A taxonomy of data types and data collection methods for building energy monitoring and performance simulation

Bharadwaj R.K. Mantha, Carol C. Menassa* and Vineet R. Kamat

Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward Street, Ann Arbor, MI 48109, USA

(Received 20 May 2015; accepted 24 September 2015)

Buildings contribute a major component of the world’s total energy consumption. During the design phase, building energy simulation models can be used to predict building performance and design energy-efficient buildings that consume an optimal amount of energy, if operated and maintained within the confines of energy model assumptions. Similarly, Dynamic Data-Driven Simulation and Analysis (DDDSA) methods can be used during the operation and maintenance phase of buildings. Such methods allow models to be seeded periodically with real-time data that influences the simulation system for better analysis and prediction of system performance. In order to optimize energy consumption in buildings and subsequently improve their energy efficiency, buildings need to be monitored and various forms of copious data need to be collected to support DDDSA. The primary objective of this paper is to classify different data types that affect the energy consumption in buildings, and provide a comprehensive review of the existing data collection methods. The taxonomy developed, relevance of each data type along with the general characteristics of different data collection methods are discussed. Finally, challenges and limitations of gathering data with existing data collection methods in the context of building energy management are described along with proposed areas for future research.

Keywords: building energy; occupant comfort; environmental monitoring; real-time data; data collection methods; building energy simulation

Introduction

The world’s energy use is growing at a rapid rate, along with serious concerns over resource depletion, water and air pollution, emissions due to global warming, endangered wild life, and climate change. Predictions show that this growing trend will continue (Pérez-Lombard, Ortiz, & Pout, 2008). Thus, reducing the consumption of energy in various sectors has become the primary agenda for sustainable development. Buildings account for 40% of the total primary energy consumption in the United States (US), with 21% consumed by the residential sector and 19% by the commercial sector (US Energy Information Administration [EIA], 2014). More importantly, 80% of the energy consumed by buildings during their life cycle occurs during the operation and maintenance phase (United Nations Environment Programme [UNEP], 2007).

The US Congress drafted an Energy Policy Act (2005) that highlights the importance of high-performance buildings as buildings designed for energy efficiency, and improved occupant
comfort, health, and productivity within the built environment (National Institute of Building Sciences [NIBS], 2008). Sustainable and high-performance building design involves the careful consideration of required building energy loads such as heating, lighting, and ventilation, which play a major role in the building energy consumption during the operation phase. Design tools such as eQuest, Energyplus, TRNSYS, and IES are the most widely used simulation tools in the industry for predicting a building’s energy profile and designing building systems. Most of the energy analysis programmes are often found to inaccurately predict building performance and energy consumption (Turner & Frankel, 2008). This is mainly due to the assumptions made about several of the input parameters (e.g. occupancy rate and after hour plug load use) during the design phase. Studies have also indicated that there are significant differences between the predicted (computed) energy performance of buildings and the actual measured energy use once buildings are operational (Bordass, Cohen, Standeven, & Leaman, 2001; Demanuele, Tweddell, & Davies, 2010; Menezes, Cripps, Bouchlaghem, & Buswell, 2012; de Wilde, 2014).

Another major challenge is that energy models are rarely re-used or revisited during operation, so no remedial action can be taken to check or modify the design assumptions. Inconsistencies between the predicted and actual performance models are mainly due to assumptions pertaining to external conditions (de Wilde, 2014), operational issues and occupancy behaviour (Azar & Menassa, 2014; Azar & Menassa, 2012; Demanuele et al., 2010; Yudelson, 2010), and oversized or underperforming mechanical/electrical systems (Bordass et al., 2001). However, if appropriate data inputs that represent actual conditions in the building are accessible, then energy models offer a significant opportunity to be used as decision support tools. They can provide simulation environments to help identify areas where inconsistencies exist, allow for retroactive analysis of design assumptions, and, more importantly, to develop and test solutions to reduce energy waste (Menassa et al., 2014; Torcellini et al., 2006).

For example, results of the Post-Occupancy Evaluation conducted by Menezes et al. (2012) show that with the help of monitored data (occupancy, lighting, and plug load as categorized in the taxonomy section of the paper) and energy modelling, the electricity consumption predictions were improved to within 3% of the actual values. Menassa et al. (2014) developed a conceptual framework that couples energy modelling with occupancy characteristics and energy use data. Information regarding the behaviour of the building occupants was provided to the model and changes in occupant behaviour were expressed in the design model through altering heating and cooling set points, lighting use schedules, and equipment use schedules, emulating the true behaviour of occupants and affecting the generated energy predictions.

Thus, in order to attain the collective objectives of a high-performance building, clear knowledge of the building’s actual performance is necessary. This can be achieved through the development of Dynamic Data-Driven Simulation and Analysis (DDDSA) methods, where continuous real-time data influence the simulation system and improve analysis and prediction of a system’s performance (Hu, 2011). Such an opportunity, in turn, emphasizes the need for credible data as the basic pillar of building energy research (Xia et al., 2014).

Although there are several data collection methods for monitoring building energy use, they are not widely adopted. Building managers still rely on their own experience to predict and manage the building energy consumption and fault detection (Lee, Painter, & Claridge, 2007). Traditionally, information related to building performance has been gathered through interviews, questionnaires, and surveys (Huizenga, Abbaszadeh, Zagreus, & Arens, 2006; Zagreus, Huizenga, Arens, & Lehrer, 2004). Gradually, advances in technology have led to the use of sensors for collecting data within buildings. Though the state-of-the-art technology deals efficiently with the flow of information, there are limitations such as power consumption of the sensor nodes, scalability issues, and spike errors.
This paper critically reviews the traditional and contemporary data collection methods currently available for monitoring building energy performance, and then discusses their relevance, advantages, and limitations. The overall goal of the study is to assess the type and quality of data obtained from these data collection methods to support DDDSA for real-time monitoring and control of buildings during the operation phase.

Research objectives and motivation

The primary objectives of this paper are to develop the taxonomy of various data types relevant in building energy management; critically review various data collection methods adopted in the industry, and illustrate how the data gathered can be relevant to buildings in different phases of their life cycle as follows:

- New buildings (about to be constructed) to improve the design decisions for more efficient performance of the building systems during the operation phase.
- Existing buildings under operation in making real-time decisions to optimize the building’s performance and building’s energy usage. In addition, for older buildings it helps in energy retrofit- and maintenance-related decisions.

The primary motivation for this study is that monitoring and measuring how, where, when, and why energy is being used plays a predominant role in energy-efficient building operation and maintenance. ‘You can’t manage what you can’t measure’ (Ambati, 2013). For example, according to the US Department of Energy (US DOE), some of the key elements in achieving energy-efficient buildings are: (a) measuring energy use in real time, (b) obtaining feedback on how buildings are actually consuming energy, and (c) revisiting/comparing the actual energy consumption with the targeted goals during post-occupancy (Chen, 2009). Several other studies emphasize that by monitoring energy consumption, studying actual occupancy patterns, and monitoring and actuating electrical loads, individual energy consumption parameters (such as heating, lighting, and cooling) can be optimized, and subsequently energy savings can also be obtained during the process (Mathews, Botha, Arndt, & Malan, 2001; Neto & Fiorelli, 2008; Von-Neida, Maniccia, & Tweed, 2001; Xia et al., 2014). Tables 1, 2 and 3 provide a summary of several similar studies that have demonstrated savings in building energy consumption resulting from the continuous monitoring and control of Heating, ventilation, and air conditioning (HVAC) system operations, electric energy use, and lighting systems.

