CIFE CENTER FOR INTEGRATED FACILITY ENGINEERING

Formalizing Assumptions to Document Limitations of Building Performance Measurement Systems

By

Tobias Maile, Martin Fischer
& Vladimir Bazjanac

CIFE Working Paper #WP125
August 2010

STANFORD UNIVERSITY

If you would like to contact the authors, please write to:

c/o CIFE, Civil and Environmental Engineering Dept.,
Stanford University
The Jerry Yang & Akiko Yamazaki Environment & Energy Building
473 Via Ortega, Room 292, Mail Code: 4020
Stanford, CA 94305-4020

Formalizing assumptions to document limitations of building performance measurement systems

Tobias Maile, Martin Fischer, Vladimir Bazjanac

1 Abstract

Building energy performance is often unknown or inadequately measured. When performance is measured, it is critical to understand the validity of the measured data before identifying performance problems. Limitations of measurement systems make adequate assessment of validity difficult. These limitations originate in the set of available data and in the functional parts of the measurement system. Previous research has used project-specific assumptions in an ad-hoc manner to describe these limitations, but the research has not compiled a list of critical measurement assumptions and a process to link the measurement assumptions to performance problems. To aid in the assessment of measured data, we present a list of critical measurement assumptions drawn from the existing literature and four case studies. These measurement assumptions describe the validity of measured data. Specifically, we explain the influence of sensing, data transmission, and data archiving. We develop a process to identify performance problems resulting from differences between measured and simulated data using the identified measurement assumptions. This paper validates existing measurement data sets based on known performance problems in a case study and shows that the developed list of critical measurement assumptions enables the identification of differences caused by measurement assumptions and exclude them from analysis of potential performance problems.

Keywords: Measurements, building energy performance, assumptions

2 Introduction

Assessing building energy performance requires adequate measured data from a building. Previous studies report problems with measurement data or missing measurements in commercial buildings. For example, Torcellini et al. (2006) describe several problems with measurement systems for six case studies. O'Donnell (2009) document measurement problems as well, but also discuss consequences of missing measurement data. These reported measurement problems are the result of limitations of measurement systems that can be described with measurement assumptions. Previous research in this area has not used a common set of measurement assumptions. In fact, assumptions have been mentioned rarely, and then only in a project-specific contexts (e.g., Salsbury and Diamond 2000).

Measurement assumptions must be articulated to link the measurement system to the real world (the actual building). When clearly articulated, assumptions can explain limitations of measurement systems and resulting measurement data. This helps building energy professionals to understand differences between measured and simulated building energy performance data. While this paper focuses on measurement assumptions, Maile et al. (2010a) detail simulation approximations, assumptions, and simplifications, and Maile et al. (2010b) discuss a comparison methodology based on measured and simulated data. A person that we call a performance assessor, who could be either a commissioning expert and/or a control system expert, would perform or supervise the tasks involved to document measurement assumptions for the comparison of measured and simulated data.

A measurement system consists of a measurement data set in the form of sensor data and control data (e.g., set points), transmission of the data, and data storage where data are archived for later use. All parts of a measurement system have a specific function that may or may not be fulfilled continuously over time. The resulting quality of the archived data depends on the use of a measurement system that works with sufficient reliability overall and in each part. Limitations or shortcomings of measurement systems have often been described as measurement errors. Reddy et al. (1999) categorized measurement errors as errors of calibration, data acquisition, and data reduction but did not mention errors in the sensor product itself. While sensor errors can be described quantitatively, assumptions provide only qualitative descriptions of a given limitation.

The basis for each measurement system is the measurement data set, which is defined by the number of sensors, sensor types, sensor placement, and sensor accuracy. Several guidelines exist that define measurement data sets (Barley et al. 2005; Gillespie et al. 2007; Neumann and Jacob 2008; O'Donnell 2009). While all of these guidelines address the questions of which physical variables should be measured by which sensor types and how often, only some contain detailed information about sensor placement and recommended sensor accuracy. In addition, only some guidelines consider control points such as set points as part of the data set. Previous validation of these guidelines has not been comprehensive, in that the guidelines were not compared to other guidelines or to actual building case studies. Measurement assumptions must be closely tied to measurement data sets and guidelines, since the characteristics of sensors (such as placement, accuracy, or product type) define the basic limitations of measurement systems. The ability to detect performance problems depends greatly on the available data set, since inconsistencies can only be detected when the related parameters are actually measured.

Transmission of the sensor output data usually occurs within the control system and/or the data acquisition system. Some possible major limitations related to transmission are related to the bandwidth (Gillespie et al. 2007), hardware, and software of the control or data acquisition system. These limitations may reduce data quality or lead to data loss.

Data storage also plays a significant role in the quality of resulting measurement data and thus is another source of potential limitations to the measurement system. Different measurement systems archive data in different formats (Friedman and Piette 2001), but all contain at least a set of data comprising a timestamp and a sensed value. The time interval and format of the timestamp define the temporal granularity of the recorded data. The archival type of the value may or may not be consistent with the resolution of the sensed and transmitted data values. These inconsistencies lead to inaccurate measurements, inconsistent temporal resolution and loss of data.

Due to the described limitations of measurement systems, the reliability of archived data varies across buildings and over time. Reliability of a measurement system describes the consistency of archived data both in the number of archived data points per time interval and in the variability of this number over time. For example, Olken et al. (1998) reported reliability issues with measurement systems. The reliability of measurement systems is an important characteristic, since more data loss leads to less information about the performance of a building.

In this paper, we discuss, summarize, and then validate the existing guidelines for measurement data sets (see section 3.1). In addition, we mention the extent to which these guidelines include information about the possible limitations of the measurement system. We illustrate how to use these guidelines to develop a measurement data set for a project (see section 3.2) to minimize limitations. We discuss in detail the limitations of the three main functions of a measurement system: sensing (section 3.3), transmitting (section 3.4), and archiving (section 3.5). We describe the process to use measurement assumptions to assess difference between measured and simulated data (section 3.6). Given all of the described sources of limitations of measurement systems, we will develop a list of critical measurement assumptions based on the existing literature and on four building case studies (section 3.7). We categorize this list of assumptions according to the three main functions of a measurement system. In view of the variety of problems and limitations with measurement systems, we illustrate how to verify a measurement system based on existing methods (section 3.8) and define a reliability measure that quantitatively describes the reliability of the archived data (section 3.9).

We based our research on case studies to provide real-life context (Yin 2003) for observing problems and limitations with measurement systems. The research method typically used in this research area is a single case study. Survey- or interview-based methods are not applicable, due to the limited existing knowledge about and lack of standard practice regarding measurement assumptions. Thus, we observed current practice with four case studies. In section 4, we provide a description of each building, the measurement data set, and the data acquisition system for each case study. We highlight the specific limitations of each measurement system and the correlated measurement assumptions.

Based on the first two case studies, we developed a concept for the measurement assumptions and observed measurement systems. One later case study was used to prospectively validate this concept. The fourth case study is in progress, but is already presented in this paper since it adds to the generality. In section 5.1, we present a validation of the existing guidelines for measurement data sets based on known performance problems as well as a validation of measurement assumptions based on the number of times when measurement assumptions can explain differences between measured and simulated data in a case study (section 5.2). The use of multiple case studies of different building types (office, mixed-use research and office, correctional) and different heating, ventilation, and air conditioning (HVAC) systems (natural ventilation, mixed natural and mechanical ventilation, and only mechanical ventilation) provides more generality for our results compared to an approach involving a single case study. After describing details of the validation, we discuss recommendations (section 6) and provide limitations and suggestions for possible future research (section 7).

3 Measurement systems and their limitations

This section details the different sources of measurement assumptions that are based in limitations of the measurement system. Since limitations depend on the types and numbers of sensors within a measurement system, we first discuss existing guidelines (section 3.1) that define measurement data sets and describe how to establish such a data set for a given case study (section 3.2). These data sets consist of mechanical sensors, which are defined as devices that respond to a physical stimulus and transmits a resulting impulse (Sensor 2010). These mechanical sensors enable an automated, continuous and effortless (excluding installation and maintenance) collection of the resulting data. Human observations may extend these data sets with useful information, but due to the related effort and subjectivity are out of focus of this paper. We discuss details of sensor accuracy (section 3.3), data transmission (section 3.4), and archiving (section 3.5). We discuss the process to identify performance problems (section 3.6) and provide a list of critical measurement assumptions (section 3.7). This paper also describes existing methods to verify the functionality of measurement systems (section 3.8) and defines a reliability measure for them (section 3.9).

3.1 Existing measurement data sets

Due to inconsistencies and differences in measurement data sets used across buildings in current practice (Brambley et al. 2005), several authors have proposed guidelines and procedures defining measurement data sets for different purposes in recent years. The standard practice of measurement in buildings is to collect data for billing (utility data set) and for ad-hoc measurements (HVAC control system data set). We describe these standard data sets as well as existing measurement guidelines reported in the literature. Through the description of these data sets we show that buildings already contain measurement data (while the data may not be archived) and that several different guidelines exists that guide building owners for selection of additional measurement data that are useful for their buildings. The resulting list of measurement data sets is:

- Utility data set
- HVAC control system (Underwood 1999) data set
- Guidelines for the evaluation of building performance (Neumann and Jacob 2008) data set
- Procedure for measuring and reporting commercial building energy performance (Brambley et al. 2005) data set
- Guide for specifying performance monitoring systems in commercial and institutional buildings (Gillespie et al. 2007) data set
- Ideal data set (O'Donnell 2009)

These different data sets exist for different purposes. The utility data set is used to determine the total energy consumption of a building to establish a basis for billing. The HVAC control system data set is needed for the operation and control of an HVAC system. The listed guidelines from the literature aim to establish data sets that support either the evaluation of building energy performance and/or the identification of performance problems. The analysis process of these guidelines ranges from benchmarking to the use of calibrated simulation models. While benchmarking is the comparison of actual building performance to comparable benchmarks such as corresponding standards or comparable buildings, calibrated simulation models try to match actual performance with simulated performance. Maile et al. (2010b) provide a more detailed discussion of existing comparison methods.

Typically, a measurement data set is characterized by:

- number of sensors
- types of sensors
- sensor placement
- sensor characteristics (accuracy and resolution)
- time interval
- installation costs
- maintenance cost

A good guideline for a measurement data set should provide information on all of the above characteristics. At a minimum, the existing guidelines for measurement data sets must describe the *types of measurements*, *number of measurements*, and recommended *time intervals*. However, the guidelines are inconsistent in providing further details. Some include information about *installation costs*, *sensor placement*, and *sensor characteristics (accuracy and resolution).* Table 1 summarizes the available characteristics for a measurement data set for each of the guidelines.

Table 1: Summary of topics covered by each guideline
("X" indicates sufficient information – "O" indicates some information – "-" indicates no information)

Guidelines	Barley et al. (2005)	Neumann and Jacob (2008)	Gillespie et al. (2007)	O'Donnell (2009)
Number of sensors	X	X	X	X
Types of sensors	X	X	X	X
Sensor placement	Refer to others	-	X	Refers to others
Sensor characteristics	X	-	X	Refers to others
Time interval	(Sub) hourly	15 to 60 min	1 min	1 min
Installation costs	O	-	-	O
Maintenance costs	O	-	-	-
Validation of data set	-	Savings shown for case studies	-	-

The next subsections highlight the differences between the guidelines, including the specific goals of each measurement set and a detailed comparison of the data points for each measurement data set (Table 2). This table is organized by different categories or levels of detail (building, system, zone, space, and floor) and compiles all the measurements mentioned in the literature that are most common in and part of at least one of the measurement guidelines.

Table 2: Summary of existing measurement data sets
(“X” indicates all, “O” indicates some, and “–“ indicates none)

Measure	**Utility set**	**HVAC control set**	**Neuman and Jacob (2008)**	**Barley et al. (2005) – Tier 2**	**Gillespie et al. (2007) – Level 3**	**O'Donnell 2009)**
	Building					
Total energy and water consumption	X	-	X	X	X	X
Total energy production	X	-	-	X	X	X
End-use electricity consumption	-	-	-	X	X	X
Domestic water usage	X	-	-	-	-	X
Outside dry bulb air temperature	-	X	X	X	X	X
Outside wet bulb temperature or humidity	-	X	X	-	X	X
Wind direction and speed	-	-	-	-	-	X
Solar radiation	-	X	X	-	-	X
	Systems (water)					
Water supply and return temperatures	-	O	X	-	X	X
Water flow rates	-	-	-	-	X	X
Water pressure	-	X	-	-	X	X
Water set points	-	X	-	-	X	X
	Systems (air)					
Air supply and return temperatures	-	O	X	-	X	X
Air supply and return humidity	-	X	O	-		
Air pressure	-	X	-	-	X	
Air flow rates	-	-	-	-	X	X
Heat exchanger temperatures	-	O	-	-	-	X
Coil valve positions	-	X	-	-	X	X
Air set points	-	X	-	-	X	X
	System components (e.g., fans)					
Component status	-	X	O	-	X	X
Fan/pump speed	-	X	O	-	X	X
Boiler/chiller water flow	-	-	-	-	X	X
Component electric consumption	-	-	-	-	X	X
	Zone					
Thermal box signals	-	X	-	-	-	X
	Space					
Space temperature	-	X	O	X	X	X
Space humidity	-	X	O	-		
Space temperature set points	-	X	-	-	-	X
Space velocity		-	-	-	-	-
Space 3-dimentional air flow	-	-	-	-	-	-
Space damper/valve positions	-	-	-	-	-	X
Space cooling/heating valve position	-	X	-	-	-	X
Space water supply and return temperatures	-	-	-	-	-	X
Space electrical sub-metering	-	-	-	-	-	X
	Floor					
End-use electricity consumption	-	-	-	-	-	X

3.1.1 Utility data set

The utility data set is needed to establish a basis for billing. It includes measurements for total building electricity and any other energy sources (such as, for example, chilled water, natural gas, or oil) that the building uses and are typically recorded at a monthly and/or annual interval. Depending on the breakdown of billing entities in a building, this data set may also include additional sub-meters. A validation of the utility data set is of little use, since it mainly includes total building-level measurements and additional sub-metering based on the breakdown of building objects. Placement of these sensors is typically at the building level and occasionally on a floor level, depending on the tenant distribution. Every building has a utility data set at the building level (sometimes multiple buildings are combined) that is archived typically on a monthly or annual basis for billing purposes.

