IMPACT ASSESSMENT

A Guide for City Governments to Estimate the Savings from Energy Benchmarking and Energy Efficiency Programs

Jayson Antonoff

PUTTING DATA TO WORK

TOOL

IMT
INSTITUTE FOR MARKET TRANSFORMATION

This material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), under Award Number DE-EE0007063.

ABOUT IMT

The Institute for Market Transformation (IMT) is a national nonprofit organization focused on increasing energy efficiency in buildings to save money, drive economic growth and job creation, reduce harmful pollution, and tackle climate change. IMT ignites greater investment in energy-efficient buildings through hands-on expert guidance, technical and market research, policy and program development and deployment, and promotion of best practices and knowledge exchange. For more information, visit imt. org.

CONTENTS

PUTTING DATA TO WORK

This resource was developed as part of *Putting Data to Work*, a three-year pilot project aimed at using building performance data and asset information to help efficiency program implementers better target their outreach to building owners and increase the number of projects executed within these programs. The project used building performance data generated by city policies to improve energy efficiency program design and delivery in the District of Columbia and New York City, and developed a toolkit of resources to enable local governments, utilities, and program implementers to learn from activities to replicate successes.

This paper describes how cities are actively using benchmarking data to evaluate the impact of their energy efficiency policies and programs, and includes best practices for other cities to conduct similar analyses.

Introduction

Mandatory programs that require benchmarking of building performance—wherein building owners track their building's energy and/or water use and report the results to local government departments—have become one of the common tools for policymakers striving to reduce energy use within their local building stock. The generally acknowledged goals of these programs are to:

- make individual building owners more aware of their buildings' performance and how it compares to other peer buildings;
- drive market transformation by generating demand for more energy-efficient buildings among potential buyers, tenants, lenders, and others involved in the real estate markets; and
- provide policymakers, program administrators, and researchers with the information needed to design and implement effective energy efficiency initiatives.

Now that these policies have been in place in some U. S. cities for more than six years,[1] the question that is rightfully being asked is whether benchmarking policies are having the intended energy- and carbon-saving impact.

This question is best considered in two parts: understanding how cities are evaluating the impact of their benchmarking programs, and understanding how benchmarking data is being used to measure the effectiveness of other energy efficiency programs. It is critical that cities are able to measure the impact of their efficiency policies. This allows them to prioritize policies and programs in order to most cost-effectively meet city climate goals; demonstrate resulting cost savings to those affected by the mandates; and use proven savings and successes to build additional policies that can glean further savings.

Several approaches have been used or proposed for evaluating the impacts of benchmarking. This paper describes some of the ways that cities are using benchmarking data to evaluate the impact of their benchmarking policies and other energy efficiency policies and programs, and presents best practices that will enable additional cities to conduct similar analyses. Our results are based on a review of annual benchmarking reports published by nine cities,[2] as well as interviews with program administrators in four of these cities—Minneapolis, New York City, Seattle, and Washington, DC.

1 The Clean and Affordable Energy Act was passed in Washington DC in 2008, with the first reporting deadline for buildings 200,000 square feet (sq. ft.) or greater being April 1, 2013. Local Law 84 was passed by the New York City Council on December 9, 2009. The first New York City reporting deadline for non-residential buildings, those of 50,000 sq. ft. or greater, was May 1, 2011.

2 The cities included Boston; Cambridge, MA.; Chicago; Minneapolis; New York; Philadelphia; San Francisco; Seattle; and Washington, DC.

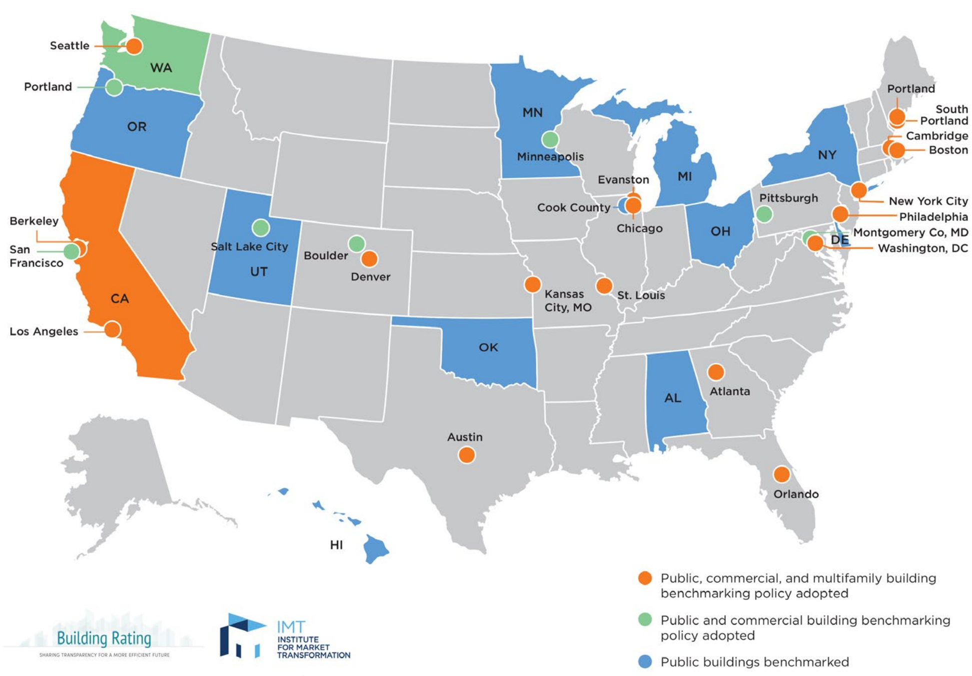
1. Methodology for Analysis

Overview

In a theoretical sense, assessing the impact of a benchmarking and transparency policy or other energy efficiency program can be done with a few simple steps:

- Analyze benchmarking data for the most recent year to determine current performance across the entire population of buildings being assessed. The key metrics often used to evaluate impacts include total energy use, median energy use intensity (EUI, often noted as energy use per square foot), total greenhouse gas (GHG) emissions, median GHG emission level, and median ENERGY STAR score.
- Analyze the benchmarking data to evaluate the same metric(s) for a comparable set of buildings in one or more previous year.
- Determine whether measureable changes have occurred by comparing the results across the years evaluated to develop a trend line for the metric(s) and calculate changes over the time period.
- If possible, identify the main contributing factors for any changes observed.

U.S Building Benchmarking and Transparency Policies



COMMONLY USED METRICS FOR EVALUATING IMPACT

- **Energy use intensity (EUI):** A measure of a building's annual energy use over time, per unit of area. It is typically measured in thousands of BTU per square foot per year (kBtu/sq. ft./yr. or kWh/m2/yr). can be calculated as either site EUI or source EUI. Site EUI is the amount of primary energy consumed at the point of use, while source EUI measures the equivalent units of raw fuel consumed to generate each unit of energy consumed on-site. Source EUI takes into account losses incurred in the production, transmission, and delivery of the energy used at the building site, and is therefore always equal to or greater than site EUI. Additionally, both site EUI and source EUI can be expressed as raw values, or as weather-normalized values. Weather normalization uses conversion factors to adjust for the impact of actual weather conditions and determine the energy a building would have used under average conditions (climate normalization), which can be especially helpful when comparing results across multiple years.
- **Greenhouse gas (GHG) emissions:** GHGs are gases that trap heat in the atmosphere. Major GHGs include carbon dioxide, methane, nitrous oxide, and fluorinated gases. GHG emissions are generally measured in metric tons of carbon dioxide equivalents (CO_2e). Because these GHGs have different impacts, CO_2e allows GHG emissions from gases other than carbon dioxide to be expressed in standard terms of carbon dioxide equivalent, based on their relative global warming potential.
- **ENERGY STAR score:** The 1-to-100 score generated by the U. S. Environmental Protection Agency's ENERGY STAR Portfolio Manager benchmarking tool helps evaluate how well a building is performing. The ENERGY STAR score assesses a building's physical assets, operations, and occupant behavior, and then compares its energy consumption to that of other similar buildings of the same space type, based on a national average. A score of 50 is the median; a score below 50 indicates it performs worse than 50 percent of similar buildings nationwide, while a score above 50 means it performs better than 50 percent of its peers.

