Not peer-reviewed version
Not peer-reviewed version

Modeling Social Equity in Energy Consumption Using Digital Twins

Aparna Kishore , Beatrice Li , Jad Atweh [*]

Keywords: Social Equity; Residential Energy Modeling; Digital Twins; Engineering

Article

Modeling Social Equity in Energy Consumption Using Digital Twins

Aparna Kishore [1], Beatrice Li [2] and Jad A. Atweh [3,*]

[1] Department of Computer Science, University of Virginia Charlottesville, VA, USA; ak8mj@virginia.edu
[2] Department of Systems Engineering, University of Virginia Charlottesville, VA, USA; blx2wj@virginia.edu
[3] Department of Systems Engineering, University of Virginia Charlottesville, VA, USA
* Correspondence: jad.atweh@virginia.edu

Abstract: This research examines the impact of social equity on energy consumption. We constructed a digital twin for residential energy consumption by enriching the synthetic population with real-world surveys and feeding them with other environ- mental and appliance data to the energy modeling framework. We analyzed household hourly energy consumption data from Albemarle County and Charlottesville City in Virginia, USA, for the year 2019. We used clustering analysis to identify patterns in social equity and energy consumption. The results demonstrated the impact of different residential attributes on energy poverty. Statistical analyses, including ANOVA and Chi-Squared tests, were conducted to test for significant differences between racial groups in quantitative and categorical variables. The study found that race is significant in determining the location and quality of housing. People of color often live in areas with higher pollution and less access to green spaces. Additionally, income levels and the age of the house are influential factors in determining energy efficiency. Future work should focus on collecting and analyzing data at the country level and using qualitative data collection methods to gain a more comprehensive understanding of social equity issues concerning energy consumption. Overall, this study provides valuable insights into the relationship between different residential attributes and energy consumption, which can inform policy development to promote more equitable and sustainable communities.

Keywords: social equity; residential energy modeling; digital twins

I. INTRODUCTION

Climate change is an existential threat. Many countries have pledged to transition to green and sustainable energy by 2050. This step ensures the global temperature rise does not exceed 1.5 2.0 degree C. Among many factors contributing to climate change, carbon emissions account to 60% of humanity's overall ecological footprint. In the United States, contributions from residential energy alone are projected at 17% of annual greenhouse gas emissions (GHG). Reducing energy demand in the residential sector is one of the crucial factors for achieving the goal of reducing GHG emissions. One of the ways of reducing residential consumption is retrofitting building stock.

Retrofitting building stock helps reduce energy consumption and provides a safe and healthy place for people to live. Many works have stressed the importance of retrofitting building stock to mitigate climate change, address energy poverty and inequity, and reach net-zero emissions [15]. Based on U.S. Energy Information Administration (EIA) [9], 1/3rd of the U.S. lives in energy poverty, and 11% keep home in unhealthy conditions just because they cannot afford to pay their electricity bills. Although we have federal policies and incentives for retrofitting, these numbers look scary.

The existing policies through federal grants and programs focus on extremely low-income households to reduce energy consumption. Other policies, such as sustainable energy loans, benefit higher-income households who can afford monthly mortgage payments. However, such strategies fail to identify and design policies for fairness and social equity. Researchers have emphasized the

importance of social equity in residentialenergy [1–3]. Some of the demographic attributes suggested in these research works are race, income, and area type.

Although the importance of social equity is undeniable,federal programs struggle to develop policies based on insightsdue to a lack of understanding of residential energy consumption at a higher granular level. In addition, there is necessary touncover social and economic demographic patterns in energy poverty at a larger scale. Combating these issues requireshigh-resolution data at the household-appliance level while addressing privacy concerns and confidentiality aspects. Adigital twin is a one-stop solution to solve this problem [16]. A digital twin is a virtual model that can be used to understandthe behavior of the real world. These models are developed using real and synthetic datasets spreading across spatial and temporal domains. Digital twins are used in various domains, such as the aviation industry, medical domain, and so on,to run simulations on a virtual platform instead of running them directly in the real world and causing havoc. Here, the synthetic population acts as a proxy for the real population.

Our contribution to this paper is as follows:

- A digital twin development of synthetic residential energyconsumption data for Albemarle-Charlottesville region.
- Analysis and modelling of social equity constraints to understand the impact of equity in energy poverty throughcluster analysis and statistical analysis.

