized schedule, there existed a considerable deviation for the graduate student (Type II) with a 45% change from the standard CPF, i.e., 2,863. This is due to the fact that Type I and Type III have similar peak power levels to those of the standard schedule, whereas Type II shows variations in peak power, its frequency, and event timing.

It is clear that with similar CPF and FPE metrics, the annual lighting energy consumption of Type I (professor) is very close to the model alternative using a standard schedule (only -0.7% deviation observed between the two). Type III (postdoc) occupancy type has a CPF and FPE of 2,437 and 0.28, respectively, for lights, which results in about 14.9% decrease in annual lighting energy consumption compared with the standardized schedule model. Simulation results for Type II (graduate student) show significant departures from the standard schedule model with deviations of up to -45.7% for lighting energy consumption.

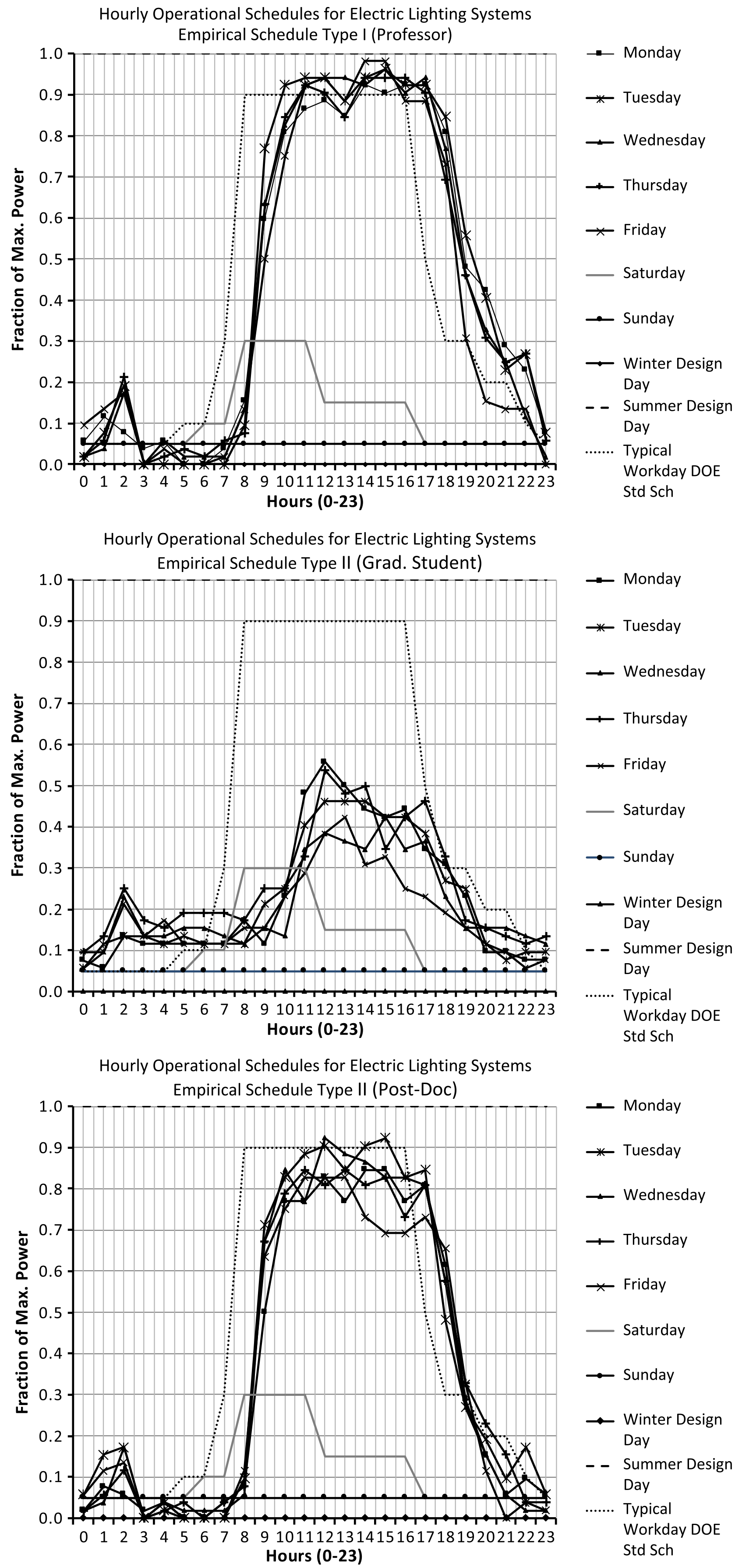


Fig. 7. Comparisons of hourly schedules for electric lights of the alternative models.

