Energy and Housing Stock Segmentation to Achieve Community Energy Savings

Rachel Scheu, Emily Robinson, and Dr. Anne Evens, Elevate Energy
Published by ACEEE
Pacific Grove, CA
August 17, 2014

ABSTRACT

Municipalities require robust energy and building data in order to meet energy efficiency and emissions reduction goals, but these data can also be used to inform and achieve broader community goals such as those related to affordable housing or economic development. These data need to be tailored to each community's unique building stock to most effectively help policymakers prioritize investments and resources. Despite increasing demand for and access to housing and energy data at the aggregate and building level, there is still no broad understanding of how particular homes types use energy differently.

This paper summarizes the replicable methodology for characterizing single family housing stock and energy use. The analysis combines the measured energy consumption in over 400,000 homes and the property assessor data of one million homes in Cook County, Illinois. Then the paper presents results from an analysis of measured energy use performance and housing characteristics segmented by construction type, age and size. The analysis provides distinct energy outcomes and provides a method to prioritize homes for maximum aggregate energy savings.

The paper then describes how housing segmentation has been applied in the Chicago region in residential energy retrofit programs, adding geo-spatial and census level household income analysis. Last, the paper discusses results and presents recommendations for how measured energy use across populations of homes can be used to examine trends that cannot be observed in single home comparisons, provide clarity in real estate transactions, and have a dramatic impact on how community-scale programs are developed and implemented.

Introduction

In order for municipalities to meet climate and energy efficiency goals it is important to move toward more targeted strategies, addressing the housing stock that is most able to provide deep savings and strategies that scale quickly and efficiently. Policymakers must address the housing stock in geographic clusters, not as individual homes. This paper presents an analysis and a methodology that marries energy use and building characteristics to identify building and community assets and opportunities. Municipalities and energy efficiency portfolios could use this approach to employ a comprehensive community-based approach toward energy and housing. Policymakers can use this approach to identify key housing types and geographies to develop outreach and efficiency programs more effectively. These methods can be used to inform community investment strategies.

Background

There are over 3.3 million homes in the Chicagoland region, which represents 63 percent of the population of homes in Illinois. This analysis focuses on Cook County, where there are nearly two million households and 1.1 million single family homes. The region's extreme weather—long, cold winters and hot and humid summers—combined with the older housing stock makes homes in this region very energy intensive. This analysis builds off of previous research that segmented half a million homes in the region with home characteristics and energy use (Spanier et al 2012.) There is significant need for energy efficiency strategies that can be deployed at scale. It is well documented that individual home energy use varies considerably, even for identical home types (Lutzenhiser et al. 2012; Jones, Taylor, Kipp 2012; Burns and Scheu, 2012; White et al. 2014). And, analyzing a population of homes removes the effect of outliers (Polly et al 2011). Other regions and states, such as Connecticut and Florida have explored complementary approaches (Jones, Taylor, Kipp 2012; NMR Group, Inc 2012.) This approach is also highly relevant to areas where residential energy disclosure is already mandated.

In the next sections, the authors explain the segmentation methodology to characterize the single family housing stock of Cook County, Illinois by energy use and the results of the analysis. It is important to note that the approach and results described here rely on measured energy consumption data obtained from utilities and home characteristics obtained from the property assessor. Therefore, due to data availability, there is no attempt to explore energy potentials, including renewables. However, other researchers are developing frameworks to explore renewable energy potential (Omitaomu, et al. 2012).

Housing Segmentation Methodology & Results

The segmentation methodology presented is intended to be replicable. Each region has variable yet distinct home characteristics (construction type, vintage, and size), weather, and energy regulatory environments that influence how much energy a home consumes and upgrades that will have highest savings potential. This replicable approach can have dramatic impact on how broad-scale retrofit implementation programs are developed, implemented, and brought to scale.

Step 1 - Data Collection: Energy and Housing Characteristics

The researchers collected data from the three utilities serving the Chicago metropolitan region. ComEd, the electric utility for northern Illinois, provided data for all residential accounts by address, including monthly kWh usage data. Integrys, the owner of Peoples Gas in Chicago and North Shore Gas in parts of Cook and Lake Counties, provided data for all residential accounts by address, including monthly therms usage data. Nicor Gas, the natural gas utility for northern Illinois that is not covered by Integrys, provided data for all residential accounts by address, including annual therms usage data.

