Performance Metric Workshop for High-Performance Commercial Buildings
Vail Cascade Hotel
Vail, Colorado
November 5-6, 2001
Version 1.0

Purpose:
The purpose of the workshop was to initiate a work plan to create a comprehensive set of Commercial Building Performance Metrics. The long-term goal of this work plan is to create a building performance framework and identify metrics related to this area.

Introduction:
The High-Performance Commercial Buildings Technology Roadmap identified a need to create a method for evaluating building performance. To begin addressing this need, the U.S. Department of Energy (DOE) and the National Renewable Energy Laboratory (NREL) held a two-day Commercial Buildings Performance Metrics Workshop to discuss the issues. Thirty experts in building energy performance as well as human health in the built environment, psychological impacts of the built environment, and building economics attended the workshop. Those attending the workshop represented both private- and public-sector entities in the United States and Canada.

The workshop is the continuation of a one-year effort than began in 1998. In response to the need identified in the Roadmap, a small group of experts met mostly through conference calls and e-mail exchanges to discuss meeting the Roadmap performance evaluation objective. This first group went as far as to produce a draft workplan. To provide continuity, there was overlap in the participants between the two groups. The results of the Vail workshop will try to incorporate the important ideas generated by this previous effort and move forward in the directions identified as being essential.

This report summarizes the activities and principal results of the workshop. Some synthesis and gathering of the ideas is presented below based on input from the participants. The agenda for the meeting is including in Appendix A and the list of participants is included in Appendix B. Paul Torcellini and Sheila Hayter of the National Renewable Energy Laboratory assembled this report. Comments are welcome and should be addressed to sheila_hayter@nrel.gov. Comments can be sent to the entire participant list by addressing metrics@mail.nrel.gov. This report and related materials can be found at www.nrel.gov/buildings/highperformance/metrics.

Synopsis of Discussion:

The Workshop began with a discussion about the **purpose** of performance metrics. Metrics create a yardstick by which building performance can be *benchmarked*. The benchmarking allows for comparisons between buildings of different types, of different designs, in varying locations, and having varying occupant use patterns. Benchmarking also helps identify best practices and helps reveal where good design intentions may not result in optimal performance. Performance metrics help building designers design better

buildings; help building owners know when their buildings are or are not performing as expected; and help building operators learn when operating changes are needed to improve performance. Workshop participants also anticipated that performance metrics will eventually encourage the manufacture of new products that use natural resources more wisely. These new products will be used both to construct new buildings and by occupants within buildings. Workshop participants agreed that ultimately, buildings should give back to the environment more than they take from the environment.

Following the discussion the purpose of performance metrics, there was some discussion about audience and metric characterization, but the primary focus was on anticipated workshop results as well as lessons learned from existing measurement tools. The group decided that a framework to describe building performance is the most important near-term outcome of DOE's performance metrics activities. The framework should be flexible enough to include building energy performance while reflecting priorities from other organizations such as human health, environmental health, productivity, and economics. It should also accept contributions and reflect the priorities of individuals or organizations representing a wide variety of differing interests. For example, a framework that works well for describing building energy metrics should also be broad enough to guide evaluation of human productivity. The group also confirmed that metrics means accurate quantitative measures.

The issue of performance *metrics* is complicated and takes form and meaning on many different levels. A pyramid illustration was presented as a starting point (see Figure 1) to show a potential hierarchy of different metrics levels from simple to more complex. The discussions revealed several important needs: the necessity of clearly defined and consistent terms; the ability to look at the interactions of data or combine data to form more complex metrics at different levels in the hierarchy.

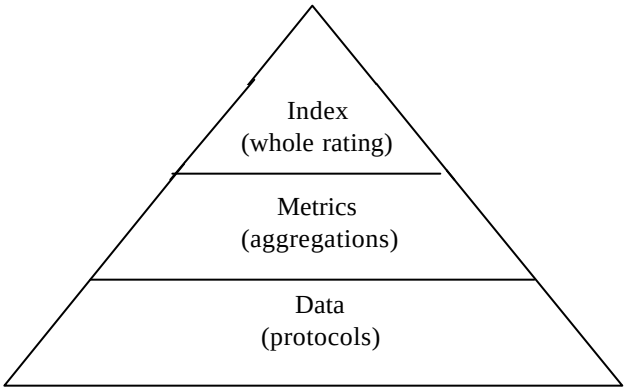

Figure 1: Hierarchy of levels for performance metrics.