3.1.2 HVAC control data set

The HVAC control measurement data set provides the feedback necessary for the control of the building (Underwood 1999). An example measurement point in this data set is an air temperature sensor in a space. Depending on that temperature measurement, the control system activates a set of components that allow for a change in the supply of heating or cooling air for that space. Typically, a corresponding set point defines the temperature or temperature range as the goal for the system to achieve. The point list of a building contains the measurements and control points of the control system. These necessary measurements for controlling an HVAC system typically include space air temperature, air loop supply temperature, system water temperature (supply and sometimes return), and absolute or differential system pressure. In addition to the mentioned set points, HVAC systems typically include a range of other control points that actuate specific components, such as valves, dampers, or fans. These control points describe the theoretical or goal state of the system, whereas sensors determine the actual state of a component, variable, or system. The HVAC control data set of measurements and control signals is the basis of operation for every HVAC system, and it is a fundamental part of a measurement data set. Typically, these data points are available in the control system, but the resulting data are not archived. Placement and accuracy goals may be included in design documentation or project specifications. The HVAC control data set is usually not validated in practice on completeness, accuracy and placement, since this has not been a major focus in buildings.

3.1.3 Guidelines for the Evaluation of Building Performance

According to Neumann and Jacob (2008), the goal of these guidelines is to enable a "rough overall assessment of the performance of the system." Neumann and Jacob use a top-down approach from benchmark comparisons to calibrated models, leading into ongoing. Their assessment is based on a so-called *minimum data set* that contains only a few measurements to keep measurement system costs to a minimum. Those are mainly building level measurements, such as primary energy consumption and some system- and space-level temperature measurements. The suggested time interval is between 15 minutes and an hour. A validation of the usefulness of these guidelines is indicated by early results that show 10-20% total energy savings in demonstration buildings (Schmidt et al. 2009). However, the authors did not validate their guidelines against other approaches. These guidelines do not contain any information about placement or accuracy of sensors.

3.1.4 Procedure for Measuring and Reporting Commercial Building Energy Performance

Barley et al. (2005) state that this procedure provides a "method for measuring and characterizing the energy performance of commercial buildings" and, therefore, defines a set of performance metrics. It also includes the necessary details on measurements for each performance metric. It defines two different tiers of measurement. Tier 1 focuses on building-level energy consumption. Tier 2 includes additional temporary sub-meters for energy consumption. Besides a difference in sensor and installation costs, the authors did not mention any other reasons to choose one tier over the other. The authors did not provide any validation of their measurement data sets other than describing an example building, its performance metrics, and the fact that the performance metrics seem to follow their expectations. The requested time interval for the data ranges from monthly (Tier 1) to (sub-)hourly (Tier 2). These guidelines contain a detailed discussion about sensor accuracy and some specific recommendations for selected sensor products. For sensor placement, Barley et al. mainly refer to manufacturer specifications.

3.1.5 Guide for Specifying Performance Monitoring Systems in Commercial and Institutional Buildings

The Guide for Specifying Performance Monitoring Systems in Commercial and Institutional Buildings (Gillespie et al. 2007) focuses on the requirements for a monitoring system that supports ongoing commissioning. It includes additional measurements compared to the procedures described above, such as measures of air and water flows as well as measures of power consumption of specific HVAC equipment, taken on a system level as a basic level of monitoring. The authors define three monitoring levels, *essential*, *progressive*, and *sophisticated*, to allow for different monitoring project scenarios. They do not provide any validation of these guidelines. Gillespie et al. (2007) recommend a one-minute interval. These guidelines contain specific information about sensor accuracy and placement but not about the costs of measurement systems.

3.1.6 Ideal data set and actor view

O'Donnell (2009) defines a so-called ideal measurement data set that includes a maximum set of measurements. While it is apparent that such an ideal set of measurements is very intensive in terms of the number of sensor products, installation costs, and control system capability, it provides a good reference for establishing a specific practical set of measurements for a particular building. Besides the measurements, this data set also includes control points and simulated data points. The control points are virtual points within the control system that either define a goal (a.k.a., set point) or are used to actuate HVAC components. Simulated data points can be generated with building energy performance simulation tools. For sensor accuracy, O'Donnell refers to Gillespie et al.'s guidelines and for the placement of sensors to Klaassen (2001) and ASHRAE (2003). O'Donnell briefly mentions the costs of sensors and their installation. He organizes the ideal data set by HVAC component and recommends collecting data at a one-minute interval. O'Donnell introduces an actor view of measurement data sets. Different actors have different requirements for measurements, depending on the purpose. For example, a financial officer requires the utility data set to calculate the total building energy consumption and, thus, the total energy costs for a building. Table 3 categorizes the mentioned measurement data sets based on this actor view concept to provide a sense for the relative magnitude of each measurement data set (from small to large).

Table 3: Categorization based on O'Donnell's actor view of measurement data sets

Actor view	Measurement data set	Time interval	Reference
Financial officer (minimal)	Utility data set	Monthly/ annually	N/A
Financial officer (advanced)	Procedure for Measuring and Reporting Commercial Building Energy Performance – Tier 1	Monthly	(Barley et al. 2005)
Control system	HVAC control data set	N/A	(Underwood 1999)
Building operator I	Procedure for Measuring and Reporting Commercial Building Energy Performance – Tier 2	(Sub-)hourly	(Barley et al. 2005)
Building operator II	Guidelines for the Evaluation of Building Performance	15 to 60 minutes	(Neumann and Jacob 2008)
Commissioning agent (essential, progressive, and sophisticated)	Guide for Specifying Performance Monitoring Systems in Commercial and Institutional Buildings – Level 1-3	1 minute	(Gillespie et al. 2007)
All	Ideal measurement data set	1 minute	(O'Donnell 2009)

One important aspect of measurement data sets is the time interval. The time intervals of the discussed guidelines range from annual to one minute, depending on the use of data. There is evidence in the literature that the one-minute time interval captures the right time scale of physical processes in buildings to detect performance problems. For example, Piette et al. (2001) report that the detection of several performance problems requires a time interval of one minute. Thus, we recommend a one-minute time interval. All of the case studies have a data set that collects data at one-minute intervals.

3.1.7 Summary of measurement data sets

Figure 1 presents a graphical overview of all data sets discussed that illustrates their relationships and the extent to which they include control signals, measured data, and simulated data. While all guidelines contain data points from measurements, only O'Donnell's ideal data set defines the data points for simulated data. Barley et al. use set points only as a reference if corresponding measurements are not available. Gillespie et al. and Neumann and Jacob include some space level information, while Barley et al. do not. Gillespie et al. and O'Donnell include set points in their measurement data sets. Set points are important, since they define the goal that the HVAC system aims to achieve. The need for simulated data points is inherent in the concept of comparison measured with simulated data. Section 5.1 compares these guidelines to the measurement data sets of the four case studies. In addition, we provide validation based on the identified performance problems of a case study.

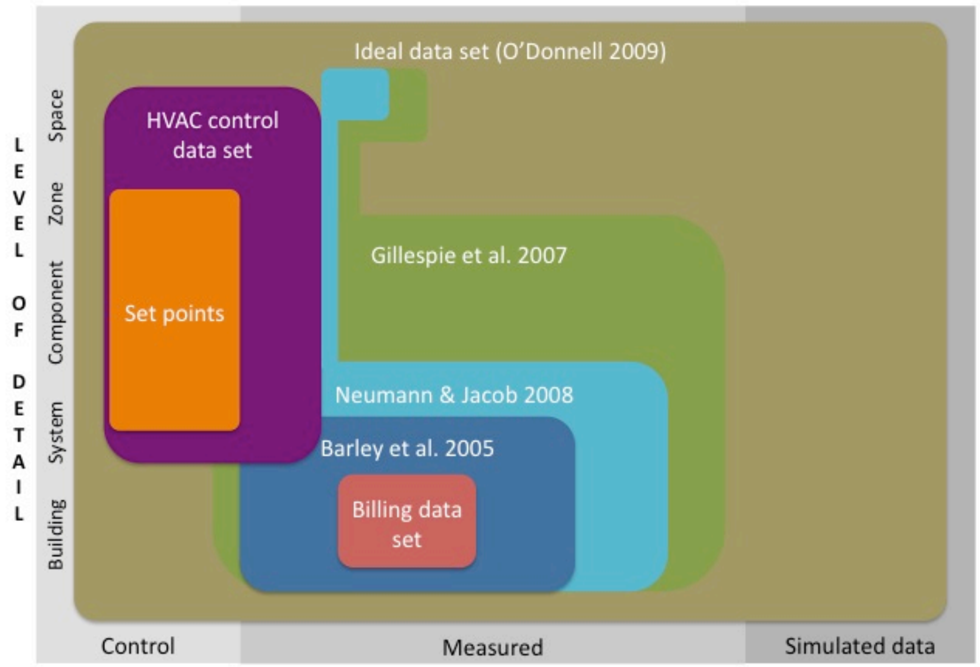

Figure 1: Overview of different measurement data sets

3.2 Selecting data points for performance evaluation

We used the following process to select the measurements for the case studies. If a point list for the building and mechanical drawings existed, the first task was to extract HVAC components from mechanical drawings and assign data points from the ideal point list to these component instances. This resulting building-specific ideal point list was the basis for the assessor to identify missing sensors and control points, by comparing the ideal point list to the existing point list. The assessor reduced this identified theoretical list to a practical point list based on project constraints. This process led to the installation of the identified additional points (Figure 2) and the addition to the data acquisition system. If no point list existed, Gillespie et al.'s guidelines provided a starting point for this first point list, since their list is the most comprehensive and practical list. We called the resulting data set the *assessor view,* based on O'Donnell's view concept (see section 3.1.6). Section 4 provides details on the resulting data sets for each case study.

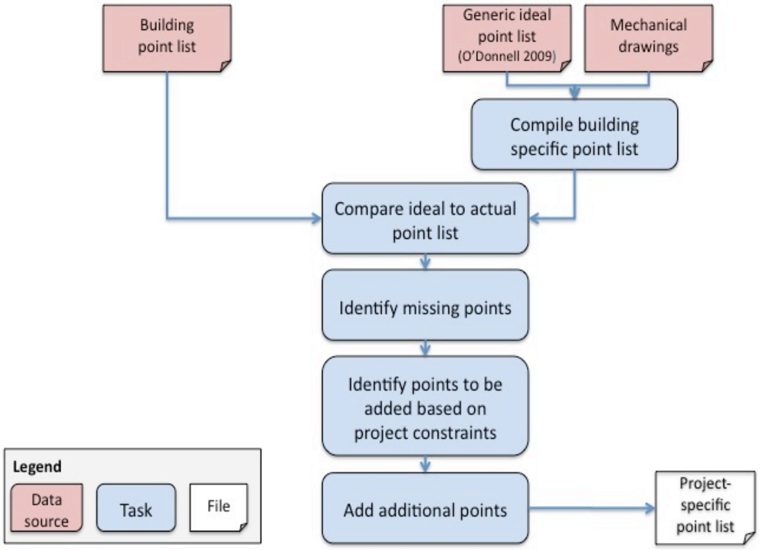

Figure 2: Process of selecting data points for performance evaluation

3.3 Sensor accuracy

Besides identifying a necessary data set, verifying the accuracy of sensors is equally important. Each sensor or measurement value has a certain error associated with it. This error is the difference between the sensor reading and the true value of the measurement (Reddy et al. 1999). The term *accuracy* also defines the same difference. Dieck (2006) argues that a measurement data point that is not described by accuracy/error and uncertainty has limited value. Thus, sensor errors are usually characterized by the error or accuracy bounds with an attached certainty. For example, we might have a measurement with accuracy bounds indicating that there is a 90% probability that the sensor reading lies within a 5% deviation of its true value.

Manufacturers define the first source of error and provide accuracy values for sensor products that typically are in the 1% to 5% range (up to 10%) for normal operating conditions. Incorrectly placed sensors or sensors that measure values that are outside of their normal operating range will have increased levels of errors, up to a level where the sensor reading does not provide any useful data. Gillespie et al. (2007) provide tables containing recommended sensor accuracy goals. We assumed that the published sensor errors provided by manufacturers are reasonably accurate, but we consider independent validation of these sensor errors as a possible future research area. For visual inspection of differences in time-series graphs, we integrated the sensor error bounds to the relevant measurement data points (see, for example, Figure 3). These error bounds support an assessor in his characterization of performance problems.

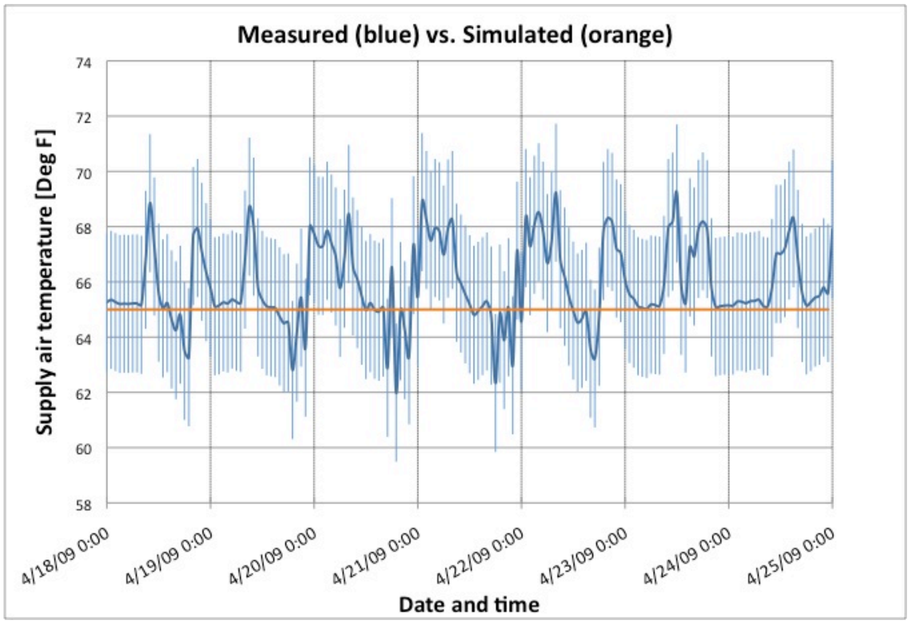

Figure 3: Example comparison graph to illustrate error margins of measurements (the measured data values are shown in dark blue, the 5% error margin is indicated with a lighter blue color, and the simulated values are shown in orange)

The most common sources of error are calibration errors, or the absence of calibration. Thus, it is important to calibrate sensors according to manufacturer specifications or, for example, procedures developed by the National Institute of Standards and Technology (NIST 2010a). For a given building project, it is important to understand the employed calibration techniques or the lack thereof and accordingly attach corresponding assumptions describing the limitations of the calibration technique to the data points.

