

A Prototype Data Archive for the PIER “Thermal Distribution Systems in Commercial Buildings” Project

Rick Diamond, Craig Wray, Brian Smith,
Darryl Dickerhoff, Nance Matson, and Skylar Cox
Indoor Environment Department
Lawrence Berkeley National Laboratory

Task Order No. 9 for LBNL Technical Assistance for PIER

CIEE Blanket Award Number C-02-01

LBNL Report Number 54191

January 2004

Summary: A prototype archive for a selection of building energy data on thermal distribution systems in commercial buildings was developed and pilot tested. While the pilot demonstrated the successful development of the data archive prototype, several questions remain about the usefulness of such an archive. Specifically, questions on the audience, frequency of use, maintenance, and updating of the archive would need to be addressed before this prototype is taken to the next level.

Legal Notice

This report was prepared as a result of work sponsored by the California Energy Commission (Commission, Energy Commission). It does not necessarily represent the views of the Commission, its employees, or the State of California. The Commission, the State of California, its employees, contractors, and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the use of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the Commission nor has the Commission passed upon the accuracy or adequacy of this information in this report.

This work was also supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Program, of the US Department of Energy under Contract No. DE-AC03-76SF00098.

Background and Purpose

The PIER “Thermal Distribution Systems in Commercial Buildings” (TDSCB) project (Diamond et al. 2003) collected a considerable amount of unique data to characterize the energy performance of an HVAC system in a large commercial building. The work reported here describes the development of a prototype web-based public-domain archive that extends the value of these data by making them available to other researchers to support a variety of analyses. The prototype serves as a model for a potential future California Energy Data Archive (CalEDA) that provides public web access to data from PIER and related projects.

Characteristics of a Building Energy Data Archive

Several groups have developed data archives in the past. The US Department of Energy supported the “Building Energy Compilation & Analysis (BECA) reports in the 1980s and others have developed data archives, both on-line and electronic (web-based) as well. The Florida Solar Energy Center (FSEC) has an excellent on-line energy data base at:

<http://www.infomonitors.com/>

An ideal data archive has the following characteristics:

1. Value.
2. Ease of Use.
3. Expandability.
4. Reliance on readily available software and hardware.
5. Ease of Maintenance.

Each of these characteristics is described below.

Value: Why a data archive for building data?

Collecting monitored field data is an enormously labor-intensive exercise. Field data collection requires careful design and execution of monitoring protocols, and substantial logistical efforts to work with the players involved in the operation and maintenance of the building. Consequently, field data are extremely expensive to collect, and efforts to do so are quite rare. As an alternative, most energy analysts rely on simulations of building performance.

The primary benefit of a data archive is that it allows multiple individuals access to data that can be used to support analyses of similar building systems and to validate building simulation tools. A secondary benefit is that the archive preserves an historic record of work that can be further explored at later times.

Ease of Use: Will it be used?

A data archive can have valuable data, but it will be of little use if it is difficult for the user to access or understand the data and be confident that the data are of the quality stated. Features that can facilitate ease of use include a simple structure that shows what data are available, and clear procedures for defining time periods for selecting data. Background information such as descriptive information (pictures and text) can help the user understand the building systems from which the data were sampled, how the data were collected, and the accuracy of the measurements. This data structure, or “metadata” is key to a well-designed data archive.

Expandability: How easy is it to add new data?

A data archive that only contains a small set of data is of limited value. Like a library, an archive needs to be designed to store new data from additional time periods, additional building systems, and additional buildings, as these data become available.

Ideally, the archive would allow data to be imported from other data sources that use data formats different from the archive. A protocol for defining the import process for each supported data structure would be included on the site so that others could submit data for inclusion in the archive.

Reliance on existing software and hardware: How much will it cost?

The main reason to use existing software and hardware is to ensure functionality at a minimal cost. If suitable products are already available, there should be no need to develop new ones specifically to support a data archive. If products don't exist, then existing ones should be modified or new ones should be developed for these needs.

Maintenance: Will it become obsolete?

Ideally, the data archive will require little maintenance. Users will provide new data following defined protocols, and the data will be accessible in the same manner as are all other data in the archive. However, just as libraries need librarians to order new books, catalogue them, store them, check them out, replace them when damaged, and answer users' questions, a data archive will also need an archivist who can maintain and update the archive to ensure its continued use and value.

Software and Hardware Selection

One of the five characteristics of an ideal data archive is the need for selecting appropriate software and hardware to ensure functionality at a minimal cost.

The following issues were identified in the software and hardware selection.

Many computers (e.g., Windows 2000-based, Windows XP-based, Linux-based) could be used for the data archive interface, as long as the following are installed:

1. A Web server such as Apache, iPlanet, Microsoft's IIS, or Zeus.
2. A computer language for creating the pages of the database and assembling data for downloading
3. An Internet connection.

