U.S. DEPARTMENT OF ENERGY

Office of ENERGY EFFICIENCY & RENEWABLE ENERGY

Building Stock Segmentation Cluster Development

Technical Reference Document

DOE/GO-102023-5835

June 2023

Disclaimer

Authors

The authors of this report are:

Harry Bergmann, U.S. Department of Energy (DOE)

Henry (Ry) Horsey III, National Renewable Energy Laboratory (NREL)

Hernan Rozenfeld, Fractal Fish

Acknowledgments

The authors would like to acknowledge the valuable guidance and input provided during the development of this report. The authors are grateful to the following list of contributors. Their feedback, guidance, and review proved invaluable.

Matthew Dahlhausen, NREL

Bing Liu, Pacific Northwest National Laboratory (PNNL)

Paul Mathew, Lawrence Berkeley National Laboratory (LBNL)

Cindy Regnier, LBNL

Amir Roth, DOE

This report was prepared by the National Renewable Energy Laboratory for the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office.

List of Acronyms

AHJ	authority having jurisdiction
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BPS	building performance standard
BTO	Building Technologies Office
DOE	U.S. Department of Energy
EERE	Office of Energy Efficiency and Renewable Energy
LBNL	Lawrence Berkeley National Laboratory
MSA	metropolitan statistical area
NBER	National Bureau of Economic Research
NREL	National Renewable Energy Laboratory
PNNL	Pacific Northwest National Laboratory

Executive Summary

The energy consumed by commercial buildings in the United States varies significantly across several high-level factors: climate zone, building type, building vintage, and building density. For example, the density, types, and sizes of buildings in Manhattan is quite different from the eastern plains of Colorado, and consequently, the energy use and emissions generated by the building stock in those areas are also quite different. Due to these differences, a single analysis that aims to understand the energy consumption and emissions generation of the U.S. building stock overall often fails to convey the wide range of outcomes and conditions across various parts of the country.

Many authorities having jurisdiction (AHJs), including municipalities, states, and counties, are exploring outcome-based policies and programs to improve the energy and emissions performance of their building stock. The development process for an outcome-based policy or program, such as a building performance standard (BPS), starts with the development of a "covered buildings list." In other words, the agency needs to define which buildings the policy does or does not address. The U.S. Department of Energy's (DOE's) Office of Energy Efficiency and Renewable Energy (EERE) Building Technologies Office (BTO) is developing a series of localized analyses designed to convey the geographic variability of the building stock, enabling interested jurisdictions to more easily understand the condition of their specific building stock—how many buildings of which types there are, how they consume energy and generate emissions, and how big, small, new, or old they may be.

This document discusses the development and implementation of a clustering algorithm that produces a rigorous, consistent, and repeatable collection of geographic clusters to serve as the basis for these localized building stock analyses. The framework considers variables across both the built and natural environments to create groups of counties that have consistent, meaningful takeaways that provide deeper insight than national averages.

The clustering algorithm described in this technical report resulted in 88 clusters across the United States. The clusters are used as the geographic basis for the "U.S. Building Stock Segmentation Series" published by DOE's Building Technologies Office (BTO). This series will provide geographically relevant insight into building stock characteristics, energy and emissions performance, and, eventually, common end use technologies. However, as the building stock changes, the clusters used in subsequent analyses may change from the set used in the initial analysis.

Table of Contents

Executive Summary v
1 Goals of the Clustering Algorithm 1
2 Clustering Algorithm Technical Implementation 3
2.1 Mathematical Representation of Counties in the United States and Clustering Algorithm 3
2.2 Test for Metropolitan Status 4
2.3 Test for ASHRAE Climate Zone 5
2.4 Test for Building Stock Type and Vintage Similarity 6
2.5 Applicability of Tests 7
2.6 Manual Alterations 8
3 County Clustering Results 11
References 12

List of Figures

Figure 1. Graph representation of counties. Counties C1, C2, C3, C4, C5, and C6 are represented as a graph. Counties (nodes) are connected in the graph (b) if they are adjacent, i.e., if they share a portion of their boundary line (a). 4
Figure 2. Graph representation of counties after removing links. (a) Example of graph in Figure 1(b) after removing some links through the tests of the clustering algorithm. This example results in two clusters: a cluster composed of counties C1, C2, and C3 (blue cluster in (b)) and another cluster composed of C4, C5, and C6 (green cluster in (b))........ 4
Figure 3. Metropolitan and micropolitan statistical areas of the United States and Puerto Rico. Metropolitan areas are represented in darker green, with dark green boarders representing the boundary of individual metropolitan statistical areas and micropolitan statistical areas are represented in lighter green........ 5
Figure 4. ASHRAE climate zones of the United States. This map shows numbered climate zones by color and lettered climate zones through the dark black lines drawn across the map. 6
Figure 5. The stock segmentation clusters developed through the process described in the report. 11