The group formulated some operating definitions. There are raw *data* points that can be collected about the building. The data requires *protocols* or specifications of the collection methods and reporting methods so that data are consistent. Consistency ensures building comparisons can eventually be done on a common basis. The data can

be combined and normalized to form additional metrics. Information obtained in both of these levels can be combined, depending on audience, to form a *rating* or *index* for the building. There might also be a top level to the pyramid that eventually will be able to assemble the ratings.

In discussing the larger issue of performance metrics framework, workshop attendees expressed that some aspects of measurement were well understood and could be gathered and manipulated (like energy performance). Other issues, like human factors, are less understood and presently difficult to characterize. A comprehensive framework would need to include both the "easy pickings" and the more challenging, less understood areas.

If the difficult things to measure are truly important, then we need to do the research to understand them and try to evaluate them. We could start by collecting information on these types of data and what is being done. Participants noted that there is much work being done but there are many questions. One question is access. The private sector has a great deal of proprietary activity underway—the problem is trying to publish some of this data. Another potential issue is comparability of the information. Examples of this proprietary information include productivity data, company data on absenteeism, and insurance industry loss and risk information. With any comprehensive effort there is a concern on the proprietary nature of the data. Some of this information may be able to be adapted for public sector use for a price.

There are also concerns about the measurement techniques employed to gather data sets. Can we determine the accuracy of the data—is it transparent and the *protocols* known and well established. Consistency becomes a key issue in trying to do comparisons. Nevertheless, the group agreed that this seems an appropriate time to start with the development of the metrics and protocols to measure them. The information can be maintained in common public databases. The combination of the data into rating systems is complex, but that could be done later or by user groups wanting to aggregate the information to suit there specific needs.

It is important to collect the data in a consistent manner. At higher levels of aggregation, users can combine data, apply the weighing factors to apply cost, environmental loading, etc. Given the availability of quality, protocol-based data, user groups would have the ability to combine and customize data to serve their needs.

Research will be required to develop accurate methods for measuring factors that are not currently as well studied as building energy performance, such as occupant productivity, and relating theses factors to building performance. A first step to developing these methods is to determine and evaluate current practices that address these factors. Workshop participants provided anecdotal information of work underway in the private sector and related how difficult it can be to get this information. Often, private-sector entities collecting this information consider the information to be proprietary and are unwilling to share the information and methods of obtaining it. A database of existing research is needed as an early outcome.

Current data allows easy performance characterization of individual buildings, especially in energy performance. It becomes more difficult when these characterizations are used to make comparisons between buildings. Part of the complexity relates to different building types and uses. This is a key issue in comparing building energy performance. However, there are unique problems associated with normalizing occupancy, weather, building use patterns, and other features that are not easily measured.

Further discussion confirmed that a metric only has meaning when it is compared against something. This raises the issue of the need for a base case and how to do the comparison. The group anticipated it could be challenging to create a base cases. This is both a method and a statistic.

The group focused on approaches to building evaluation as the first data to gather. For example, one useful construct is to evaluate the whole building first, then move to investigate systems level issues and then components. Using this model may help to prioritize which metrics to study first and how to start at the highest level and drill down to the lowest level in order to effectively evaluation a building.

A significant amount of work has been done on developing *tools* to help rate the performance of the building. Examples of these tools include the U.S. Green Building Council's (USGBC) LEED rating system, the Green Building Challenge (GBC) assessment, and others such as BREEAM and the AIA Committee on the Environment Top Ten Buildings application database. All agreed that these tools support the Roadmap objectives in some way, but that all have gaps preventing them from being able to comprehensively describe and benchmark a building's performance.

LEED has been successful in starting to penetrate and change the market. For example, the LEED criteria can stimulate interest and guide design teams in high-performance buildings. The tool creates performance criteria and evaluates many aspects of the building design, but does not measure the actual performance of the building. Finally, LEED is not a consensus-based tool, so the building industry as a whole does not embrace LEED. LEED is a good staring point for an important sub-set of the buildings audience. It works in the design world, but is not a tool used to measure financial or operational performance.