Some attempts have also been made to look at early market transformation indicators, to determine if benchmarking is leading to qualitative changes in awareness and practices that will result in quantifiable improvements in building performance in the future. The most notable example is the methodology developed by the U. S. Department of Energy (DOE) in its *Benchmarking & Transparency Policy and Program Impact Evaluation Handbook*[3]. This methodology considers the impact benchmarking and transparency policies can have in areas such as:

3 Navigant Consulting, Inc. and Steven Winter Associates, Inc. "Benchmarking & Transparency Policy and Program Impact Evaluation Handbook" (May 2015), https://energy. gov/sites/prod/files/2015/05/f22/DOE%20Benchmarking%20and%20Transparency%20Policy%20and%20Program%20Impact%20Evaluation%20H. . . . pdf, accessed September 17, 2017.

- market transformation progress
- direct energy impacts
- non-energy impacts (GHG emissions, job creation and economic growth)
- changes in real estate valuations

The methodology was applied in a test case to evaluate the impact of New York City's Local Law 84 (LL84), which requires owners of all buildings in the city over 50,000 square feet to benchmark their performance annually[4]. However, while this approach is attractive for its technically sound and robust methodology, other cities have found it to be too resource-intensive to implement.

Portfolio-Wide EUI

The metric that cities most commonly use to evaluate changes in the energy efficiency of covered buildings is weather-normalized site energy use intensity (EUI), measured in terms of energy use per square foot. Using a weather-normalized factor allows for energy performance over multiple years with varying weather conditions to be compared to one another, as it adjusts for heating and cooling conditions that may affect energy consumption.

Most cities evaluating impact to date have chosen to use the median, or middle value, when evaluating the EUI across a set of buildings. The median is generally lower than the mean (average), as it is less affected by extreme outlier values. It therefore gives a better representation of the energy use of a typical building where there are potential anomalies caused by bad data or atypical buildings.

Though the examples cited in this paper generally refer to median EUI values, reflecting current practice, cities should instead consider using mean EUI values in lieu of the median for future analyses. As measures to improve data quality become more widely adopted, the likelihood of extreme outliers skewing mean EUI values will diminish. By using mean values, analysts can weight individual EUI values based on the size of the buildings, ensuring that buildings with a larger gross floor area have a higher influence on energy impacts than smaller buildings. If the benchmarking data is clean, a weighted mean EUI value provides the best representation of the overall energy performance of a portfolio of buildings, and would encourage cities and their partners to focus on improving performance in those buildings that have the largest impact on overall energy use.

4 Ibid

The recommended approach for calculating the changes in EUI between different years is:

- For each building, (total weather-normalized energy use) = (weather-normalized EUI) x (building area)
- Total portfolio-wide, weather-normalized energy use for the year = sum of the total weather-normalized energy use across all buildings in that dataset
- Mean portfolio-wide, weather-normalized EUI = (total portfolio-wide, weather-normalized energy use for the year) / (total area of all buildings in the portfolio for that year)
- Percentage change in mean weather-normalized EUI for the entire portfolio at year n = 100 x ([mean weather-normalized EUI for year n] – [mean weather-normalized EUI for year 1]) / (mean weather-normalized EUI for year 1)

Analyses to determine changes in mean GHG emissions should use a similar approach, replacing energy use with emissions in these calculations.

As discussed later in this report, the term “portfolio wide” can refer to all benchmarked buildings across an entire city, or to a smaller, targeted subset of those buildings. Note that the total area of the buildings being considered should generally be constant for all years, as explained in Section 3, Scope of Metrics Assessed.

ENERGY STAR Score

A building’s ENERGY STAR score is another metric commonly used to assess whether benchmarking is impacting building performance. However the ENERGY STAR score is a non-linear metric. As it measures where a building’s performance falls on the distribution curve of its national peers, a 10 percent change in a building’s energy use does not equate to a 10 percent change in its ENERGY STAR score. Therefore, weighting by building size is not appropriate when evaluating the changes in ENERGY STAR scores, and median ENERGY STAR score is the most useful indicator for evaluating trends across a set of buildings. In this case, the calculation would simply be:

- Median portfolio-wide ENERGY STAR score for the year is the mid-point of all ENERGY STAR scores (the same number of buildings lie above and below the median value)
- Percentage change in median ENERGY STAR scores for the portfolio of buildings = 100 x ([median ENERGY STAR score for year n] – [median ENERGY STAR score for year 1]) / (median ENERGY STAR score for year 1)

Since these results are not weighted by the size of buildings, changes in the median ENERGY STAR score represent how the performance of the local building stock has changed compared to national distributions. It can provide insights into the number of buildings that are moving to higher or lower performance levels, and the magnitude of that movement, but it does not represent the percent improvement in overall energy use or GHG emissions.

Comparing Metrics

The importance of understanding how to correctly interpret these different metrics becomes clear when looking at the results observed in Philadelphia from 2012 to 2014[5]. Note that the first benchmarking reporting deadline for buildings in Philadelphia was in November 2013, so these results were from the initial inception of the program, and should be viewed as reflecting baseline conditions of the building stock rather than the impacts of the benchmarking program. **Over that period:**

- The average weather-normalized site EUI of buildings benchmarking throughout that period dipped only slightly, from 101. 2 kBtu/sq. ft. in 2012 to 100. 8 kBtu/sq. ft. in 2014. This indicates that there was no noticeable change in the energy efficiency of the buildings during this period.

- From 2012 to 2014 the median ENERGY STAR score of eligible buildings reporting all three years decreased from 66 to 59, which means that the performance of these buildings actually degraded as compared to national peers. Since their EUI values improved slightly over this period, this suggests that the relatively poorer performance may have been due to changes in some other factors that can impact ENERGY STAR scores, such as lower occupancy rates or reduced operating hours.

- Total energy usage in large non-residential buildings increased from 22,300 million kBtu in 2012 to 24,500 million kBtu in 2014, matching citywide energy trends. Staff at the Office of Sustainability believe that this increase can be attributed to growth in population, jobs, and development, as well as more extreme summers and winters. This implies that the data and analysis was not normalized for weather or for changes in building area.

- Although total GHG emissions grew from 2. 35 million metric tons of carbon dioxide equivalent ($MtCO_2e$) to 2. 48 million $MtCO_2e$, there was a 7 percent reduction in carbon emissions from large buildings between 2013 and 2014, indicating that even though buildings were not becoming more energy efficient, they were relying on less carbon-intensive energy fuel sources such as natural gas and steam.

These observations show that each metric can provide a slightly different perspective on how building performance is changing, and the progress that a city is making toward any goals it may have established. A combination of a few key metrics is generally needed to provide a robust understanding of the progress that is being made.

However, in practice, the evaluation of impacts has proven to be even more complicated than these examples might indicate. For instance, how does one ensure that the data is accurate enough to allow for meaningful analysis? How does one ensure that the datasets for the years being evaluated are comparable? How does one evaluate whether impacts are more pronounced within different subsets of the buildings? These challenges, along with best practice solutions, and some examples of how individual cities have addressed these issues, are described in the remainder of this report.

5 City of Philadelphia, "2016 Energy Benchmarking Report," 2016, http://www.buildingrating.org/file/1794/download.

2. Developing Usable Datasets for Measuring Impacts

Improving Data Quality for Annual Compliance

In order to perform meaningful analysis of benchmarking data, the City must first ensure that the data collected is of sufficiently high quality. A high-quality dataset is both complete and accurate. Complete datasets are those in which a high percentage of the records are complete, with information entered into each data field; accurate datasets are those in which repeated data fields contain the same information in each instance, the data recorded in each field is correctly formatted using the correct units, and the data accurately reflects the characteristics and performance of the buildings being reported[6].