In deciding which software to use for the data archive, the primary goal was to use off-the-shelf products that are freely available, commonly used, or likely to come with the computer hosting the data archive.

With that in mind, the clear choice to handle the CGI (common gateway interface) used to process information specified by the user from the page was to use the Perl scripting language. Perl is bundled with most Linux and Unix systems, and is freely available to install on a Windows-based computer. JavaScript was the clear choice to enhance the limited capabilities of the HTML choice buttons. JavaScript is freely available and is available in most browsers.

To generate images for use in the archive (e.g., the building floor plan), any image editing software may be used. Examples include proprietary products such as Jasc PaintShop Pro and Adobe PhotoShop, and freely-available products such as Gimp and xfig.

For the data archive, we used both xfig and PaintShop Pro. Because the floor plans of the building are available in AutoCAD format, we exported the areas of interest as bitmap images to xfig so that image maps and annotations could be added. It is useful to use an image editor such as xfig that can create "image maps". Image maps describe the areas (e.g., polygon, circle) that when clicked on by the mouse will follow a URL link. The image editing programs mentioned above all have this capability, although PaintShop Pro doesn't have a way to save the information to easily recreate an image map if the original image is modified.

Data Selection

One of the issues for developing a prototype data archive is to identify a set of useful data.

For this project, we considered several criteria:

1. Completeness (i.e., a lack of missing data).
2. Accuracy (confidence that the sensors were well calibrated and correctly installed).
3. Value (the data represented unique and useful information).
4. Representative of different system operating conditions.
5. Ease of documentation.

Using these criteria, we selected specific subsets of the data collected by Diamond et al. (2003) for inclusion in the prototype archive.

Prototype Data Archive

The prototype archive is located at the following website: <http://epb4.lbl.gov/CalEDA/>

This archive is divided into five different sections that include:

1. Data locators, descriptors, and selection for downloading.
2. Descriptions of buildings and systems.
3. Data collection methods, sensor placements, and calibrations.
4. Protocols and tests for system performance.
5. Photos and other images of the buildings and systems.

The archive contains data from one representative time period: June 2, 2002 to August 3, 2002. Data are included for selected parts of the air handling system in the following four zones:

1. Northeast zone.
2. Southeast zone.
3. West mechanical room.
4. East mechanical room.

Data selections that are available in the archive include air temperatures, relative humidity, airflows, air pressures, damper positions, fan power, and heater power in both Metric and English units. Depending on the zone selected, the specific data available include:

1. Return air temperature and relative humidity at air-handler inlet.
2. Outdoor air temperature and relative humidity at air-handler inlet.
3. Return airflow entering air-handler.
4. Outdoor airflows entering air-handler (minimum and total).
5. Return and outdoor air damper positions.
6. Air-handler mixing box air temperature and relative humidity.
7. Cooling coil outlet air temperature.
8. Supply air temperature and relative humidity at the air-handler.
9. Airflow through supply fan.
10. Supply fan power (electric).
11. Static pressure in the main duct.
12. Supply airflow and static pressure entering VAV box from main duct.
13. Supply air temperature upstream of VAV boxes.
14. VAV box inlet supply air temperatures (four points on a grid).
15. Ceiling plenum air temperature at VAV box induction fan entrances.
16. VAV box heater power.
17. VAV box induction fan power.
18. Supply air temperature and static pressure downstream of VAV boxes.
19. Controlled duct leakage airflow downstream of VAV box.
20. Supply air temperature at supply grille farthest from VAV box.
21. Room air temperature (thermostat reading).
22. Relief fan power.
23. Indoor-outdoor static pressure difference.

As an example of the data available in a zone, Figure 1 shows the sensor layout for one mechanical room.