Proxies are sometimes used as substitute measures in the rating systems. That is, something is measured because the actual "metric" cannot be measured. An example is the desire to measure indoor air quality (IAQ), but Carbon Dioxide and ventilation rates are measured and conclusions are drawn about IAQ based on those proxy measurements.

The extended discussion of LEED and other existing tools guided workshop participants to identify that there are two important pieces to performance metrics. The first piece is evaluating a building design, such as is the direction of LEED. The other piece is to evaluate the building operation, such as is the direction the GBC database. Participants also began to differentiate or categorize the areas in which a building could be evaluated during both the design and operation. For the purposes of the discussion at this

workshop, it was decided that prospective metrics for the spectrum of building design and operation performances could be divided into five areas. These areas are: 1) resource consumption and environmental loading, excluding energy and human factors; 2) resource consumption and environmental loading of energy consumption; 3) human factors; 4) economics; and 5) service quality and other. It is important to note that these five areas were created as a tool to facilitate discussion during the workshop and are not necessarily areas among which performance metrics will be eventually categorized.

The workshop participants then brainstormed a list of possible data or metrics ideas that fit into each of the five (5) categories for both the building design and operation. Participants divided themselves into working groups to focus on each of the 5 areas. The work of the subgroups was to try to describe an objective or objectives in category measurement, what should be measured, and how to measure data.

Each breakout group reported its major findings. In the discussion that followed the reports, the group raised fundamental definitional issues: how far should performance metrics be developed? We can agree that they should be a useful, widely accepted method of evaluation. The question is, how to achieve this.

In an earlier discussion, the group described many possible target audiences/users with differing needs. One challenge is how to balance those needs or desired characteristics to serve those many prospective users. One useful analogy was distinguishing the "signal from the noise." When will the signal be strong enough that the noise will no longer be a distraction? For example, as stated earlier in this report, successful inroads have been made to measure building energy performance, but methods for measuring building occupant productivity and relating that productivity to building design and operation decisions are not currently established. At this point, it is extremely difficult to quantify measures such as productivity, leading to challenges in determining total building performance.

Traditionally, DOE's role has involved working on the underlying science and not the market application. Doing the research to support the *standards*—but not the writing of standards. DOE would provide the data and the models. This group can use that data and market and deploy this information as appropriate. In some ways, standards are minimum benchmarks.

Reporting of Groups

Workshop attendees chose five major categories as possible areas for measurement: site, water, energy, materials, and IEQ (indoor environmental quality). Within these categories, five areas were identified as existing areas of concern and breakout groups were formed around these areas: a non-energy resource consumption and environmental loading group; an energy-related environmental loading group; a human factors group; an economics group, and a group for service quality and other factors.

Group 1. Resource consumption and environmental loading, excluding energy and human factors

Participants: Sandy Mendler, . Joel Todd, Sheila Hayter, Bruce Hunn, Otto van Geet, Wayne Trusty

Discussion:

It would be useful to have a comprehensive database that shows all the material flows in and out of nature by each transformation process. We have relationships between source/site energy. We have good information about CO_2 and emissions from combustion. The piece that is missing is the impacts from mining, building a power plant, including water flows, etc. What is the ecological footprint of the building—it is bigger than the footprint of the building?

There was discussion as to where the system boundary is located. If you only consider the design and operation of the building, then there is no consideration made for the industrial processes or material uses by the occupants in the building. If we are talking about high-performance buildings, then what applies to the building? Does productivity work here?

If we are looking at methods for normalization, then one should look at the functional units of the building in question. If the functional unit is a restaurant, then the fundamental unit relates to food served. If the functional unit is education, then number of student contact hours is the important fundamental unit.

Group 2. Resource consumption and environmental loading of energy consumption

Participants: Ron Judkoff, Mike Brambley, Dale Sartor, Mary Ann Piette.

Discussion:

This group distinguished between a set of data and an attempt to normalize the information. A distinction between estimated and measured data was established. The concept of a "reference building" for purposes of different kinds of comparisons was introduced. The reference building could either be from statistical information, or derived from a real building based on a set of modeling rules.