A variety of practices are employed by cities to ensure that the benchmarking data that is reported to them meets reasonable criteria for data completeness and accuracy. These measures to improve data quality can occur during any of three distinct phases:

- Policy and program design: Define data verification requirements that must be met for all reports submitted.
- Implementation: Provide adequate training and support resources for users, and streamline/automate the benchmarking and reporting process.
- Data cleansing and initial validation: Review data as it is submitted, and conduct standardized checks for reasonableness.

The specific practices that can be employed in each of these areas are described in more detail in Chapter 4 of "Putting Data to Work: How Cities are Using Building Energy Data to Drive Efficiency."

Producing Datasets for Multi-Year Trend Analyses

The data quality checks described in Section 3. 1 help ensure that the data being reported as part of the current year's benchmarking submission is as accurate and complete as possible, and that the data is in compliance with any standards established by the City. However, ensuring that the current dataset meets the basic criteria for compliance with the City's policy is a necessary, but not sufficient, condition for a City to perform an assessment of benchmarking impacts. There are additional steps that must be taken before the City can compare individual datasets from multiple years to produce an accurate understanding of trends over time.

Identifying Outliers

By looking at changes in values over multiple years, cities can conduct additional screening—beyond what can be done by looking at the data for only the current reporting cycle—to identify potential errors. These additional data-cleansing checks are generally viewed as being independent of the process to validate whether records meet a quality standard sufficient to be

6 David Hsu, "Improving Energy Benchmarking with Self-Reported Data" (21 February 21, 2014), accessed September 5, 2017. http://www. tandfonline. com/doi/full/10. 1080/09613218. 2014. 887612.

considered in compliance. Washington, DC views as invalid any records w where the building has reduced its EUI by more than 50 percent in a single year[7], while Seattle flags all records where the EUI for any individual fuel type has changed by more than 50 percent compared to the previous year[8]. In both cases, these records are removed from any further analysis of impacts, unless further examination and discussion with the building owner indicates that the data was correct as entered. Additional examples of how cites have identified and removed potential outlier values can be found in the main report of this toolkit, in Chapter 4 of "Putting Data to Work: How Cities are Using Building Energy Data to Drive Efficiency."

Historical Data Coverage

To analyze year-over-year changes in building performance, it is necessary to have access to a dataset that covers multiple years. Typically, at least three years of data is needed to evaluate changes in energy performance. Data from only two back-to-back years could be influenced by any number of external factors and is insufficient to demonstrate performance trends. Per Lawrence Berkeley National Laboratory's (LBNL) assessment of U. S. benchmarking and transparency programs, 13 jurisdictions have now implemented a policy for at least three years: Austin, TX; Boston; Cambridge, MA.; Chicago; Minneapolis; Montgomery County, MD; New York; Philadelphia; San Francisco; Seattle; Washington, DC; and the states of California and Washington. These are the only jurisdictions with sufficient data to consider performing an impact assessment at this time, and only six of these cities (and none of the states) have themselves calculated and publicly reported energy impacts over time[9].

The City of Chicago noted that results from other cities with energy transparency ordinances and conversations with local management firms indicate that properties tend to first start seeing decreases in energy use and cost savings after the first one to two years that they have benchmarked[10]. Many property managers could still be focusing on compliance, rather than on making operational and physical improvements, during the first one to two years that they are required to benchmark under a new energy transparency policy. Therefore, any assessment of impacts that is attempting to determine if there has been an uptick in energy efficiency improvements should include data that extends beyond the first two to three years of a benchmarking program.

7 Marshall Duer-Balkind (District of Columbia Department of Energy and Environment), interview with the author, September 21, 2107.

8 Terry Sullivan and Rebecca Baker (City of Seattle Office of Sustainability & Environment), interview with the author, September 21, 2017.

9 Natalie Mims, Steven R. Schiller, Elizabeth Stuart, Lisa Schwartz, Chris Kramer, and Richard Faesy. "Evaluation of U. S. Building Energy Benchmarking and Transparency Programs: Attributes, Impacts, and Best Practices" (April 28, 2017), https://emp. lbl. gov/sites/default/files/lbnl_benchmarking_final_050417_0. pdf, accessed September 19, 2017.

10 "City of Chicago Energy Benchmarking Report 2016," https://www. cityofchicago. org/content/dam/city/progs/env/EnergyBenchmark/2016_Chicago_Energy_Benchmarking_Report. pdf, accessed September 18, 2017.

Similarly, the City of Minneapolis found that private commercial buildings greater than 100,000 square feet (which had, at the time, reported for two years) exhibited no major change in their weather-normalized EUI performance from 2013 to 2014. The City viewed this period as providing a consistent baseline before the effect of performance transparency was underway[11]. Since the reporting date lags behind the performance year, in calendar year 2014 building owners and managers had little opportunity to react to the benchmarking results and improve performance. In addition, since the transparency elements of the benchmarking policy did not begin to take effect until August 2015, building owners had not yet experienced the complete roll-out of the benchmarking policy. The full market cycle of understanding benchmarking results and then planning, making decisions, and investing in efficiency projects is expected to require additional time. All of these factors combine to indicate that as much as four to five years of data should ideally be available in order to determine trends in energy performance that can be attributed to the impacts of a local benchmarking policy.

> “As much as four to five years of data should ideally be available in order to determine trends in energy performance that can be attributed to the impacts of a local benchmarking policy.”

In fact, the DOE’s Benchmarking and Transparency Policy and Program Impact Evaluation Handbook[12], which proposed indicators to show when building owners, tenants and investors are beginning to overcome the barriers preventing broad scale investment in energy efficiency, projected that the long-term market transformation needed to drive major improvements could take up to 10 years. During that period, building owners could be expected to progress through the following representative stages:

1. Short term (1 to 3 years): Building owners are aware of the annual energy performance of their buildings.
2. Intermediate term (3 to 6 years): Building owners include energy performance as a component of retrofit/renovation planning.
3. Long term (7 to 10 years): Building owners incorporate benchmarking and transparency data into energy management decisions as a matter of standard practice.

11 City of Minneapolis, “2014 Energy Benchmarking Report” (February 2016),http://www. minneapolismn. gov/www/groups/public/@health/documents/images/wcmsp-176597. pdf, accessed September 18, 2017.

12 Navigant Consulting, Inc. and Steven Winter Associates, Inc., “Benchmarking and Transparency Policy and Program Impact Evaluation Handbook” (May 2015), https://energy. gov/sites/prod/files/2015/05/f22/DOE%20Benchmarking%20and%20Transparency%20Policy%20and%20Program%20Impact%20Evalua-tion%20H. . . . pdf, accessed January 5, 2018.

Consistency of the Building Population

To perform an accurate analysis of benchmarking impacts, not only must the dataset cover at least three to five years, but the data to be analyzed should be limited to those individual buildings that have consistently reported throughout those years. As many cities phase in their reporting requirements over two or three years, even if a policy has been in place for three or more years it is likely that only those buildings that were subject to the earlier reporting deadlines have submitted data for all of those years. Buildings that met allowable exemptions, such as construction or demolition of properties, or financial hardship, may not have reported during some years. Furthermore, some buildings, even if they were required to report for all three years, may show gaps during years where they failed to comply, or where their reported data did not meet the established data standards. Any buildings that fail to have complete results for all of the years being analyzed should be eliminated from the dataset.