Resolution is another key parameter of a sensor product. Resolution is the smallest change that the sensor can account for, and one can usually observe this change in the smallest significant digit of the resulting value. The resolution of a sensor may be reduced within data transmission and archiving. An assessor needs to document resolution issues in the same manner as other accuracy issues, with the help of assumptions.

In addition to sensor errors from the device itself, the placement of the sensor also plays a key role. For most of the sensors used in HVAC systems in buildings, the sensors provide spot measurements of temperature, flow, or pressure. If the placement is not appropriate, the reading of the sensor will not provide a representative value. A well-known example of this problem is the space temperature measurement. This measurement typically resides within the space thermostat, which is usually located next to one of the doors in a space. Avery (2002) illustrates a similar problem with mixed air temperature measurements in air ducts. He argues that the resulting measurement result can be inaccurate and influence the behavior of the system dramatically, even with multiple measurements that are averaged. We refer to existing placement guidelines such as those provided by Klaassen (2001), who specifically recommend placement guidelines for temperature sensors. ASHRAE (2007) provides more generic and comprehensive guidelines for a wider range of sensor types.

In the context of a comparison with simulated data, errors of measurement have three key influences. The first influence of measurement error is its relationship to assumptions. It is important to highlight the limitations of sensors and document them as assumptions. These assumptions help one to understand where differences between measurements and simulation results can originate. For example, we encountered such a bottleneck in the form of the connection of the electrical sub-meters to the data acquisition system, where multiple signal transformations led to numerous problems with our data. Measurement assumptions are discussed in section 3.7. Since we used a set of measurements as input for the updated simulation, the simulation results depend on the accuracy of the measurements. Maile et al. (2010a) discuss the accuracy of simulation results. The second influence of measurement errors occurs during the comparison process. If a measurement has a high error margin, the difference between the measurement and simulation values becomes less important, and performance problems may be hidden within the error margin.

3.4 Data transmission

While the sensor defines the first level of accuracy and limitations, the transmission of data from the sensor into storage may introduce additional shortcomings. A controller connects a number of sensors as a part of the control system and may need to use some transformation to adjust for differences between the sensor output signal and the controller input (typically analog to digital). The control system exposes the controller input signal and allows corresponding data logger software to archive the data. Through the use of additional data acquisition systems (such as specialized electrical sub-metering systems), more transformation steps may be introduced and, therefore, more sources of errors may exist. With each transformation step, the likelihood of errors increases. The use of different standards and practices for measurements and data communication add to the problem of data transmission. Typically, in a building, a number of different technologies and networks for measurements exist, creating a need to integrate these different systems. A possible solution to problems of data transmission in buildings is using an IP-based building system. If each sensor is directly integrated into an IP-based network, only a single transformation is necessary (sensor signal to IP). Maile et al. (2007) discuss this vision of IP-based building systems. In the context of a comparison with simulated data, it is important to understand these shortcomings of the data acquisition system and document them with assumptions.

While the resolution of a sensor defines the basic resolution level, data transmission may decrease the resolution level. If the right variable type is not assigned to a point within the control system, it may dramatically change the resolution. Small changes in value stored in the smallest digit can get lost. For example, a space temperature sensor normally has a resolution of a tenth of a degree; however, an integer variable assigned to a temperature sensor value can change the resolution to one full degree. It is common that control systems use integer variables for temperature measurements; however, these values are multiplied by a factor of ten to retain the resolution level. Obviously, one needs to perform a backwards conversion to return to the actual measured value. The resolution of the sensor and of the data transmission needs to be documented and described with assumptions, in case limitations exist.

3.5 Data archiving

Data archiving is the last step of a measurement system, storing data values and corresponding timestamps. Each sensed value has a timestamp attached, which becomes another potential source of error. Problems with these timestamps or biases, in time, will pose a limitation to sensor networks where the relationships between sensors are of interest. To ensure the proper date and time on storage servers and data logging hardware, it is beneficial to synchronize time based on a shared time source. Standard practice is to use a Coordinated Universal Time (UTC) timestamp to circumvent problems with daylight saving time changes (e.g., Olken et al. 1998). Limitations of data storage, in particular with respect to timestamps, should be documented with assumptions. Our recommendations for the design and functionality of data archiving are summarized in section 6.3.

3.6 Process of identifying performance problems with the knowledge of measurement assumptions

The previously described limitations of measurement systems in buildings are the reasons that specific data points in a building may not match simulation results. To identify performance problems from differences, we introduce the concept of measurement assumptions. Previous work has mentioned project-specific measurement assumptions only sparsely (e.g., Salsbury and Diamond 2000) and has not provided a process for assigning assumptions to building objects and data points.

The process for using assumptions to detect performance problems from differences between simulated and measured data is illustrated in Figure 4. The starting point for this process is a data graph containing the measured and simulated data values of the same variable. The first step is to detect differences between the simulated and measured data. We use simple statistical variables, the root mean squared error (RMSE) and the mean bias error (MBE), to detect differences. The RMSE gives an overall assessment of the difference, while the MBE characterizes the bias of the difference (Bensouda 2004). If an assessor finds a difference, he uses the assigned assumptions to determine whether the difference is in fact due to a performance problem. If he cannot explain the difference with assumptions, he classifies the difference as a performance problem. Otherwise, if the assumptions explain the difference or if there is no difference, the assessor moves on to the next data pair of measured and simulated data.

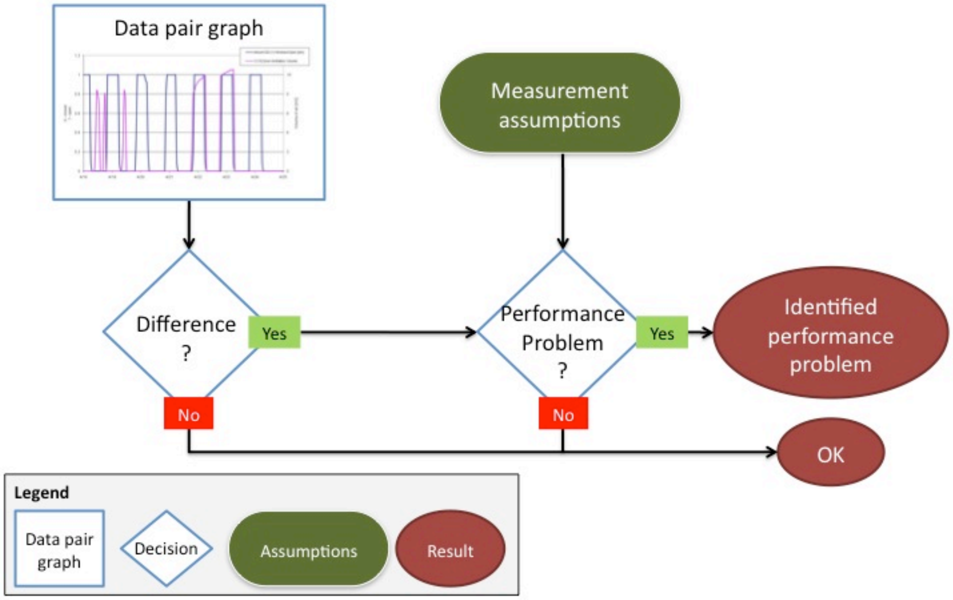

Figure 4: Process for using measurement assumptions to detect performance problems from differences between pairs of measured and simulated data

Since this process requires measurement assumptions to be linked to data graphs, we developed a formal representation of building objects that provides this link and is described in Maile et al. (2010b). For a given project, an assessor develops a project-specific list of assumptions based on a generic list of critical assumptions.

3.7 Measurement assumptions

This generic list of critical measurement assumptions is based on existing literature and four case studies. One example of a measurement assumption found in the literature is that of air leaking through a damper. Salsbury and Diamond detected a difference between the measured and simulated energy transferred across a mixing box and explain this difference by assuming air leakage in the return damper. We categorize the list of measurement assumptions by the three main functions of the measurement system in the list below. For each assumption, we identify the buildings from our case studies where this assumption is relevant, together with any additional references.

Sensor assumptions:

1. Direct solar beam values are derived from a solar model (SFFB, GEB; Soebarto and Degelman 1996)
2. Measurement is one-dimensional (SFFB)
3. Measurement is a spot measurement (SFFB, GEB, Y2E2, SCC; Avery 2002)
4. Surrounding medium causes temperature sensor drift (GEB, Y2E2)
5. Air is leaking though damper (Y2E2; Mills 2009; Salsbury and Diamond 2000)
6. Temperature sensor is influenced directly by the sun (SFFB, Y2E2, SCC)
7. Solar radiation measurement is not local (SFFB, GEB)

8. Sensor operating range is not appropriate (following guidelines) for application (Y2E2)
9. Manufacturer accuracy is not correct (SFFB, GEB, Y2E2, SCC)
10. Sensors are not or insufficiently calibrated (SFFB, GEB, Y2E2, SCC; Bychkovskiy et al. 2003)
11. Resolution is not sufficient (compared to guidelines) or reduced (SFFB, Y2E2; Reddy et al. 1999)
12. Diffuse solar radiation measurement is adjusted manually once a month (SFFB)
13. Sensor is oversized (Y2E2)
14. Sensor or physical cable is disconnected (SFFB, Y2E2)
15. Set point is valid for occupied hours only (Y2E2)

Data transmission assumptions:

16. Timestamps are from different sources (SFFB, GEB, SCC, Y2E2)
17. Bandwidth does not support data transmission of all points at the specified time interval (SFFB, GEB, Y2E2, SCC)
18. Data cannot be temporarily cached (Y2E2, SCC)
19. Hardware and software are not capable of handling the data transmission load (GEB, SCC)

Data archiving assumptions:

20. Data archiving software does not run continuously (SFFB, GEB, Y2E2, SCC)
21. Data are file-based (SFFB, GEB, SCC)
22. Data are overridden (GEB)
23. Daylight saving time change is not accounted for (SFFB; Olken et al. 1998)

Assumptions that occurred at all four case studies illustrate critical limitations of today's measurement systems. Those include spot measurements, insufficient manufacturer accuracy, insufficient calibration, timestamps from different sources, and data archiving software that does not run continuously. The latter two indicate reliability problems with data acquisition systems. Insufficient accuracy and calibration indicate problems with the process of selecting and verifying sensors, and spot measurements indicate problems with sensor placement or the number of sensors.

To illustrate the use of assumptions from case studies, we provide two examples. The first one is shown in Figure 5 and illustrates the window status of automated windows for natural ventilation (in this case atrium A&B 2nd floor at Y2E2). The graph shows a difference on the first four days as well as on the last day. The intermediate two days show a match of simulated (orange) and measured (blue) data. While the measured data correspond to a binary signal that indicates either an open or a closed status (left axis), the simulated data show the airflow (right axis) through the corresponding window. The small difference between the data on these two intermediate days is based on the difference in data types (binary versus airflow) and could be eliminated with a conversion of the airflow to a binary signal. However, the actual problem is the first four days and the last day, where the simulation predicts closed windows (no airflow) but the measured data indicate open windows at night. This difference indicates a performance problem, since there is no assumption that can explain the difference. The building operator later verified this performance problem.

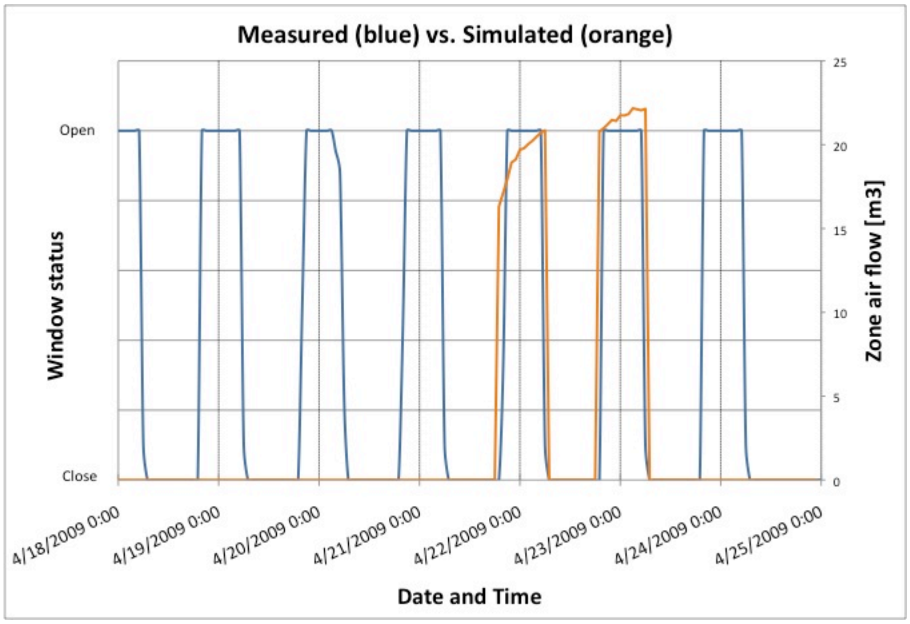

Figure 5: Comparison data pair graph: Window status atrium A&B 2nd floor

The second example illustrates a comparison of a space heating temperature set point (Figure 6). The correlated assumption (assumption no. 15) is that these two time series should only match during occupied hours. The set point used in the simulated model combines the occupied and unoccupied set points, whereas the set point in the control system shows only the occupied set point. The unoccupied set point is a different data point in the control system. Taking this assumption into account, the differences between the two time series all fall within unoccupied hours, and the two time series match otherwise. Thus, we can explain this difference based on the corresponding assumption and identify it as not a performance problem.