For building performance metrics, normalization parameters used in the equation energy/X can also be in the form of energy cost/X. In other words, energy cost can serve as a surrogate for source energy, which includes environmental concerns. When source energy concerns such as environmental impact from power plants, etc. are included in energy cost, it can speak to both the owner and the environmental greater good.

Normalized information included topics such as normalized site energy consumption, normalized source energy consumption, energy cost, and environmental impact. Normalization can be made with respect to people, weather, unit area, or product produced. This is similar to the discussion on defining the building block or purpose of the building. One conclusion is that different user groups have different needs for normalization.

The group created a non-comprehensive list of normalized metrics. Each of these requires a set of data to create them. This model supports the pyramid model for performance metrics.

The group also determined that a diagnostic set of metrics can be used for evaluating sub-system performance that helps determine how well various building systems are operating. Is the building good/bad on a number of dimensions? These become quick-checks to do a preliminary quantification: KW/ton, W/cfm. All of these could/would have a reference point (similar to the concept of the reference building) to determine if something is good/bad.

To create metrics and base cases, we will need simulations and will need this to normalize the information. We have many different base cases. Data needs: building characteristics that are much more granular (detail about the buildings) Engineering details. Calibrated models for the buildings.

We need a common database for precombustion effects for using energy, otherwise, benchmarking does not make sense. Some of these data exist. The big problem is consistency and making these data publicly available.

Different constituencies have different needs and interests. Owners tend to be interested in evaluating the building at the site. This has to do with where the meter is and the financial incentive/burden at this point. Researchers are interested in source numbers for energy because of global impacts.

To change people's thinking on energy efficiency, you not only need compelling economic justification, but relate to other values in the decision making process.

Group 3: Human Factors:
Participants: Merle McBride; Mark Mendell, Judi Heerwagen, Jim Hill, Rick Fedrizzi.

This group looked at the metrics that effect human outcomes as a result of the building. It is quite often necessary to use proxies to substitute for actual metrics. The result is having something that you know how to measure, but it is really not measuring what you want to measure. For example, measuring outside air rates to quantify "clean" air. CO_2 to measure amount of OA introduced and control it to relate back to air quality.

The group produced a framework for defining types of quantities and interrelationship between these quantities. Many of the measurements from this group need to be strongly related to good/best practices based on our current understanding of Indoor Environmental Quality (IEQ).

This group made a distinction between the outcomes that they want and how close they could come to measuring that outcome. A Real Outcome is a direct measurement by objective means. Related factors become direct measurement by subjective means.

Another concept from this group was loss prevention related to disasters—this is similar to the work that we were doing on using energy efficiency and renewable technologies to minimize disaster issues. The insurance industry has some of this data, perhaps by building type.

Many of these areas are very difficult to quantify and we don't understand what data to collect. There are severe limitations related to this. On productivity, we are able to quantify the negative aspects of productivity, but we don't know the links to develop the positive aspects of productivity.

It is important to recognize that productivity has multiple possible data collection points that have a complex interrelationship. It may be that productivity is captured or arrayed at a higher level on the framework pyramid because these interrelationships become complex and controversial when establishing interrelationships.

There is very little research on the quality of lighting relating to human satisfaction. IESNA does recognize this and is trying to get work going on this effort. As a starting point, the Center for the Built Environment (CEC PEER) has research on lighting.

The human factors area is also complex because it is difficult to have real time measurements related to this area.

Some of the metrics become thresholds while some need to be measured along a continuum. We need to know the relationship of how much/little pollution affect health and productivity, not just a threshold level. For example, the measurements that are made for IAQ (Indoor Air Quality) have little impact before understanding the actual health impacts. As another example, many buildings have Legionella, but there is a threshold for it before it becomes a health risk.

Group 4: Economics:
Participants: David Hanson (DOE); Ellen Franconi, Mike MacDonald

Discussion:
It is important to make the economic case for high-performance buildings. There are many relationships between the human factors and the economics. They give a starting point, but not a comprehensive set of measurements.

We need to define the mission of a building to associate "value" with it.

Examples: total cash flow through facility divided by the total cost of operations (annually); cost vs. value of education of kids.