Table 1. Energy savings from monitoring and controlling HVAC system operations in buildings.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Savings</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occupancy-based demand response HVAC control strategy</td>
<td>20%</td>
<td>Erickson and Cerpa, (2010)</td>
</tr>
<tr>
<td>Fully integrated building, HVAC, and control simulations</td>
<td>60%</td>
<td>Mathews et al. (2001)</td>
</tr>
<tr>
<td>HVAC control principles</td>
<td></td>
<td>Sezgen and Koomey (2000)</td>
</tr>
<tr>
<td>Interaction between lightning and space conditioning</td>
<td>Not quantified</td>
<td>Erickson et al. (2009)</td>
</tr>
<tr>
<td>Occupancy-based system</td>
<td>42%</td>
<td>Gorter (2012)</td>
</tr>
<tr>
<td>HVAC Equipment right-sizing</td>
<td>Not quantified</td>
<td>Freire et al. (2008)</td>
</tr>
<tr>
<td>Predictive controllers for energy savings</td>
<td>Not quantified</td>
<td>Spyropoulos and Balaras (2011)</td>
</tr>
<tr>
<td>Regulating the indoor set point temperature</td>
<td>56 kWh/m²</td>
<td></td>
</tr>
</tbody>
</table>
Several studies were conducted to prove that significant savings in lighting energy consumption can be generated by employing various strategies as shown in Table 2. Williams, Atkinson, Garbesi, Page, and Rubinstein (2012) provide a meta-analysis of lighting energy savings, and report 240 savings estimates from 88 papers and case studies. All of the strategies stated require data acquisition and analysis analogous to the DDDSA discussed in this paper.

### Table 2. Energy savings from monitoring and controlling lighting in buildings.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Percentage of savings</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daylighting</td>
<td>20–80</td>
<td>Kapsis, Tzempeliko, Athienitis, and Zmeureanu (2010) and Bodart and De Herde (2002)</td>
</tr>
<tr>
<td>Using occupancy sensors</td>
<td>30</td>
<td>Chung and Burnett (2001)</td>
</tr>
</tbody>
</table>

### Table 3. Energy savings from monitoring and controlling plug loads and per appliance power consumption in real time.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Savings</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor-based intelligent system</td>
<td>Not quantified</td>
<td>Anastasi, Corucci, and Marcelloni (2011)</td>
</tr>
<tr>
<td>Serious game interventions</td>
<td>7–23%</td>
<td>Orland et al. (2014)</td>
</tr>
<tr>
<td>RECognition of electrical Appliances and Profiling in real time (RECAP)</td>
<td>Not quantified</td>
<td>Ruzzelli et al. (2010)</td>
</tr>
<tr>
<td>Plug load energy management solution</td>
<td>43–67%</td>
<td>Ghosh et al. (2013)</td>
</tr>
<tr>
<td>ViridiScope</td>
<td>Not quantified</td>
<td>Kim, Schmid, Charbiwala, and Srivastava (2009)</td>
</tr>
<tr>
<td>Energy-efficient PPL equipment and design strategies</td>
<td>47%</td>
<td>Lobato, Pless, Sheppy, and Torcellini (2011)</td>
</tr>
<tr>
<td>Appliance activity monitoring using wireless sensors</td>
<td>Not quantified</td>
<td>Schoofs et al. (2010)</td>
</tr>
</tbody>
</table>

Successful building projects always begin with identifying and considering all the goals and objectives that are primarily of interest to the building owners, designers, and occupants. Increasingly, energy-efficient building operation; minimal environmental impact; life cycle analysis; occupant health, comfort, and productivity; building functionality; adaptability; durability; and sustainability are becoming important priorities for stakeholders (O’Sullivan, Keane, Kelliher, & Hitchcock, 2004).

For efficient management and better performance of buildings, building managers require large amounts of data (Wang & Xie, 2002). Buildings comprise copious data (Jiao et al., 2013; Vanlande, Nicolle, & Cruz, 2008) such as occupant comfort parameters, air flow, hot water supply, plug loads, outdoor temperature, and indoor air quality (IAQ) data. Examples of how these data can be used include the fault detection of air handling units (AHUs) which requires
information regarding several parameters such as supply air temperature, return-air temperature, return-air relative humidity, and outdoor air relative humidity to identify whether AHUs are working in the most efficient manner (Schein, Bushby, Castro, & House, 2006).

During the design phase of a building, the accurate prediction of parameters such as equipment loads, size, specifications, and occupant comfort plays a crucial role in designing the various building systems that in turn determine to a large extent the energy performance of the building during the operation phase (Crawley et al., 2001). The accuracy of the assumptions made during the design phase is directly related to the amount and extent of data available for analysis. However, these data (e.g. accurate weather conditions and occupancy) are generally not available, which leads to a decoupled approach to design, simulation, modelling, and subsequently energy management.

This further emphasizes the critical need for DDDSA where the design phase models can be updated and seeded with real-time data during the operation phase to enhance the understanding of the impact of the design assumptions, and more importantly to be able to use these models to predict energy performance due to technical and behavioural interventions (Menassa et al., 2014). For this reason, the taxonomy of data types presented in the paper follows a similar approach in classifying the data types that directly or indirectly influence the building energy use. Parameters that affect building energy consumption or building performance are broadly classified into static and dynamic parameters. Further classifications of these parameters are shown in Figures 1 and 2, and are discussed next.

Figure 1. Taxonomy of data types (static parameters) in a building.
Static parameters

As the name suggests, static parameters are those that typically do not change with time and are fixed for a given building throughout its life cycle once design and construction are completed unless the entire building is renovated or reconstructed.

Primary building information

Some of the key factors governing the energy consumption of a building during the operational phase are building location, source of energy, function, and use of the building (United Nations Environment Programme [UNEP], 2009). Function and use of a building are fixed for an extended period of time and hence considered static parameters. Since occupants tend to maintain comfort levels in extreme weather conditions, huge variations in building energy sources and
consumption can be observed across different weather zones. Similarly, heating and cooling requirements can significantly vary between developed and developing nations, commercial and residential buildings, single-storey and multi-storey facilities, and urban and rural areas. For example, studies have demonstrated that the cooling and electricity loads of urban buildings can be two to three times higher than those in rural areas when considering the impact of the urban climate (Santamouris et al., 2001).

**Building orientation**

The orientation of the building simply means what the compass coordinates of the building are. It is one of the key elements concerning passive visual and thermal comfort in buildings. Successful orientation can help improve energy efficiency by maximizing free solar energy and wind energy. Building orientation also plays a crucial role in the design of renewable energy systems especially for high-performance buildings that rely on renewable energy to balance their energy demand (Shi & Chew, 2012).