The researchers collected data on residential housing stock from the Cook County Assessor. The researchers focused on Cook County because the assessor data on the residential housing stock is comprehensive and robust. The Cook County property assessor collects 35 data points for each home, including housing characteristics that contribute to energy usage such as

exterior construction, type of attic, basement, heating and cooling systems. Table 1 below presents a snapshot of the data collected by the Cook County Assessor.

Table 1: Cook County, IL property assessor residential data

Variable	Description
PIN	13-digit unique identifier
Address	
City	Mailing city
ZIP	5-digit zip code
Township	Assessor township within Cook County
Assessor class	Class is based on age, square footage, and number of units
Number of units	Number
Square footage	Measured as finished space
Year built	
Bedrooms	Number
Bathrooms (full)	Number
Bathrooms (half)	Number
Exterior Construction	Type of exterior construction
Roof	Type of roof construction
Basement	Type of basement
Attic	Type of attic
Heating System	Type of heating system
Air Conditioning	Type of air conditioning system
Fireplace	
Garage	Number of spaces available
Garage (exterior construction)	Exterior construction of garage

Step 2 - Data Cleaning and Manipulation

The researchers separated baseload energy use from heating and cooling load energy use using a regression model based on heating and cooling degree days. The heating and cooling energy use was then weather normalized[1] to enable accurate comparisons with future energy use or other normalized models. Other normalizations, such as for occupancy, were not performed due to data being unavailable at the household level.

Housing characteristics, electricity usage, and natural gas usage data were compiled and matched by address to create a complete dataset by individual home. Because of the unique address formatting styles particular to each dataset (ex. Ave vs. Avenue), address cleaning was required. Cleaning was performed through a series of Perl scripts to identify, split, and extract complex character sequences. The output files from the cleaning process were imported into a MySQL database and queries to that database were used to match addresses across data sources. The addresses were also geocoded for GIS analysis.

[1] Chicago has a 30 year average of 6,362 HDD and 827 CDD. Source: http://cdo.ncdc.noaa.gov/cgi-bin/climatenormals/climatenormals.pl

Homes with missing home characteristics or energy data were excluded from analysis. Outliers were removed using the following criteria:

- Missing energy use data. If a home was missing data or had zero usage for at least one utility, it was excluded from analysis
- Missing home characteristics data. If a home was missing data for square footage, year built, or number of stories, that home was excluded
- Homes with assessor codes 211 and 212 were deleted, to exclude
    - Two to Six Apartments, Over 62 Years
    - Mixed commercial/residential building, 6 units or less, square footage less than 20,000
- Homes were sorted into their groups and ranked into percentiles within their groups for each of therms, kWh, EUI, and square footage
    - Homes in the bottom 1 percent for therms, kWh, or EUI were excluded
    - Homes in the top 1 percent for therms, kWh, EUI or square footage were excluded

The Population

Datasets: 919,103 homes were included in the housing characteristics dataset after eliminating missing data and outliers. 423,956 homes were included in housing characteristics + energy use analysis after eliminating missing data and outliers. A summary of the characteristics of the two general populations are presented below.

Table 2: Characteristics of the housing characteristics general population

Variable (n=919,103)	Mean	Std Dev	Median	Interquartile Range
Square footage	1567.6	795.7	1313.0	728.0
Year built	1954.6	27.5	1957.0	30.0

Table 3: Energy use characteristics of the energy + housing characteristics general population

Variable (n=423,956)	Mean	Std Dev	Median	Interquartile Range
Energy Use Intensity (kbtu/sf/yr)	132	45	126	57
kBtu	173,806	53,806	165,625	64,141
Therms	1,411	460	1,338	550
kWh	9,588	4,606	8,761	5,539

Step 3 - Segmenting the Housing Groups

The researchers explored three segmentation methodologies to create the housing groups: statistical clustering, construction characteristics, and heating characteristics. After evaluating output results using the three methodologies, the researchers chose segmentation by construction characteristics to define the housing groups. The housing groups are based on a three tiered system of characteristics, which were chosen by contribution to energy use:

- Tier 1) Construction type – either masonry or frame (includes frame and “frame and masonry” and stucco houses)

- Tier 2) Year built: based on three categories: Pre-1942; 1942-1978 (1978 is the year that many building codes changed to include insulation), and Post-1978
- Tier 3) Number of stories; 1 to 1.9 stories, and 2 to 3 stories

Number of stories was used to segment groups as proxy for home size. The assessor categorizes stories as the number of floors above ground level with finished space. This includes 1 to 1.9 stories, split-levels, and 2 to 3 stories classifications. For example, the vernacular Chicago Bungalow, built with an unfinished attic on the second floor with a pitched roof and knee walls, is classified as a one-story building in the assessor data.