This requirement can be complicated by the fact that some buildings may have undergone a change in ownership, or may for some other reason have reported under different ENERGY STAR Portfolio Manager[13] building IDs during the time period being assessed. Within Portfolio Manager, a unique building ID is assigned to each building within a user's portfolio of buildings. If a building changes ownership, the new owner will typically have to create a new account for the building[14], which means the building will receive a new Portfolio Manager ID. Though these Portfolio Manager records may appear to represent different properties, in actuality they are referencing the same building. Although no city is doing this yet, there appears to be a significant opportunity to reduce the gaps caused by changes in ownership. At a minimum, each city should attempt to identify any mismatched records that, when properly linked together, could be used to evaluate the performance of additional buildings over time. Ideally cities, perhaps in collaboration with the EPA's Portfolio Manager team, could develop a mechanism to link together the data for all historical submissions for a single physical building, even if those reports were filed by different owners under different Portfolio Manager building IDs.

Three of the cities interviewed—Minneapolis, Seattle, and Washington, DC—remove any buildings that did not report consistently for each of the years being analyzed. However, this means that, because of gaps, the dataset can become more and more limited. For example, in New York City a building is not required to report for the first year after a change in ownership, a common occurrence in that market. To increase the size of the valid dataset, New York City is now using a different methodology; it looks at six years of data, and includes all records where a building reported for at least five of those years. The City then interpolates missing data points by looking at the values for adjacent years. Staff estimate that this gave them about 4,200 buildings to work with during their last evaluation, as compared to the roughly 2,200 buildings that had complete data with no gaps for all six years[15].

13 The U. S. EPA's ENERGY STAR Portfolio Manager, a free web-based application for benchmarking building performance, is the common platform through which building owners across the country benchmark their buildings and, where required, report results to their cities.

14 Though it is technically possible to share a building's Portfolio Manager record with another party, thereby maintaining the building ID and all historic data, in practice this is rarely done when a building changes ownership.

15 Ufei Chan (NYC Department of Buildings), interview with the author, September 21, 2017.

Because of the large amount of brownfield construction now underway in Seattle, many existing buildings there are being retired and replaced by new ones with a limited reporting history. These new buildings would all be excluded if the City's analysis screened out buildings that had not reported consistently since 2011, the performance year that buildings first began to report under the ordinance. Therefore, the City has opted instead to use a three-year rolling window for its trend analysis, starting with a more recent baseline each year rather than always using the same, consistent baseline year. It is seeing a steady increase in the quality of submissions, and anticipates that fewer and fewer building records will have to be excluded from the analyses because of data quality issues. In addition, Seattle has had very high annual compliance rates—approaching 100 percent participation. Because of these factors, staff feel confident that they will not see many gaps in their data, and thus the dataset of validated buildings they will be able to work with during future years, based on a three-year rolling window, will actually be increasing each year[16].

Unusual External Factors

Although the ENERGY STAR score attempts to normalize for key factors such as weather and occupancy levels, there may be other significant external factors that could influence energy consumption. These must always be considered when assessing the impact that the benchmarking requirements may have had on energy use. Economic downturns (or economic growth) can play a role in driving the occupancy levels and intensity of use within a building, and these changes are frequently not accurately captured within Portfolio Manager records.

For example, Minneapolis recently experienced a major construction boom, spurred on in part by development related to the city's preparations to host the NFL Super Bowl in February 2018.

16 Terry Sullivan and Rebecca Baker (City of Seattle Office of Sustainability & Environment), interview with the author, September 21, 2017.

The City is hopeful that these new buildings will, in general, perform better than the historic building stock. If true, this means that Minneapolis should expect to see a noticeable uptick in its building performance metrics within the next two to three years, when those new properties will have reported for enough years to be included within the dataset used for impact analysis[17].

Unusual weather events can also affect the quality and quantity of the data that can be analyzed. In 2012, Hurricane Sandy generated the biggest storm surge to strike New York City in recorded history. It had a significant impact on waterfront properties, with a particularly acute effect on office properties. Over 2,400, or roughly 16 percent, of the properties covered by the City's benchmarking ordinance were located in areas affected by prolonged power losses due to utility damage. Some of the largest covered properties were located within the inundation area, and many of these properties were without electrical service for more than three weeks. Because fewer properties that were affected by Hurricane Sandy benchmarked in the following year, the cleaned dataset only included 25 percent of the properties as compared to the previous year, limiting the City's ability to perform a meaningful analysis of changes in building performance over that period.

Sometimes, actions undertaken by the City itself can lead to changes which could skew the assessment of benchmarking impacts. For example, in April 2017 New York City Mayor Bill de Blasio announced that by 2022 the City will install air conditioning in 11,500 classrooms that currently lack this equipment, to increase the comfort levels in schools. New York City realizes that this will increase the energy load within schools, but has not yet done an analysis to determine the anticipated impacts[18].

17 Katie Schmitt (Center for Energy and Environment), interview with the author, September 20, 2017.

18 City of New York, "Mayor de Blasio, Chancellor Fariña and City Council Announce Every Classroom Will Have Air Conditioning by 2022," http://www1. nyc. gov/office-of-the-mayor/news/261-17/mayor-de-blasio-chancellor-fari-a-city-council-every-classroom-will-have-air, accessed September 22, 2017.

Comparability of Data

Several cities emphasized that the rules they apply to eliminate potential errors may vary depending on the purpose for the analysis. For example, in Minneapolis a trend analysis will only include buildings with complete reports for each year over the multi-year span, while a graphic of annual EUI distribution will show essentially all buildings for that year, including some outliers that would likely have been screened out before performing a trend analysis.

Similarly, for its quartile analyses Seattle strips out only the top/bottom 1 percent of buildings, but removes additional outliers before performing other trend analyses. Before conducting these other analyses, City staff looks at each building type and manually tries to establish a reasonable cut-off value for outliers based on the distribution of data points.

Because of these variations in approach and in composition of the underlying datasets, it is not necessarily possible to compare the results of individual years to the results derived as part of a year-over-year trend analysis of impacts.

Tools

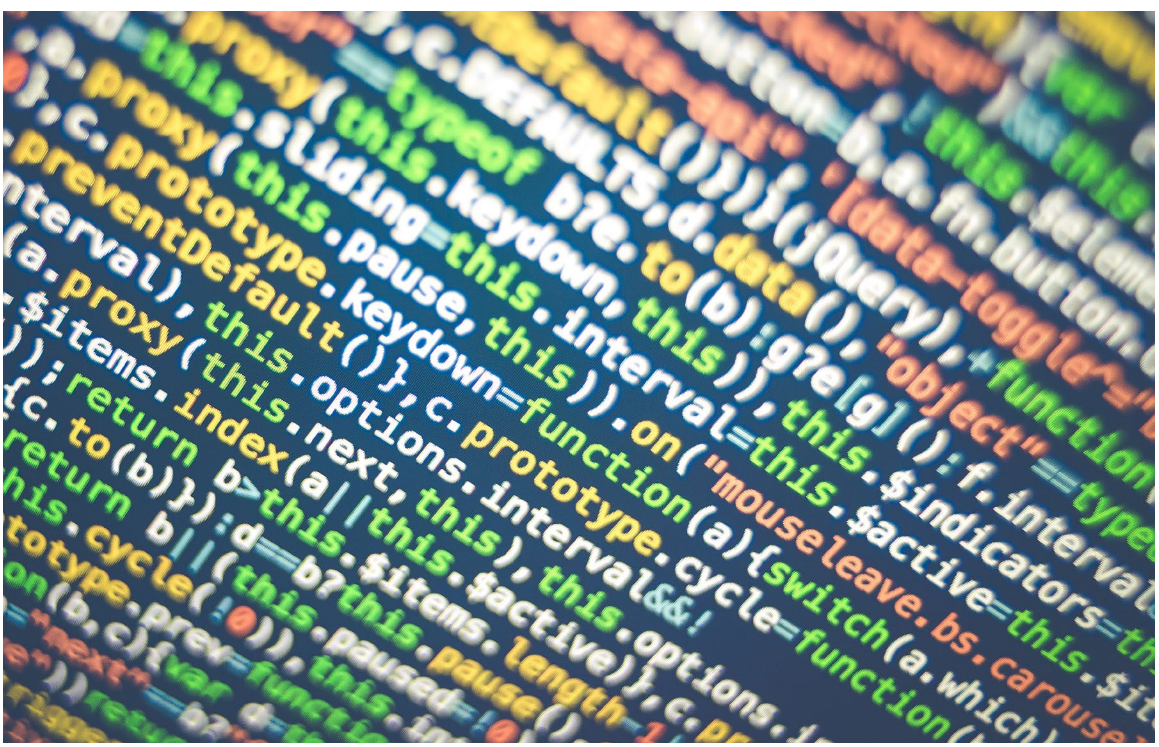
Performing the analyses needed to generate year-over-year metrics on changes to large populations of buildings is not a trivial undertaking. Therefore, every city that is doing this on a regular basis is using some combination of software tools to automate the process to the greatest extent possible. Some examples of the solutions being used include:

- Minneapolis uses its custom-designed Microsoft Access-based Benchmarking Management System to automatically perform many of the initial data quality checks. All compliant records are then exported to Tableau, a business intelligence software product designed to provide interactive data visualizations, where additional manual analyses are performed.