Geographic Information Systems (GIS) data have the geographic positioning information incorporated in them, such as a road map. Analogously, in the context of buildings, the neighbouring buildings, abutting roads, street configuration, wall features, and roof features can be included. The energy consumption of the building is significantly affected by the direct and diffused sunlight, and also reflected light from neighbouring facades and the ground (Kesten, Tereci, Strzalka, & Eicker, 2012). The main factors that affect the daylight on buildings are the distance between buildings, the height of the adjacent buildings, the orientation of and the reflectance from the adjacent buildings, the size of openings, and the size of the shading devices (Santamouris & Asimakopoulos, 2001). Hence, the orientation of the building plays a very important role in passive design and thereby has a notable impact on the energy efficiency throughout the building’s life cycle.

**Building materials**

Typically, materials used for the structure of buildings contribute to more than 50% of the total embodied energy in a building (Asif, Muneer, & Kelley, 2007). In addition, the choice of the material significantly affects the heating and cooling demands (Zabalza Bribián, Valero Capilla, & Aranda Usón, 2011). Significant research has been done in developing advanced and sustainable materials for buildings (e.g. wool bricks, solar tiles, and sustainable concrete). Meticulous selection and installation procedures of building envelope components – such as different types of walls, insulation materials, thermal mass, and phase change materials – generally lead to more energy-efficient building performance (Sadineni, Madala, & Boehm, 2011).

**Building zones**

Building zones are defined as those areas in an indoor built environment which are periodically monitored and subjected to variations in thermal loads. For example, temperature near windows varies continuously because of the variations of heat from the sun. In addition, areas such as basement stores and open penthouses generally have low/high temperatures, respectively, because of varying sunlight conditions in those areas, and are typically designed to receive ventilation from different air-handling units (McDowall, 2006; US Department of Energy [DOE], 2014). For an energy-efficient operation of building zones and respective HVAC components, determining the optimal operation variables such as building zone temperature, minimum zone air flow rate, and optimal supply air temperature of the AHU is important. Hence, the primary objective
of a building manager is to optimize and/or minimize the energy consumption by the HVAC system, while at the same time maintaining thermal comfort conditions within the various building zones.

**Dynamic parameters**

Dynamic parameters are those that vary with time, location, and various other factors. Several parameters assumed during the design phase of the building can vary during the operation phase resulting in significant deviations from initial energy estimates. Having the capability to measure these parameters in real time and continuously update the design phase energy model can support DDDSA for improved decision-making during the building operation phase (Menassa et al., 2014). The most significant of these parameters are discussed in the subsections below.

**Weather data**

Weather data refers to changes in temperature, incoming solar radiation, precipitation, and humidity. According to the National Oceanic and Atmospheric Administration (NOAA), there are 344 climatic divisions in the CONtiguous US (CONUS). Monitoring region and nationwide variations in temperature, precipitation, heating/cooling degree days was the primary objective of CONUS (National Climatic Data Center [NCDC], 2014). The ability to predict weather and the ability to respond appropriately to required changes in building operation as a result of weather related data would significantly improve the energy efficiency of an operating building. For example, Oldewurtel et al. (2010) illustrated how Model Predictive Control and weather predictions can increase energy efficiency without compromising occupant comfort.

**Occupancy data**

The actions, attitudes, and needs of building occupants have a significant impact on the total building energy consumption. In particular, having the ability to collect data about occupant schedules, perceived comfort levels, and occupant characteristics can significantly improve the energy efficiency of the operating buildings. Each of these categories is discussed in detail in the subsections below.

**Schedule information.** The state-of-the-art building technologies use occupancy information mostly in the form of schedules for building energy control during the operation phase (where maximum occupancy is typically assumed for a given zone), and also for assumptions in the energy modelling software during the design phase (Erickson et al., 2009; Oldewurtel, Sturzenegger, & Morari, 2013). This means that in most cases there are rooms that are heated/cooled needlessly, resulting in wasted energy. Having a better understanding of how the building is used and the related user activities can lead to better control and possibly improved user comfort while reducing energy consumption.

For example, studies have shown significant energy-saving potential by engaging in strategies such as night-time setback, which means either relaxation of standard indoor occupant comfort parameters or even shutting down the building systems during the night (Bloomfield & Fisk, 1977; Erickson, Carreira-Perpiñán, & Cerpa, 2011; Murphy & Maldeis, 2009). This also leads to an obvious conclusion that information regarding long-term vacancies such as holidays and vacations have significant energy-saving potential. The offline sensor data that are gathered can offer better assistance regarding the actual hours of occupancy of a building, and can
develop better predictive control routines. Based on occupancy data, a savings in space and water heating of approximately 14% was reported by Boait and Rylatt (2010). Karjalainen and Lappalainen (2011) suggest that occupancy detection plays a crucial role in improving energy efficiency, as occupants’ presence and level of activity need to be known for better building energy control.

**Occupant comfort information.** Comfort, in the simplest of terms, can be defined as the absence of discomfort. People do not feel comfortable if the indoor temperature is too hot or too cold; there is lack of fresh air; it is too humid; too many noisy disturbances; and also when other factors related to human comfort stand out in a negative way (Bradshaw, 2010, p. 154). Enhancing occupant comfort, optimizing energy use, and reducing operation costs have always been primary objectives of building operations. At the same time, studies proved that energy consumption optimization is possible without compromising the indoor occupant comfort (Freire, Oliveira, & Mendes, 2008). However, the state-of-the-art building controls are also not able to meet the standards of the occupant level comfort and energy efficiency goals simultaneously (Singhvi, Krause, Guestrin, Garrett, & Matthews, 2005). In addition, optimal environmental conditions required for comfort and efficiency of the occupants are very narrow (Bradshaw, 2010). The factors that affect comfort are categorized and the respective comfort measures are noted in Table 4.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Units</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interior/indoor factors</td>
<td>Thermal comfort</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input air temp (AHU)</td>
<td>°C</td>
<td>Mathews et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Return-air temp (From room/hall)</td>
<td>°C</td>
<td>Mathews et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>CO₂</td>
<td>ppm</td>
<td>Mathews et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Ventilation</td>
<td>cfm/person</td>
<td>Azar and Menassa (2012)</td>
</tr>
<tr>
<td></td>
<td>Indoor temperature</td>
<td>°C</td>
<td>Bradshaw (2010) and Olofsson, Andersson, and Sjögren (2009)</td>
</tr>
<tr>
<td></td>
<td>Ventilation type and air flow</td>
<td>–</td>
<td>Bradshaw (2010)</td>
</tr>
<tr>
<td></td>
<td>Heat loss coefficient</td>
<td>–</td>
<td>Olofsson et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Floor area</td>
<td>m²</td>
<td>Bradshaw (2010)</td>
</tr>
<tr>
<td></td>
<td>Air velocity (Indoor)</td>
<td>m/s</td>
<td>Bradshaw (2010)</td>
</tr>
<tr>
<td></td>
<td>Humidity (Indoor)</td>
<td>%</td>
<td>Bradshaw (2010)</td>
</tr>
<tr>
<td></td>
<td>Mean radiant temperature (MRT)</td>
<td>°C</td>
<td>Azar and Menassa (2012)</td>
</tr>
<tr>
<td></td>
<td>Visual comfort</td>
<td>Lighting</td>
<td>Lux</td>
</tr>
<tr>
<td></td>
<td>Acoustic comfort</td>
<td>Noise/sound levels</td>
<td>dB</td>
</tr>
<tr>
<td></td>
<td>External comfort</td>
<td>Wind Speed (Outdoor)</td>
<td>m/s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wind direction (Outdoor)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outdoor temperature</td>
<td>°C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Humidity (outdoor)</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Solar radiation</td>
<td>W/m²</td>
</tr>
</tbody>
</table>
During the maintenance and operations phase of a building, facility managers strategically determine standard set points for most of the aforementioned comfort parameters because of the unavailability of real-time occupancy information. The utilization of predetermined set points without taking into effect the real-time variations of occupancy levels can cause building occupants’ dissatisfaction and can also lead to low efficiency of the building systems. Several studies have shown the dissatisfaction of the occupants with pre-defined set point and scheduled operation of the HVAC building systems (Guo & Zhou, 2009). Furthermore, thermal and visual comforts also have a significant impact on occupants’ productivity (Codreanu, 2013; Huizenga et al., 2006; Tsuzuki, Arens, Bauman, & Wyon, 1999; Wyon, 2004).