Some of the data needs include: capital cost information, measurement of asset value, total cash flows (difficult), O&M costs, churn costs, leasing implications, contributions to mission objectives, productivity and performance, percent rented space, percent rented.

Many of the metrics are based on a single point of time and need to be extracted over some period of time such as the life of the building. Future financial impacts such as medical insurance, liability issues need to be considered. Profitability also fits into this category.

There are currently methods to calculate the cost of turn-over. What components are building related and which components have nothing to do with the building.

Those that need to be convinced are awarded and can quantify the payoff.

There is also a trade off between the energy-efficiency aspects and other values of the decision making process. The energy cost is a small part of the overall capital and operational costs of the building. However, it can sometimes be the largest impact that can be affected in the building. Another analogy is that a mutual fund sells an under-performing stock even though it is a small piece of the total portfolio.

The credibility of data and resulting assessments of possible impacts is a function of perceived risk by building owners. Risk is the bigger word in determining the importance of data. It determines whether an owner takes time to assess the credibility of any measure or data stream. What motivates someone from deviating from their standard practice? Return on investment is one factor, although there are many more intangible "soft" factors. It is important to separate the soft numbers from the hard numbers. This is what creates the analytic basis and motivation to deviate from standard practice.

Research is needed on how to get over market barriers. One example would be a predetermined budget level or formulas for designing and constructing a high-performance building. This would include measures such as the dollars per square foot of expected costs to complete a project. If a proposed project does not have a certain level of funding from the start, as reflected in the formula, then the building is not likely to achieve a "high performance" level. For example, by the time a design team is assembled, the team may already be in an economic box. We need to investigate and establish ranges and determine parameters for construction budgets (certain constraints within the building disciplines are understood). It should be noted that many important costs may be related to the aesthetic costs. These aesthetic or "image-" related expenses can overwhelm an otherwise "dollar-driven" decision process. An example is that people in Colorado are willing to spend more for electricity options that are green.

Being able to identify the relative importance of different factors is important. Identifying key values and mission critical elements are central to an organization and how it functions. Most organizations do not feel that energy and environmental issues are mission critical. Some of this changed in the last year with the California energy crisis.

Group 5: Service Quality:
Participants: Mark Hydeman, Kristin Heinemeier, Chris Early

Discussion:
What is the quality of the service that is provided by a building. The main purpose is to satisfy the occupants and make it a productive atmosphere.
Most issues here came down to facility management.

1. The goal is a building that is easy to manage and understand. The level of documentation on the building is important—if you have good documentation and training, then you can manage a complicated system. There is a tradeoff between the following three categories. Projected and planned O&M costs. Specific things to measure: (1) salary for maintenance personnel. (2) costs of parts and contracted maintenance. Some existing database that can provide a baseline for this type of information. (3) methods of tracking building occupant satisfaction: occupant surveys, number of calls to maintenance. Give tests to the staff on the knowledge of the building: testing the designers based on how well the maintenance staff can respond as to the aspects of the building.
2. Adaptability of the building—does it serve the mission of the organizations/owners (occupants). Is it adaptable to changes over time. Building flexibility and system flexibility indices. Performance penalties for turn-down. (for unoccupied areas turn-down parts of the building—what is the penalty for doing this.) Building scorecard for building mission. Churn statistics. Level of disruption for renovation (small, lack of access, moving people multiple times).
3. Life Safety: failure statistics. Biohazard, etc.

There are sources of failure data within the insurance industry—catastrophic failure. Lawsuits. Federal Facilities Council has done some statistics on this. There is a BOMA Experience Exchange that deals with why tenants change. Most of this is currently done with user surveys. Real-estate agents also have information about why tenants leave buildings or stay in them.