- New York City outsourced the initial data cleansing process to New York University's Center for Urban Science & Progress (CUSP), which applied a statistical methodology to clean the data and remove any outliers. Both CUSP and the Urban Green Council (UGC) then worked with this clean dataset in parallel to conduct different types of impact analyses. CUSP and UGC each developed their own proprietary processes for screening and analyzing the data. CUSP created a Python-based platform to automate the process, while UGC wrote custom scripts in the R programing language.

- Seattle developed a custom internal solution—its Energy Benchmarking and Reporting platform—which automates most of the initial data-cleansing steps, including identifying records that are missing required data or have EUI values outside of acceptable limits. This data is then exported, and the subsequent level of analysis is conducted by an external consultant, using Tableau.

- Washington, DC relies on Microsoft Access as its central data storage repository. The data is then exported to Excel, where District staff perform their initial analyses. Some of the work is also done using Python scripts, based on data cleansing and outlier identification procedures that CUSP created—the same scripts that are used for New York City's data validation.



3. Scope of Metrics Assessed

Current practices in cities with benchmarking programs show that the metrics they have chosen for evaluating impacts fall into three categories, of increasing specificity:

1. Citywide: Used to evaluate year-over-year changes in citywide performance, and the City's progress in achieving any stated energy efficiency or carbon reduction targets.

2. Sector or program level: Generally the same as the citywide metrics, but with the analysis broken down by subsets of buildings. These metrics are used to evaluate progress within a specific real estate sector, or to examine how participation in different energy efficiency related programs or activities has impacted changes in performance.

3. Building level: These metrics can be used to provide insights on why individual buildings are under performing or over performing as compared to the performance of the building stock as a whole.

Each of these categories is described in more detail below, while Table 1 lists the key metrics that each City we interviewed collects within each category.

Evaluating Citywide Impacts

Citywide metrics are used to evaluate year-over-year changes in citywide performance, and the city's progress in achieving any stated energy efficiency or carbon reduction targets. Cities could, in theory, consider any number of metrics for evaluating impacts, including: energy savings, non-energy savings (improvement to air/water quality, and ensuing health benefits), market transformation, and job creation[19]. However, not only is direct energy savings the easiest metric to measure and quantify, but numerous economic analyses have shown that this is the largest contributor to the positive impact of benchmarking policies. Although this may change as more data becomes available, in practice we found that, when assessing the citywide impact of their benchmarking policies, cities have chosen to focus exclusively on metrics related to energy savings and related reductions in GHG emissions energy savings.

19 These categories are considered in the DOE's proposed impact evaluation methodology. See "Impact Evaluation Handbook Benchmarking & Transparency Policy and Program Impact Evaluation Handbook" (May 2015), https://energy. gov/sites/prod/files/2015/05/f22/DOE%20Benchmarking%20and%20Transparency%20Policy%20and%20Program%20Impact%20Evaluation%20H. . . . pdf, accessed September 17, 2017.

Table 1: Metrics Analyzed by the Cities Interviewed

CITY	CITYWIDE METRICS	SECTOR- OR PROGRAM-LEVEL METRICS	BUILDING-LEVEL METRICS
Minneapolis	• Median site EUI • Median weather-normalized EUI • Median ENERGY STAR score	Median levels and sums for each of the three citywide metrics, broken down by property types (especially offices and hotels)	None
New York City	• Primary: Total GHG emissions • Secondary: weather-normalized energy use, ENERGY STAR scores	• Total GHG emissions per building sector • Median GHG emissions per building sector • Sectorwide energy use by fuel type (electricity, district steam, natural gas, heating fuel oil and diesel fuel oil)	Performance vs. type of heating system; cooling system configuration (centralized vs. separated systems); and presence of sub-metering.
Seattle	Total site energy use	• Median EUI by building type • Quartile ranges by building type for site EUI • Distribution across defined buckets for ENERGY STAR scores	Change in EUI for municipal buildings, with an explanation of contributing factors
Washington, DC	Weather-normalized site EUI	None	None

Citywide metrics can be used by cities to understand the direct impact of a benchmarking and transparency policy. The primary citywide metrics that a City chooses to focus on are often determined at least in part by the specific goals that the City has publicly established. For example, since New York City has a goal of reducing GHG emissions by 80 percent by 2050, the primary citywide metric that it monitors is GHG reduction. New York City also looks at year-over-year changes in weather-normalized energy use and ENERGY STAR scores.

Washington, DC looks at trends in weather-normalized site EUI as its primary means for evaluating citywide impacts. It selected site EUI because the fuel mix factors within Portfolio Manager change each year, which could skew year-over-year comparisons based on source EUI. Although the District has ambitious plans to be carbon neutral by 2050, it has opted not to look at the GHG emissions data that could be collected through benchmarking reports. Staff determined that because buildings can report green power purchases within Portfolio Manager, this complicates the use of GHG emissions as a metric. Although the District encourages green power purchases, GHG reductions gained through Renewable Energy Certificates (RECS)[20] do not directly affect their overall emissions, so the Portfolio Manager outputs are not an accurate representation of the values the District wants to track. Furthermore, the eGRID[21] factors used by Portfolio Manager for the carbon intensity of the energy supplied to buildings change constantly. The District has achieved a 24 percent drop in GHG emissions over the last 10 years, but estimates that two-thirds of this has been from decarbonizing the grid, not making buildings more efficient, so evaluating changes in GHG emissions as provided from Portfolio Manager benchmarking data is not an accurate indicator of the impact of the District's energy efficiency efforts.

For citywide impacts, Seattle has chosen to look at total site energy use, rather than EUI, to avoid any potential issues with changes in the area reported for individual buildings. Since the City works with a consistent set of buildings for the entire period being evaluated, it can reasonably expect that there was no change in the total building area during that time. City staff have determined that measuring total site energy use is the most accurate way to determine if the performance of benchmarked buildings is improving.

Evaluating Sector or Program Level Impacts

Evaluating impacts at the sector or program level appears to be gaining an increasing amount of attention. Every City we interviewed is analyzing some aspect of this, or would like to. Tracking sector- or program-level metrics can be an effective way to use benchmarking data to evaluate the success of other, more targeted energy efficiency efforts within a city. In some cases, cities felt that these analyses are even more important than assessing citywide impacts, as the information and insights that can be derived are more directly actionable.

20 Renewable Energy Certificates (RECs) are tradeable commodities that represent electricity that has been generated from an eligible renewable energy resource and fed into the grid. RECs are used to track the ownership of the environmental and social benefits of the renewable energy, and are sold separately from the electricity itself. Therefore, although the use of RECs does encourage renewable energy generation, it is not necessarily associated with any reduction in energy use or direct changes in GHG emissions at the location of the party that purchased the RECs.

21 eGRID (Emissions & Generation Resource Integrated Database) is a comprehensive source of data on the environmental characteristics of almost all electric power generated in the United States, maintained by the US EPA.