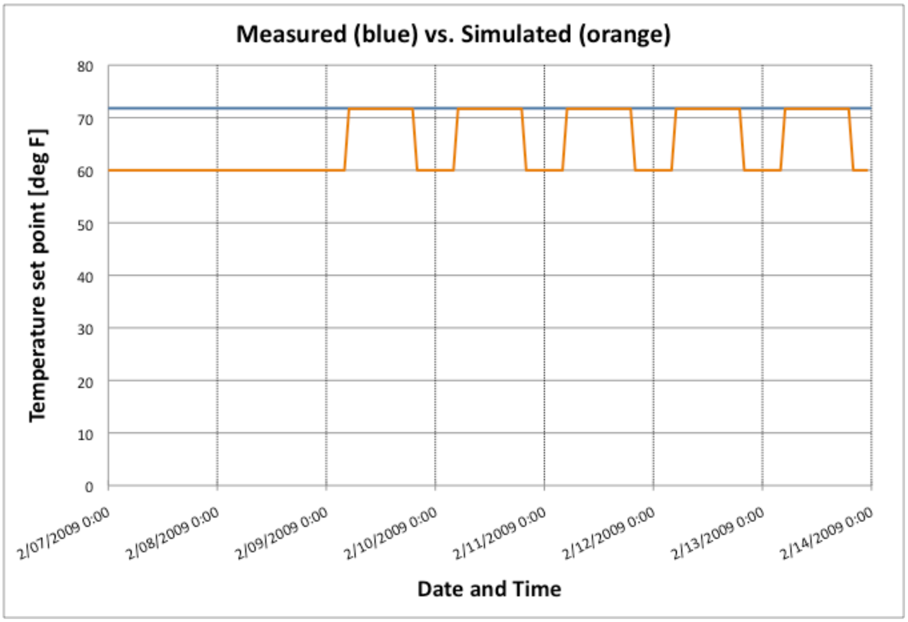

Figure 6: Comparison data pair graph: Temperature heating set point for space 143

3.8 Verification of the functionality of a measurement system

After the implementation of the data acquisition system, one needs to collect initial data, validate them, and crosscheck the initial data to flesh out potential and unanticipated problems with sensors and archived data. We based this process (Figure 7) on existing data analysis techniques (Friedman and Piette 2001; Scientific Conservation 2009; Elleson et al. 2002; Seidl 2006). Verification of data is especially important for weather and other data used as input for the simulation, since the results of a simulation model can be only as accurate as its input data. We performed four manual validation tests that included testing values against typical bounds (minimum and maximum limits), crosschecking values with other sources, validating daily patterns of measurements, and verifying continuous data collection. We developed an automated routine only for the latter, the full automation of these and additional data validation techniques was outside of the scope of this work, but would be a very fruitful area of future research.

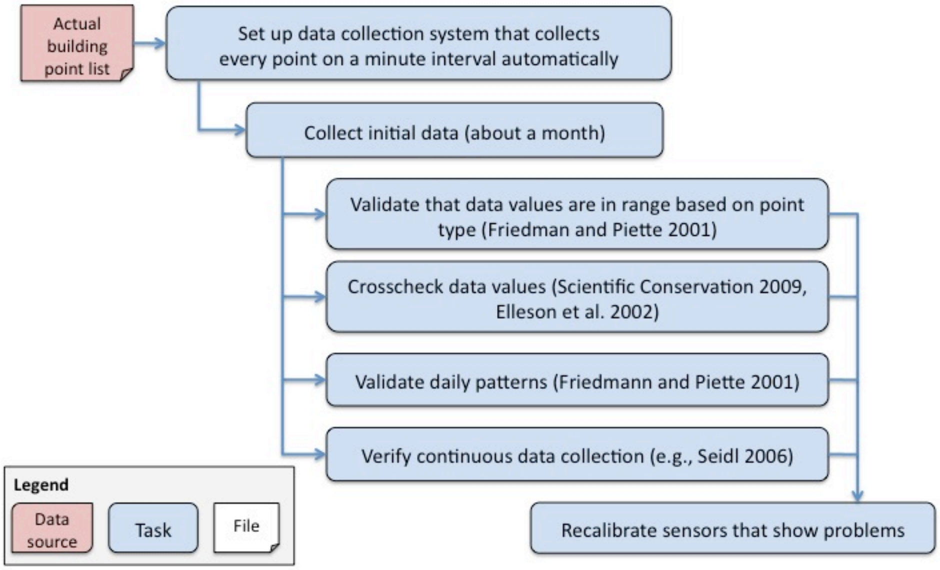

Figure 7: Process of setting up and verifying a data acquisition system

Validation can be accomplished by comparing each data point with its expected range (Friedman and Piette 2001). The sensor types define the acceptable value ranges. For example, a space temperature reading is typically between 60 and 80 °F, whereas a solar radiation reading is between 0 and 1,000 W/m^2.

Crosschecking specific measurements with other sources is a useful technique to verify sensor readings. For example, comparing the building-specific outside air temperature to the nearest weather station temperature measurements (Scientific Conservation 2009) can highlight either problems with the temperature sensors or unknown local sun effects. In addition, aggregating sub-measurements to the total and comparing the two will demonstrate the accuracy of measurements or show problems with measurements. For instance, the sum of adequate electrical sub-meters should equal the total building electricity measurement. Elleson et al. (2002) describe this principal to crosscheck data of different related measurements in the context of cooling systems.

A third option to verify measurements is to look for specific patterns. Typically, each sensor type follows a specific pattern. For instance, a solar radiation measurement should be zero at night and show a curve that increases in the morning, peaks sometime around noon, and decreases into the early evening.

The initial data set can also be investigated for missing data by identifying and resolving issues with the data acquisition system that would have led to data loss. Seidl (2006) mentions a simple technique that counts available data values within a specific period and verifies this count with the expected number of values. For example, based on a one-minute time interval, there should be 60 values within one hour for each data point.

Based on the results of these validation techniques, one needs to recalibrate sensors that show problems. Our focus for the manual validation of sensors was on weather and other input data for the simulation to minimize the time effort and focus on the most important sensors. This manual verification process of critical sensors took about half a day per case study. We automated the validation of the number of archived data points per period to continuously control the reliability of archived data. Other automated data validation techniques (e.g., Hou et al. 2006) require correct sensor readings to train algorithms and, thus, are not useful for an initial data validation.

3.9 Reliability of a measurement system

A data acquisition system needs to provide some level of reliability. We defined a reliability measure that illustrates how many data values are present in the archive compared to the theoretical number of data values (the reliability measure equals the number of actual archived data values divided by the theoretical total number of data values for a given time interval and timeframe). We used this reliability measure to compare the reliability of the data acquisition systems of the four case studies as shown in section 4.6.

4 Case studies

We chose four case studies to observe the current practice of measurement assumptions and limitations of measurement systems. The case studies provide real-life context (Yin 2003) for these topics. Case studies are commonly used in this research field, since surveys or questionnaires rely on the existence of sufficient knowledge and standard use. Based on our observations with the first two case studies, we developed concepts for measurement assumptions and a process to detect performance problems from differences. The two later case studies were used to prospectively validate these concepts and compare them to the methods used in practice to illustrate the power of our approach. In this section, we provide a brief description of the case studies, details on the measurement data set, and information about the data acquisition system. We also show example findings; for detailed validation results, please see Maile et al. (2010b).

For each of the case studies, we provide a short general description of the building (including floor area, location, and start of occupancy). We briefly describe the air conditioning strategy and special characteristics of the building, as well as details about the measurement data set. All of the measurement data sets of the four case studies are compared in Table 4 and put into the perspective of existing measurement guidelines. We describe the data acquisition systems and highlight the advantages and disadvantages of the particular setup. The selection process for these four case studies is described in detail in Maile et al. (2010b). The four case studies have different HVAC systems and different usage patterns to provide a range of buildings and HVAC systems that show the generality of our concepts.

4.1 Case study 1: San Francisco Federal Building (SFFB)

4.1.1 Building description

The SFFB is an office building in downtown San Francisco (Figure 8). Its main tower is 18 floors high, and the total facility has approx. 575,000 square feet (approx. 53,000 square meters). Occupants moved into the building in the spring of 2007.

Figure 8: East view of the SFFB

The building's main conditioning concept is natural ventilation, with the exception of core spaces and the lower floors. The lower floors are mechanically ventilated for noise and security reasons as well as less favorable conditions for natural ventilation because of surrounding buildings. The building has a long and narrow profile to facilitate cross-ventilation. The typical layout of the floors and the section view (Figure 9) reveals that the conference rooms and private offices are in the center of the floor and leave a gap zone around the cabin zones to allow natural cross-ventilation (McConahey et al. 2002).

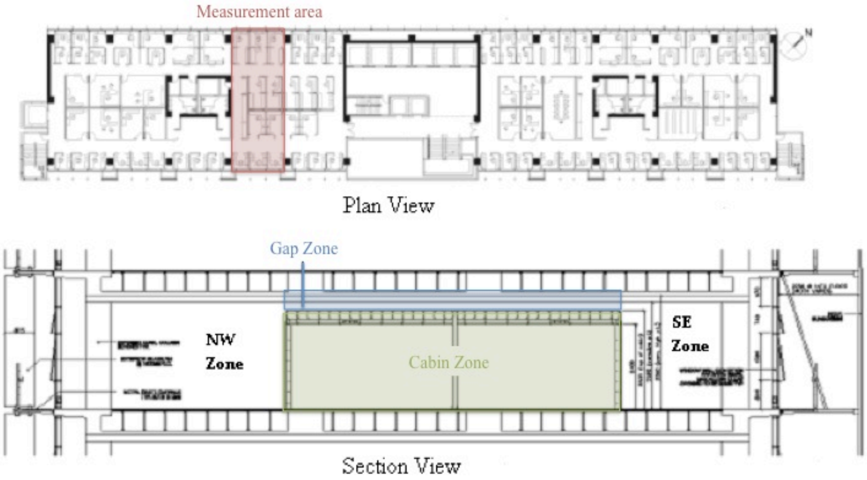

Figure 9: Plan and section view of the SFFB
(highlighting the "Gap Zone" and the "Cabin Zone" in the section view, and highlighting the measurement area in red in the plan view)

This case study is somewhat special compared to the other three case studies, since the opportunity to save additional energy through improved operation is insignificant, and performance problems relate only to thermal comfort.

4.1.2 Measurement data set

We set sensors on a part of the sixth floor of the building (Figure 10) and collected data from October 10, 2006 to February 8, 2007, a preoccupancy period. Unfortunately, the control system of the building was not fully functional yet; thus, the measurements were taken for two periods with different configurations. During the first period, all windows were open all of the time. The second window configuration was characterized by a regular open/close schedule (open from 8:00 p.m. to 8:00 a.m.; otherwise closed). The actual control strategy operates the windows based on various rules and conditions around the building.

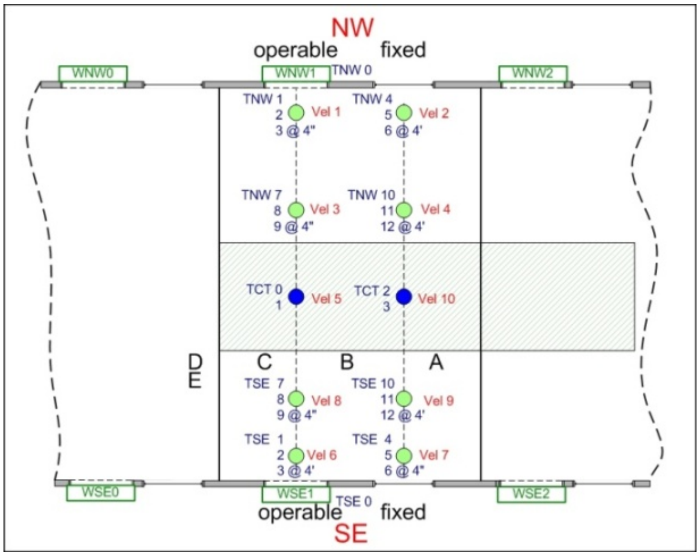

Figure 10: Plan view of sensor layout in the SFFB

The measurement setup included poles (indicated by colored circles in Figure 10) that each host three temperature sensors (occupant zone air, below-ceiling air, and ceiling temperature) and one one-dimensional velocity sensor (just below the ceiling). We equipped the automated operable windows in this part of the building with window opening ratio measurement devices. In addition, we installed outside air temperature sensors on the northwest and southeast façades of the building. Two pressure sensors (one for high and one for low pressure) measured the pressure difference across two opposite facades. Finally, five sonar anemometers (indicated with letters in Figure 10) provided more detailed data about the airflow within the space through their 3-dimensional airflow measurements. Table 4 summarizes the available measurements for this and the other following case studies and provides a comparison with existing measurement guidelines. Since this was the first case study and we participated in the design of the measured data set, no additional sensors needed to be identified.

4.1.3 Data acquisition system

The project team created a custom data acquisition system (Figure 11) to collect measurement data. It consisted of sensors, a custom-made hardware interface, and a designated PC that contained a project-specific Labview (National Instruments 2009) data logger script. This script captured the sensor data via the hardware interface and archived them into text files.