1 Goals of the Clustering Algorithm

There is significant spatial diversity in the American building stock—simply put, buildings in Arizona are not the same as buildings in Minnesota. Variability in the building stock is a result not only of climate and geography, but also of when they were built and what needs and people they were designed to accommodate. Depending on those conditions, different technologies, building materials, building codes, and construction practices were used and followed. For example, buildings dating to the 18th century or before are not uncommon along the East Coast but are nearly nonexistent in the American West. In cities and urban cores, buildings far higher than five stories are a common sight, but in the suburbs and more rural areas, such tall structures are exceptionally rare. The factors of where, when, for what, and for whom are all critical elements that drive variability across the American building stock. However, this same variability also limits the value that a single, national-level building stock analysis can bring to regionally or locally focused organizations, including state and local governments.

This report presents a methodology that leverages those variables of where, when, what, and who to generate a series of geographic clusters that enable more regionally and locally relevant analyses to be conducted. Core to this approach are the concepts of climate zones and new construction codes. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) recognizes 16 different climate zones across the United States. California is an exception; the state uses the California Climate Zones (established by the California Energy Commission), as they are more responsive to the state's unique building code (California Energy Commission 2022). There are also at least five different versions each of the residential and commercial building codes adopted across the United States (State Code Adoption Tracking Analysis 2022). The combination of different climate zones and building codes (which vary across time) present a complex matrix of competing requirements and seemingly bespoke conditions. Much of the work conducted by the U.S. Department of Energy, national laboratories, and other organizations that are working to improve energy efficiency and reduce emissions in the built environment produced solutions that are broadly applicable across these geographies, leading to solutions that are marginally less unique than the building stock we currently face. The clustering or geographic aggregation approach outlined in this report represents these overlapping variables in a model to enable useful, actionable insights by regionally or locally focused organizations.

The clustering algorithm presented in this document is designed to create a statistically representative set of clusters that adequately conveys the spatial diversity of the U.S. building stock. Given common data aggregation practices, counties serve as the foundational geographic unit, and metropolitan statistical areas (MSAs) serve as the common nucleus around which clusters are constructed. MSAs are groups of counties defined by the U.S. Census Bureau as a function of contiguous dense metropolitan area. Some MSAs may be smaller, more discrete areas around small or medium cities, but larger cities or urbanized areas may have a large collection of counties in their MSA due to the many layers of development around them (e.g., suburbs, exurbs, and adjacent cities). We considered other geographic units but using MSAs to anchor geographic clusters provided the most pragmatic balance between increasing geographic granularity and minimizing cluster count.

The algorithm considers three factors when evaluating whether two counties or geographic extents (except in the case of California) will be joined: whether counties are part of a shared MSA, whether they share a climate zone, and whether they share similar densities of building types, sizes, and vintages. Section 2 considers each of these factors in detail. The goal of this process is to produce geographic aggregations, or clusters, that will provide sufficiently specific characterizations of the U.S. building stock for local or regional analysis and decision-making, while recognizing that they are not exact representations of the building stock.

2 Clustering Algorithm Technical Implementation

This section describes the algorithmic process used to determine which counties (outside of California) are clustered together. In Section 2.1, we discuss the mathematical representation of counties in the United States and present an overview of the clustering algorithm. Three tests are used to determine whether counties should or should not be clustered together:

1. Metropolitan status (density)
2. ASHRAE climate zone (weather)
3. Distribution of buildings (similarity).

The details of each test are described in Sections 2.2, 2.3, and 2.4, respectively, and the logic for determining test applicability is described in Section 2.5. Finally, Section 2.6 describes the three classes of manual alterations that were used to fine-tune the results of the clustering algorithm.

2.1 Mathematical Representation of Counties in the United States and Clustering Algorithm

In the clustering algorithm, we represent counties as nodes in a graph. Two nodes (counties) are connected through a link in the graph if they are geographically adjacent (Figure 1). Counties are considered to be adjacent if they share any portion of their boundary lines or if they are connected by major transit infrastructure such as bridges, tunnels, or ferry lines (mostly relevant around the Great Lakes).

The central idea of the clustering algorithm is to determine whether to keep or remove the links between adjacent counties according to the three tests described in Sections 2.2, 2.3, and 2.4. Removing links produces discrete clusters whose counties are internally linked but lack links to other clusters, resulting in unique spatial units (Figure 2).

Adjacency was specified according to the National Bureau of Economic Research (NBER; County Adjacency Dataset, 2017), which created a representation of all adjacencies as of the 2010 census. In a limited set of cases, updates were required to adjust the NBER model to align with 2022, the year of this study.

The final step of the algorithm is to manually audit the resulting clusters and make minor adjustments (i.e., remove and add links) according to Section 2.6. After the manual process, each cluster is assigned a unique number.