New York City has active programs to drive action within specific sectors, such as separate carbon challenges for universities, hospitals, commercial owners and tenants, multifamily buildings, and hotels. One such program, the Retrofit Accelerator, is administered through the Mayor's Office of Sustainability and offers building owners and decision makers free, personalized advisory services that streamline the process of making energy efficiency improvements to buildings. For additional detail about the program, reference the *Putting Data to Work* case study, "Successful Partnerships to Accelerate Efficiency: NYC Retrofit Accelerator. "

Because the City wants to be able to evaluate the impact of each of these sector focused programs, benchmarking data is used to conduct sector-specific analyses. However, Portfolio Manager did not historically provide enough detailed information on building characteristics and use to meet the City's needs. Therefore, the City matched each building record reported from Portfolio Manager with data from internal City datasets—Primary Land Use Tax Lot Output and Real Property Assessment Database—which contain the more granular information on building characteristics that the City needs to perform these detailed, sector-specific analyses. As Portfolio Manager now includes more than 80 property types[22], City staff may no longer need to link to these external data sources to determine specific building types, though these sources can still provide additional information on other building characteristics—for example, number of floors, type of heating and cooling systems, and envelop construction—that would not be available from Portfolio Manager.

Like New York City, Seattle does a deeper dive into specific building types, particularly within multifamily properties, where it uses the number of floors to categorize buildings as low rise, medium rise and high rise. This allows the City to evaluate how those different typologies affect changes in median EUI values.

22 U.S. Environmental Protection Agency, "Property Types, Definitions, and Use Details," last modified April 2017, https://www.energystar.gov/sites/default/files/tools/Property%20Use%20Details%20and%20Definition%20updated%204-21-17.pdf

Seattle also monitors changes in ENERGY STAR score for all buildings, by building type. Each building is assigned to one of four categories—poor (1–50); fair (50–75); good (75–90), and excellent (above 90)—allowing the City to measure the percentage of buildings that falls under each category every year, and track movement over time between these categories. Finally, the City is monitoring the impacts of other variables that could be influencing changes in building performance. This includes evaluating whether acquiring a public certification such as LEED, or participating in Seattle City Light's incentive programs, affects actual building performance.

Washington, DC has not yet done sector-level analyses. Though the District sees the value in this, it currently doesn't have the capacity to do this level of analysis, and also feels that only four sectors—office, multifamily, schools, and hotels—have enough data to be meaningful.

Evaluating Building and System Level Impacts

One of the most valuable analyses that one can perform with benchmarking data is to look across a complete portfolio of buildings to identify best and worst performers, and then attempt to understand the contributing factors that may have led to those results. This exercise can be done by individual owners and property managers responsible for their own privately owned buildings, by municipal general services departments responsible for managing the properties owned or operated by the City, or by the City department that collects data from all public and private buildings subject to a citywide mandatory reporting requirement. Looking at these datasets to identify buildings that are performing significantly better or worse than their peers is one of the fundamental goals for benchmarking buildings, and is a relatively simple analysis to perform. However, determining the underlying causes for these variations is a much more challenging and resource intensive task, though this step may ultimately provide important insights and lessons learned which can then be applied to help other buildings.

For its municipal buildings, Seattle completed a robust evaluation of changes in performance at the building level, along with determining the factors that may have contributed to those observed changes. This detailed level of impact assessment, a subset of which is shown in Table 2, is an excellent way to show trends across a large portfolio of buildings, while allowing users to drill down to understand what is happening at the individual building level[23].

23 "Seattle Municipal Buildings 2013-2014 Energy Performance Report" (September 3, 2015), http://www.seattle. gov/Documents/Departments/OSE/EnergyPerformanceReport_2014_Final. pdf, accessed September 19, 2017.

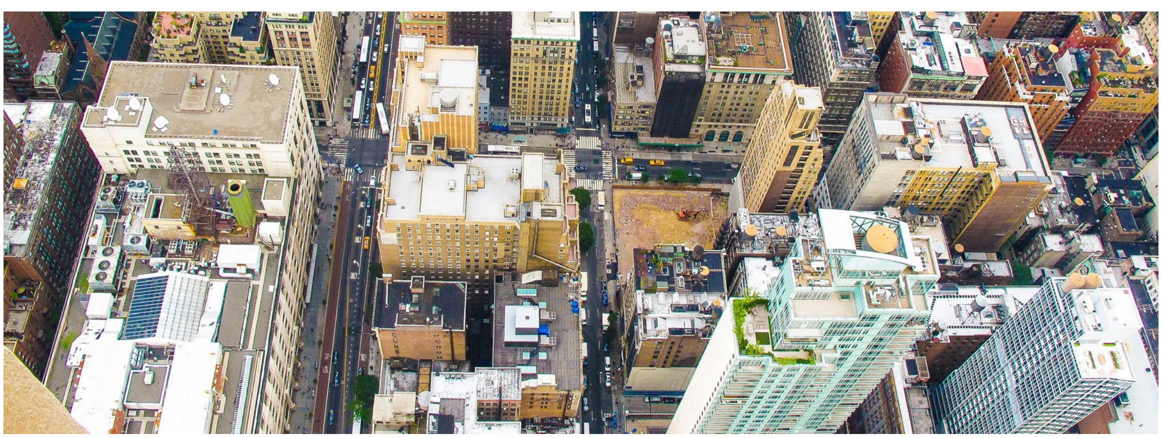
Table 2: Excerpt from City of Seattle's Listing of Benchmarked City-Owned Buildings

BUILDING	MGMT GROUP	TENANT	AREA (SQ. FT.)	2014 EUI (KBTU/SQ. FT.)	CHANGE FROM 2013	YEAR BUILT/ RENOVATED	COMMENTS
West Court Building	CEN	CEN	10,596	57.4	28.1%	1962	Additional tenant
Fairview Building	SDOT	SDOT	8,488	42.6	39.7%	1959	Increased occupancy/use
Seattle Children's Theatre (SCT)	CEN	Arts Group	33,000	74.5	-8.1%	1992	Reduced occupancy – fewer shows
Exhibition Hall	CEN	CEN	52,000	50.4	10.8%	1962	Additional occupancy, ~20% more bookings
Garfield Community Center	Parks	Parks	20,050	90.9	-18.8%	1994	2013 gym lighting upgrade
South Park Community Center	Parks	Parks	14,101	46.9	-14.1%	1912	2013 boiler upgrade
Medgar Evers Pool	Parks	Parks	20,740	356.8	-18.0%	1971	Temporary closure
Broadview Library	SPL	SPL	15,000	58.5	-19.9%	2007	O&M controls optimization
Fire Station 18 – Ballard	FAS	SFD	16,319	82.3	-10.6%	1974	2013 exterior lighting upgrade, corrective maintenance, improved O&M
Fire Station 41 – Magnolia	FAS	SFD	5,664	131.8	23.3%	1936	Equipment/ control issue – resolved
SDOT Traffic Shop	FAS	SDOT	41,939	43.2	-14.5%	1970	Reduced heating set point/heated area
South Transfer Station	SPU	SPU	138,602	38.2	26.5%	2011	Increased use w/ North Station closure

In addition to gathering benchmarking data, New York City has also been collecting more granular building data for several years under Local Law 87, the City's local law that requires buildings over 50,000 sq. ft. to undergo an energy audit every 10 years. The City now does its analysis of audit data and benchmarking data together; these process are completely interwoven and cannot be easily separated.

This combined use of audit data (providing detailed information on the physical and operational characteristics of the systems in each building) and benchmarking data (providing annual energy use data) allows the City to conduct detailed analyses on the impacts of different system configurations. In New York City's Energy and Water Use 2014 and 2015 Report[24], the City looked at linked benchmarking and audit data to evaluate impacts in three areas: heating systems (space and domestic hot water), cooling systems (central vs. separated systems), and metering configurations. The City found, for example, that when controlled for other variables such as building size, location, age, and type of system (hydronic vs. steam), separate heating systems are more efficient than central systems. This is incredibly valuable and actionable information, which is now being used to help owners understand the importance of rightsizing systems, and the benefits of separating heating systems. The City has a wealth of granular data that is allowing staff to make these sorts of determinations. For additional information about the City's flagship program deploying this approach, the Retrofit Accelerator, reference the *Putting Data to Work* case study, "Successful Partnerships to Accelerate Efficiency: NYC Retrofit Accelerator."