Step 4 – Geocode Homes for Geographic Analysis

Homes were geocoded for geographic information system (GIS) analysis, then mapped by housing group to locate clusters of homes by municipality and neighborhood clusters. This is described in more detail in a subsequent section.

Results of Housing Segmentation: 15 Housing Groups

Using the segmentation methodology above, we initially segmented the Cook County single family home population into 18 housing groups. Three similar groups were consolidated due to small percentage of the population for a total of 15 housing groups. The 15 groups defined represent homes with similar energy use and construction characteristics. See Table 4 below.

Table 4: Cook County housing stock segmented into 15 groups

Construction	Year	Stories	Group	N Obs[2]	% of Population	Therms Median	kWh Median
Masonry	Post-1978	< 2 (no split level)	1	12,078	1.3%	1173	8962
Masonry	All	split level	2	18,507	2.0%	1445	8729
Masonry	Post-1978	≥ 2	3	30,629	3.3%	1775	15530
Masonry	1942 - 1978	< 2 (no split level)	4	176,352	19.2%	1229	7849
Masonry	1942 - 1978	≥ 2	5	42,009	4.6%	1412	9036
Masonry	Pre-1942	< 2 (no split level)	6	67,594	7.4%	1583	8396
Masonry	Pre-1942	≥ 2	7	20,437	2.2%	1978	11220
Frame	Post-1978	< 2 (no split level)	8	14,864	1.6%	1033	8132
Frame	Post-1978	split level	9	20,199	2.2%	1173	9527

[2] Based on housing characteristics dataset n = 919,103

Frame	Post-1978	≥ 2	10	80,557	8.8%	1402	11881
Frame	1942 - 1978	< 2 (no split level)	11	168,731	18.4%	1163	8201
Frame	Pre-1979	split level	12	59,543	6.5%	1297	9035
Frame	1942 - 1978	≥ 2	13	61,873	6.7%	1507	11082
Frame	Pre-1942	< 2 (no split level)	14	100,910	11.0%	1428	8390
Frame	Pre-1942	≥ 2	15	44,820	4.9%	1740	10476

Nine of the 15 housing group homes have gas use that is greater than the county median of 1,338 annual therms. Eight of the 15 housing group homes have electricity use greater than the county median of 8,761 annual kWh.

The researchers calculated the interquartile ratios of energy use for the housing groups. There was an approximately 60 kBtu/sf/year (EUI) spread between the 25th and 75th percentiles. The interquartile ratio for gas = 550 therms. The interquartile ratio for electricity = 5,539 kWh. These are important findings because the results describe the population by fuel use, but also characterize the fuel use of fifteen distinct populations of similar homes. Once characterized an individual home can be compared to similar homes. Small, old homes can be compared to other small homes; and large, new homes can be compared to other large new homes, regardless of where they are located across the county. The results describe what 'typical' looks like, and how typical differs by home type. The researchers removed the effect of outliers to describe the energy use of 15 distinct single family home populations.

The following section describes the geographic distribution of the housing groups and how the housing group segmentation and characterization methodology has been applied in Cook County to inform energy efficiency program delivery.

Housing Segmentation Applied in the Chicago Region

Large datasets allow municipalities and program implementers to approach housing and energy comprehensively. These data should be used to provide strategic outreach by geography and by housing type in order to target homes and communities for greatest savings potential. This section describes how these data are being applied in the Chicago region to identify homes to target, and how this approach can be used in energy programs at a local level. The City of Chicago is the largest municipality in Cook County. The researchers recommend different strategies for the County as a whole (inclusive of Chicago), as well as a Chicago-only strategy. The researchers identified six home types to target in Cook County, and four home types to target in the City of Chicago, though three home types are common to both the County and the City. Both of these strategies are discussed below.