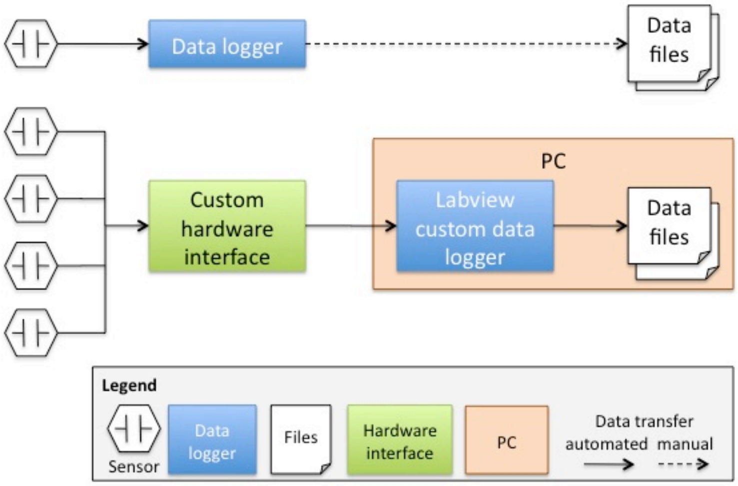

Figure 11: Data acquisition system at the SFFB

The biggest advantage of the data acquisition system used at the SFFB is its simplicity and efficiency. It did not use any additional or unnecessary intermittent controllers, communications, or protocols. Its simplicity even enabled data collection from the anemometers at one-second intervals. Use of such a simple data acquisition system was possible due to the proximity of the sensors and the temporary nature of the experiment. The difficulties and problems with file-based data (assumption no. 21) with this case study led to the use of a database in further case studies.

Regardless of the simplicity of the data acquisition system, we faced some minor problems. Single sensors were accidently disconnected (there was still construction in the space; assumption no. 14). The data acquisition PC crashed twice (assumption no. 20), which led to data loss for the time that the PC was not operational. While one crash occurred due to a full hard drive, the reason for the second crash remains unknown. The switch from daylight saving time to standard time caused some difficulties (particularly in the autumn when the same hour exists twice; assumption no. 23). The use of time intervals for the anemometers of less than a second led to some invalid data. The integration of data from this main data acquisition system and data from a data logger for solar data increased the effort of data analysis, due to the use of different data formats and time stamps (assumption no. 16). The average reliability measure (defined in section 3.9) for this data set is 96.96 %.

In addition, we had to collect solar data at a different site in San Francisco, which led to uncertainty about the validity of such measurements for the SFFB building. The use of a standalone data logger for the solar measurements, the need for manual download of data from that data logger, and the need to adjust the solar band on one of the two pyranometers caused data loss and additional uncertainty in the solar data. The assumption (assumption no. 12) about increased uncertainty of the solar measurements was an important finding of the project and allowed the explanation of differences in the space temperatures between measured and simulated data. Measured space temperatures that were lower than simulated occurred if the SFFB had cloud cover and the solar measurement side did not, and vice versa if the measured temperatures were higher than simulated. This finding initiated the assumption concept as described in section 3.6.

4.2 Case study 2: Global Ecology Building (GEB)

4.2.1 Building description

The GEB is a research and office facility of 11,000 square feet (approx. 1,000 square meters) on the Stanford campus. It is a two-story building (Figure 12, left). The first story (mainly lab area) has a mechanical air-conditioned system (see plan view in Figure 12, right), while the system on the second level is natural ventilation. One specific feature of the building is its lobby space. With three large operable glass doors, an evaporative cooling tower, and radiant floors, it has an interesting HVAC system combination that intends to allow smooth transition between the outside and the inside environment. Another innovative feature of the building is its water-spraying roof. At night when the outside air is cooler than the chilled water loop temperature, the system sprays water into the air above the roof. Evaporation cools the water, a tank stores it, and during the day, the chilled water system uses it for cooling inside the building (Carnegie Institution for Science 2010). Building occupancy started in April of 2004.

Figure 12: Southeast view of the GEB (left) and plan view of the first floor of the GEB (right) (Carnegie Institution for Science 2010)

4.2.2 Measurement data set

The GEB case study consists of a data set collected during the summer of 2006. An extended measurement period of one week included 20 additional temporary measurements, such as air temperature and plug loads, manually observed window positions and manually observed occupancy, and approximately 60 data points from the control system (available for a period of one month). This additional data set was identified using the process described in section 3.2. The control points included air and water temperatures, flow rates, valve positions, fan status, and set points. In addition, a second data acquisition system measured electric consumption on a building level for lighting, plug loads, server use, and HVAC components. The available measurements for the GEB are summarized in Table 4.

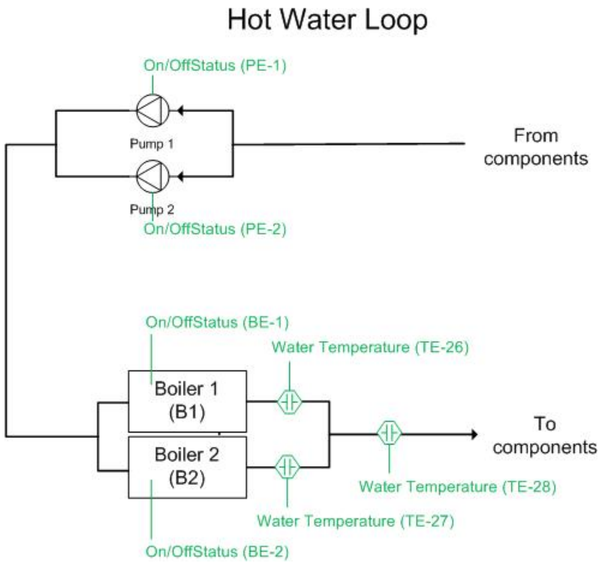

Figure 13: Example schematic for hot water loop data points in the GEB

(a schematic of the main hot water loop and the corresponding data with unique identifiers in brackets)

4.2.3 Data acquisition system

Most of the data from the GEB was collected via the control system; however, a designated Campbell data logger (Campbell Scientific 2010) archived the electrical sub-meter data. A project-specific Labview (National Instruments 2009) script extracted data from the data logger and united it with data files generated by the control system (Figure 14). The control system connected sensors to controllers. A modem allowed communication between the control system and the corresponding control software tool to archive data in comma-separated data files.

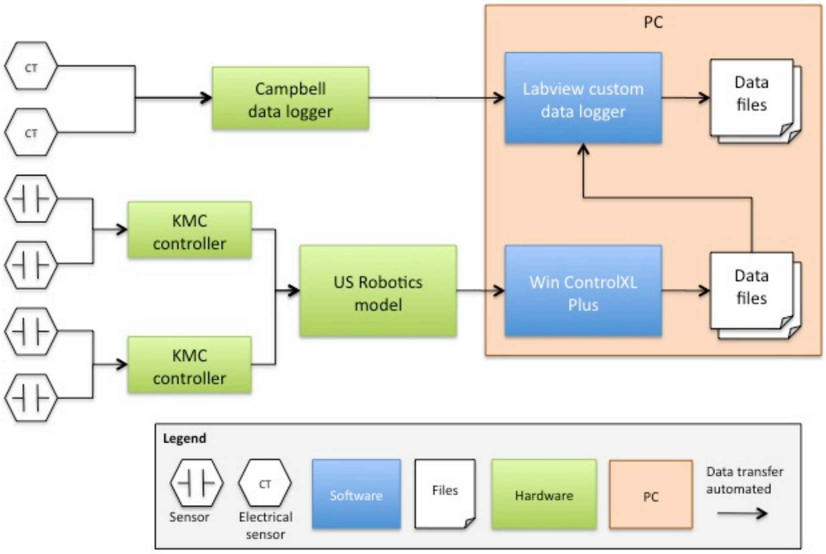

Figure 14: Data acquisition system at the GEB

The use of two separate data acquisition systems caused a number of problems during the project. In particular, the modem connection between the control system and the data acquisition PC was very unreliable and caused loss of data over hours and days. The modem would regularly stop working, and one could only reset it manually at its physical location. The control system was unable to archive all available data points at a one-minute time interval in a reliable fashion (assumption no. 19). The limited bandwidth of the main building control system (assumption no. 17) hindered the collection of data at a one-minute time interval. The limited data acquisition capabilities of the main control system and its software were other reasons for the unreliable data collection. Accidentally overwritten data files (assumption no. 22) due to a software bug just after our initial data collection period added to the loss of data. In spite of close attention by the research team, the average reliability measure of this system was 34.74% considering all data points (including the temporary data).

4.3 Case study 3: Yang and Yamazaki Environment and Energy Building (Y2E2)

4.3.1 Building description

The Y2E2 building is a research building of about 166,000 square feet (15,000 square meters) located on the Stanford campus (Figure 15). Its basement hosts mostly laboratories, and the three upper floors contain offices, meeting rooms, and classrooms. Its occupancy started in December of 2007.

The building has a hybrid HVAC system consisting of natural ventilation and mechanical air conditioning. The four atria, one of its key architectural features, facilitate the use of natural light but also play an important role in natural ventilation. They provide a stack effect that exhausts air naturally. The building uses active beams to supply air to the spaces. Its extensive thermal mass enables night flushing to cool down the building at night in hot summer weather and keep it cool during the day. The building also includes some radiant floors, radiators, ceiling fans, and fan coil units (Graffy et al. 2008). The Stanford Cogeneration Plant provides the chilled water and steam to serve the cooling and heating needs, respectively.

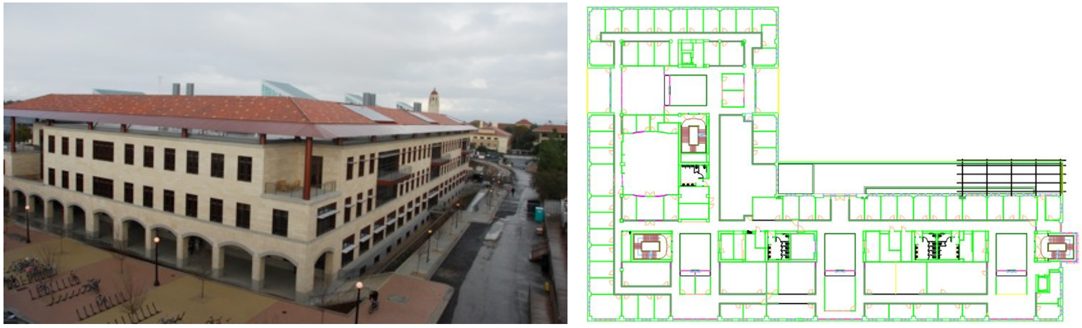

Figure 15: Illustration (left) and floor plan of second floor (right) of the Y2E2 building

4.3.2 Measurement data set

The measurement data set contains 2,231 data points that include water and air temperatures, valve positions, flow rates, and pressures for the air and water systems. As an example system, the main hot water loop and its related data points are illustrated in Figure 16. The building also has electrical sub-meters that measure plug loads on a half-floor basis and lighting consumption on a floor basis as well as the electricity used by mechanical equipment. There are three photovoltaic units, and each has its own power-generation sensor. The dataset for Y2E2 also includes a set of control points such as set points (for temperature, airflow, and pressure), occupancy indicators, and valve and damper position signals. Its most unique features are the four so-called representative offices that contain more sensors compared to the majority of the office spaces. For example, measurements of plug loads, lighting electricity, ceiling fan electricity (if present), supply, and return water temperature for the active beams or radiators are also collected in these representative offices. We installed a solar radiation sensor that measures both diffuse and total radiation on the roof. This solar sensor was the only additional sensor identified by the process described in section 3.2. Table 4 compares the available data set for the Y2E2 with the other case study data sets.

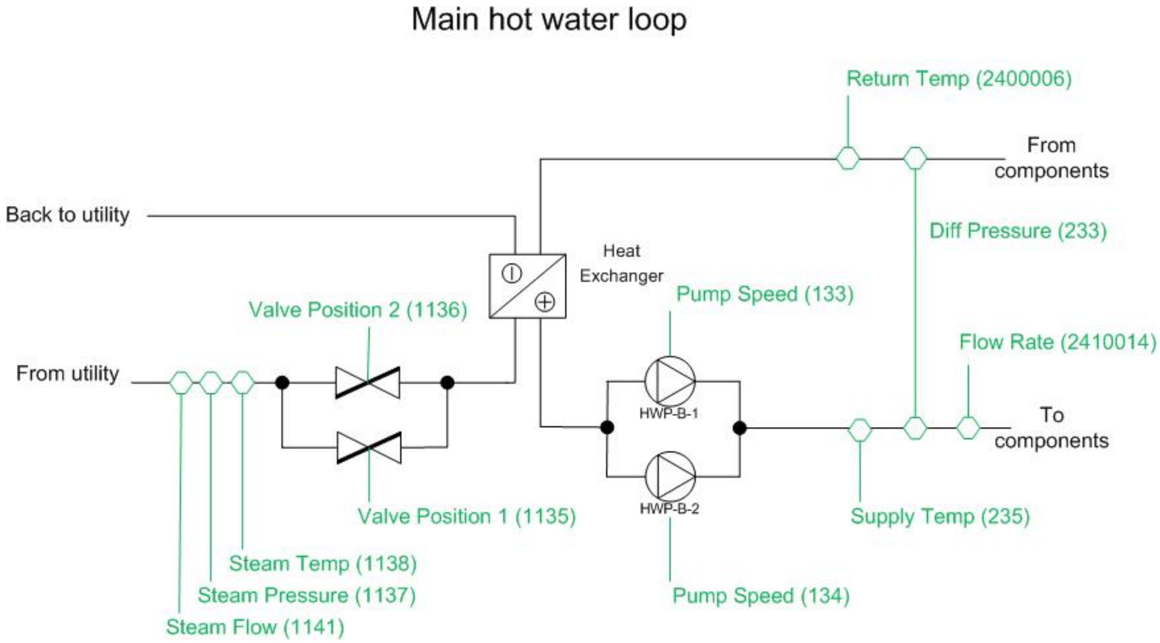

Figure 16: Example schematic for hot water loop data points in the Y2E2
(a schematic of the main hot water loop and the corresponding data points including unique identifiers in brackets)(Haak et al. 2009)

4.3.3 Data acquisition system

The HVAC control system of the Y2E2 uses a LONwork's protocol-based network (LonMark International 2010) consisting of six subnets (Figure 17). The sensors communicate directly with the controllers. The controllers, as part of one sub-network, connect to one iLON server (Echelon 2010). Each iLON server has its own data logger that stores the data temporarily on the iLON server. A Windows logging service connects to the iLON servers and fetches the available data from it via the Ethernet. This logging service archives the received data into a MySQL database. We implemented this setup since it was the only feasible solution to archive all sensor data at one-minute intervals. Since the bandwidth of the control system was limited, the solution with an iLON server on each subnet enabled us to reduce network traffic compared to using a solution in which all network traffic goes through one interface.