Next Steps and Action Items:

Item	Who	Date
Create a web-based database for recording reference information	NREL	
Complete draft report of meeting and send out for review	NREL	
Create Template for reporting and flushing out information from the breakout groups	NREL	
Fill out template to complete process from breakout groups	Group Leaders	
Integrate with work plan that was done 2 years ago.	DOE	
Identify champions for various sections and establish accountability.	DOE	
Write Executive Summary describing progress and direction	DOE/NREL	

Appendix A – Meeting Agenda

Preliminary Agenda

Performance Metrics Planning Workshop

Vail Cascade: Vail, Colorado
November 5-6, 2001

Purpose: To Define the Framework, Key Elements and Initiate Planning to Create a Comprehensive Set of Commercial Building Performance Metrics

Monday, November 5th

8:00-8:30 Continental Breakfast

8:30-9:15 Welcomes, Introductions, Purpose of Workshop, Agenda Review

9:15-10:30 Creating a Shared Vision for Building Performance Metrics (BPM)

1. What are BPM intended to accomplish? What will they be used for?
2. What would be the BPM's core/ key characteristics?
3. Who are the users/ target audience?

10:30-10:45 Break

10:45-12:15 Creating a 'Representation' of BPM: What Major Features or Elements of the Landscape (or Broad Outline) are Identifiable at this Time?

II. Are there other useful models? What can we learn from them?
III. Is there an obvious 'architecture' for BPM?
IV. What are the key categories and sub-elements within those categories?

12:15-1:00 Lunch

1:00-1:45 Discussion: Confirming the Categories and Major Sub-elements

1:45-3:00 Small Groups: Fleshing-out the Categories and Major Sub-elements

V. Identify what to measure and, if possible, how to measure
VI. Is there other work already underway that can contribute?
VII. Describe how the work in this category should be undertaken; e.g., which elements to initiate/ develop first

3:00-3:15	Break
3:15-4:00	Report Back From Small Groups VIII. Note key developmental steps and major gaps
4:15-4:30	Preview Day 2, Parked Issues, Clarify Travel Logistics, Evaluate Session
4:30	Adjourn

Performance Metrics Planning Workshop

Vail Cascade: Vail, Colorado

Tuesday, November 6th

8:00-8:30 Continental Breakfast

8:30-9:00 Reflections, Insights from Yesterday

IX. Additions/ changes to architecture/ landscape?

9:00-10:45 Creating a Work Plan

1. How do we launch and sustain this effort?
2. What are the key things to initiate now?
3. Who else might contribute to the effort?
4. Capture key first steps

10:45-11:00 Break

11:00-12:15 Brainstorm Deployment Opportunities

X. How can we convince the industry to use BPM?

XI. Are new capabilities or talents needed?

XII. Are there incentives or other hooks?

12:15-1:00 Lunch

1:00-2:45 Develop a Time Line, Assignments and Roles for the Action Planning

2:45-3:00 Action Items, Next Steps, Parked Issues, Closing Remarks

3:00 Adjourn (Shuttle Departs for Airport)

Appendix B – List of Attendees

Performance Metrics Workshop Participants:

Brambley, Michael R - Pacific Northwest National Laboratory

Crawley, Dru - U.S. Department of Energy

Early, Chris - U.S. Department of Energy

Fedrizzi, Rick - Green-Think

Franconi, Ellen - Schiller Associates

Hansen, David G. - U.S. Department of Energy, EE-41

Hayter, Sheila - National Renewable Energy Laboratory

Heerwagen, Judith - J H Heerwagen & Associates Inc

Heinemeier, Kristin - Brooks Energy and Sustainability Laboratory

Hill, James - National Institute of Standards and Technology

Hull, Julie - Honeywell Laboratories

Hunn, Bruce - American Society of Heating, Refrigerating, and Air-Conditioning

Hydeman, Mark - Taylor Engineering

MacDonald, Michael - Oak Ridge National Laboratory

Sandy Mendler - HOK Architects

McBride, Merle F - Owens Corning

Mendell, Mark - Lawrence Berkeley National Laboratory

Piette, Mary Ann - Lawrence Berkeley National Laboratory

Judkoff, Ron - National Renewable Energy Laboratory

Van Geet, Otto - National Renewable Energy Laboratory

Poole, Lauren - National Renewable Energy Laboratory

Sartor, Dale - Lawrence Berkeley National Laboratory

Todd, Joel Ann - Joel Todd Consulting

Torcellini, Paul - National Renewable Energy Laboratory

Trusty, Wayne - Athena Sustainable Materials Institute

Appendix C – Version Control

This report was assembled by Paul Torcellini (NREL), Dru Crawley (DOE), Sheila Hayter (NREL), and Lauren Poole.
Version 1.0 11/14/01