**Occupancy characteristics.** Several studies were conducted to identify the impact of occupant behaviour on energy consumption in buildings. For example, actions performed by occupants and building managers are some of the major causes of excessive energy use in buildings (Azar & Menassa, 2012). Furthermore, Augenbroe, Castro, and Ramkrishnan (2009) observed that the differences between desired and actual energy levels even when ‘technological’ strategies are implemented in the building are due to the lack of understanding and account of human actions. Petersen, Shunturov, Janda, Platt, and Weinberger (2007) studied the effect of feedback, education, and incentives on the occupants’ behaviour. Results showed an alarming reduction of 32% in electricity use and a 3% reduction in the total water use. Yu, Fung, Haghighat, Yoshino, and Morofsky (2011) developed a data mining technique to evaluate the building energy-saving potential by improving the behaviour of building occupants. Thus, data pertaining to occupant characteristics such as the motivation, involvement, and interest of the occupants along with their connection to their peers in the building can be crucial for energy conservation in buildings.

**Energy use**

The primary energy source for water heating, space heating, cooling, ventilation, lighting and plug loads, and the equipment can be either natural gas or electricity depending on the standard practices followed in that region. The subdivision of categories is done by the general industry standards rather than by the primary source of the energy use.

**Hot water.** A significant portion of the energy consumed in residential buildings is for heating the ambient cold water supply. According to US EIA (US Energy Information Administration [EIA], 2013), in 2009, 37% of the natural gas consumed by US homes was for water heating, cooking, and miscellaneous uses. In addition, Cheng (2002) stated that water use consumes significant energy and predominant savings can be generated by creating awareness among the public. To that end, researchers studied the effect of several energy-efficient techniques, such as lowering the hot water temperature, or using ambient heat for preheating the supply water to reduce demand for hot water in buildings and to conserve energy (Park & Jacobi, 2009). Hot water data referred to as ‘service hot water’ in (RETScreen, 2014) or ‘hot water use’ in (eQUEST, 2014) are the data pertaining to the total volume (gallons) or volume per occupant (gallons/occupant/day) of hot water required at the desired temperature for the entire building or facility, and are necessary for both designing the hot water systems as well as its efficient operation.

**Heating, ventilation, and air conditioning.** Energy consumption by HVAC systems in buildings is generally calculated in kWh/unit area but it can also be expressed as kWh/occupant. Some of the primary data types that affect the HVAC energy usage are indoor temperature, temperature set points, number of occupants using the indoor space, and indoor humidity levels. For decades, more than half of the total residential building energy consumption was due to heating and
cooling loads (US Energy Information Administration [EIA], 2013). Most of the current energy-efficient solutions for HVAC systems still function based on occupancy schedule information, typically assuming maximum occupancy conditions rather than actual occupancy. Research has shown that energy savings are achievable by controlling HVAC systems based on the actual number of occupants using the building, the movement of occupants inside the building, occupancy schedules, and predictions regarding the external weather conditions such as temperature and humidity (Erickson et al., 2009). Hence information regarding dynamic occupancy, real-time indoor and outdoor temperature and humidity levels is the key for optimizing and controlling the HVAC energy usage in buildings. In addition, significant savings ranging from 14% to 60% of the total HVAC energy usage were obtained by having an optimal control strategy with the help of an effective data collection and subsequent analysis (Erickson et al., 2009; Freire et al., 2008; Jazizadeh, Kavulya, Klein, & Becerik-Gerber, 2011; Mathews et al., 2001; Neto & Fiorelli, 2008; Sezgen & Koomey, 2000; VonNeida et al., 2001).

Lighting. The total electrical energy consumed by the entire lighting system in a building is expressed in kWh/unit area of the building or kWh/occupant. Lighting is responsible for approximately one-third of electricity use in commercial buildings, and more than one-half in lodging and retail (US Energy Information Administration [EIA], 2003). Lighting systems have significant potential to conserve and reduce energy use in buildings (Desroches & Garbesi, 2011). Information regarding occupancy, availability, quality of daylighting (natural light levels), and individual preferred light levels play a crucial role in saving the lighting energy consumption in buildings (Williams et al., 2012). Though efficient technologies such as daylight sensors and/or motion sensors, photo sensors, dimmers, and wireless on-off switches exist, they are not yet widely adopted.

Plug load. In the USA, 33% of electricity use is consumed by plug and process loads – more than lighting, space heating and cooling, or ventilation. By 2030, the energy consumption of plug loads is predicted to increase to 49% of total electricity use. This is not only because of the increasing efficiency of building envelopes, but also because of the drastic increase in office automation equipment such as desktop computers, tablets, and mobile phones (National Renewable Energy Laboratory [NREL], 2013). For example, the energy wastes due to electrical appliances in standby mode contribute to about 10% of the total building energy consumption (International Energy Agency [IEA], 2003). All the minor pieces of equipment such as refrigerators, desktops, laptops, and microwaves are considered appliances and related data are considered in the plug load data. Hence, gathering and analysing data regarding the state of the appliance (whether it is on, off, or in standby) and usage (whether it is being used or not) plays a crucial role in achieving savings in the plug load consumption.