The LONwork-based control system also integrates two additional systems. The campus-wide energy management control system (EMCS) connects via a field server to the LONwork network. This EMCS system controls the main systems in the building. Therefore, the sensors of the main systems are integrated into the data acquisition system via the EMCS and field server. The iLON server also integrates the electrical sub-meters directly via Modbus (Modbus 2009).

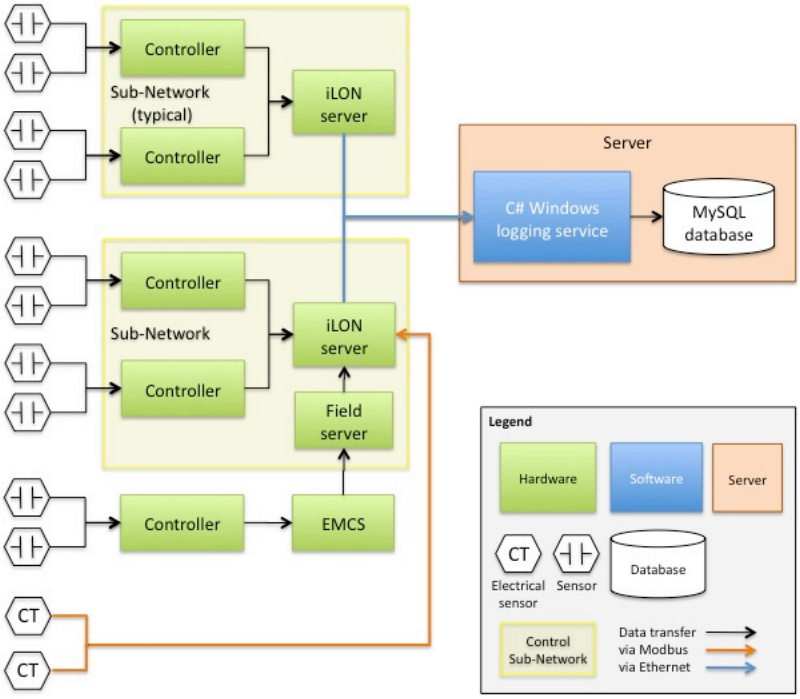

Figure 17: Data acquisition system of the Y2E2

The temporary saving of data on the iLON servers makes the data acquisition system more reliable (assumption no. 18). When the server with the MySQL database is temporary unavailable, the iLON servers queue data temporarily. The data logging service picks up these data as soon as it returns to normal operation. Depending on the internal disk capacity of the iLON server, data can be queued for 6-15 hours. However, once the iLON servers have a full data queue, they become unresponsive and the data logging service cannot keep up with the pace of newly accumulated data. This results in data loss. To minimize data loss due to hard drive failures, we installed a server with a mirrored hard drive to protect the collected data.

Due to the limited bandwidth of the HVAC control system, it was necessary to install six iLON servers (one for each subsystem) to reduce the data transfer load on the control network. Even though the six network subnets share the communication load, occasionally data point values are lost. We observed that, on average, between 100 and 200 out of a total of 2,231 data points do not have exactly 60 data values in an hour, which results in an average monthly reliability measure of this data set of 96.41%.

The measurement concept also contains a set of electric sub-meters on a floor level or even on partial floors. While it is beneficial to have more rather than fewer sub-meters, electrical sub-meter divisions should, ideally, coincide with the zoning of air-handling units. This would enable encapsulation of the different air-handling units so that sub-meter measurements correspond to the configuration of air-handling units.

The control system configuration did not expose all sensor and control points automatically. We could only expose points after understanding the architecture of the control network and the behavior of different controller types within it. The actual implementation of this control system dramatically increased the effort and difficulty of exposing points by several days. The controller inputs and outputs of hardware and software did not correlate for some controllers. While it was possible to directly access the hardware inputs and outputs and expose them for archiving, this workaround circumvented data conversion routines within the control system. We had to reapply conversion routines within the MySQL database.

In the Y2E2, the central university EMCS system connects to the control system, and several times; someone accidently disconnected the physical connection (assumption no. 14). As a result, some data were lost several times over multiple days. In addition, the connection between the electrical sub-meters and the data acquisition system was so complex that it took over 24 months to identify and fix the problems with electrical sub-metered data.

Another problem is that the water flow sensors for domestic water provide unrealistic (and inaccurate) data because they are oversized (assumption no. 13). The same was true for the electric sub-meters for the first two years because of numerous problems of the sensors and data transmission. New sensors and improved data transmission resolved these problems. If the true measured value is not within the typical operating ranges (assumption no. 8), the error of measurement grows large enough that the data collected is not trustworthy.

4.4 Case study 4: Santa Clara County Building (SCC)

4.4.1 Building description

The SCC is a 10-story building (Figure 18) of approx. 350,000 square feet (32,500 square meters). This correctional facility mainly houses cellblocks but also includes some office spaces and a kitchen. The building is about 30 years old. It has a traditional mechanical HVAC system. Most air systems are constant volume systems with 100% outside air. It has hot water boilers that serve the building's heating needs. Chiller and cooling towers provide the necessary cooling for the building.

We selected the building to be one of the case studies because it seemed to be relatively simple in terms of its architecture, its HVAC systems, and its occupancy. The majority of floors have the same layout, which simplifies the modeling of the geometry, even though it is the largest case study building. While there have not been any dramatic changes to the building, the documentation in the form of drawings and specifications is 30 years old and partially inconsistent. Compared to the other case studies, the HVAC systems are very typical and relatively straightforward to model. Its occupancy is mostly controlled and not as dynamic as in a typical office building. In addition, it is an existing building compared to the other case studies, which are all new constructions.

Figure 18: Northeast view of the SCC (County of Santa Clara 2010)

4.4.2 Measurement data set

The measurements for this case study contain typical measurements that are available to control the HVAC system in the building. These include space air temperatures, water and air temperatures within the HVAC system, total building electricity and gas consumption, flow rates, damper and valve positions, and on/off status of constant-volume fans and pumps. In the context of this study, we installed a set of new sensors that include water flow at the condenser loop, some system temperatures at a small number of typical air handlers, and a solar anemometer on the roof of the building. This additional set of measurements was identified following the process described in section 3.2. The resulting data set for the hot water loop is illustrated as an example in Figure 19. Data collection started on March 15th, 2009 for the existing measurements and on October 21st, 2009 for the newly installed sensors. The data set does not include set points due to limitations with the data acquisition system. Table 4 shows this measurement data set, among others.

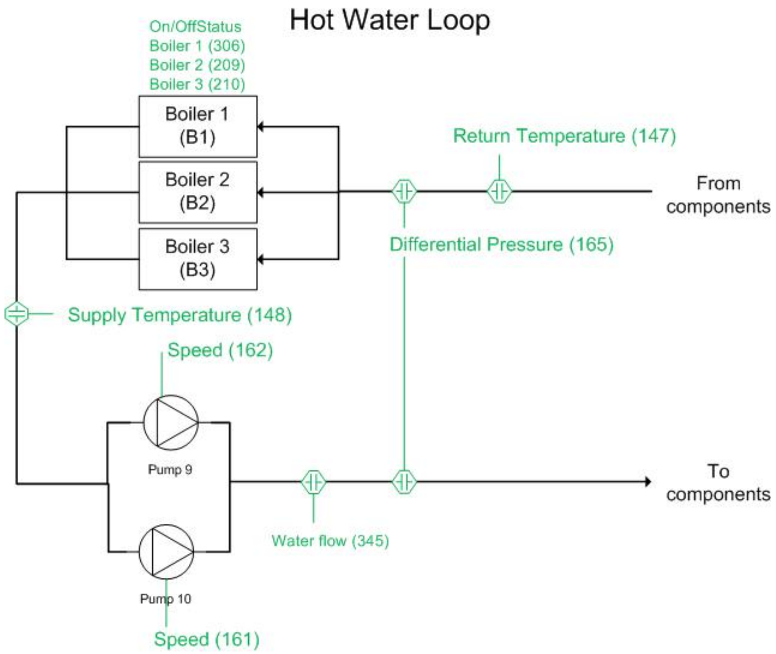

Figure 19: Example schematic for hot water loop data points in the SCC
(a schematic of the main hot water loop is shown in black and the data points are in green, including unique identifiers in brackets)(Kim et al. 2010)

4.4.3 Data acquisition system

The building control system is relatively simple due to its age. According to the building operator it is very reliable but lacks the capability to collect measured data in an automated and continuous fashion. Thus, we installed a control system upgrade to enable acquisition of the available measurement data.

This facility has an existing pneumatic control system that has no data archiving functionality. Within the project, we added a universal network controller that provides access to the global control module (GMC) and, thus, the control system via Ethernet. The GMC connects to the Microzone controllers, which, in turn, interface with the sensors. A software archiving tool based on the Niagara Framework (Tridium 2009) archives the data to an onsite server. To transfer the data to an off-site server, a script has been put in place to push the data files onto a file server, and a data import service downloads these files from that file server and imports them into a MySQL database (Figure 20).

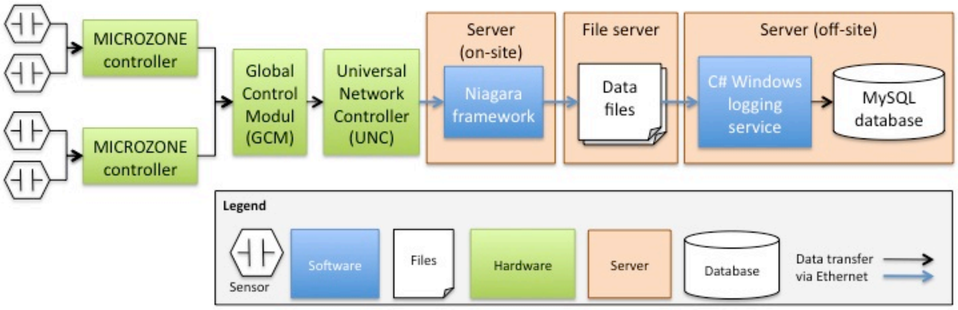

Figure 20: Data acquisition system of the SCC

The archiving of data with the Niagara Framework provides a redundant database along with the MySQL database. This setup adds to the reliability of the data acquisition system and minimizes data loss. The average monthly reliability measure of this data set is 74.34%.

Due to the architecture of the rather old original control system that transfers all data over the GCM, the system is prone to failures and/or temporary unresponsiveness of the GCM. In addition, it is not possible with this system to log any set points (these are hardcoded on the Microzone level). The limited bandwidth of the control system also restricts the time interval for the logged points (2- to 5-minute intervals).

4.5 Summary of measurement data sets

The following table (Table 4) summarizes the measurement data sets for all four case studies and relates them to the existing guidelines discussed in section 3.1. For reference, we include only the two guidelines with the highest number of data points for comparison, since the case study measurement data sets mostly fit in between these two data sets. The data sets focus on technical data and do not include automated occupancy sensors.

Table 4: Summary of measurement data sets of case studies and guidelines

Measure	**SFFB**	**GEB**	**Y2E2**	**SCC**	**Gillespie et al. 2007**	**O'Donnell 2009**
Building						
Total consumption	-	X	X	X	X	X
End use electricity	-	X	X	X	X	X
Domestic water usage	-	-	X	-	-	X
Outside dry bulb air temperature	X	X	X	X	X	X
Outside wet bulb temperature or humidity	-	-	X	-	X	X
Wind direction and speed	O	X	X	-	-	X
Solar radiation	O	-	X	X	-	X
Systems (water)						
Water supply and return temperatures	-	X	X	X	X	X
Water flow rates	-	X	O	X	X	X
Water pressure	-	-	X	X	X	X
Water set points	-	-	X	-	X	X
Systems (air)						
Air supply and return temperatures	-	O	X	O	X	X
Air pressure	-	O	X	X	X	
Air flow rates	-	O	X	O	X	X
Heat exchanger temperatures	-	-	X	O	-	X
Air set points	-	-	X	-	X	X
System components (such as fans, pumps, boilers, and chillers)						
Component status	-	O	X	X	X	X
Fan/pump speed	-	O	X	X	X	X
Boiler/chiller water flow	-	O	-	-	X	X
Component electric consumption	-	-	X	-	X	X
Zone						
Thermal box signals	-	-	X	-	-	X
Space						
Space temperature	X	X	X	O	X	X
Space temperature set points	-	-	X	-	-	X
Space velocity	X	-	-	-	-	-
Space 3-dimentional air flow	X	-	-	-	-	-
Space damper/valve positions	-	-	O	-	-	X
Space water supply and return temperatures	-	-	O	-	-	X
Space electrical sub-metering	-	-	O	-	-	X
Floor						
End-use electricity consumption	-	-	X	-	-	X

As illustrated in Table 4, the measurement data sets from the four case studies typically exceed the guidelines of Gillespie et al. (2007). This table also shows that while the SFFB case study is special in terms of using measurements to assess the natural ventilation only, the granularity of measurement increases from the GEB to the SCC to the Y2E2.

Besides the measurement data sets, these four case studies also included four different measurement systems with their advantages and disadvantages. Based on the experience of these case studies we summarize a number of recommendations in section 6.

4.6 Reliability measure of case studies

Figure 21 shows the average monthly reliability measures of the two case studies over time. The available data for GEB and SFFB consist only of a few months. Thus, no trends emerge from these small datasets. However, for Y2E2 and SCC datasets exist that exceed a year worth of data.