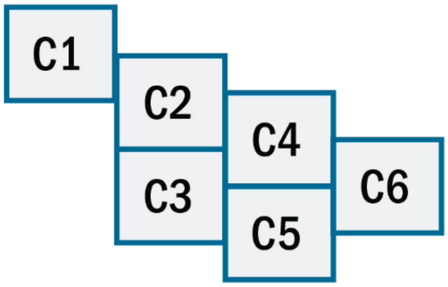

Figure 1a

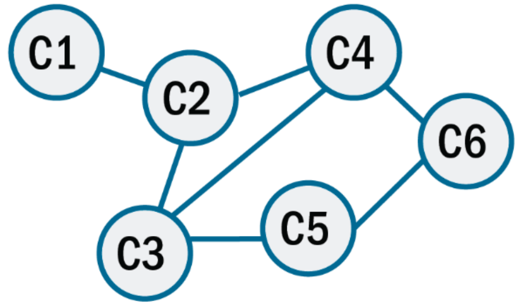

Figure 1b

Figure 1. Graph representation of counties. Counties C1, C2, C3, C4, C5, and C6 are represented as a graph. Counties (nodes) are connected in the graph (b) if they are adjacent, i.e., if they share a portion of their boundary line (a).

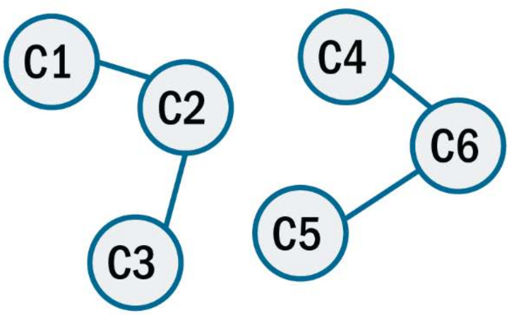

Figure 2a

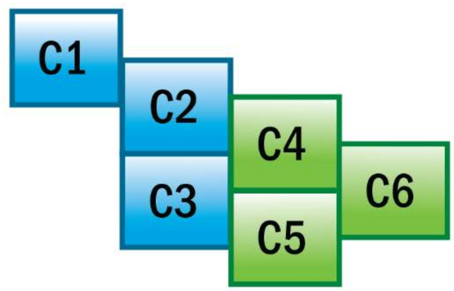

Figure 2b

Figure 2. Graph representation of counties after removing links. (a) Example of graph in Figure 1(b) after removing some links through the tests of the clustering algorithm. This example results in two clusters: a cluster composed of counties C1, C2, and C3 (blue cluster in (b)) and another cluster composed of C4, C5, and C6 (green cluster in (b)).

2.2 Test for Metropolitan Status

The relative urbanity or density of an area is a major factor in correctly representing the energy consumption of its buildings. An obvious example the contrast between a small rural city and a large metropolis. In the rural example, not only is the density of the building stock significantly lower, but the need for certain building types is also quite different. Rural areas might have only one school or one hospital serving three or four cities, whereas a sizeable city likely has multiple schools and hospitals. In many cases, the density of a geography is correlated with the building types found there. These differences can result in drastically different building stocks.