Target Six Home Types in Cook County

The 15 housing groups represent a total of 919,103 single family homes in Cook County. Comprehensive energy and housing programs should focus on six home types in the County. These six most prevalent housing groups represent 71.4 percent of homes in the County. See Table 5 below.

Table 5: Six home types represent 71.4% of the population of homes in Cook County, IL

Construction	Year	Stories	Group	N Obs	% of Population	Therms Median	kWh Median
Masonry	1942 - 1978	< 2 (no split level)	4	176,352	19.2%	1229	7849
Masonry	Pre-1942	< 2 (no split level)	6	67,594	7.4%	1583	8396
Frame	Post-1978	$\geq$ 2	10	80,557	8.8%	1402	11881
Frame	1942 - 1978	< 2 (no split level)	11	168,731	18.4%	1163	8201
Frame	1942 - 1978	$\geq$ 2	13	61,873	6.7%	1507	11082
Frame	Pre-1942	< 2 (no split level)	14	100,910	11.0%	1428	8390

Geo-spatial analysis was performed on each housing group to identify clusters of homes. One home type is illustrated as an example. The map below (Figure 1) shows the geographic distribution and density of the most prevalent home type in the county: *Masonry 1942-1978 less than 2 stories* by municipality/neighborhood. There are 176,352 of these type masonry homes in Cook County, representing 19.2 percent of County's population. The construction of these homes followed common metropolitan land use development patterns. These mid-century brick homes were built on the outskirts of the City of Chicago, inner ring suburbs, and along transportation corridors (in this case expressways and rail corridors.) Hypothetically, if an energy efficiency program were to target this home type only, it would have a significant market sample. But the geo-spatial analysis allows for more specific targeting. Approximately 25 percent of these 176 thousand homes are clustered in 10 suburban municipalities or Chicago Community Areas. The Ashburn community in Chicago, for example, has more than 7,000 of these homes in that neighborhood alone. See Table 6 & Figure 1. These community level analyses allow programs to launch a focused community level outreach and marketing strategy.

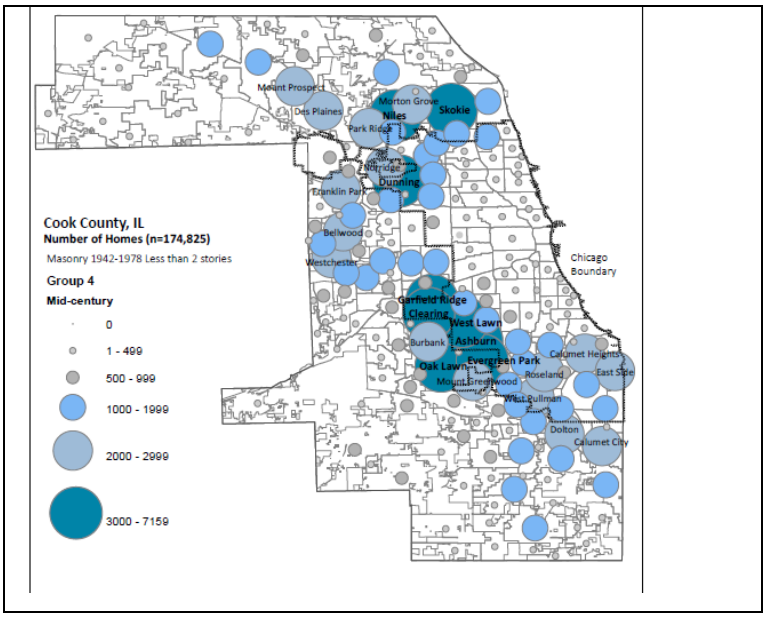

Figure 1: Geographic clustering and density of masonry homes built between 1942-1978 less than 2 stories in Cook County, IL (group 4)

Figure 2: Photo of typical masonry home built between 1942-1978 less than 2 stories in Cook County, IL (group 4)

Table 6: Highest community prevalence of masonry homes built between 1942-1978 less than 2 stories in Cook County, IL (group 4)