As this type of information can directly influence the City's policies and programs, staff anticipate an even greater emphasis on building and system level analytics on in the future. Though citywide metrics such as job creation and health benefits will continue to be of interest to New York City, these will require measurement of benefits not traditionally collected through the benchmarking or audit programs[25].

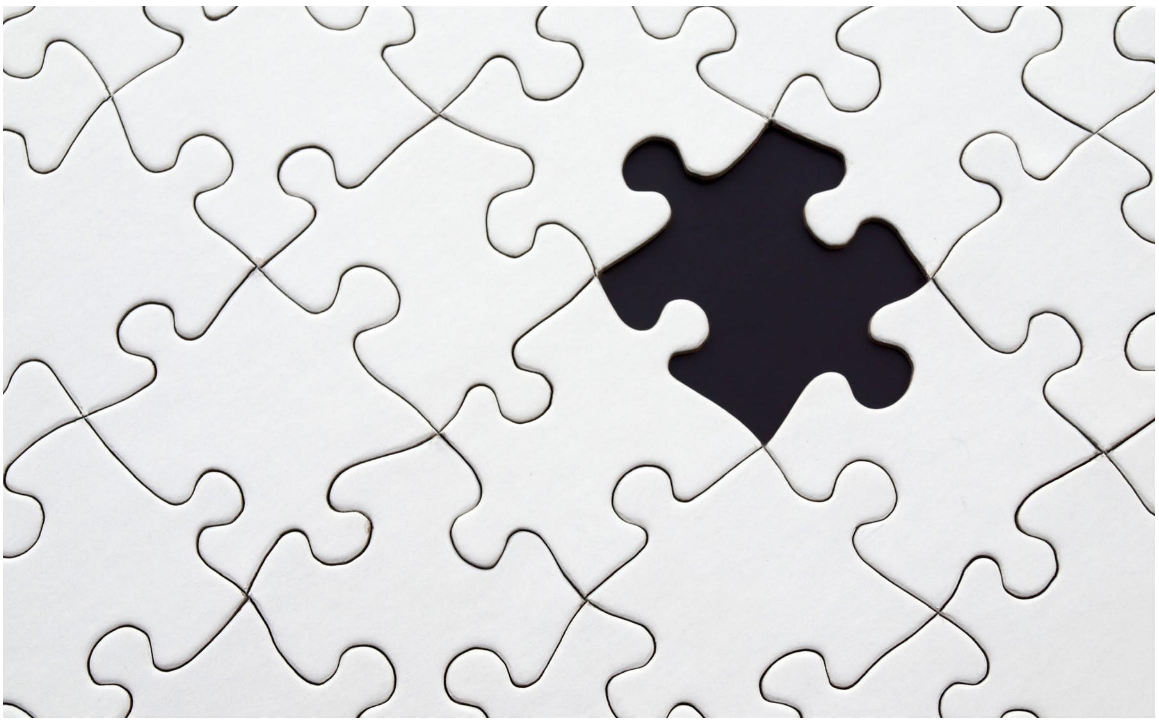


24 Urban Green Council, "New York City's Energy and Water Use 2014 and 2015 Report" (October 2017), http://www. nyc. gov/html/gbee/downloads/pdf/UGC-Benchmarking-Report-101617-FINAL. pdf, accessed December 20, 2017.

25 Ufei Chan (NYC Department of Buildings), interviews with the author, September 21 and 26, 2017.

4. Attribution of Impacts

As efficiency policies and programs do not exist in isolation, attribution remains a significant challenge when attempting to evaluate the impact of a benchmarking program. A benchmarking program is often implemented within a complex environment where energy codes, utility incentives, and other energy efficiency initiatives may also influence building performance. At least 11 of the 25 cities and counties with benchmarking requirements couple their benchmarking and transparency policy with complementary policies, such as building energy audits and retrocommissioning requirements[26], and all have energy code requirements that affect buildings that undergo major renovations. A few of the jurisdictions encourage building owners who are benchmarking their buildings to also take advantage of ratepayer-funded energy efficiency retrofit programs and incentives[27].



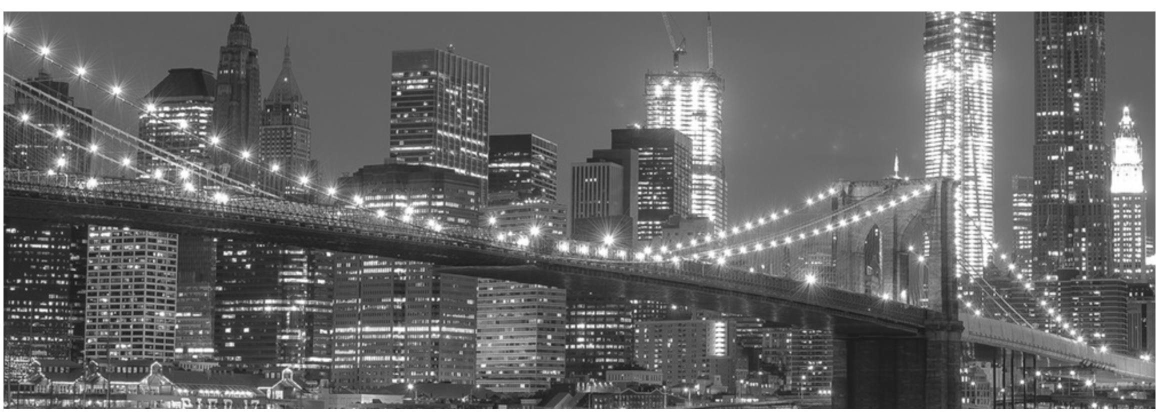
26 "U. S. City Policies: Building Benchmarking, Transparency and Beyond", BuildingRating. org, last updated September 1, 2017, https://www. buildingrating. org/file/1908/download.

27 Natalie Mims, Steven R. Schiller, Elizabeth Stuart, Lisa Schwartz, Chris Kramer, and Richard Faesy. "Evaluation of U. S. Building Energy Benchmarking and Transparency Programs: Attributes, Impacts, and Best Practices" (April 28, 2017), https://emp. lbl. gov/sites/default/files/lbnl_benchmarking_final_050417_0. pdf, accessed September 17, 2017.

In an ideal scenario, a control set of buildings not subject to the benchmarking requirements could be compared to the buildings that are required to benchmark and report results. This would provide a direct comparison for evaluating the impact of the benchmarking requirements. However, this scenario rarely exists in the real world. Instead, cities that have evaluated the impact of their benchmarking programs have generally elected to simply measure how energy performance has changed within those buildings subject to benchmarking requirements.

Because of this, most cities are very careful to describe the performance improvements that are being observed within those buildings that are subject to benchmarking requirements, without implying that benchmarking was directly responsible for those changes. Results of an impact assessment are typically presented in terms such as "there has been an x percent change in average energy use within the buildings that have been consistently benchmarking over the past three years." Regardless, cities such as New York City continue to view benchmarking as an essential measurement tool to see how well other programs are working.

Since understanding how to accurately attribute impacts to benchmarking and transparency policies is so challenging, this is an issue that demands the attention of skilled academics and researchers. A recent Resources for the Future study[28] found that enactment of benchmarking laws led to about a three percent reduction in quarterly utility bills in buildings covered by the laws in the four early-adopter cities, while a report by researchers at the Massachusetts Institute of Technology and the University of Pennsylvania[29] found that New York City's benchmarking and transparency policy led to energy savings of about 6 percent three years after implementation began, and 14 percent four years after the policy took effect. These two studies are only the first of what will likely be a number of research efforts to attempt to evaluate how much of the improvement in building performance can be attributed directly to the act of benchmarking performance, and making such information publicly transparent and accessible.