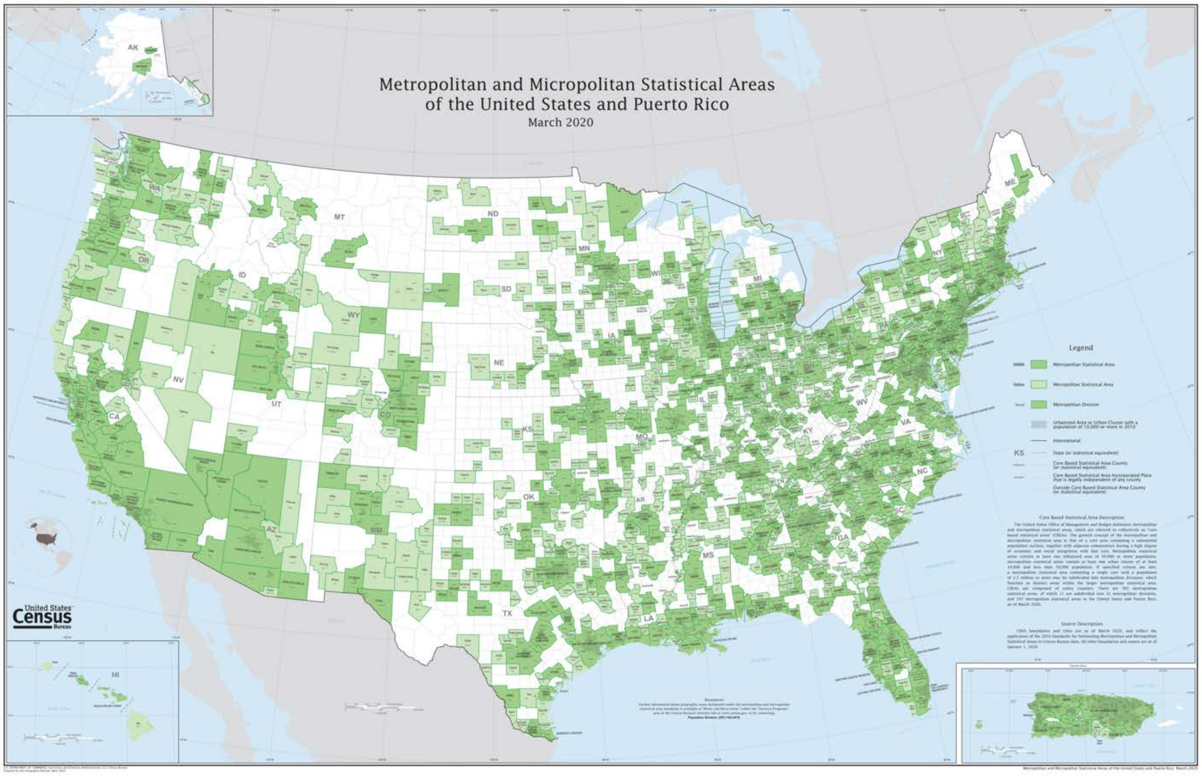

Figure 3. Metropolitan and micropolitan statistical areas of the United States and Puerto Rico. Metropolitan areas are represented in darker green, with dark green boarders representing the boundary of individual metropolitan statistical areas and micropolitan statistical areas are represented in lighter green.

Image courtesy of the United States Census Bureau. Data tables can be found on the Census Bureau website.

MSAs are exclusively made up of counties. The U.S. Census Bureau creates a clear distinction between MSAs, micropolitan statistical areas, and other counties. These definitions provide excellent differentiation between counties that have major urbanized populations and those that do not.

In the clustering algorithm, all links connecting counties within an MSA are preserved. Links between pairs of counties not in an MSA are passed through the ASHRAE climate zone test described in Section 2.3. Finally, all links connecting a county in an MSA and a county not in an MSA are removed.

2.3 Test for ASHRAE Climate Zone

The energy use of buildings in the United States is highly dependent on the climate in which they are built. An obvious example is to compare Miami, Florida, and Fairbanks, Alaska. In Miami, the built environment is designed to manage summer heat and year-round humidity. In Fairbanks, the built environment is designed to provide effective heating in frigid winters, and in many cases, no cooling is required. In this example, it is easy to see why the two cities result in separate clusters, but less extreme differentiation, e.g., Indianapolis vs. Chicago, also deserves attention.