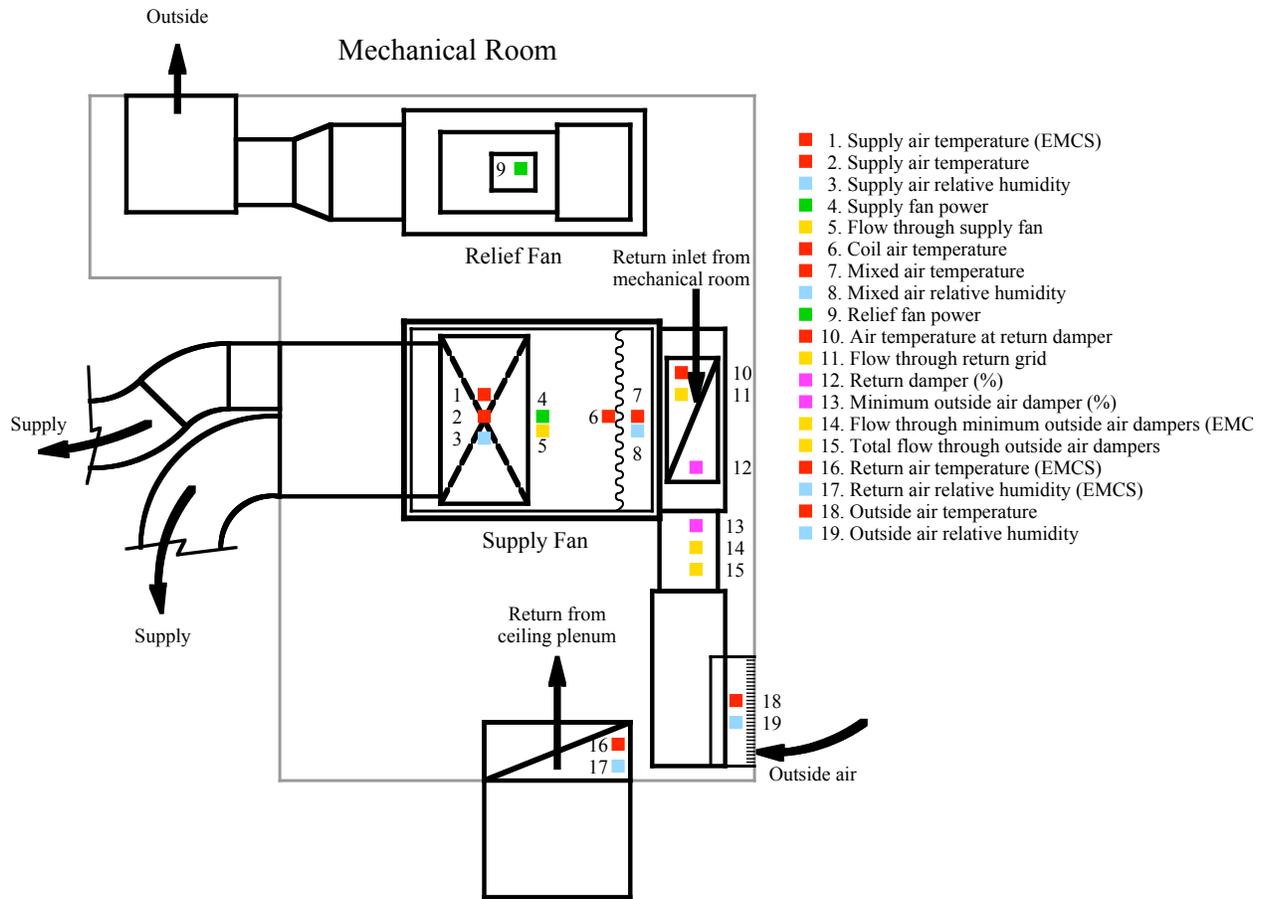


Figure 1. Sensor layout for a mechanical room.

All data were collected using a one-minute time interval. Many separate data loggers were used to collect the data, which means that some data have different time references than others. To facilitate analyses using these data, we carried out a concerted effort to correct the small discrepancies in the timing of the data streams so that all the data have the same time reference. An example of this correction is shown in Figures 2 and 3.

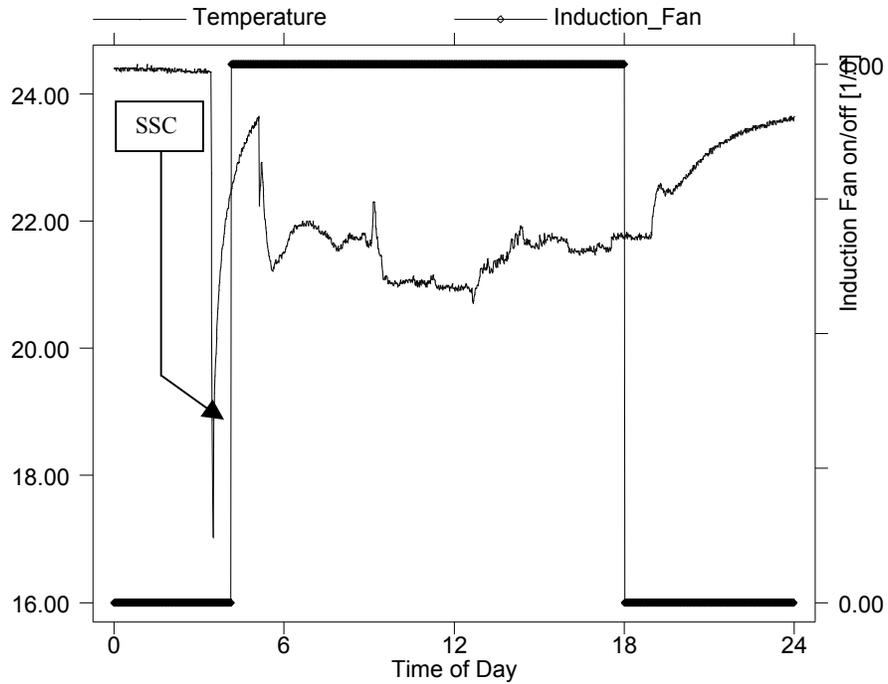


Figure 2. This plot illustrates the time shift in the data due to the two different data collecting system clocks. Notice the low spike on the left side of the graph; this spike is due to the supply fan turning on during the early morning smoke-control system check (SSC).