Equipment data. A building is defined jointly by its components, equipment, and systems (Grondzik, Kwok, Stein, & Reynolds, 2011). In the context of buildings, major equipment comprises chillers, diesel generators, and AHUs. Successful equipment commissioning does not guarantee the energy-efficient potential and adequate occupant comfort (Ma, Cooper, Daly, & Ledo, 2012; Pisello, Bobker, & Cotana, 2012). Conditional monitoring techniques have been developed to monitor the operational characteristics and thereby the efficiency of the machines. These monitoring techniques require data regarding the aging of the insulation, electrical faults (such as insulation problems and burning of transition resistors), and mechanical faults (such as springs, bearings, shafts, and drive mechanisms) (Han & Song, 2003). In totality, equipment data have a great potential to influence the energy consumption in buildings.
Water use

Water consumption pertaining to domestic and public use—such as drinking water, sanitation, landscaping, cooling, heating, laundry, restroom, kitchen, and others— is generally measured in gallons per day (gal/d) or Million gallons per day (Mgal/d) or litres/day. Residential and commercial buildings in the US account for 12% of fresh water consumption (Barber, 2009). Water has significantly high embedded energy and involves very high economic costs in supplying to the utilities (California Energy Commission [CEC], 2006; Cheng, 2002). Studies show that raising awareness amongst the occupants for conserving water will result in significant energy savings. For example, Petersen et al. (2007) showed that by installing an automated data monitoring system that provides real-time information and feedback to occupants on energy and water use, significant reduction in total water use can be obtained. To achieve savings, information regarding the real-time, location-specific usage of water in the buildings needs to be collected.

Data collection methods

In order to better understand building energy performance and to improve the efficiency of buildings to reduce waste and optimize resources, the data discussed above need to be collected and analysed. Information regarding how the aforementioned data types are gathered using different methods is presented in this section of the paper. Although there are no established standards that specify the frequency with which the data needs to be collected, several studies propose varied time intervals ranging from 5 to 15 minutes for real-time monitoring (Xia et al., 2014), to over 60 minutes for weather data analysis (Santamouris et al., 2001). Table 5 provides a summary linking the five primary data collection methods discussed in this paper to the specific data types introduced earlier.

Table 5. Data collection methods mapped to specific types of data.

<table>
<thead>
<tr>
<th>Data types</th>
<th>Questionnaires and surveys</th>
<th>Manual data collection</th>
<th>Wired sensor networks</th>
<th>Wireless sensor networks</th>
<th>Other methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Static parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary building information</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Building orientation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Building materials</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Building zones</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Dynamic parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weather data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Occupancy data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Schedule information</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Occupant comfort information</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Occupancy characteristics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Energy use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water heating</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HVAC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lighting</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Plug load</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Equipment data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Water use</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Figure 3 conceptually illustrates how the data collection methods progressed with advancements in technology, and the corresponding improvement in data quality. The quality of data used in the analysis is very important for DDDSA, because, better the data quality, more credible the results are. Technology advancements have made it easier to deploy a large number of data sources and have also improved the frequency of data collection, which has resulted in larger data sets.

In the early stages, data collection only involved humans collecting necessary information manually. With technological advancement, sensors are now used for monitoring and data collection and this has improved the data size and subsequently the data quality. The current state-of-the-art technology is moving towards wearable technology and miniature devices embedded with sensors. Despite the opportunities for analytics afforded by these large data sets, questions still remain regarding how these data can be used to make effective decisions that reduce the building’s overall energy and carbon footprint.

**Questionnaires and surveys**

Questionnaires have a wide variety of data collection applications ranging from determining a small opinion, obtaining subjective information of public perception, and analysing quantitative opinions, to developing an advanced user interface tool. Survey is a specific type of questionnaire where only some opinions or beliefs are analysed which make them suitable for gathering information about occupants’ perception of comfort in a building (Stawarski & Phillips, 2008). For example, Ng and Akasah (2013) used a building performance survey framework, Energy-efficient
Building Environmental Quality Questionnaire (EBEQ2), to identify the problems affecting occupants' comfort and buildings' Indoor Environmental Quality. Wilkinson, Reed, and Jailani (2011) examined the satisfaction levels and expectations of sustainable building users about their workplace through a questionnaire survey. Contrary to the traditional paper based surveys, researches also proposed smart phone application based on real-time data collection of occupant comfort in indoor environments (Jazizadeh et al., 2011; Weiss, Helfenstein, Mattern, & Staake, 2012). In addition, primary building information (such as function, use, geographical location of the building) and the type of building materials used can be gathered using online-based surveys or questionnaires to the original building designers, contractors, or current facility managers. Though, technology progression has improved the process of data collection, data related to occupant comfort and occupant characteristics can only be collected using questionnaires and surveys. However, most of the times, too few responses during a survey make the survey statistically insignificant for further analysis. In addition, survey responders may lack a unified understanding of the survey questions, leading to huge variability in survey responses (Greening, Greene, & Difiglio, 2000).

Manual data collection

Many building operation tasks require a lot of data for analysing the efficiency and performance of building systems. Traditionally, these data were manually collected and documented with the help of multiple observers monitoring the spaces with data logs or data sheets noting down all of the information. This involves a lot of paperwork. For example, data regarding primary building information such as building type, location of the building, function and use of the building used to be manually collected and stored in the paper records. In addition, Ploeger et al. (2015) collected primary building information such as building location, function, use, as well as information regarding building materials manually using human inspectors and a smart phone/tablet/desktop-based application. Though it is tedious, time consuming, and not entirely accurate, data regarding energy use parameters such as hot water, plug loads, and HVAC can be indirectly or directly calculated with the help of installed meters that measure the respective usage quantities. Furthermore, the aforementioned parameters can also be obtained by going through historical data related to building energy use and subsequently document the corresponding data. Even

Table 6. Limitations of manual data collection.

<table>
<thead>
<tr>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hence, manually monitoring different parameters in a building can become</td>
<td></td>
</tr>
<tr>
<td>highly expensive</td>
<td></td>
</tr>
<tr>
<td>It is also highly time consuming and error prone especially in larger</td>
<td></td>
</tr>
<tr>
<td>buildings because a human observer has to gather information in each and</td>
<td></td>
</tr>
<tr>
<td>every room</td>
<td></td>
</tr>
<tr>
<td>The manual data collection process has limitations on the frequency (larger</td>
<td></td>
</tr>
<tr>
<td>the building and lower the monitoring frequency) at which the data can be</td>
<td></td>
</tr>
<tr>
<td>collected. Therefore, it becomes difficult to obtain real-time information</td>
<td></td>
</tr>
<tr>
<td>which is a key for DDDSA.</td>
<td></td>
</tr>
</tbody>
</table>
this process is tedious, time consuming, and can be error prone. The limitations of this particular type of data collection approach are listed in Table 6.

**Wired sensor networks**

Sensors are devices that measure or respond to a physical phenomenon and generate a corresponding output. For example, a thermometer is a sensor which responds to the ambient temperature (physical phenomenon) and produces the corresponding temperature value (output). In the context of wired sensors in buildings, an actuator is responsible for controlling any of the building systems or mechanisms. A wired sensor network consists of several interconnected sensors and actuators which are ultimately wired to the central node (Bharathidasan & Ponduru, 2002).