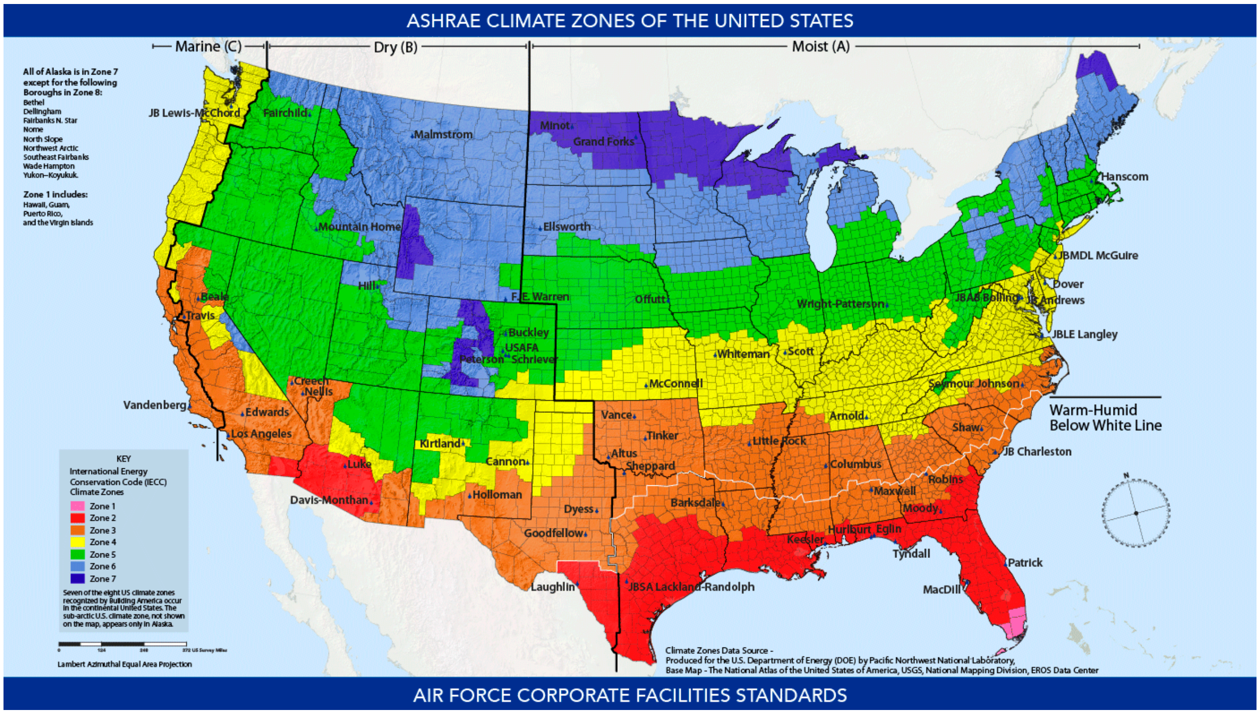

Figure 4. ASHRAE climate zones of the United States. This map shows numbered climate zones by color and lettered climate zones through the dark black lines drawn across the map.

Image courtesy of the Air Force Corporate Facility Standards organization. Data tables can be found in ANSI/ASHRAE Standard 169-2006.

ASHRAE climate zones, as shown in Figure 2, are used to segment clusters for the climate zone test. As discussed in Section 2.5, for non-MSA edges, the ASHRAE climate zone is used as a criterion for edge removal. A link between two non-MSA counties is preserved if they share the same climate zone and is removed otherwise. In the case of edges within a single MSA, the ASHRAE climate zone is not used, because many MSAs cross one or more ASHRAE climate zones. It should be noted that climate zones are defined not only by number (e.g., 3 vs. 4), but also by letter (e.g., 2A vs. 2B) for a total of 16 climate zones. The letters differentiate areas driven by latent vs. sensible loads. The heating and cooling strategies—and therefore energy use—can differ significantly depending on the location and building type in question.

2.4 Test for Building Stock Type and Vintage Similarity

A key factor in clustering counties or MSAs is whether building stocks are sufficiently similar. So far, we have discussed tests that describe whether the climatic or density metrics associated with two counties are similar. Although both of these tests can be useful on their own, neither ensure that the two geographies being connected have a similar mix of building types or vintages.

The similarity test for clustering performed in this study compares both type and vintage on an MSA basis. That is, the test compares two adjacent aggregated MSAs, rather than individual counties.

The first test compares stock vintages to determine whether there is a significant difference in the proportion of building square footage built before 1980. Although 1980 is not the universal year in which building codes and standards programs were implemented across the United States, it is a useful year for comparing buildings built "pre-code" and "post-code" on a national basis. This test intervenes to ensure that the MSAs that share edges will not be combined if their building stocks were built in different eras, from a construction perspective. For the purposes of clustering, if two different MSAs have a >30% difference in the percentage of building square footage built before 1980, then any existing link between them is removed.

To differentiate MSAs by their mix of building types (e.g., hotels vs. offices vs. warehouses), we use a similar approach. Because energy use can differ dramatically between building types, significant differences in the mix of building types and their relative energy use can alter policy recommendations from stock-based analysis. If two MSAs have radically different building type mixes (e.g., if one is dominated by offices and big-box retail stores, and the other is dominated by warehouses and strip malls), any link between the MSAs is removed, and the geographies are not joined.

Automating judgements for the building type-specific qualification test is nontrivial. To simplify the process enough to allow automation, we utilized a cutoff mechanism. If two or more building types between relevant geographic extents displayed a >10% absolute (as opposed to relative) difference, then the edge in question was removed. In some cases, a single building type showed a difference between counties exceeding 10%, but the remaining building categories absorbed the building type differences in such a way that the edge was not removed.

2.5 Applicability of Tests

The tests described above, when initially applied in all cases to all edges, resulted in several hundred distinct clusters. In particular, the tests for building type and vintage, when applied to edges between counties with less dense urban environments, often removed edges and resulted in many small clusters with very little square footage.

Three primary issues surfaced following the initial clustering outcome: First, the overall number of analyses was not significantly reduced compared to the total number of counties or other relevant geographic units, failing to make a compelling case for aggregation at all. Second, the data aggregation platform underpinning the stock analysis to be performed across these clusters is geographically limited and unable to provide coverage for every cluster produced by this first pass due to insufficient data. Third, climate zones often split major MSAs and metropolitan areas, creating nonsensical splits in clusters that otherwise would not be treated separately.

To address these issues, we applied two, rather than three, tests at a time. In all cases, the metropolitan area test was applied first. Next, if there was an edge between two MSAs, the building type and vintage tests were applied; however, if the two counties belonged to the same MSA, the link was preserved, and no additional test was applied. If there was an edge between two counties that did not belong to any MSA, the ASHRAE climate zone test was applied.

This sequential approach to applying the clustering algorithm mostly solved all three challenges described above. It resulted in a manageable number of representative geographic clusters, eliminated most clusters with insufficient data, and minimized the number of manual interventions required to avoid nonsensical divisions (i.e., divisions of a single political jurisdiction and maintain a more intuitive cluster design.