Table 6. Comparison of annual lighting schedules types and occupancy

Empirical schedule	Type I (professor)	Type II (graduate student)	Type III (postdoctoral)	DOE Std. Sch.
Cumulative power fraction (CPF)	2,843	1,553	2,437	2,863
Full-power equivalent (FPE)	0.32	0.18	0.28	0.33
Occupied hours	3,094	1,633	2,676	5,097
Unoccupied hours	5,503	6,964	5,921	3,663
Input/output timeout hours	163	163	163	N/A

Note: DOE Std. Sch. = Department of Energy Standard Schedule; FPE = CPF/8,760; and input/output timeout = September 12, 2016, 7:00 p.m. to September 19, 2016, 9:00 a.m. and May 25, 2016, 7:00 a.m. to May 26, 2016, 9:00 a.m.

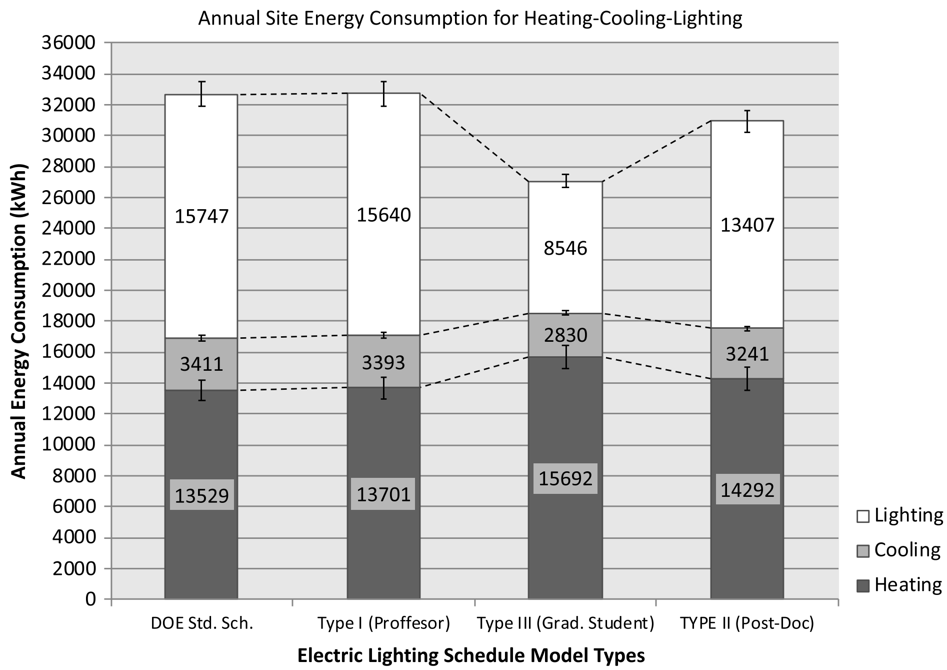


Fig. 8. Comparisons of annual site energy consumptions.

To demonstrate the dynamic effects of alternative hourly lighting schedules on cumulative energy performance, annual energy simulations were performed, facilitated by the VIFI platform. Annual site energy consumption levels for the end uses of space heating, cooling, and electric lighting (kWh) were extracted from annual EnergyPlus simulation runs for comparison purposes (Fig. 8). Because the lights in a building energy model contribute to internal heat gains, deviations in the predictions of the energy models are also found in space heating and cooling energy consumptions. For example, results in Fig. 8 indicate that due to reduced usage of lights in the Type II (graduate student) model, annual space heating energy was increased by 16%, whereas annual space cooling energy was reduced by 17% compared with the standard schedule model. Although such increments and decrements appear to balance each other out in the particular climate zone (Pittsburgh, Pennsylvania) where heating dominates, such effects would have important implications for cooling energy savings estimates in cooling-dominated climates.