Paper organization. First, we introduce our methodology in Section II. Second, we provide our results and insights in Section III. Next, we discuss our limitations and future workin Section IV. Finally, we pitch our discussion and concludein Section V.

II. METHODS

Our methodology is divided into three steps: (i) Data Extraction, (ii) Cluster Analysis, and (iii) Statistical Analysis as in Figure 1.
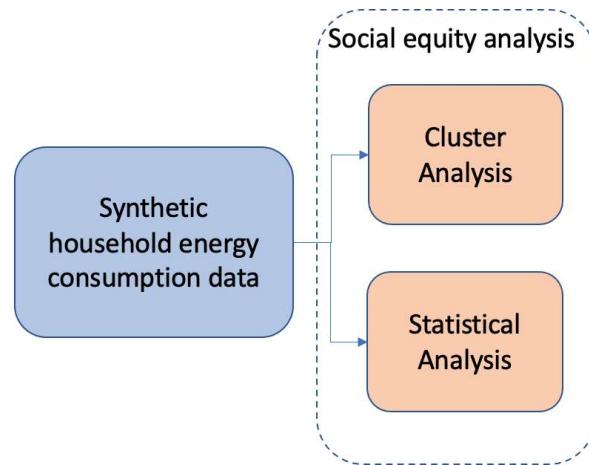
Figure 1. Hourly residential energy consumption data at the household level is used to develop insights on social equity constraints through cluster analysis and statistical analysis.

A. Data Extraction

We begin by collecting the data for our analysis as in Figure 2. Here, we use the digital twin of the US populationfrom the Network Systems Science and Advanced Computing (NSSAC) synthetic population [8]. The base population of the digital twin is developed using a 5% sample of completerecords from Public Use Microdata Sample (PUMS) [11]. In this work [8], each person P from household H in a residential location R_A is assigned an activity sequence A from the list of activities, where each activity has a type T_A, start time$s$, duration d and activity location L_A using two-stage fitted value method [5]. The activity model is constructed based on the National Household Travel Survey (NHTS) [12], theMultinational Time Use Study (MTUS) [6], and the AmericanTime Use Survey (ATUS) [10]. We use the base populationdata of the household along

with the activity information that captures the demographic attributes, household characteristic features, and in-house activity. This information also provides the occupancy patterns in the home.

Next is the enrichment step for energy information, where we map the attributes from the Residential Energy consumption survey (RECS) [9] to the synthetic population. This survey consists of real samples representing the US population for household characteristics and attributes, appliance information, and other energy-related information. The problem formula- tion for this mapping is given in Problem 1.

Problem 1: Let H be the set of households in the syn- thetic population. Let S and R be the set of RECS survey households. Let $A_{iv} = \{a_1, a_2, \ldots, a_k\}$ represent the common demographic and household attributes between S and R. The objective is to enrich S with appliance and residential at- tributes from R, $A_{dv} = \{b_1, b_2, \ldots, b_j\}$, where the distribution of A_{iv} closely matches between S and R.

The mapping between S and R helps to enrich S with appliance information and residential attributes from R, where the distribution of A_{iv} is similar between S and R. The common attributes A_{iv} are called the independent variables, and the mapped appliance information and residential at- tributes, A_{dv}, are called the dependent variables. We enrich S by using multivariate conditional inference trees in R, where the conditional tree is built for multiple independent and dependent variables simultaneously. Conditional inference trees are developed by recursive binary partitioning where the node split is based on the statistical significance between the independent and dependent variables [13].

We modify the energy model by Swapna et al. [4] to develop high-resolution hourly energy consumption data for each household. The energy demand modeling framework consists of individual appliance models that simulate hourly energy use of different end-uses such as HVAC, lighting, refrigerator, hot water use, major appliances such as dishwasher, clothes washer, clothes dryer, and miscellaneous plug load. This model also takes in other information from other datasets such as (i) weather data, (ii) irradiance data, (iii) appliance rating and other information, (iv) insulation levels, (v) water temperature, and (vi) water usage data. Depending upon household occupancy and behavior, energy use is generated hourly for each household-level appliance.

B. Cluster Analysis

The problem formulation for the first set of analysis is given in Problem 2.