2.6 Manual Alterations

At this stage of the process, the results of the clustering algorithm were all generated using automated processes. The clusters, while generally responsive to the goals described above, still contained three issues, including some already discussed in Section 2.5.

The first issue was that some clusters were exceedingly small, both in terms of the number of counties and the covered square footage. These clusters had too few samples to produce a statistically meaningful characterization of the building stock. This issue was resolved by releasing the ASHRAE climate zone constraint for counties outside of MSAs, allowing them to join larger clusters in an adjacent climate zone.

The second issue was that MSAs with similar characteristics but no shared boundary were treated as separate areas. This increased the number of clusters unnecessarily. To resolve this issue, similar clusters were grouped under a single analysis, even though the geography itself was not contiguous.

The third issue was that for continuously dense geographies such as the East Coast, the tests resulted in too few clusters geographically. Edges were removed to support the goal of localized analysis.

The following counties were impacted by exceedingly small clusters, requiring a release of the ASHRAE climate zone constraint: Gila County, Arizona; Graham County, Arizona; Greenlee County, Arizona; Fremont County, Colorado; Huerfano County, Colorado; Summit County, Colorado; Chattooga County, Georgia; Jackson County, Georgia; Oneida County, Idaho; Clark County, Illinois; Coles County, Illinois; Cumberland County, Illinois; Douglas County, Illinois; Edgar County, Illinois; Moultrie County, Illinois; Washington Parish, Louisiana; Gratiot County, Michigan; Tuscola County, Michigan; Pearl River County, Mississippi; Cheshire County, New Hampshire; Colfax County, New Mexico; Harding County, New Mexico; Hidalgo County, New Mexico; Luna County, New Mexico; Mora County, New Mexico; Rio Arriba County, New Mexico; San Miguel County, New Mexico; Taos County, New Mexico; Cayuga County, New York; Columbia County, New York; Cortland County, New York; Genesee County, New York; Greene County, New York; Seneca County, New York; Alleghany County, North Carolina; Ashe County, North Carolina; Avery County, North Carolina; Lee County, North Carolina; Mitchell County, North Carolina; Watauga County, North Carolina; Yancey County, North Carolina; Gallia County, Ohio; Washington County, Ohio; Bradford County, Pennsylvania; Sullivan County, Pennsylvania; Custer County, South Dakota; Fall River County, South Dakota; Shannon County, South Dakota; Duval County, Texas; Hardeman County, Texas; Page County, Virginia; Shenandoah County, Virginia; Ferry County, Washington; Island County, Washington; Okanogan County, Washington; Braxton County, West Virginia; Calhoun County, West Virginia; Gilmer County, West Virginia; Mason County, West Virginia; Pleasants County, West Virginia; Ritchie County, West Virginia; Roane County, West Virginia; Tyler County, West Virginia; Goshen County, Wyoming; Lincoln County, Wyoming; Natrona County, Wyoming; Platte County, Wyoming; Sublette County, Wyoming; Teton County, Wyoming.

The following counties were dissolved together to address physically disconnected but otherwise similar clusters: Kenai Peninsula Borough, Alaska; Bethel Census Area, Alaska; Wade Hampton Census Area, Alaska; Denali Borough, Alaska; Fairbanks North Star Borough, Alaska; Haines Borough, Alaska; Hoonah-Angoon Census Area, Alaska; Skagway Municipality, Alaska; Sitka City and Borough, Alaska; Petersburg Census Area, Alaska; Prince of Wales-Hyder Census Area, Alaska; Wrangell City and Borough, Alaska; and Ketchikan Gateway Borough, Alaska, were dissolved with Anchorage Municipality, Alaska. Santa Cruz County, Arizona, was dissolved with Pima County, Arizona. La Paz County, Arizona, was dissolved with Yuma County, Arizona. Sherman County, Kansas, was dissolved with Kit Carson County, Colorado. Delta County, Colorado, was dissolved with Mesa County, Colorado. Litchfield County, Connecticut, was dissolved with New Haven County, Connecticut. Putnam County, Florida, was dissolved with Hardee County, Florida. Terrell County, Georgia, was dissolved with Stewart County, Georgia. Kauai County, Hawaii; Hawaii County, Hawaii; and Kalawao County, Hawaii, were dissolved with Honolulu County, Hawaii. Cape Girardeau County, Missouri, and Posey County, Indiana, were dissolved with Jackson County, Illinois. Delaware County, Indiana, was dissolved with Allen County, Indiana. Woodbury County, Iowa, and Grundy County, Iowa, were dissolved with Boone County, Iowa. Grant Parish, Louisiana, was dissolved with De Soto Parish, Louisiana. St. Mary Parish, Louisiana, was dissolved with St. Martin Parish, Louisiana. Lincoln Parish, Louisiana, was dissolved with Union Parish, Louisiana. Piscataquis County, Maine, was dissolved with Penobscot County, Maine. Kent County, Maryland, was dissolved with Cecil County, Maryland. Dukes County, Massachusetts, was dissolved with Barnstable County, Massachusetts. Cass County, North Dakota, and St. Louis County, Minnesota, were dissolved with Polk County, Minnesota. Butler County, Kansas; Greene County, Missouri; and Benton County, Arkansas, were dissolved with Jasper County, Missouri. Dallas County, Missouri, was dissolved with Moniteau County, Missouri. Missoula County, Montana; Jefferson County, Idaho; Stillwater County, Montana; and Natrona County, Wyoming, were dissolved with Cascade County, Montana. Greeley County, Nebraska, was dissolved with Howard County, Nebraska. Erie County, Pennsylvania, was dissolved with Erie County, New York. Tompkins County, New York, was dissolved with Onondaga County, New York. Beaufort County, North Carolina, was dissolved with Lenoir County, North Carolina. Meade County, South Dakota, was dissolved with Morton County, North Dakota. Putnam County, Ohio, was dissolved with Allen County, Ohio. Lincoln County, Oregon, was dissolved with Douglas County, Oregon. Jackson County, Oregon, was dissolved with Lane County, Oregon. Columbiana County, Ohio, and Lawrence County, Pennsylvania, were dissolved with Beaver County, Pennsylvania. Charleston County, South