28 Karen Palmer and Margaret Walls, "Resources for the Future; Discussion Paper: Does Information Provision Shrink the Energy Efficiency Gap?" (April 2015).

29 T. Meng, D. Hsu, and D. Han. "Measuring Energy Savings from Benchmarking Policies in New York City" (August 2016), 2016 ACEEE Sumer Study on Energy Efficiency in Buildings.

5. Findings

Results available to date from cities that have building benchmarking requirements, and have published their analysis of citywide impacts, show that cities are seeing consistent and measureable improvements in energy efficiency across buildings participating in the benchmarking programs. The following are examples of the trend data that has been publicly released by cities.

Chicago

- Properties that have reported consistently for three consecutive years have reduced energy use by 4 percent, leading to an estimated savings of $11. 6 million per year. This group of buildings has also improved its ENERGY STAR scores by 6. 6 percent[30].

Minneapolis

- Three-year analysis for the 99 public buildings consistently benchmarked from 2012 to 2014 shows a 7 percent reduction in total weather-normalized site energy use intensity[31].

New York City

- In year three, the median office EUI went down 13 percent to 191 kBtu/sq. ft. , as compared to the office median EUI of 220 kBtu/sq. ft. in year two. This is a dramatic decline, especially as compared to the relatively consistent EUI for office reported samples in Years One and Two, 234 kBtu/sq. ft. and 220 kBtu/sq. ft. , respectively[32].

- For multifamily properties, the median EUI in the year three reported sample set was also 12 percent lower at 121 kBtu/sq. ft. , as compared to the median EUI 137 kBtu/sq. ft. and 138 kBtu/sq. ft. , in years one and two reported sample sets, respectively[33].

- Total GHG emissions reduced 14 percent from 2010 to 2015, while weather-normalized source energy dropped 10 percent over the same 6-year period[34].

30 Urban Green Council, "New York City's Energy and Water Use 2014 and 2015 Report" http://www.nyc.gov/html/gbee/downloads/pdf/UGC-Benchmarking-Report-101617-FINAL.pdf, accessed September 18, 2017.

31 City of Minneapolis, "2014 Energy Benchmarking Report," February 2016, http://www.minneapolismn.gov/www/groups/public/@health/documents/images/wcmsp-176597. pdf, accessed September 13, 2017.

32 City of New York, "PLANYC: New York City Local Law 84 Benchmarking Report," February 2016,, https://www. google. com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwi-A8oSu26nYAhUr9YMKHSh6C-AQFgg6MAA&url=http%3A%2F%2Fwww. nyc. gov%2Fhtml%2Fplanyc%2F-downloads%2Fpdf%2Fpublications%2F2014_nyc_ll84_benchmarking_report. pdf&usg=AOvVaw1ro2c-yc-CwFmhog4ALWJTh, accessed December 21, 2017.

33 Ibid.

34 Ufei Chan (New York City Department of Buildings), interview with the author, September 21, 2017.

San Francisco

- The 176 properties that benchmarked energy use consistently over the 5-year period from 2010 to 2014 demonstrated regular year-over-year savings, with a 7. 9 percent overall reduction[35].

These results show that improvements in energy use in these cities have ranged from 1. 3 percent to 4. 3 percent per year, over a period of three to six years. Though more cities have evaluated their annual results, and may have even conducted an internal analysis of year-over-year trends, most cities are not yet consistently publishing this level of information.

One troubling trend is that even amongst those cities that have published results, few have released detailed analysis reports after the first year or two of their program, making it challenging to assess long-term patterns in building performance. In response to public demands for open access to government datasets, several cities including Austin, Boston, and Seattle are publishing results for each benchmarked building on a public website, often augmenting these with data maps that show the performance of participating buildings in a more visual format. This trend is also likely a response to the non-trivial level of effort that cities found was required to produce an annual benchmarking report of the quality referenced in this paper.

Publishing data on individual buildings serves an important purpose, but it is not a substitute for performing the types of citywide or sector-specific analyses described in this paper. A raw dataset with records for individual buildings does not provide a clear indication as to effectiveness of a benchmarking program and other related energy efficiency efforts, and the progress a City is making toward achieving its goals. Although it may be difficult for cities to justify the effort to produce formal, highly formatted annual reports, the energy efficiency community would benefit from regular sharing of these summary results by all cities that have benchmarking programs, with the citywide impact analyses ideally coalescing onto a consistent set of key metrics and the methodologies to calculate them.

35 SFEnvironment and ULI Greenprint Center for Building Performance, "San Francisco Existing Commercial Buildings Performance Report, 2010-2014," https://sfenvironment. org/sites/default/files/fliers/files/sfe_gb_ecb_performancereport. pdf.

6. Recommended Best Practices

From the examples described above, we can summarize the following best practices that cities should consider when using benchmarking data to assess impacts:

- Start with accurate and complete datasets for each year under consideration.
- To establish trends, work with at least three years of data, ideally extending to at least five years beyond the date when buildings were first required to report.
- Ensure that the same set of buildings is used throughout all years of the impact assessment
- Define metrics for evaluation that align with the publicly stated goals of the City.
- Make sure that calculations to measure changes in energy intensity use weather normalized mean EUI values that are properly weighted to take the impact of building size into account.
- Automate the data cleansing and validation process or work with external consultants when possible to streamline the upfront activities needed to be able to perform impact analyses.
- Augment benchmarking data with information from other available data sources to provide visibility into building details that may not otherwise be accessible.
- Consider evaluating not only citywide impacts, but also the impacts within individual building sectors.
- Include analyses that provide insights into the impact of system-level and program-level variables that the City can directly influence.
- Even if a formal report is not produced, share summary results with the public on a regular, annual basis.

Conclusions

Our research clearly shows that mandatory benchmarking program are playing an important role in helping cities achieve their energy efficiency and GHG reduction goals. Benchmarking programs are serving both as direct drivers of change, and as a source of data that is being used to evaluate the impacts of other energy efficiency efforts. Though there is still much work to be done to determine how much of a direct impact benchmarking and transparency has on citywide performance improvements, the results to date show that those buildings that have been consistently participating in benchmarking programs are experiencing average reductions in energy use of between 1. 3 and 4. 3 percent per year. Equally important, using benchmarking results to examine targeted subsets of buildings is helping cities develop important insights into the variations in performance across different market sectors, and the impact of external factors such as the type of building systems installed or the building's level of participation in utility incentive programs. This information is just beginning to be used by cities to help them better understand how to direct their resources, and refine their energy efficiency programs and policies, to more effectively achieve the important energy efficiency and GHG reduction goals that they have established.

Through this paper we have surveyed the different ways that cities are using benchmarking data to conduct impact assessments, and have laid out recommendations for the best practices that jurisdictions should consider adopting when doing these analyses, so that the results can be interpreted and shared in a more consistent manner. Our hope is that the information in this paper has illustrated the effectiveness of the analyses that have been conducted by leading cities, and will serve to dramatically reduce the challenges faced by other cities that may choose to follow in their footsteps.

Acknowledgements

IMT wishes to acknowledge the following individuals for contributing their time and expertise to this effort. Analysis and conclusions herein were informed by conversations with these experts, but do not necessarily represent their opinions or conclusions and should be attributed to the author alone.

Rebecca Baker, City of Seattle Office of Sustainability & Environment

Ufei Chan, New York City Department of Buildings

Marshall Duer-Balkind, District of Columbia Department of Energy & Environment

Bridgett Neely, Firefly Consulting

Katie Schmitt, Center for Energy and Environment

Terry Sullivan, City of Seattle Office of Sustainability & Environment

PUTTING DATA TO WORK

Institute for Market Transformation | imt.org/PuttingDatatoWork