Problem 2: Let H be the set of households in the enriched synthetic population. Let E be the corresponding energy consumption values and let $A = \{a_1, a_2, \ldots, a_k\}$ represent the demographic and household attributes for H. The objective is to group H based on similarities in E and A to identify the set of attributes A_e, where $A_e \subseteq A$, that contribute to energy poverty.

To extract insights from the data, we explored various cluster-ing algorithms, including K-prototype and HDBSCAN. Prior to clustering, an exploratory analysis was conducted with Principal Component Analysis (PCA) to see if there is enough merit in using Principal Components since the ratio of features($f = 33$) to observations (n=58,283) is more than satisfactory in mitigating the curse of dimensionality. As there were no clear loadings on the principal components, we decided not to use principal components in reducing the dimensionality of the data as no additional value was added.

K-prototype was the clustering algorithm decided on due to its capability in handling numeric and categorical data. The algorithm is an extension of the popular K-means algorithm by utilizing a matching dissimilarity measure for categorical data.
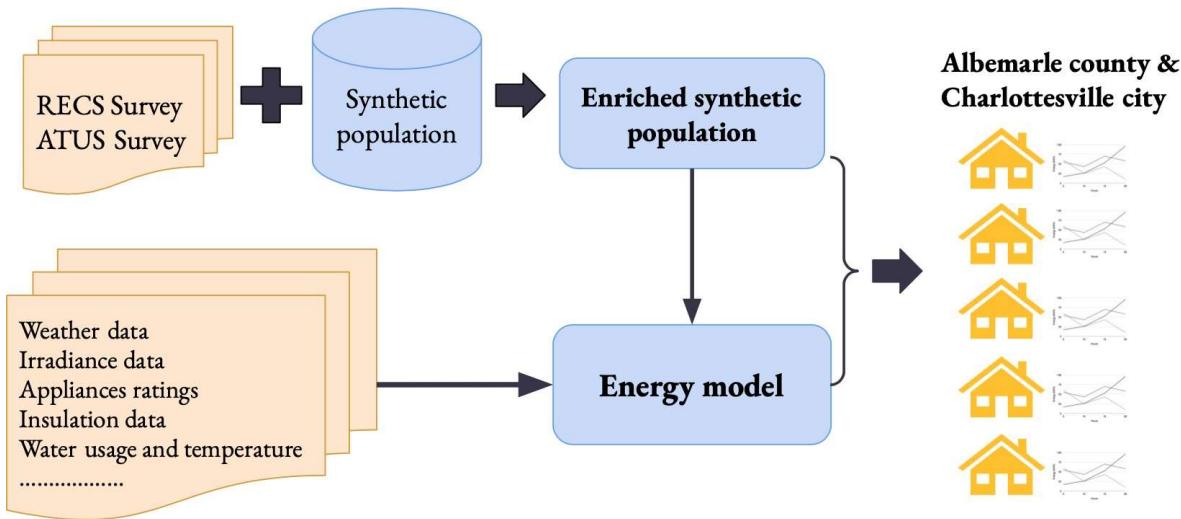
Figure 2. Framework to extract hourly residential energy consumption data at the household level. For any given county and date, the enriched syntheticpopulation, along with environmental and appliance data, is fed to the energy model to produce highly granular synthetic residential energy consumption data.

We conducted the Elbow method and calculated the Calinski- Harabasz index as a clustering validation metric to determine the optimal number of clusters k. Repeated analysis of the original data and subsets of the data yielded optimal cluster values of 3 and 4. By clustering households based on their energy consumption and a set of household attributes, we can identify patterns and characteristics that contribute to energy poverty. The k-prototype algorithm is well-suited for this task as it allows us to capture the complex relationships between household attributes and energy consumption. The resulting clusters can provide insights into the key factors contributing to energy poverty.

This can be achieved by using the k-prototype clustering algorithm to identify the clusters and analyzing the attribute distributions within each cluster to identify the attributes that contribute to energy poverty.

C. Statistical Analysis

Problem 3: Let H be the set of households in the enrichedsynthetic population. Let E be the corresponding energy consumption values and let $A = \{a_1, a_2, \ldots, a_k\}$ represent the demographic and household attributes for H. The objectiveis to find the statistical significance of a particular attribute a_i with other attributes $A \setminus a_i$, for understanding social equity.