Carolina, was dissolved with Beaufort County, South Carolina. Roane County, Tennessee, was dissolved with Bradley County, Tennessee. Crockett County, Tennessee, was dissolved with Tipton County, Tennessee. Nueces County, Texas, and Goliad County, Texas, were dissolved with Cameron County, Texas. Webb County, Texas, was dissolved with La Salle County, Texas. Reeves County, Texas, was dissolved with Loving County, Texas. Lynn County, Texas; Taylor County, Texas; Sterling County, Texas; and Armstrong County, Texas, were dissolved with Martin County, Texas. Hanover County, Virginia, was dissolved with Louisa County, Virginia. Surry County, Virginia, was dissolved with Sussex County, Virginia. Franklin County, Washington; Asotin County, Washington; Douglas County, Washington; and Deschutes County, Oregon, were dissolved with Spokane County, Washington. Marathon County, Wisconsin, was dissolved with Chippewa County, Wisconsin. Wabasha County, Minnesota, was dissolved with Eau Claire County, Wisconsin. Manitowoc County, Wisconsin, was dissolved with Sheboygan County, Wisconsin.

The following counties were manually disconnected to support localized analysis in exceedingly dense geographies: Yuma County, Arizona; Pima County, Arizona; and Yavapai County, Arizona, were disconnected from Maricopa County, Arizona. Pima County, Arizona, was disconnected from Pinal County, Arizona. Westchester County, New York, and Putnam County, New York, were disconnected from Fairfield County, Connecticut. Floyd County, Georgia, was disconnected from Bartow County, Georgia. Harris County, Georgia, and Talbot County, Georgia, were disconnected from Meriwether County, Georgia. Newton County, Indiana, and Jasper County, Indiana, were disconnected from Benton County, Indiana. Forrest County, Mississippi, and Perry County, Mississippi, were disconnected from Stone County, Mississippi. Mohave County, Arizona, was disconnected from Clark County, Nevada. Ocean County, New Jersey, was disconnected from Atlantic County, New Jersey. Monmouth County, New Jersey, and Ocean County, New Jersey, were disconnected from Burlington County, New Jersey. Hunterdon County, New Jersey; Somerset County, New Jersey; Middlesex County, New Jersey; and Monmouth County, New Jersey, were disconnected from Mercer County, New Jersey. Pike County, Pennsylvania; Sussex County, New Jersey; Morris County, New Jersey; and Hunterdon County, New Jersey, were disconnected from Warren County, New Jersey. Putnam County, New York, was disconnected from Dutchess County, New York. Putnam County, New York; Westchester County, New York; Rockland County, New York; Passaic County, New Jersey; Sussex County, New Jersey; and Pike County, Pennsylvania, were disconnected from Orange County, New York. Berkshire County, Massachusetts, was disconnected from Rensselaer County, New York. Linn County, Oregon, and Lane County, Oregon, was disconnected from Deschutes County, Oregon. Hunterdon County, New Jersey, was disconnected from Bucks County, Pennsylvania. Pike County, Pennsylvania, and Sussex County, New Jersey, were disconnected from Monroe County, Pennsylvania. Chester County, South Carolina, and Lancaster County, South Carolina, were disconnected from Fairfield County, South Carolina. Lancaster County, South Carolina, was disconnected from Kershaw County, South Carolina. Madison County, Virginia, was disconnected from Greene County, Virginia. Spotsylvania

County, Virginia, was disconnected from Hanover County, Virginia. King County, Washington; Snohomish County, Washington; and Skagit County, Washington, were disconnected from Chelan County, Washington. King County, Washington; Pierce County, Washington; and Skamania County, Washington, were disconnected from Yakima County, Washington.