Discussion

The conducted case study demonstrates three key attributes of the VIFI computing infrastructure. First, the case study shows the technical feasibility and application potentials of integrating data-centric and analytic-centric sharing strategies. In the building industry, data-centric sharing has been the norm, and technologies such building information models are vehicles allowing stakeholders to directly share data created by others. The case study shows that VIFI complements the conventional approaches by utilizing both data and analytics for sharing purposes. This capability may allow building designers and engineers to access more data than what conventional methods such as building information models can provide. Second, the case study shows the VIFI platform enables the use of empirically driven operational building schedules for energy performance simulations instead of relying on predetermined and generic schedules, which have no apparent contextual and occupational relevance to building design and operation characteristics. Such an approach can significantly affect the predictive ability and representativeness of energy performance models and can be a critical element of model calibration and validation procedures. Third, the case study has demonstrated the technical feasibility of the creation, sharing, and execution of analytics that can be executed across distributed sites (e.g., BIM server, data server, and energy server). This capability is at the core of the VIFI platform and its proposed computing infrastructure.

Furthermore, the case study demonstrates that the VIFI platform can be instrumental to the automation of specific labor-intensive energy modeling tasks (e.g., repeated development of annual hourly schedules for different building spaces with list lengths as high as

8,760 elements). Such automation not only reduces human errors but also speeds up the process of model development and interpretation of results. Additionally, the VIFI computing infrastructure can be instrumental to model-based building control applications, which carry building energy models from the design stage through to the operational stage in order to perform predictions across a specific time horizon, allowing optimal control of building mechanical and electrical systems for increased energy efficiency and occupant comfort.

Conclusions and Future Research

This study has demonstrated the potential of VIFI, an open computing infrastructure, to offer a novel analytics-based data-sharing strategy that complements conventional data-centric strategies. By sharing analytics on data, the new strategy points toward a future where empirical data from existing building operations are closely connected with decision-making processes that drive building design and operation through additional cyberinfrastructure. The study has demonstrated many possibilities for using this novel infrastructure to enhance building performance simulations. For example, the study has shown the potential for the VIFI platform to support both data-centric and analytic-centric sharing. This new strategy bears significant implications for the building design and engineering community, where data fragmentation has been recognized as a major challenge to data sharing. As shown in the case study, designers and engineers can access both a building information model and empirical data in an integrated manner through VIFI. This strategy enables designers and engineers to access a larger pool of data that are not only more relevant to improving their simulations, but may also set the foundation for future applications of advanced methods such as machine learning to transfer knowledge that is gained during building operations back to new building designs.

This study only examined the functional capabilities of the open computing infrastructure, and many other aspects related to the infrastructure require further research. For example, the presented case study was implemented with the help of computer scientists who are technical experts in deploying and implementing applications using VIFI. Looking ahead, it is critical that the VIFI platform is useful to those in design and engineering domains who do not have the technical background to understand technical details of the proposed infrastructure such as the PAC and registry. Therefore, a user-friendly workflow process is important to develop for future VIFI implementations. In addition, although the VIFI platform does not need data owners to share data per se, a certain level of access to data needs to be granted to outside users in order for the VIFI analytics to work on the data sets. This requirement goes beyond the issue of addressing data privacy and security issues, bringing forth more fundamental questions about the motivation and benefits of sharing data. Rich data sets are needed for access by VIFI analytics in order to realize the full benefits of the proposed infrastructure, and the incentives for creating and sharing such data sets must be improved. For example, "information as commodity" (Smith 1983) provides an interesting idea to drive data creators and users to share data. Such an idea can be pilot-tested in a small community, such as those working on the development and initial applications of VIFI. Experimentation with this idea may shed more light on how to increase the sharing of data, which is important to other VIFI functions.

Another important issue to be explored is the quality of metadata, which plays a significant role in searching for, reusing, and executing analytics on the VIFI platform. Although metadata are not addressed in this case study, this topic has been discussed for decades in the buildings research community and as a result, many metadata schemas are available, for example gbXML, industry foundation classes (IFC), obXML, and CityGML. The VIFI research community must develop a strategy for building from such existing metadata schemas and integrating new concepts as needed to support VIFI applications. Finally, there is a need to develop a set of performance metrics to measure the benefits and risks of using VIFI; these metrics should include technical, economic, and social measures at the individual, company, and society levels.