We conducted multiple statistical analyses to test for significance between race groups in terms of other variablesto shed light on the importance of accounting for social equity moving forward. We included 4 quantitative variables, where we performed one-way ANOVA tests for each, and 6 categorical variables, where Chi-Sqaured tests were performedfor each. All tests were followed by a post-hoc test to see where the statistical difference, if any, lie.

III. RESULTS AND INSIGHTS

A. Data Extraction

We collected daily residential energy consumption data using the framework as described in section II-A. We performedthis simulation for 12 random days, one day from each month for Albemarle County (County code: 003) and Charlottesville City (County code: 540). We have $\approx$ 58, 000 householdsin this region. We then averaged energy consumption from these 12 days in our optimization models. This setup helps us capture the diversity in energy consumption due to seasonal differences.

The total energy consumption data from the energy model from section II-A is the sum of energy consumed by HVAC, refrigeration, water heater, dishwasher, clothes washer, clothesdryer, lighting, and some miscellaneous plug loads (cleaning, cooking, computer, television and so on). Output used foranalysis and modelling is explained in Table I.

The six major household attributes, along with their pro- portions in the collected data, are shown in Figure 3. These attributes, such as race, ownership of the house, dwelling type, household income, age of the constructed house, and square footage of the house, play a crucial role in developing social equity constraints. While 93.5% of the households are White alone and Black/African-American alone, 6.5% of the population in this region is constituted by races belongingto Asians, Native Hawaiian/Pacific Islander, American Indian/Native Alaskan, and multi-household races. The major contributors in house ownership and dwelling type include owned houses (66.2%) and five-plus apartment dwellings (64%), respectively. It is also important to note that 12.8% of households have income less than the Virginia area median indicator (AMI) of $60,000, and 32.7% of houses have aconstruction age greater than 50 years.

Table I. CODE TABLE FOR THE SELECTED OUTPUT COLUMNS USED FOR ANALYSIS.

Variable	Variable description	Response set and explanation
race	Householder race	1- White Alone 2 - Black or African/American Alone 3 - American Indian or Alaska Native Alone4 - Asian Alone 5 - Native Hawaiian or Other Pacific Islander Alone6 - Some Other Race Alone 7 - 2 or More Races Selected
areatype	Census area type	R - Rural U - Urban
ownRent	Owned/rented house	1. - Own 2. - Rent
dwellingType	Type of housing unit	1 - Mobile home 2 - Single-family detached house3 - Single-family attached house 4 - Apartment in a building with 2 to 4 units 5 - Apartment in a building with 5 or more units
hasElectricH2oHeater	Has electric water heater	1. - Yes 2. - No
hasElectricHeat	Has electric heater	1. - Yes 2. - No
hasAC	Has air conditioner	1. - Yes 2. - No
hasCw	Has clothes washer	1. - Yes 2. - No
hasCd	Has clothes dryer	1. – Yes 2. - No
hasDW	Has dishwasher	1. – Yes 2. - No
oldInsulation	poor or not adequately insulated	1. - Yes 2. - No
OldDW	dishwasher ≥ 10 years	1. - Yes 2. - No
OldCW	clothes washer ≥ 10 years	1. - Yes 2. - No
OldCD	clothes dryer ≥ 10 years	1. - Yes 2. - No