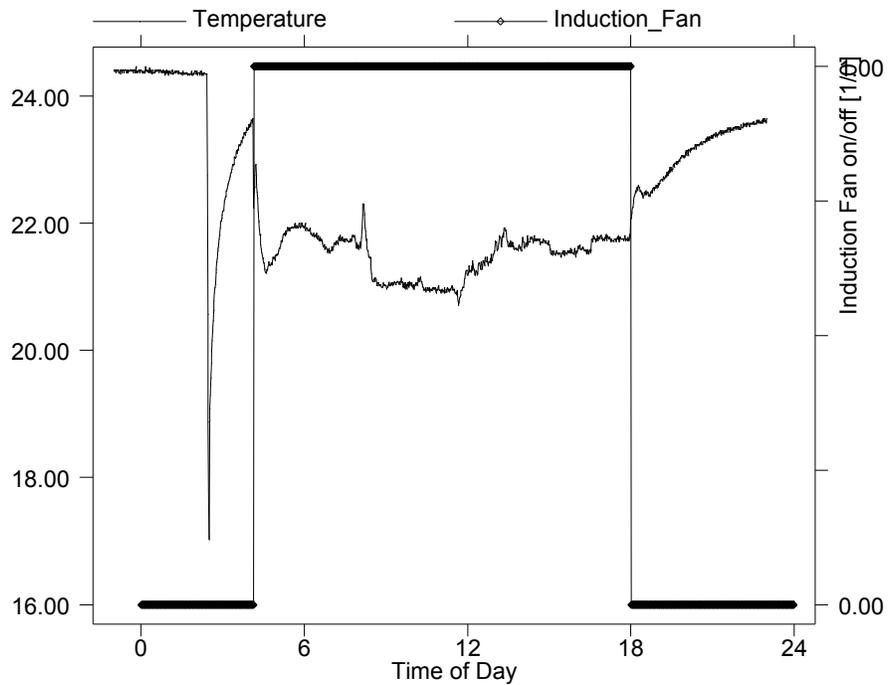


Figure 3. This plot shows the same data shown in Figure 2 after the correct time shift has been applied to the data. The plots represent one day from midnight to midnight.

Conclusions & Next Steps

The prototype data archive that we created is a successful demonstration of the proposed concept. Users reported that they found it easy to navigate and to download data.

One user had several suggestions on how to improve future versions of the archive:

1. Data downloaded from the archive should be provided to the user as two separate files:
 - a. A comma-separated-value (CSV) file that serves as a legend for the points selected.
 - b. A CSV file that only contains the data of interest.
2. Data file names generated by the archive should never include the period character ".". Using this character causes problems on some computer platforms.
3. There should be an option to select individual points, rather than only groups of data.
4. The archive should provide the capability of previewing the data in graphical form. Data plots could allow two to four data selections to be graphed on a 24 hour basis, with navigation buttons to scroll through it if needed.
5. All photographs and schematics need detailed captions to enable the user to understand what is being shown.

Several questions still need to be answered before the prototype should be taken to the next level. These include:

1. Would other researchers and practitioners make use of the data available in the archive? Related to this question are the issues of whether analysts would both understand the meaning of the data and trust the accuracy of the measurements?
2. Would new data be added to make the archive an on-going activity? The archive was designed so that new data could be easily added, but the cost of maintaining the archive may be prohibitive.
3. Who should operate the data archive? Data cleaning, entry and maintenance, is a labor-intensive activity. Availability of students for this work would make universities a natural choice for such a role.

Ultimately, the decision that needs to be made is whether the benefits of developing and maintaining the data archive outweigh the costs. Although the costs of collecting building data are high, the continued demand for the use of these data remains unknown.

Acknowledgments

We appreciate the thoughtful comments from Mark Hydeman (Taylor Engineering) on the prototype website, and the reviewers of the report, Max Sherman and Richard Brown (LBNL).

References

Diamond, R., C. Wray, D. Dickerhoff, N. Matson, and D. Wang. 2003. "Thermal distribution systems in commercial buildings," Lawrence Berkeley National Laboratory Report, LBNL-51860.