Municipality/Community Area	Sum of Group 4 Homes	SF Homes in Municipality/Area	% Of Municipality
Ashburn	7159	11777	61%
Skokie	6663	14913	45%
Garfield Ridge	4898	10967	45%
Norwood Park	4633	11172	41%
Oak Lawn	4245	17641	24%
Evergreen Park	3888	7620	51%
West Lawn	3527	6620	53%
Dunning	3213	10338	31%
Clearing	3118	6161	51%

Niles	3050	7789	39%
sum	44394	104998	42%

Target Four Home Types in Chicago

The 15 housing groups represent a total of 299,373 single family homes in the City of Chicago. Comprehensive energy and housing programs should focus on four home types in Chicago. These four most prevalent housing groups represent 69.6 percent of homes in Chicago. In addition to prevalence, we chose home types that have significant energy savings opportunities as a group. These home types have either higher than area median gas use, higher than area median electricity use, or higher than area median energy use intensity (kBtu/sf/year). See Table 7 below. Within their housing group, efficiency programs can also prioritize high users. An analysis performed on 300 historic Chicago bungalows homes (group 6) found that the higher gas use homes pre-retrofit saved more than those homes whose gas use was lower than the median. (Bailey and Scheu 2013).While this finding seems intuitive, these large, local datasets are valuable to provide energy use information in context with similar home types.

Table 7: Four home types represent 69.6% of the population of homes in Chicago, IL

Construction	Year	Stories	Group	N Obs	% of Population	Median	Median
Masonry	1942-1978	< 2 (no split level)	4	70,478	23.5%	1248	7849
Masonry	Pre-1942	< 2 (no split level)	6	47,899	16.0%	1611	8396
Frame	1942-1978	< 2 (no split level)	11	24,687	8.2%	1257	8201
Frame	Pre-1942	< 2 (no split level)	14	65,332	21.8%	1465	8390

Similar to the results from the County, we see geographic concentrations and clustering of the prevalent home types in Chicago.