totalDailyKWH	total daily energy consumption in kWh	float values
hvacDailyKWH	hvac energy consumption in kWh	float values
refrigeratorDailyKWH	refrigerator energy consumption in kWh	float values
dwDailyKWH	dish washer energy consumption in kWh	float values
washerDailyKWH	clothes washer energy consumptionin kWh	float values
dryerDailyKWH	clothes dryer energy consumption in kWh	float values
hhSize	household size	numeric value
income	household income	numeric value in dollars
houseAge	construction age of house	numeric value
sqft	House square footage	numeric value
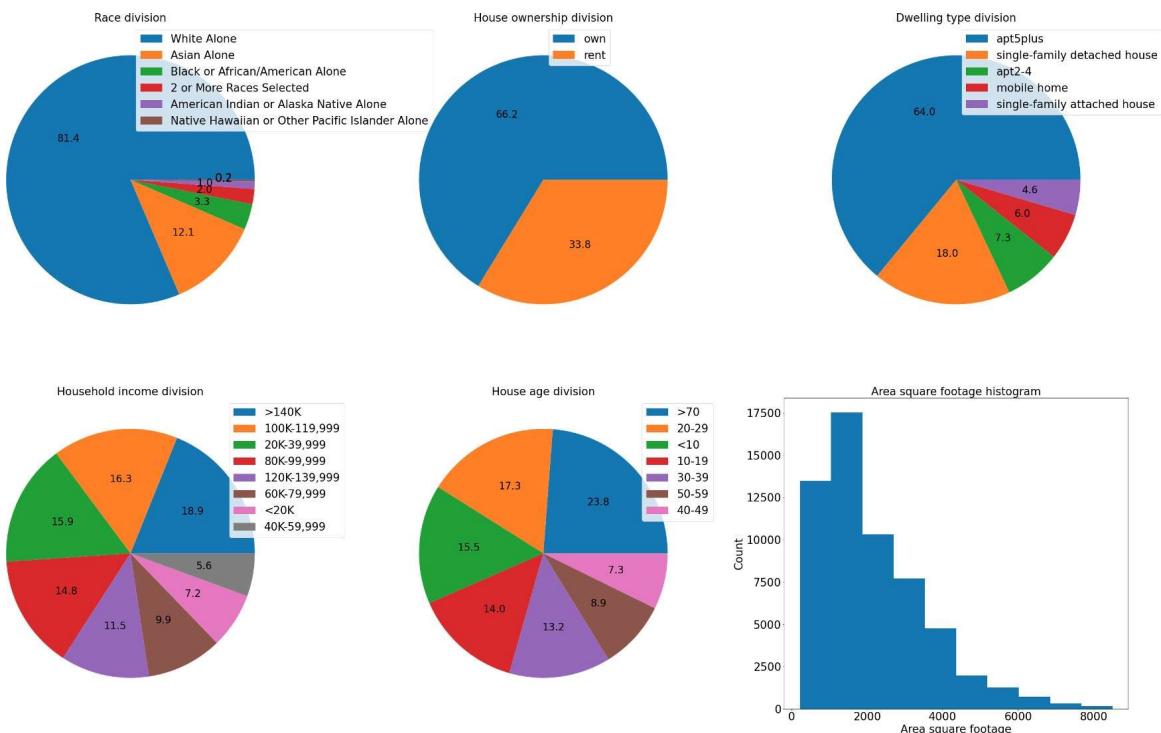
Figure 3. Data distribution in the Albemarle-Charlottesville region for the household attributes such as race, house ownership, dwelling type, household income, house age, and square footage. These attributes are major contributors to consider for social equity constraints. While the first five attributes are demonstratedin a pie chart, the area square footage of the household is shown in a histogram.

B. Clustering

We conducted clustering for the entire dataset with *k=3* clusters where the second cluster emerged as the most energy burdened, as pictured in Table II. The summary table forthe clusters is composed of the mean values of the numeric features and the mode of the categorical features. The first cluster seems to be the most energy-burdened compared tothe others as it has the lowest income bracket (<$20K), and older housing (20-29 years old). It also has less square footage yet about the same HVAC KWH. It should also be noted that the numeric values in Table II are scaled according to the minimum-maximum. However, moving forward, the other table on the later results will not be scaled for easeof interpretability.

It is difficult to extract more conclusive insights into energy-burdened households in the general clustering; thus, we intro- duce specifications for equity constraints.

1) (HouseAge > 20) ∧ (Income < 60k) ∧ (hhSize 2)
2) (HouseAge 20) ∧ (Income   60k) ∧ (hhSize   2)
3) Apartments
   a) (HouseAge > 20) ∧(Income < 60k)
   b) (HouseAge ≥20) ∧(Income < 60k)
4) Single-family Houses
   a) (HouseAge > 20) ∧(Income < 60k)
   b) (HouseAge ≥20) ∧(Income < 60k)

Table II. CLUSTER ANALYSIS OF THE ENTIRE DATASET WITH k=3 CLUSTERS.