Wired networks are highly secured and are really fast in transmitting large amounts of data. Though wireless sensor networks are growing in demand and importance, applications such as video surveillance (security and privacy) in commercial or residential buildings need a wired set-up because of higher bandwidth and power needs (Chandramohan & Christensen, 2002). Gathering information pertaining to the occupancy data as discussed in the dynamic parameters section of the paper is one of the major applications of wired data collection in buildings. In addition, Park and Hong (2009) proposed an experimental model based on the Building Automation and Control network, which is a type of wired controlled system that can achieve a 40% reduction in average power consumption in lighting loads (lighting energy use as categorized in the Dynamic Parameters section of the paper).

Wired networks are more feasible and cheaper for smaller buildings than for larger buildings. This is because the cost of wiring greatly increases with the increase in area that needs to be monitored (Agarwal et al., 2010). Most of the current state-of-the-art actuators are generally wired to the physical system or process that is being controlled. For example, the actuator controlling the position of a HVAC damper is generally integrated into the damper linkage. This implies that wiring needs to be installed between every actuator and controller. This process is highly error prone, tedious, and expensive in existing buildings (Brambley et al., 2005).

In the context of buildings, the wires used for data transmission and power supply must be designed carefully for avoiding disturbances to the ambient indoor environment. In addition, a wired approach needs significant effort for deployment and maintenance which comes at a cost and is a perpetual issue for building managers. Finally, threats from rodents make it very difficult to protect the wiring in buildings (Wang, Liu, & Sun, 2010). Listed in Table 7 are some of characteristics of monitoring and collecting data from wired sensor networks in buildings.

**Wireless sensor networks**

With increasing technology and advancements in the field of micro-electro mechanical systems and wireless communications, much advancement in wireless sensor networks are being observed. Wireless sensor networks comprise low-cost, low-power, multifunctional sensor nodes/motes that can communicate wirelessly across short distances (Akyildiz, Su, Sankarasubramaniam, & Cayirci, 2002). They also constitute wired actuators at requisite locations as discussed in the wired sensor networks section of the paper.

With the help of on-board computing capabilities, the deployed sensor nodes in a wireless sensor network are capable of gathering, analysing, processing, and wirelessly transmitting the necessary information (such as temperature, humidity, and light intensity) from the environment (indoor or outdoor) to the targeted location (Yick, Mukherjee, & Ghosal, 2008). The size and cost of the sensors vary drastically depending on the type of application. Data collected at the sensor nodes are transmitted to the targeted location using the gateway...
application (Nourollah, 2009). Since the need of wires is eliminated, wireless sensors have a lot of potential to drastically reduce the cost of conveying information or collecting data in comparison to the wired counterparts. There are a wide range of applications ranging from health monitoring, home awareness, area monitoring, military applications, home entertainment and control, security, indoor environment monitoring, building energy monitoring, and industrial monitoring.

Wireless solutions also have lot of advantages over wired networks in cases of building retrofits because the additional wiring costs much more than in new installation and also it is highly inconvenient for the building occupants. Building energy monitoring applications where wireless systems have significant value are relocating and adding additional temperature sensors, air quality sensors (such as CO₂, oxygen, and carbon monoxide), humidity sensors, and occupancy sensors (for lighting and HVAC control) (Vlissidis, Charakopoulos, Kolokotsa, & Boian, 2008). Studies also show that the number (using two temperature sensors instead of one) and placement (placing sensor at non-standard height) of wireless sensors significantly impact the energy-saving potential in buildings (Wang, Arens, Webster, & Shi, 2002).

Wireless systems have been widely used for monitoring and/or controlling parameters such as weather, occupancy, occupancy comfort, and water use (Vlissidis et al., 2008), which have a direct or indirect effect on the building energy use (by HVAC, lighting, and plug loads) (Ghosh, Patil, & Vuppala, 2013; Lanzisera et al., 2013; O’Connell et al., 2011; Vlissidis et al., 2008). For example, a rollup door switch tied into the room HVAC shuts off the HVAC if the door remains open for at least five minutes. In addition, water-saving fixtures such as motion sensor lavatory faucets conserve water and lower domestic hot water heating energy (Vlissidis et al., 2008). Wireless nodes can remotely communicate and actuate the building systems (such as heating, lighting, and ventilation) in accordance with occupancy presence. Replicating these solutions with wired networks is not possible and also not practical in almost all the cases.
Most of the systems described above are a star-based sensor network with only one route of communication. One of the main disadvantages of such a network is that each sensor node in the network does not support communication on behalf of any other sensor node (Vlissidis et al., 2008). Consequently, if a node stops working because of a physical damage or internal failure, all the information of that particular sensor node will be lost. The loss of information can have serious consequences while performing a whole building energy audit.

Overcoming the disadvantages of the aforementioned star-based network, Dr Sollacher has developed self-organizing sensor networks (Ruth, 2011). These networks contain intelligent sensors which can organize themselves in a network, pinpoint their own location in the network, autonomously assign radio channels for communication, synchronize their internal clocks, and interpret data. In addition, sensors take reading at designated time intervals, communicate with their neighbours, and go back to sleep. This helps in the energy-efficient operation of the sensor networks. The interesting part is the recorded average (by a sensor) also includes the value measured by the nodes unknown to them (Ruth, 2011). For example, the sensors can determine mean temperatures by comparing their measurements with those of their neighbours and use this value to estimate the average of the overall system inside a built environment. This not only ensures a better data storage but also offers a better communication.

The advent of wireless sensors allowed researchers to develop and experiment with various systems. For example, Ruzzelli, Nicolas, Schoofs, and O’Hare (2010) developed a system for profiling and identifying individual sources of electricity consumption in buildings. Erickson et al. (2009) used wireless sensor data and agent-based model to detect occupancy pattern in buildings which is a key aspect for energy consumption in buildings. Schoofs, Ruzzelli, and O’Hare (2010) used machine learning algorithms to determine appliance load monitoring system to identify the per appliance energy consumption. Currently, a lot of research is focused on building energy monitoring, control, and simulation which is briefly discussed in the further sections of the paper. With the rapidly growing energy demand and increasing technology, there is a lot of scope and need for innovation in the field of building energy optimization and control.

In summary, the general characteristics of collecting and monitoring wireless sensor data regarding various parameters in a building which directly or indirectly influence building energy consumption are briefly stated in Table 8.

**Other methods**

There exist some other methods of data collection in buildings that cannot be solely classified into the above categories. An example includes implicit sensing methods for detecting occupancy using existing network infrastructure. Implicit occupancy sensing uses existing IT infrastructure to replace and/or supplement traditional dedicated sensors to determine building occupancy. They are based largely on monitoring MAC and IP addresses in routers and wireless access points (Melfi, Rosenblum, Nordman, & Christensen, 2011). As already discussed, occupancy data can be used to control lighting, HVAC, and other building functions to reduce energy use in buildings. Another example can be Nonintrusive Appliance Load Monitoring to determine the energy consumption of different appliances based on the detailed analysis of the total loads’ current and voltage. This approach simplifies the data collection of energy consumption by utilities and also other appliances. This approach is called ‘nonintrusive’ because it does not require retrofitting the internal hardware of the appliance with sensors to gather the appliance load data or plug load data (Hart, 1992). However, there are still a lot of challenges that need to be addressed in terms of monitoring real-time plug load/appliance state (Zoha, Ghuak, Nati, & Imran, 2013).