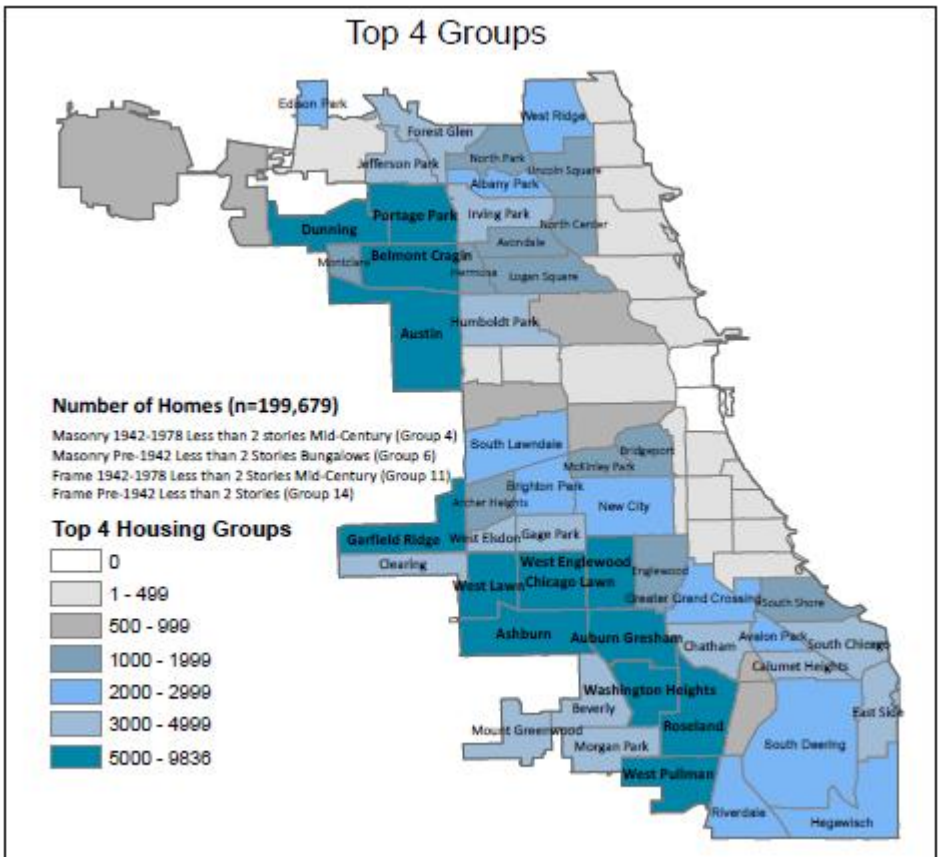

Figure 3: Geographic clustering and top four target homes in Chicago, IL

Census Tract Income Data with Housing Characteristics

There are many applications for applying housing characterization at a community scale. For example, the researchers performed an analysis of the Illinois Home Performance (IHP) program participants in Cook County[3]. They characterized 504 IHP homes into the 15 housing groups. The analysis determined that IHP has had the largest participant numbers near large cities and in census tracts with low-to-moderate median household incomes. Overall, 93 percent (n=469) of the IHP homes in Cook County are in the middle three income census tract quintiles ($20,000 to $100,000) annual household income. These findings support Lawrence Berkeley National Laboratory recommendations to target middle income households (Zimring et al 2011).

Discussion and Recommendations

The ability to describe both the structural characteristics and energy use of a population of homes helps policymakers make informed housing decisions. In this section, the researchers present recommendations for how measured energy use across populations of homes can be used

[3]The analysis was completed for Midwest Energy Efficiency Alliance for US DOE Building America project. The analysis is referenced in the not yet published report: Yee, S., Milby, M., & Baker, W.
DRAFT Technical Report: Evaluation of Missed Energy Saving Opportunity Based on Illinois Home Performance (IHP) Program Field Data: Homeowner Selected Upgrades vs. Cost-Optimized Solutions."

to streamline audits and reduce the cost of a retrofit, target limited efficiency resources to housing that is in most need of and can provide the most energy savings, and provide clarity in real estate transactions. By approaching housing and energy together, municipal goals can be achieved resulting in a more comprehensive approach to community investment.

Streamline Audits and Reduce Retrofit Costs

Combining housing segmentation with measured energy data analysis for a large population of homes reduces the effect of outliers and can significantly reduce program costs by enabling modeling of a population of homes (Polly, Kruis, Roberts 2011). Although customization of retrofit packages is required to retrofit an individual home, being able to use data based on the population helps service providers more quickly design retrofit packages that deliver cost-effective energy savings (Spanier et al 2012; Jones, Taylor, Kipp 2012; Gates and Neuhauser 2014). This approach will have the most impact if done at a geographic scale large enough that it includes utility service territories, or municipal or regional geographies.

Target Scarce Resources Most Efficiently

Housing characterization can be used to inform community investment strategies. Policymakers can use this approach to identify key housing types and geographies to design efficiency programs and to target outreach strategies more effectively. The ability to describe both the structural characteristics and energy use of a population of homes helps policymakers make informed decisions about the residential housing stock. These methods can be used to inform community investment strategies. Housing and energy segmentation can be used to identify housing types with high energy use, and can be used to target homes that will have the highest savings potential. For example, Chicago bungalows are an important component of Chicago's residential infrastructure and its vernacular architecture. They are prevalent and they are energy intensive; they are 10 percent more energy intensive per square foot than other single family homes in Chicago and approximately 25 percent more energy intensive per square foot than the median home in the region (Spanier et al. 2012). Per the results of the bungalow analysis that found that the higher gas use homes pre-retrofit saved more (Bailey and Scheu 2013), programs such as the Historic Chicago Bungalow Association could target its highest users to achieve deeper savings. In addition to focusing on specific housing types, policymakers may focus resources on specific geographies/communities in order to meet other strategic goals.

Provide Clarity in Real Estate Transactions

In July 2013, the City of Chicago became the first municipality in the country to disclose residential energy costs when a home is listed for sale via a multiple listing service (MLS). When a home is listed for sale on the MLS serving Chicago, REALTORS can access natural gas and electricity costs for the property from an online, third-party database, in near real-time. This first-of-its-kind disclosure happens during the real estate selection process in the MLS so buyers can review energy cost information *before* purchase.

Using the Chicago bungalow again as an example, certified historic bungalows are a searchable field in the region's MLS. The energy cost disclosure allows potential buyers to view energy costs and an optional energy use report. The bungalows' median annual therm use was 1653, with an interquartile range of 500 therms. In the future, a potential buyer could be able to view a home and understand how that bungalow compares to other bungalows, similar to the miles-per-gallon comparison for automobiles. This information broken out by housing type can play an important role in a prospective buyer's decision making.

