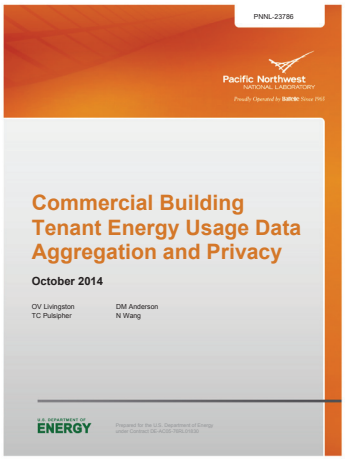
Better Buildings®
U.S. DEPARTMENT OF ENERGY

Briefing Document: Statistical Analysis of Data Access and Privacy

ENERGY DATA ACCELERATOR

This briefing document summarizes the findings of the study *Commercial Building Tenant Energy Usage Aggregation and Privacy*, prepared by the Pacific Northwest National Laboratory (PNNL) for the U.S. Department of Energy. The study was released in October 2014.



Available online at: www.pnnl.gov/main/publications/external/technical_reports/PNNL-23786.pdf

Background and Research Objectives

Many utilities are now providing building owners with energy usage information for an entire building by aggregating the usage from individual meters within that building. This method provides building owners with the information they need to assess energy efficiency opportunities without having to secure consent from every tenant within the building to access their energy usage data. To ensure that aggregation helps protect tenant privacy, many utilities have established "aggregation thresholds" requiring the combination of a minimum number of meters before aggregation can take place.

The objective of this study was to determine the impacts of different aggregation thresholds both on tenant privacy, and on whole-building data access program delivery. The findings can help utilities, utility regulators, and other parties make informed decisions related to setting aggregation thresholds.

The study analyzed non-residential energy consumption data provided under non-disclosure agreements from six utilities representing different climates and geographies. The total number of data points comprised approximately 715,000 meters within approximately 129,000 buildings. Using these data points, the study answered three questions related to aggregation thresholds:

- What is the likelihood that the energy usage of an individual meter can be estimated if a party knows the total number of meters within the building and the total building energy consumption?

KEY TAKEAWAYS

- The "aggregation threshold" – the minimum number of meters that must be combined to enable whole-building data access – has implications on both tenant privacy and on whole-building data access program delivery.
- As the aggregation threshold increases (as more meters are aggregated together), the probability that the energy usage of any tenant resembles the Average Building Meter Profile (ABMP) decreases. This increases privacy for tenants.
- As the aggregation threshold increases, the number of buildings that are eligible to receive whole-building energy data decreases. This limits the effectiveness of providing whole-building energy data.
- Many utilities that are providing whole-building aggregated energy data are setting aggregation thresholds in the range of 2 to 5 meters.

Threshold (# of meters)	% of Meters Similar to ABMP *	% of Buildings Eligible to Receive Data**
2	44.4	100
3	29.2	56.6
4	22.9	40.3
5	20.4	31.0
6	17.3	25.4
7	16.0	20.0
8	14.7	14.6

* Weighted avg. by number of meters from each contributing utility
**Weighted avg. by number of buildings from each contributing utility

- What percentage of individual meters have consumption values that are similar to the Average Building Meter Profile (ABMP), defined as the total building energy consumption divided by total number of meters?
- As the aggregation threshold increases, how does this impact the number of buildings that are eligible to receive whole-building aggregated energy data?

This study does not prescribe or recommend a specific aggregation threshold, or directly measure the risk of meter- or tenant-level re-identification of aggregated energy consumption information based on an aggregation threshold.

Study Results

- As the aggregation threshold increases from 2 meters, the percentage of meters at a given building that are similar to the ABMP decreases. A lower percentage of meters similar to ABMP represents a lower likelihood that the consumption of any single meter can be estimated.
    - As shown in Figure 1, the percentage of meters similar to the ABMP decreases sharply as the aggregation threshold moves from 2 meters to 4 meters. At a 5-meter threshold, each subsequent increase in the aggregation threshold begins to level off.
- As the aggregation threshold increases from 2 meters, the percentage of buildings eligible to receive aggregated data decreases. This is because as the aggregation threshold increases, fewer buildings have enough individual meters to meet the threshold.
    - As shown in Figure 2, the number of buildings eligible to receive aggregated data decreases with each incremental increase in aggregation threshold. This rate of decrease flattens as the aggregation threshold continues to increase.
- In general, as the aggregation threshold increases from 2 to 6 meters, the decrease in building eligibility is sharper than the decrease in percentage of meters similar to AMBP.
- In policymaking, decision makers must consider how to balance gains in privacy protection against losses in building coverage.

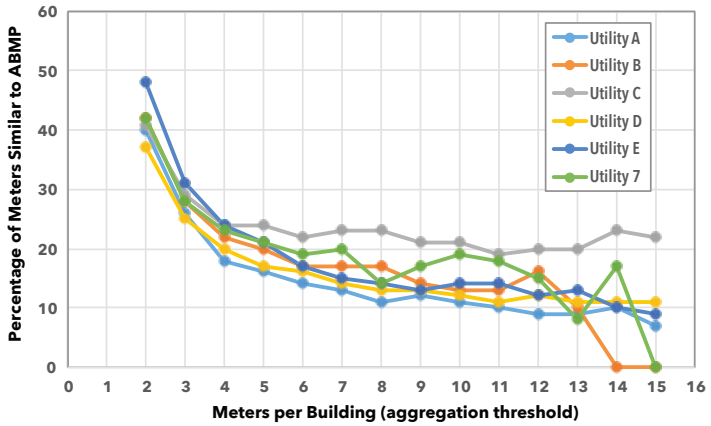


Figure 1: Percentage of meters with consumption values similar to their average building meter profile (total building energy consumption / # of meters) at different aggregation thresholds

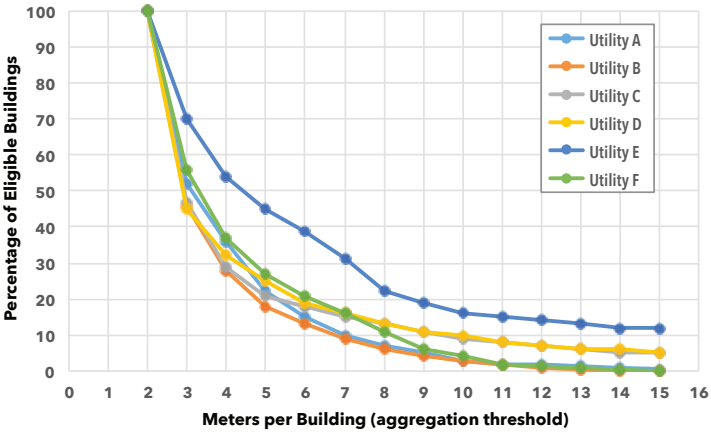


Figure 2: Percentage of buildings eligible to receive aggregated data at different aggregation thresholds

Notes on the Study Methodology

- This study uses individual meters as a proxy for individual tenants. However, there is not always a one-to-one relationship between meters and tenants within a building (a tenant space could be associated with one or more meters).
- The measure for privacy protection in this study is the likelihood that any single meter at the building is similar to the ABMP. The ability to estimate the energy usage of any specific meter based on the ABMP is not the same as "re-identifying" a tenant. Identifying the tenant based on energy consumption would involve linking a specific tenant with a specific meter consumption figure. Even where the study results indicate similarity between meter consumption and the ABMP, this does not imply that a specific tenant can be "re-identified".