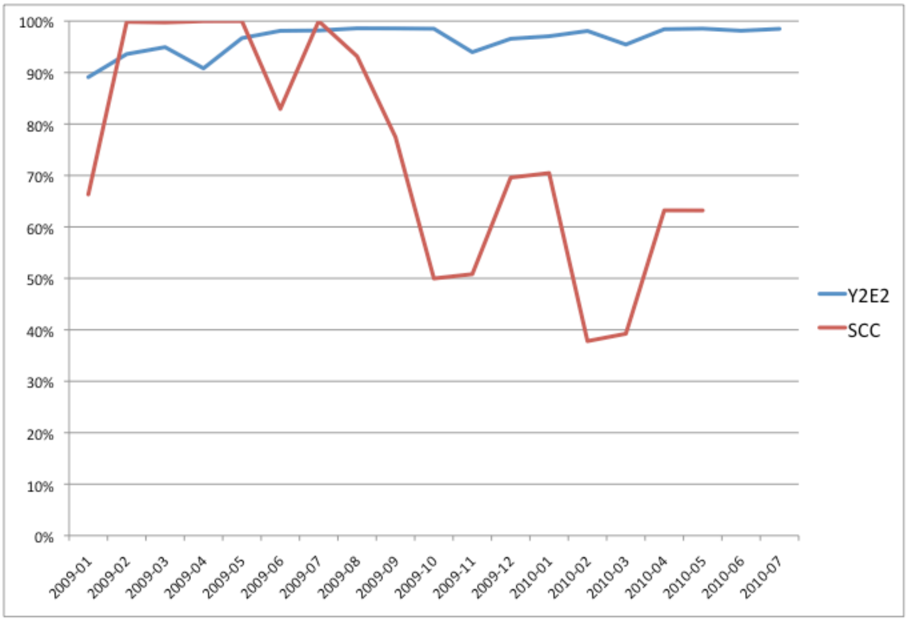

Figure 21: Monthly reliability measures of two case studies over time

For Y2E2 the reliability increased dramatically over the first two months during which the system was setup. After this setup period the reliability measure stayed pretty consistently around 97%, except of a couple of months where it dipped slightly because of electrical building shutdowns or server problems. At SCC, a very different picture emerged, the initial data collection quickly produced 100% reliable data, however, once the data set doubled in size, the reliability drastically dropped and varied somewhere between 40 and 70%.

5 Validation

In this section, we show the validation of existing measurement data sets as well as the validation of the measurement assumptions listed above. The validation is based on the performance problems found at the Y2E2 case study. Detailed information about the performance problems is included in Maile et al. (2010b) and Kunz et al. (2009).

5.1 Validation of measurement data sets

To show the value of the measurement data sets, we validated the existing data set guidelines based on identified and known performance problems at the Y2E2 building (Kunz et al. 2009). For each performance problem, we identified the data points that are necessary to detect it. Figure 22 shows the number of performance problem for each guideline, categorized by how many data points each existing guideline includes (none, some, all). Appendix A in Maile (2010) provides details about the analysis.

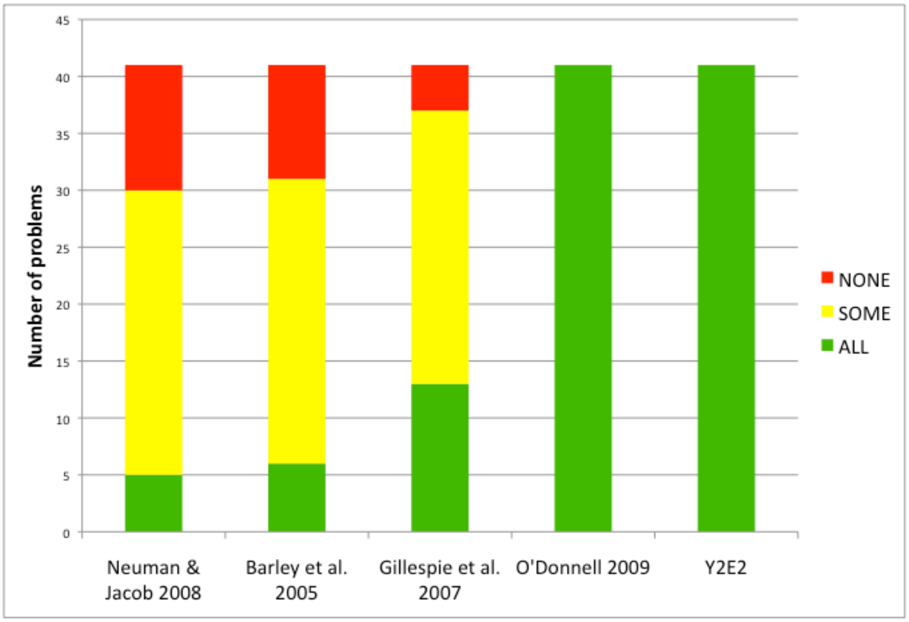

Figure 22: Results from validation of existing measurement data sets with the Y2E2

Not surprisingly, O'Donnell's ideal set contains all necessary sensors for all identified performance problems. Figure 22 also shows that the first three guidelines contain all the sensors for at most 30% of the performance problems, some sensors for about 60% of the problems, and no sensors for the remaining problems. While this figure shows the number of problems, it does not include an indication of the severity of each problem. For example, problems with particular sensors may have little influence on overall energy consumption, whereas incorrect control strategies can have a large impact on total energy consumption. If the problematic sensor is not used as input for a control strategy but, rather, a sensor that observes some conditions, it may have no influence on the building controls but only on observing the performance. If a control strategy is incorrect, its effects on building performance can be dramatic. Future research could investigate which guidelines capture which problems based on the problem severity. These findings indicate that using extended measurement data sets for buildings, at least at the level of Gillespie et al., can help in finding performance problems that otherwise cannot be detected.

5.2 Validation of measurement assumptions

The value of measurement assumptions lies in their use in evaluating differences between simulated and measured data. For the identification of performance problems at Y2E2, 29 differences between measured and simulated data could be explained with corresponding measurement assumptions. From a total of 109 differences, these 29 (26.6%) could be eliminated as false positives with the use of measurement assumptions.

Figure 23 illustrates the occurrences of measurement assumptions (indicated by measurement assumption numbers, see section 3.7 or Appendix B in Maile (2010)) in the four case studies and literature. Most of the measurement assumptions occur in more than one case study, are mentioned in existing literature or are used to eliminate false positives in the validation case study. The measurement assumption list does not include assumptions that are mentioned once in literature but did not occur in the four case studies.

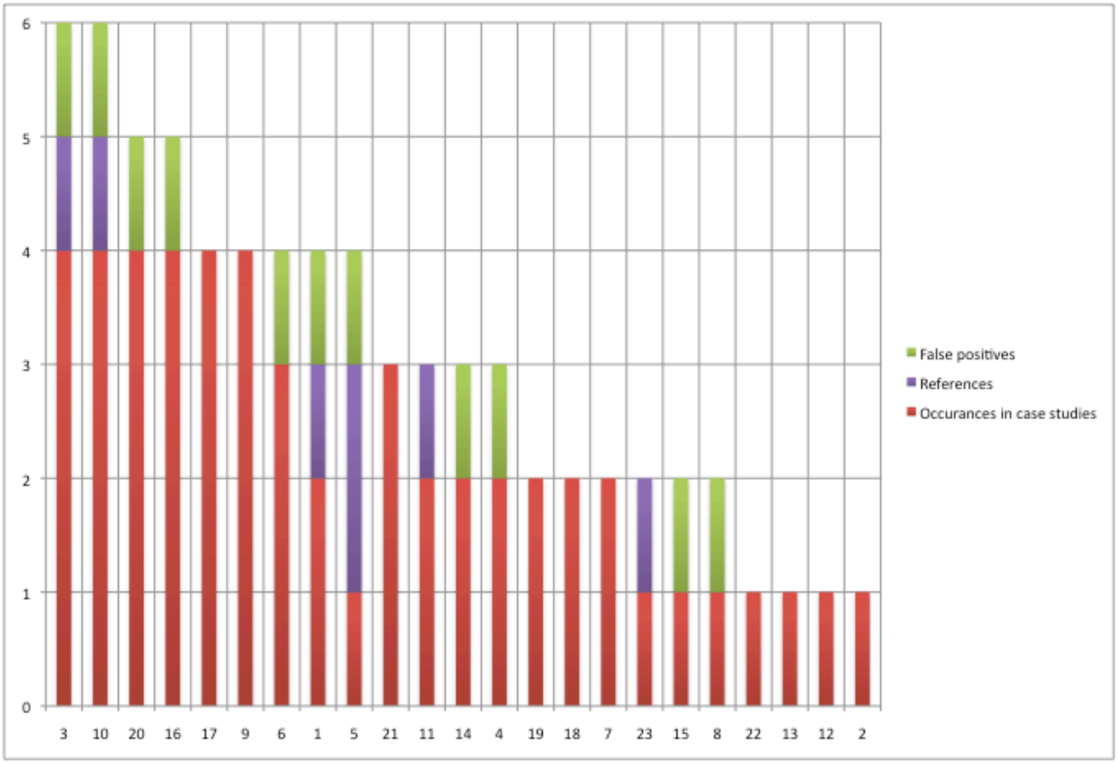

Figure 23: Occurrences of measurement assumptions in case studies and literature

This list of measurement assumptions highlights areas where measurement systems have limitations. Thus, these assumptions also indicate possible future research areas where limitations can be eliminated. Some areas of possible improvement are described via recommendations in section 6. The documentation of these areas provides a first step towards developing less limited measurement systems in buildings.

6 Recommendations

This section summarizes our recommendations based on the limitations of the measurement systems we encountered during the case studies. These recommendations aim to improve the quality and consistency of measurement systems by using existing measurement guidelines, calibrating sensors more thoroughly, improving data archiving, and using local solar measurements.

6.1 Select appropriate sensors based on existing measurement guidelines

The quality of data starts with the sensor products. While a number of guidelines and protocols exist (e.g., Gillespie et al. 2007; Barley et al. 2005), our experience indicates that these guidelines are not used to design data acquisition systems and to select appropriate sensor products. The accuracy, resolution, and operating range of sensor products are key metrics that designers need to consider during the process of planning data acquisition systems. It is important to select sensors of the right size that reflect the anticipated flow or electrical loads, rather than a size based on the worst-case scenario (see section 4.3.3). If the actual flow is difficult to estimate during design or varies dramatically during measurement, one can install two sensors with different operating ranges that provide accurate readings over an extended data scale. We used this concept at the SFFB (see section 4.1.2) for differential pressure measurements. Thus, we recommend the use of existing guidelines to develop measurement data sets.

6.2 Use more thorough sensor calibration

We found several indications that sensor calibration is insufficient in practice during commissioning and operation. The only case study where sensors were calibrated regularly was the SFFB case study, where the solar sensor shading band needed adjustment once per month and the pressure sensor was calibrated once per month. The insufficient calibration is illustrated by one example where we found that the electrical sub-meters were off by a factor of eight after these sensors had been operating for about a year. Any performed calibration effort should have identified this difference. Thus, we recommend that sensor calibration needs to be more thorough during the commissioning or operation phase of the building.

6.3 Design control systems that support continuous data archiving

For the three case studies (except for the custom solution needed at SFFB), we used an existing control system as the basis for data collection. It became apparent that those control systems, including hardware, software, bandwidth, and interfaces, are not well suited for continuous and reliable data collection of all data points at one-minute intervals. Various studies (Brambley et al. 2005; Torcellini et al. 2006; O'Donnell 2009) report similar problems with data collection.

Bandwidth limits were a problem in all three control systems. The lack of bandwidth made it difficult to accommodate the communication necessary to archive all measurements on a one-minute basis. This bandwidth problem is clearly an issue that originates in the control system design. The effort to increase bandwidth at the design stage is minimal compared to that of retrofitting a building control system to increase the bandwidth (e.g., at Y2E2 the costs of initially increasing the bandwidth from 78 kbit/s to 1.2 Mbit/s for the subnets and 1.2 Mbit/s to 100 Mbit/s for the backbone would have been several thousand dollars, but the retrofit costs totaled several hundreds thousand dollars). The latter is a major remodeling effort due to the need for rewiring.

Additionally, we encountered a number of problems with installed hardware that provided significant limitations or led to data loss. For example, the modem at the GEB limited the data collection quite dramatically (see section 4.2.3). Other examples include the iLON server’s inability to respond to requests while having full data queues (see section 4.3.3) or the use of various sensor products that did not perform as needed. The data storage capacity is another example of hardware limitations. We encountered problems with exceeding storage capacities and data loss due to overridden data files.

A range of software limitations and problems reduced the reliability of the data collection. Instability of operating systems (e.g., see section 4.1.3) and unresponsiveness of controller and iLON server software led to data losses. In addition, limitations in the functionality of data collection software increased our efforts to configure these systems and created a need for additional software development.

Another problem with the data acquisition systems is the interface between different types of control systems. We encountered interface problems with two case studies, as described in the previous sections (4.2.3 and 4.3.3). While communication between different control systems is needed to allow for a central data acquisition system, it is important that integrating these systems does not reduce the quality and reliability of data. For example, to ensure the data acquisition of set points and other control variables, one should integrate data acquisition systems with the control systems. A missing integration leads to the use of separate systems to archive data with corresponding challenges to synchronize the data. A tight integration of the control system with the data acquisition system also allows better understanding of the control system (Torcellini et al. 2006). A completely separate data acquisition system may not be able to archive existing sensor and control signal data.

These bandwidth, hardware, software, and interface limitations create an environment in which reliable continuous data collection is hard to achieve. Torcellini et al. (2006) report similar problems with the reliability of data archiving with control systems. Thus, we recommend increasing the bandwidth of control systems early in design to eliminate later problems with data acquisition, as well as selecting appropriate data acquisition software that is capable of archiving all data points at one-minute intervals. Data should be archived in a common database format with at least a timestamp and value for each sensor reading. While current control systems have significant shortcomings in archiving data in an efficient, continuous, and reliable manner, control system developers should improve these capabilities in the future.

6.4 Use local solar data measurements

The SFFB case study indicates that local solar data measurements are an important input parameter for performance comparison. We collected solar measurements for the SFFB at a different building in San Francisco. One pyranometer measured the total radiation, whereas the second one with an attached shading band measured a reduced diffuse radiation. Our data analysis indicated some uncertainty in the measured solar data. One observation from the two different solar measurement sensors we installed at Stanford and in San Jose (20 miles away) showed that the peak total solar radiation was about 10% higher in Stanford than in San Jose on a cloudless day. In addition, the measured radiation intensity during the last hour of sunshine was dramatically different for these two locations due to the different hilltop geography west of both locations. These differences within only a 20-mile distance clearly indicate the importance of using local solar measurements. Thus, it is difficult to compare actual performance with simulated performance without local solar data. This is especially true for buildings with large windows and/or sophisticated shading devices.