With ubiquitous computing, wearable technology mainly focuses on bringing the technology into the day-to-day lives. Wearable devices comprise highly efficient and advanced electronic
components and computer capabilities. Google glass is one of the very good examples of wearable devices. Researchers at the University of California, Berkeley are developing an instrument called Rapid Building Energy Modeler (RAPMOD) (carried by a human), which maps the three-dimensional geometry to the respective heat map (“EETD’s Rapid Building Energy Modeler”, 2014). It is equipped with a Light Detection and Ranging (LiDAR) scanner, an inertial

Table 8. Characteristics of wireless sensor networks.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wireless sensor networks are untethered networks and hence the deployment is very easy. This can be very useful in monitoring vast areas such large public buildings, giant office spaces, and retail stores by eliminating complex wiring</td>
<td>Ceriotti et al. (2009), Dermibas (2005) and Kim &amp; Lynch (2011)</td>
</tr>
<tr>
<td>Advancements in the technology lead to very low-cost sensor nodes. Saves lot of initial capital investment required for building automation systems (BAS)</td>
<td>Dermibas (2005), Menzel, Pesch, O’Flynn, Keane, and O’Mathuna (2008) and Swartz et al. (2012)</td>
</tr>
<tr>
<td>Wireless sensor nodes monitor, transmit, and process data locally and very precisely.</td>
<td>Kim and Lynch (2011) and Menzel et al. (2008)</td>
</tr>
<tr>
<td>Functionality is relatively better in comparison to the traditional wired networks.</td>
<td>Swartz et al. (2012)</td>
</tr>
<tr>
<td>In wireless sensor networks, data collected by the individual sensor nodes are transmitted to the base station through relay nodes. For monitoring larger building networks (such as office, retail, educational, and public), many relay nodes are required which might potentially increase the cost of energy monitoring.</td>
<td>Bhadauria, Tekdas, and Isler (2011), Jea, Somasundara, &amp; Srivastava (2005) and Tekdas, Isler, Lim, and Terzis (2009)</td>
</tr>
<tr>
<td>One of the major problems in the low-cost wireless sensor networks is power consumption, since the sensor nodes are battery powered. Hence there might be issues in achieving the desired longevity of the network, and also considerable operation costs such as replacing batteries, and the involvement of manual labour increases the complexity of building manager in the process.</td>
<td>Al-Karaki and Kamal (2004), Anastasi, Conti, Di Francesco, and Passarella (2009), Bhadauria et al. (2011), Chakrabarti, Sabharwal, and Aazhang (2003), Diehl, Curtis, Rodriguez, Tosun, and Zhu (2007), Egea López (2006), Estrin (2001), Goerner, Chakraborty, and Sycara (2013), Juang et al. (2002), Min et al. (2000), Shah, Roy, Jain, and Brunette (2003), Singh, Singh, and Singh (2010) and Tekdas et al. (2009)</td>
</tr>
<tr>
<td>Larger wireless sensor networks generally suffer from scalability issues due to the involvement of a large number of sensor nodes. Though sensor nodes have on-board computing capabilities, they have limited processing and information storage capacity.</td>
<td>Alazzawi, Elkateeb, and Ramesh (2008), Al-Karaki and Kamal (2004), Bhadauria et al. (2011), Cerpa et al. (2001) and Singh et al. (2010)</td>
</tr>
<tr>
<td>A denser deployment of multiple sensor nodes typically have correlation or redundancy due to their proximity. This might be a major concern in case of the built environment.</td>
<td>Akkaya and Younis (2005), Al-Karaki and Kamal (2004); Swartz et al. (2012) and Singh et al. (2010)</td>
</tr>
<tr>
<td>The low costs of wireless sensor nodes render them more susceptible to faults and failures. Precautionary measures need to be taken to make it reliable over longer periods of service. In the perspective of building manager, considerable amount of money and man-hours needs to be spent for this.</td>
<td>Lo et al. (2013)</td>
</tr>
</tbody>
</table>

Downloaded by [University of Michigan] at 19:24 21 February 2016
measurement unit, and other sensors for generating a photorealistic 3D model of the building interior. The model thus generated can be inputted into an energy simulation model to get a better understanding in terms of building energy performance. Since the data gathered are automatically uploaded online for processing, RAPMOD can be operated by technicians without the help of energy experts. Hence it is relatively cheaper to produce an energy model of the building (Abaffy, 2014). On the other hand, Mantha, Feng, Menassa, and Kamat (2015) developed and tested a tele-operated robotic platform that gathers energy related data by navigating in an indoor environment. The results of the study indicate that there is a great potential for this application to make the data collection process more economical and efficient. This in turn will provide decision makers with quality data for more effective control and improved performance of buildings.

The aforementioned approaches and several other methods exist, which cannot be solely classified into any of the data collection methods discussed in this paper. Hence all the methods that use techniques to indirectly measure or gather data regarding specific building functions are categorized into the other methods.

**Post-processing of building energy data**

Once information regarding some or all of the data types discussed in the taxonomy are collected using various data collection methods, the question that follows is how the data can be utilized. The data gathered henceforth can be utilized in two ways; namely, online simulation and offline simulation. Data for online simulation are being generated, transformed, managed, and analysed in real time; while for offline simulation data are being generated, transformed, and stored for future analysis (Kamat, Menassa, & Lee, 2013).

Offline simulation is typically used for three main purposes such as energy profiling (weekly and daily energy use patterns), energy benchmarking (comparison of energy uses among selected buildings to understand differences in buildings’ performance and establish good operation practices), and energy diagnostics (to analyse the building systems, identify operating issues, and recommend necessary retrofit measures) (Xia et al., 2014). For example, Zhao, Lasternas, Lam, Yun, and Loftness (2014) studied the office appliance power consumption data with the help of offline simulation to learn the occupants’ ‘passive’ behaviour, which has a significant impact on building energy consumption. Traditionally, most of the building energy simulation models are not coupled with the real building processes which make offline simulation a popular method for analysing building performance to implement potential corrective measures and evaluate their effectiveness again through the same process (Mirdamadi, Fontanili, & Dupont, 2007). In addition to helping facility managers to identify deficiencies in a complex dynamic system on a real-time basis and take corrective actions based on multiple simulation results, this approach can also help designers and engineers improve their design assumptions. For example, information obtained from sensors (i.e. temperature, humidity, and CO₂) and meters in a building can be automatically updated in the energy simulation model to predict energy required by the building systems for retrofit and commissioning analyses (Kamat et al., 2013).

On the other hand, the online building energy simulation involves collection of data from a real world process, immediately used for analysis in real time or quasi real time using building energy simulation tools and arrives at a corrective action or policy decision that can be implemented directly in the building or system (Kamat et al., 2013; Rao, He, Shao, & Zhang, 2008). The primary objective of online simulation is to assist facility manager to run several near-future simulations for a small number of alternative actions (or decisions), and select the option that best optimizes the objective function relevant to that context rather than taking a controlling action based on pre-defined policy. This entire process is graphically shown in Figure 4.
Examples of data-driven simulation studies

Several case studies have been published in the last 10 years that employ the method of data-driven simulation for decision-making in buildings. Table 9 highlights a few of these studies and provides information about the data types collected, data collection methods utilized, and the post-processing methodology employed. Two of these case studies are discussed in more detail in this section of the paper.