Acknowledgments

This paper was partially supported by the US National Science Foundation—Data Infrastructure Building Blocks (DIBBS) program's funding for the project "Virtual Information Fabric Infrastructure (VIFI) for Data-Driven Decisions from Non-Shareable Distributed Data" (NSF DIBBs Award 1640818). This work was also supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, US Department of Energy, under Contract No. DE-AC02-05CH11231. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or US Department of Energy.

References

ASHRAE (American Society of Heating, Refrigeration and Air Conditioning Engineers). 2004. *Energy standard for buildings except low-rise residential buildings*. ANSI/ASHRAE 90.1. Atlanta: American Society of Heating, Refrigeration and Air Conditioning Engineers.

Bouktif, S., and W. K. Ahmed. 2015. "Monitoring system for residential energy consumption." *Int. J. New Comput. Archit. Appl.* 5 (4): 156–164.

BTO (Building Technologies Office). 2018. "The U.S. Department of Energy—Builing Technodlogies Office (BTO)." Accessed July 01, 2019. https://www.energy.gov/eere/buildings/building-technologies-office.

Cao, Y., H. Li, W. Dong, Z. Tang, and G. Zhang. 2011. "Energy-saving performance analysis of distributional building energy system." In *Proc., Asia-Pacific Power Energy Eng. Conf.*, 1–4. New York: IEEE.

Celik, N., S. Lee, K. Vasudevan, and Y.-J. Son. 2010. "DDDAS-based multi-fidelity simulation framework for supply chain systems." *IIE Trans.* 42 (5): 325–341. https://doi.org/10.1080/07408170903394306.

Chen, Y., T. Hong, and X. Luo. 2018. "An agent-based stochastic occupancy simulator." *Build. Simul.* 11 (1): 37–49. https://doi.org/10.1007/s12273-017-0379-7.

Coakley, D., P. Raftery, and M. Keane. 2014. "A review of methods to match building energy simulation models to measured data." *Renewable Sustainable Energy Rev.* 37 (Sep): 123–141. https://doi.org/10.1016/j.rser.2014.05.007.

Crawley, D. B., et al. 2001. "EnergyPlus: Creating a new-generation building energy simulation program." *Energy Build.* 33 (4): 319–331. https://doi.org/10.1016/S0378-7788(00)00114-6.

Darema, F. 2005. "Dynamic data driven applications systems: New capabilities for application simulations and measurements." In *Proc., Int. Conf. on Computer Science*, 610–615. Berlin: Springer.

Deru, M., et al. 2011. *US Department of Energy commercial reference building models of the national building stock*, 1–118. Lakewood, CO: National Renewable Energy Laboratory.

D'Oca, S., and T. Hong. 2015. "Occupancy schedules learning process through a data mining framework." *Energy Build.* 88 (Feb): 395–408. https://doi.org/10.1016/j.enbuild.2014.11.065.

DOE-EERE (Department of Energy-Office of Energy Efficiency and Renewable Energy). 2018. "Commercial reference buildings." Accessed June 24, 2018. https://www.energy.gov/eere/buildings/commercial-reference-buildings.

Elshambakey, M., M. Khalefa, W. J. Tolone, S. Das Bhattacharjee, H. Lee, L. Cinquini, S. Schlueter, I. Cho, W. Dou, and D. J. Crichton. 2017.

"Towards a distributed infrastructure for data-driven discoveries & analysis." In *Proc., 2017 IEEE SmartWorld, Ubiquitous Intelligence Computing Advanced and Trusted Computed, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation*, 1–8. San Francisco, CA: IEEE.

Feng, X., D. Yan, and T. Hong. 2015. "Simulation of occupancy in buildings." *Energy Build.* 87 (Jan): 348–359. https://doi.org/10.1016/j.enbuild.2014.11.067.

Field, K., M. Deru, and D. Studer. 2010. *Using DOE commercial reference buildings for simulation studies*. Rep. No. NREL/CP-550-48588. Golden, CO: National Renewable Energy Laboratory.