In addition to the need for local solar measurements, there is also great value in automatically measuring total and diffuse solar radiation. The conversion to direct normal and horizontal diffuse radiation (as needed for the simulation) is straightforward and does not depend on complicated solar models. Sunshine pyranometers automatically determine total and diffuse horizontal radiation without the need to manually adjust a shading band. This drastically reduces maintenance and errors in the solar measurements. Local measurements also allow integration into the local data acquisition system and archiving at one-minute time intervals, which is not available from most weather stations.

7 Limitations and future research

This section summarizes the limitations of our research and describes possible areas for future research. These include the validation of measurement guidelines based on the number of performance problems, the identification of an expanded measurement assumption list to meet future needs, the development of additional case studies with more measurements, and the testing of sensor accuracy.

7.1 Validation of measurement guidelines

As mentioned in section 5.1, the validation of existing measurement guidelines is limited to the number of sensor points and one case study. While this validation provides an indication about the number of problems that could have been identified with the corresponding data set, it does not provide any information about the severity of the problems. Also we performed no cost analysis of the corresponding sensors and did not associate costs with specific problems. A fruitful future research area is the validation of the measurement data sets using additional case studies with more focus on the severity and cost of specific problems. With the severity of performance problems and costs of each sensor a cost benefit analysis is possible to identify the actual value of a sensor. With more data on actual performance problems in buildings, the cost benefit analysis and likelihood of a given problem would provide a well-founded quantitative assessment of the need for particular sensors.

7.2 Identification of an expanded assumption list

The list of assumptions (see section 3.7) is based on a literature review and the four case studies. Future case studies and research may identify more assumptions. The development of new HVAC components, systems, control strategies, data acquisition systems, and sensors may eliminate the need for some measurement assumptions as well as create a need for additional measurement assumptions. While the list of measurement assumptions is likely to change in the future due to the mentioned reasons, the related processes will still be the same.

7.3 Development of case studies with more measurements

While the four mentioned case studies have extended measurements above the typical level, future research could use additional case studies that use even more measurement data points. With sensors becoming more affordable and buildings more complex, additional sensors may provide additional value for performance evaluation. In particular, occupancy sensors or dimming and luminance sensors (via the lighting control system) may provide additional information regarding performance problems in a building. Specifically, more measurements that provide more details about the usage of the building such as occupant counts and window and door positions, would help eliminate some of the unknowns in buildings. Without installing, testing and analyzing additional measurements the value and usefulness of additional measurements may never be known. Thus, case studies that exceed the sensor level of the four mentioned case studies could result in more insights about measuring and analyzing building energy performance.

7.4 Testing of manufacturer accuracy of sensors

As mentioned in section 3.3, we did not question the accuracy of sensors as provided by the manufacturers. Future research could test the validity of claimed sensor accuracy. For example, a ongoing research project at NIST focuses on the issue of sensor accuracy for commissioning and fault detection of HVAC systems (NIST 2010b).

8 Conclusion

The two contributions of this paper are the list of critical measurement assumptions and the comparison and validation of existing guidelines for measurement data sets. Measurement assumptions document limitations of measurement systems according to their three functions (sensing, transmitting, and archiving) and can support the assessor in evaluating the quality of measured data. We compiled a list of critical measurement assumptions, categorized these measurement assumptions according to the three functions of the measurement system, and developed a process to determine performance problems from differences between simulated and measured data. This measurement assumption concept allowed us to deal successfully with limitations we encountered with measurement systems in the four case studies. These measurement assumptions are crucial for assessing measured data and understanding the difference between measured data and the energy performance of the actual building.

The limitations of measurement systems start with the data set of sensors and control points; thus, we compared and validated existing measurement data set guidelines and showed the value of using additional sensors beyond those recommended in guidelines from Gillespie et al. (2007) or Barley et al. (2005). We provided an overview and summary of existing guidelines for measurement data sets, compared them to data sets of the case studies, and validated the data sets based on the known performance problems of one case study. The results indicated that even extended measurement data sets did not capture all sensors that were needed to identify our known set of performance problems. Based on the results, we recommend designing measurement systems at least on the level of detail that Gillespie et al. (2007) describe. Additionally, we recommend paying more attention to possible limitations of measurement systems to decrease limitations and increase the quality of data. Particular conclusions from the case studies include the need for local solar measurements due to possible variations of local climates. We also learned that a key parameter of sensors is their operating range, which can dramatically reduce the quality of the resulting data if it is not correct.

Based on these limitations of measurement systems, we found it difficult to achieve high reliability in a measurement system. Even with upgraded data acquisition systems, the best achievement in terms of average monthly reliability was about 97% of data at one-minute intervals over a time period of 15 months. We see large potential for improvement of data acquisition systems in the future to enable a smoother analysis of measured data.

9 Bibliography

ASHRAE. (2003). *HVAC applications handbook*. Atlanta, GA: ASHRAE Publications.

ASHRAE. (2007). *2007 ASHRAE handbook - Heating, ventilating, and air-conditioning applications (I-P edition)*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

Avery, G. (2002). Do averaging sensors average? *ASHRAE Journal*, *44*(12), 42–43.

Barley, D., M. Deru, S. Pless, and P. Torcellini. (2005). *Procedure for measuring and reporting commercial building energy performance*. Technical Report #550. Golden, CO: National Renewable Energy Laboratory.

Bensouda, N. (2004). Extending and formalizing the energy signature method for calibrating simulations and illustrating with application for three California climates. Master Thesis, College Station, TX: Texas A&M University.

Brambley, M., D. Hansen, P. Haves, D. Holmberg, S. McDonald, K. Roth, and P. Torcellini. (2005). *Advanced sensors and controls for building applications: Market assessment and potential R & D pathways*. Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830. http://apps1.eere.energy.gov/buildings/publications/pdfs/corporate/pnnl-15149_market_assessment.pdf last accessed on March 13, 2010.

Bychkovskiy, V., S. Megerian, D. Estrin, and M. Potkonjak. (2003). A collaborative approach to in-place sensor calibration. *Lecture Notes in Computer Science*, *2634*, 301–316.

Campbell Scientific. (2010). *Campbell Scientific: Dataloggers, data acquisition systems, weather stations*. http://www.campbellsci.com/ last accessed on March 15, 2010.

Carnegie Institution for Science. (2010). *Carnegie Department of Global Ecology*. http://dge.stanford.edu/about/building/ last accessed on March 15, 2010.

County of Santa Clara. (2010). *Main Jail Complex - Correction, Department of (DEP)*. http://www.sccgov.org/portal/site/doc last accessed on March 16, 2010.

Dieck, R. H. 2006. *Measurement uncertainty: methods and applications*. ISA - The Instrumentation, Systems and Automation Society.

Echelon. 2010. i.LON SmarServer - Smart Energy Managers. http://www.echelon.com/products/cis/.

Elleson, J.S., J.S. Haberl, and T.A. Reddy. 2002. Field Monitoring and Data Validation for Evaluating the Performance of Cool Storage Systems. In *ASHRAE Transactions*, 108:1072-1084. Vol. 108. 1. Atlantic City, NY.

Friedman, H., and M.A. Piette. 2001. Comparison of Emerging Diagnostic Tools for Large Commercial HVAC Systems. *Proceedings of the 7th National Conference on Building Commissioning*. 1-13. Cherry Hill, New Jersey: Portland Energy Conservation Inc (PECI).

Gillespie, K.L., P. Haves, R.J. Hitchcock, J.J. Deringer, and K. Kinney. 2007. *A Specifications Guide for Performance Monitoring Systems*. Berkeley, CA.

Graffy, K., J. Lidstone, C. Roberts, B.G. Sprague, J. Wayne, and A. Wolski. (2008). Y2E2: The Jerry Yang and Akiko Yamazaki Environment and Energy Building, Stanford University, California. *The Arup Journal*, *3*, 44-55.

Haak, T., K. Megna, and T. Maile. (2009). *Data manual of Stanford University's Yang and Yamazaki Environment & Energy Building*. Internal document. Stanford, CA: Stanford University.

Hou, Zhijian, Zhiwei Lian, Ye Yao, and Xinjian Yuan. 2006. Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Conversion and Management* 47, no. 15-16 (September): 2479-2490. doi:10.1016/j.enconman.2005.11.010.

Kim, M.J., K. Megna, and T. Maile. (2010). *Data manual of Santa Clara County Main Jail North*. Internal document. Stanford, CA: Stanford University.

Klaassen, C.J. 2001. Installing BAS sensors properly. *HPAC Engineering* (August): 53-55.

Kunz, J., T. Maile, and V. Bazjanac. 2009. Summary of the Energy Analysis of the First Year of the Stanford Jerry Yang & Akiko Yamazaki Environment & Energy (Y2E2) Building. CIFE Stanford University. http://cife.stanford.edu/online.publications/TR183.pdf.

LonMark International. 2010. LonMark International. *LonMark International*. http://www.lonmark.org/.

Maile, T. (2010). *Comparing measured and simulated building energy performance data*. Ph.D. Thesis, Stanford, CA: Stanford University. http://purl.stanford.edu/mk432mk7379 last accessed on August 31, 2010.

Maile, T., V. Bazjanac, M. Fischer, and J. Haymaker. (2010a). *Formalizing approximation, assumptions, and simplifications to document limitations of building energy performance simulation*. Working Paper #126. Stanford, CA: Center for Integrated Facility Engineering, Stanford University.

Maile, T., M. Fischer, and V. Bazjanac. (2010b). *A method to compare measured and simulated building energy performance data*. Working Paper #127. Stanford, CA: Center for Integrated Facility Engineering, Stanford University.

Maile, T., M. Fischer, and R. Huijbregts. 2007. The vision of integrated IP-based building systems. *Journal of Corporate Real Estate* 9, no. 2: 125–137.

McConahey, E., P. Haves, and T. Christ. (2002). The integration of engineering and architecture: A perspective on natural ventilation for the new San Francisco Federal Building. *Proceedings of the ACEEE 2002 Summer Study on Energy Efficiency in Buildings*, *3*, 239-252. Pacific Grove, CA: American Council for an Energy-Efficient Economy (ACEEE).

Mills, E. (2009). *Building commissioning: A golden opportunity for reducing energy costs and greenhouse gas emissions*. Berkeley, CA: Lawrence Berkeley National Laboratory. http://cx.lbl.gov/2009-assessment.html last accessed on April 25, 2010.

Modbus. 2009. The Modbus Organization. http://www.modbus.org/.

National Instruments. (2009). *NI LabVIEW - The software that powers virtual instrumentation - National Instruments*. http://www.ni.com/labview/ last accessed on November 26, 2009.

Neumann, C., and D. Jacob. (2008). *Guidelines for the evaluation of building performance*. Workpackage 3 of Building EQ. Freiburg, Germany: Fraunhofer Institute for Solar Energy Systems. http://www.buildingeq-online.net/fileadmin/user_upload/Results/report_wp3_080229_final.pdf last accessed on March 15, 2010.

NIST. (2010a). *NIST Calibration Program*. http://ts.nist.gov/MeasurementServices/Calibrations/index.cfm last accessed on March 18, 2010.

NIST. (2010b). *BFRL Project: Commissioning, fault detection, and diagnostics of commercial HVAC*. http://www.nist.gov/bfrl/highperformance_buildings/intelligence/com_fault_detect_diag_hvac.cfm last accessed on July 30, 2010.

O'Donnell, J. (2009). *Specification of optimum holistic building environmental and energy performance information to support informed decision making*. Ph.D. Thesis, Department of Civil and Environmental Engineering, University College Cork (UCC), Ireland.

Olken, F., H. Jacobsen, C. McParland, M.A. Piette, and M.F. Anderson. (1998). Object lessons learned from a distributed system for remote building monitoring and operation. *Proceedings of the 13th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*, *33*, 284–295. Montreal, Canada: Association for Computing Machinery, Special Interest Group on Programming Languages (ACM SIGPLAN).

Piette, M.A., S. Kinney, and P. Haves. (2001). Analysis of an information monitoring and diagnostic system to improve building operations. *Energy and Buildings*, *33*(8), 783-791.

Reddy, T., J. Haberl, and J. Elleson. (1999). Engineering uncertainty analysis in the evaluation of energy and cost savings of cooling system alternatives based on field-monitored data. *Proceedings of the ASHRAE Transactions*, *105*, 1047-1057. Seattle, WA: American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE).

Salsbury, T., and R. Diamond. (2000). Performance validation and energy analysis of HVAC systems using simulation. *Energy & Buildings*, *32*(1), 5–17.

Schmidt, F., N. Andresen, and U. Jahn. (2009). Report on energy savings, CO2 reduction and the practicability and cost-benefit of developed tools and ongoing commissioning. Stuttgart, Germany: Ennovatis GmbH.

Scientific Conservation. (2009). The arrival of automated continuous commissioning: How to optimize operational and energy efficiency for commercial facilities. White Paper. Berkeley, CA: Scientific Conservation.

Seidl, R. (2006). Trend analysis for commissioning. *ASHRAE Journal*, *48*(1), 34-43.

Sensor. (2010). In Merriam Webster's online dictionary. http://www.merriam-webster.com/ last accessed on July 19, 2010.

Soebarto, V., and L. Degelman. (1996). *A calibration methodology for retrofit projects using short-term monitoring and disaggregated energy use data*. Technical Report #96. Energy Systems Laboratory. http://handle.tamu.edu/1969.1/6667 last accessed on October 26, 2009.

Torcellini, P., S. Pless, M. Deru, B. Griffith, N. Long, and R. Judkoff. (2006). *Lessons learned from case studies of six high-performance buildings*. Technical Report #550. Golden, CO: National Renewable Energy Laboratory.

Tridium. (2009). *NiagaraAX Tridium*. http://www.tridium.com/cs/products_/_services/niagaraax last accessed on November 26, 2009.

Underwood, C.P. (1999). *HVAC control systems: Modelling, analysis and design*. London: Taylor & Francis Group.

Yin, R.K. (2003). *Case study research: Design and methods*. 3rd ed. Thousand Oaks, CA: Sage Publications.