Conclusion

This paper summarized a replicable methodology for characterizing single family housing stock and energy use and presented the results of an analysis that combined the measured energy consumption in over 400,000 homes and the property assessor data of one million homes in Cook County, Illinois. The analysis provided distinct energy outcomes and recommends replicating this methodology in order to achieve for maximum aggregate savings at the community level.

The authors presented recommendations for how measured energy use across populations of homes can be used to examine trends that cannot be observed in single home comparisons, provide clarity in real estate transactions, and have a dramatic impact on how community-scale programs are developed and implemented, and give examples of how this approach has been applied in the Chicago region in residential energy retrofit programs by adding geo-spatial and census level household income analysis.

There are many metrics that could be used to demonstrate the future impact for communities who employ this approach. A significant and quantifiable metric would be a shift of the median energy use of a population of homes and to narrow the interquartile ratios of energy use community. Municipalities could set different targets for each home type, depending on cost and prevalence, or perhaps advocate for variable cost effectiveness thresholds (TRCs) by housing type. Other impacts could include: quantifying the real estate sales premium for energy efficient homes; these data being used for appraisal comps; greenhouse gas reduction goal progress; the number of eco-districts utilizing as a residential community development/investment strategy. This replicable approach could have dramatic impact on how broad-scale energy efficiency retrofit programs are developed, implemented, and brought to scale.

References

Bailey-Burns, E. and R. Scheu. 2013. "400 Retrofitted Bungalows: An Energy Use Analysis." Presented at 2013 ACI National Home Performance Conference. http://aci.sclivelearningcenter.com/index.aspx?PID=6963&SID=177370

Gates, C., and K. Neuhauser. 2014. "Performance Results for Massachusetts & Rhode Island DER Pilot Community." Washington, DC: U.S. Department of Energy.

Jones, P., N. Taylor, and J. Kipp. 2012. "Housing Stock Characterization Study: An Innovative Approach to Measuring Retrofit Impact." Washington, DC: U.S. Department of Energy.

Lutzenhiser, L., H. Hu, M. Moezzi, A. Levenda, J. Woods. 2012. "Lifestyles, Buildings and Technologies: What Matters Most?" In *ACEEE Summer Study on Energy Efficiency in Buildings*, 2: 256-270. Washington, DC: ACEEE.

Omitaomu, Olufemi A., et al 2012. "An Integrated Geovisual Analytics Framework for Analysis of Energy Consumption Data and Renewable Energy Potentials." Oak Ridge National Laboratory (ORNL); Center for Computational Sciences.

NMR Group, Inc. 2013. "Characterization of Single-Family Housing Stock in the State of Connecticut." http://www.energizect.com/sites/default/files/R6%20FINAL%20CT%20Housing%20Characterization%20study%20061213.pdf

NOAA (NOAA Satellite and Information Service). "U.S. Climate Normals." Accessed March 6, 2014. http://cdo.ncdc.noaa.gov/cgi-bin/climatenormals/climatenormals.pl

Polly, B.; Kruis, N.; Roberts, D. 2011. Assessing and Improving the Accuracy of Energy Analysis for Residential Buildings. DOE/GO-102011-3243. www.nrel.gov/docs/fy11osti/50865.pdf

RESNET (Residential Energy Services Network). "Home Energy Audits FAQs." Accessed March 6, 2014. https://www.resnet.us/home-energy-audits-faqs

Spanier, J., R Scheu, L. Brand, and J.Yang 2012." Chicagoland Single-Family Housing Characterization." Golden, CO: National Renewable Energy Laboratory, NREL/SR-5500-55089.

White S., L. Potter, J. Jordan, L. Valencia, C. Valenzuela. 2014. "Residential Energy Consumption Using a Housing Unit Archetype Framework" Presented at the 2014 Applied Demography Conference. https://idserportal.utsa.edu/Resources/Presentations/IDSER/2014/ADC/ADC2014_7B_Potter.pdf

Zimring, M., M. Goggio Borgeson, I. Hoffman, C. Goldman, E. Stuart, A. Todd, and M. Billingsley. 2011. “Delivering Energy Efficiency to Middle Income Single Family Household.” Berkeley, CA: Lawrence Berkeley National Laboratory.