3 County Clustering Results

The process described above results in 76 clusters outside of California and 12 clusters in California. These clusters are shown in Figure 5.

Considering the impact of built environment density, diversity, and climate when creating groupings of counties enables far more nuanced energy analyses compared to using national averages. The stock clusters, combined with an analysis tool engine, will provide data for building stocks at sub-state geographic resolution. These geographic clusters can be used as the basis for localized analysis. Moreover, they may be updated as more data becomes available, and our understanding of the building stock improves.

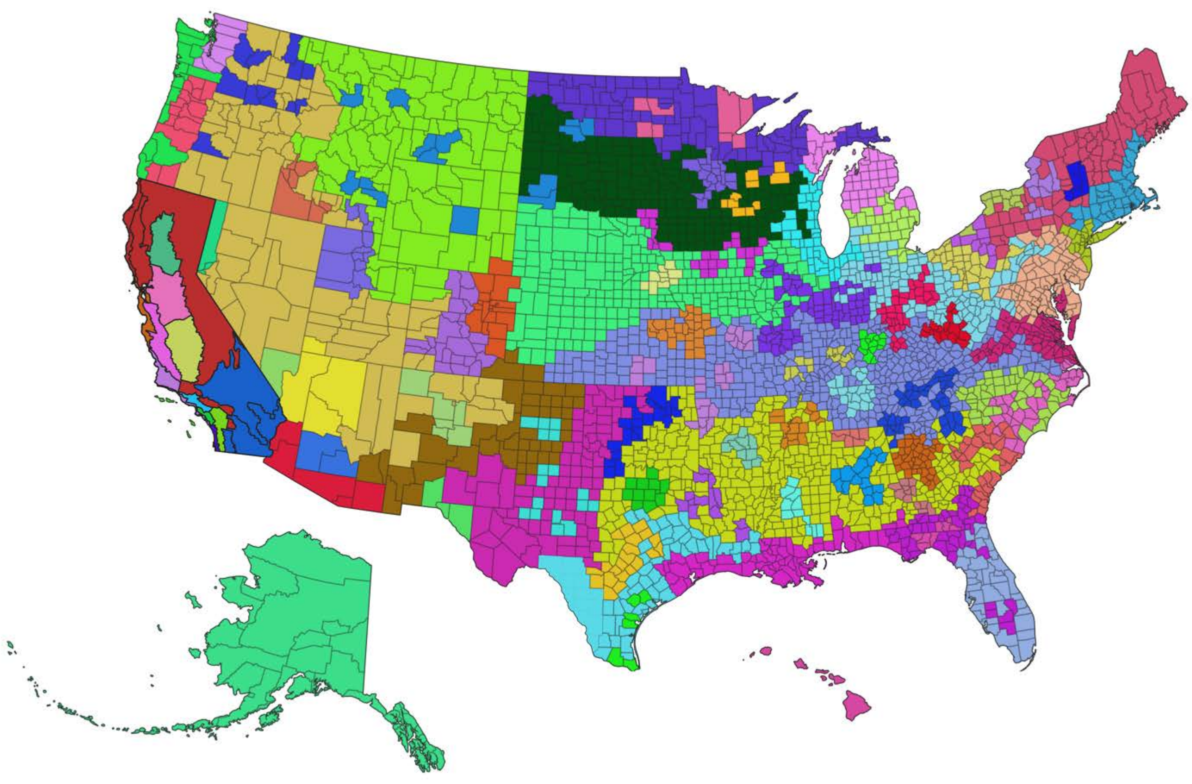

Figure 5. The stock segmentation clusters developed through the process described in the report.

References

"State Code Adoption Tracking Analysis." 2022. United States Department of Energy, Energy Efficiency and Renewable Energy Office, Building Technologies Office, Building Energy Codes Project. January 13, 2022. https://www.energycodes.gov/sites/default/files/2022-01/StateLevelCodesEnergyUseIndex_FY2022Q1_20220113.xlsx.

California Energy Commission. 2022. "Title 24 2022: Reference Appendices, Appendix JA2-1." CEC-400-2022—010-AP, August, 2022. https://www.energy.ca.gov/sites/default/files/2022-08/CEC-400-2022-010-AP.pdf.

"County Adjacency Dataset." 2017. National Bureau of Economic Research. May 8, 2017. https://www.nber.org/research/data/county-adjacency.

U.S. DEPARTMENT OF ENERGY

Office of ENERGY EFFICIENCY & RENEWABLE ENERGY

For more information, visit:
energy.gov/eere/buildings

DOE/GO-102023-5835 • June 2023