Menassa, Taylor, and Nelson (2013) developed an automated Hybrid Ventilation (HV) control for the public spaces of complex commercial buildings with the help of real-time data collection and analysis using sensors; this control identifies a commissioning methodology that determines the best way to operate a HV system in an occupied building. Data regarding humidity, temperature, energy use, and IAQ were collected with the help of BAS (wired data collection methods). Initially offline simulation was performed for estimating the energy usage during Traditional Mechanical Ventilation mode. Furthermore, online simulation was performed for validating the proposed methodology of automatically controlling the HV. The step-wise methodology of the entire monitoring process involving each step discussed is figuratively shown in Figure 5. In conclusion, the research successfully showed that the automated control of HV with the help of real-time monitoring and data collection can provide significant energy savings in complex buildings.

Zhao, Zhang, and Liang (2013) developed an internet-based, real-time energy monitoring system for monitoring the energy consumption of large public buildings in China. Data regarding office appliance power consumption (plug load as described in this paper), Hot water usage, Indoor temperature (Occupancy comfort information as described in the paper), HVAC, and water use were monitored every 5 minutes with the help of wired smart meters (i.e. using
wired sensor networks data collection method) and uploaded to the internet with authorized access to respective personnel. In addition, the data display system consists of graphs, pie charts, or tables with either instant 5-minute values or hourly, daily, and monthly data. Though the data aggregated was used to create awareness among the occupants, stakeholders, and

<table>
<thead>
<tr>
<th>Reference</th>
<th>Category of the data type(s) monitored</th>
<th>Data type(s) monitored</th>
<th>Data collection method(s) used</th>
<th>Post processing of the data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engvall, Lampa, Levin, Wickman, and Öfverholm (2014)</td>
<td>Dynamic parameters</td>
<td>Occupancy data (occupant comfort information and Occupancy characteristics)</td>
<td>Questionnaires and surveys</td>
<td>Offline simulation</td>
</tr>
<tr>
<td>Valovcin, Hering, Polly, and Heaney (2014)</td>
<td>Static and dynamic Parameters</td>
<td>Primary building information, building orientation, building materials, weather data, energy use (hot water and HVAC)</td>
<td>Questionnaires and surveys, wired/wireless sensor networks</td>
<td>Offline Simulation</td>
</tr>
<tr>
<td>Erickson et al. (2009)</td>
<td>Dynamic parameters</td>
<td>Occupancy data (occupancy schedule information)</td>
<td>Wireless sensor networks</td>
<td>Online simulation</td>
</tr>
<tr>
<td>Tsai and Lin (2012)</td>
<td>Dynamic parameters</td>
<td>Energy use (plug loads)</td>
<td>Other methods</td>
<td>Online simulation</td>
</tr>
<tr>
<td>Balaji, Xu, Nwokafor, Gupta, and Agarwal (2013)</td>
<td>Dynamic parameters</td>
<td>Occupancy data (occupancy schedule information) and energy use (HVAC)</td>
<td>Other methods</td>
<td>Online simulation</td>
</tr>
<tr>
<td>Menassa et al. (2013)</td>
<td>Static and dynamic parameters</td>
<td>Occupancy data (occupant comfort information), weather data, and energy use</td>
<td>Wired sensor networks</td>
<td>Offline and online simulation</td>
</tr>
<tr>
<td>Zhao et al. (2013)</td>
<td>Dynamic parameters</td>
<td>Energy use (plug loads, HVAC, hot water usage), occupancy data (occupant comfort information), and water use</td>
<td>Wired sensor networks</td>
<td>Offline simulation</td>
</tr>
</tbody>
</table>
building managers to conserve energy, the data could also be utilized in real time to make immediate decisions for savings in energy use. The step-wise methodology of the entire monitoring process involving each step discussed is figuratively shown in Figure 6.

The example case studies discussed above focus on specific data types, with an emphasis on energy savings on the respective data types that were considered. Some other similar studies in the recent past are summarized along with the data types monitored, data collection methods utilized, and the post-processing methodology employed and are tabulated in Table 9. There is also a significant opportunity for future research by extending the same approach to all of the data types discussed in this paper. For example, significant contributors to energy consumption in buildings – such as lighting, hot water, and water usage – can also be monitored and controlled in real time. In addition, hybrid sensors that can simultaneously measure data regarding different data types such as temperature, humidity, and CO₂ levels – can be used to improve the efficiency of the data collection process.

Discussion and conclusions
This paper presents the taxonomy of the data types that directly or indirectly affect the building energy consumption and also presents an overview of the data collection methods along with their advantages and disadvantages. The challenges associated with data collection from the perspective of decision maker (i.e. building manager or design/engineer) are discussed. In summary, the
extensive review of literature showed that wired networks provide more reliable transmission, but building managers and users might find them aesthetically displeasing and cost prohibitive to install in certain type of facilities (e.g. old historic buildings). In addition, installing wiring is a critical challenge for building managers in the case of existing buildings or buildings that need to be retrofitted. On the other hand, wireless networks are flexible, cost effective, and easy to install compared to their wired counterparts. They do not need to be connected to the building wiring and hence are immune to the electrical damage caused inside the buildings. However, since wireless sensor nodes are battery powered, power consumption issues and subsequent maintenance (which comes at a cost) are major areas of concern for the building manager. Future research should focus on addressing these issues and making the availability and accessibility of data more streamlined and easy to collect and analyse.

Apart from the aforementioned challenges, there are other barriers that hinder the wide adoption of energy-efficient technologies in buildings. These include higher up-front costs for more efficient equipment, a lack of access to financing, a lack of energy subsidies, and a lack of internalization of environmental, health, and other external costs (Metz, 2007). In addition, different characteristics of individuals and groups in an organization can hinder energy-efficient technologies and practices (Carbon Trust, 2005). A lack of training information on the energy-savings potential and some of the characteristics of political system – such as incomplete information,
lack of leadership interests, and insufficient social control of the standard procedure – also makes it difficult to invest in energy-efficient technologies (Yao, Li, & Steemers, 2005).

With increasing awareness and responsibility towards sustainable living, there is a lot of ground to be broken in the fields of building energy monitoring and building automation systems. The collaboration of academic and industry experts to delve more into dealing with the issues pertaining to data collection, energy monitoring, and subsequent savings is thus of continued importance.

Acknowledgments
The authors would like to acknowledge the financial support for this research received from the US National Science Foundation (NSF) CBET 1407908 and CMMI-BRIGE 1414855 awards. Any opinions and findings in this paper are those of the authors and do not necessarily represent those of NSF.

Disclosure statement
No potential conflict of interest was reported by the authors.

References
at the proceedings of the 11th ACM conference on embedded networked sensor systems (pp. 17). Roma: ACM.