Guglielmetti, R., D. Macumber, and N. Long. 2011. *OpenStudio: An open source integrated analysis platform*. Rep. No. NREL/CP-5500-51836. Golden, CO: National Renewable Energy Laboratory.

Hong, T., S. D'Oca, S. C. Taylor-Lange, W. J. N. Turner, Y. Chen, and S. P. Corgnati. 2015a. "An ontology to represent energy-related occupant behavior in buildings. Part II: Implementation of the DNAS framework using an XML schema." *Build. Environ.* 94 (P1): 196–205. https://doi.org/10.1016/j.buildenv.2015.08.006.

Hong, T., S. D'Oca, W. J. N. Turner, and S. C. Taylor-Lange. 2015b. "An ontology to represent energy- related occupant behavior in buildings. Part I: Introduction to the DNAs framework." *Build. Environ.* 92 (Jan): 764–777. https://doi.org/10.1016/j.buildenv.2015.02.019.

Hong, T., H. Sun, Y. Chen, S. C. Taylor-Lange, and D. Yan. 2016. "An occupant behavior modeling tool for co-simulation." *Energy Build.* 117 (Apr): 272–281. https://doi.org/10.1016/j.enbuild.2015.10.033.

Kwak, Y., J.-H. Huh, and C. Jang. 2015. "Development of a model predictive control framework through real-time building energy management system data." *Appl. Energy* 155 (Oct): 1–13. https://doi.org/10.1016/j.apenergy.2015.05.096.

Lomas, K. J., H. Eppel, C. J. Martin, and D. P. Bloomfield. 1997. "Empirical validation of building energy simulation programs." *Energy Build.* 26 (3): 253–275. https://doi.org/10.1016/S0378-7788(97)00007-8.

Madey, G. R., et al. 2007. "Enhanced situational awareness: Application of DDDAS concepts to emergency and disaster management." In *Proc., Int. Conf. on Computer Science*, 1090–1097. Berlin: Springer.

Miell, I., and A. H. Sayers. 2016. *Docker in practice*. New York: Manning.

Pan, Y., Z. Huang, and G. Wu. 2007. "Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai." *Energy Build.* 39 (6): 651–657. https://doi.org/10.1016/j.enbuild.2006.09.013.

Pang, X., T. S. Nouidui, M. Wetter, D. Fuller, A. Liao, and P. Haves. 2016. "Building energy simulation in real time through an open standard interface." *Energy Build.* 117 (Apr): 282–289. https://doi.org/10.1016/j.enbuild.2015.10.025.

Roth, A. 2017. "Building Energy Modeling 101: What is it and what is DOE's role?." Accessed June 24, 2019. https://www.energy.gov/eere/buildings/articles/building-energy-modeling-101-what-it-and-what-does-role.

Smith, K. A. 1983. "Information as a commodity or public good." *Natl. Forum* 63 (3): 27.

Spitler, J. 2006. "Editorial: Building performance simulation: The now and the not yet." Supplement, *HVAC&R Res.* 12 (S1): 711–713. https://doi.org/10.1080/10789669.2006.10391194.

Talukder, A., M. Elshambakey, S. Wadkar, H. Lee, L. Cinquini, S. Schlueter, I. Cho, W. Dou, and D. J. Crichton. 2017. "VIFI: Virtual information fabric infrastructure for data-driven discoveries from distributed earth science data." In *Proc., 2017 SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, 1–8. New York: IEEE.

Vohra, D. 2017. *Docker management design patterns: Swarm mode on Amazon web services*. New York: Apress.

Wang, L., D. Chen, W. Liu, Y. Ma, Y. Wu, and Z. Deng. 2014. "DDDAS-based parallel simulation of threat management for urban water distribution systems." *Comput. Sci. Eng.* 16 (1): 8–17. https://doi.org/10.1109/MCSE.2012.89.

Wetter, M. 2011. "Co-simulation of building energy and control systems with the building controls virtual test bed." *J. Build. Perform. Simul.* 4 (3): 185–203. https://doi.org/10.1080/19401493.2010.518631.

Zhu, Y. 2006. "Applying computer-based simulation to energy auditing: A case study." *Energy Build.* 38 (5): 421–428. https://doi.org/10.1016/j.enbuild.2005.07.007.