Segment	First	Second	Third
Count	21229	24183	12871
race	1	1	1
areatype	U	U	U
ownRent	rent	own	own
dwellingType	5	2	2
hasElectricH2oHeater	Yes	No	Yes
hasElectricHeat	Yes	No	No
hasAC	Yes	Yes	Yes
hasCw	No	No	Yes
hasCd	No	No	Yes
hasDw	No	No	Yes
oldInsulation	No	No	No
OldDW	No	No	No
OldCW	No	No	No
OldCD	No	No	No
income cat	<20K	>140K	>140K
houseAge cat	20-29	<10	<10
totalDailyKWH	0.067	0.098	0.125
hvacDailyKWH	0.116	0.112	0.125
refrigeratorDailyKWH	0.807	0.818	0.818
lightsDailyKWH	0.172	0.377	0.382
dwDailyKWH	0.062	0.134	0.256
washerDailyKWH	0.021	0.003	0.517
dryerDailyKWH	0.015	0.002	0.4
cookDailyKWH	0.127	0.215	0.28
tvDailyKWH	0.136	0.222	0.271
computerDailyKWH	0.028	0.041	0.058
cleanDailyKWH	0.086	0.112	0.208
dailyHotWaterKWH	0.024	0.036	0.05
hhSize	0.068	0.156	0.189
adultNum	0.128	0.179	0.201
sqft	0.114	0.313	0.292

To investigate the extremes of the households in terms of house age and income levels, the first two specifications are drawn up. The first two specifications use 20 years as a bar forhouse age and $60K as a bar for income, where the extremes are greater or less than while constraining the household size to be greater or equal to 2. Specification 3 and 4 aimsto investigate the lower-income households while comparing dwelling type (3-4) and house age (a-b). The specifications

for equity constraints are used to create subsets of the original data for the cluster analysis to be performed on each subsetof data that meets each specification.

Table III shows the clustering results concerning Specifi- cations 1 and 2, which provide further evidence of energy- burdened households. In Specification 1, the majority rentlarge apartments (5 plus units) with a mean income range of $6K to $52K. In comparison, most own single-family detachedhouses with a mean income range of $88K to $512K. The cluster with the lowest mean income bracket comprises a majority of Black or African-American (Specification 1). It also appears that as income increases, the square footage ofthe dwelling also increases.

Through Specifications 3a-b and 4a-b, we consider the lower-income households (less than $60K) and older dwellings (> 20 years) while comparing dwelling types (apartments and single-family houses). A composite metric of the ratio of annual energy consumption to square footage (RAE) was created to evaluate the level of energy consumption better, as shown in Formula 1. In comparing 3a (Single-Family Houses)and 4a (Apartments), the average RAE was 3.163 and 5.131, respectively. The lowest income cluster (in 4a) comprises Black or African-American and White with a mean incomeof $5K but is also found to have the highest total daily energy consumption. At the same time, the second lowest income cluster is found in 3a with a mean income of $8K. It comprises mainly of Black or African-American and Native American or Alaskan Native, respectively.

$$RAE = (totalDailyKWH * 365)/sqft \qquad (1)$$

The final set of specifications considers the lower-income households (less than $60K) and newer dwellings (≤20 years)while comparing dwelling types (apartments and single-familyhouses). In comparing 3b (Single-Family Houses) and 4b (Apartments), the average RAE was 4.039 and 6.613, respectively. The lowest income cluster (in 4b) comprises Whiteand Black or African-Americans with a mean income of $3K, which is even lower than that of 4a. In all four specifications, 3a-b and 4a-b, this cluster has the lowest income and thehighest RAE of 8.225. Regardless of age, the RAE is higher for apartments than single-family houses, but newer dwellings (b) tend to have a higher RAE. The low-income clusters with comparable or higher RAE to the higher-income clustersindicate those clusters as energy-burdened.

Table III. CLUSTER ANALYSIS FOR SPECIFICATION 1 AND SPECIFICATION 2.

	Specification 1				Specification 2			
Segment	First	Second	Third	Fourth	First	Second	Third	Fourth
Count	1832	2532	2652	2436	1658	3924	5694	638
race	2	1	1	1	1	1	1	1
ownRent	Rent	Own	Rent	Rent	Own	Own	Own	Own
dwellingType	5	2	2	2	3	2	2	2
oldInsulation	No	No	No	No	No	No	No	No
OldDW	No	No	No	No	No	No	Yes	No
OldCW	No	No	No	No	No	Yes	Yes	Yes
OldCD	No	No	No	No	No	Yes	Yes	Yes
totalDailyKWH	18.669	18.848	18.844	20.275	23.876	25.228	25.771	24.168
hvacDailyKWH	3.018	3.538	3.942	4.047	4.876	5.708	6.579	5.111
refrigeratorDaily KWH	2.466	2.457	2.461	2.462	2.491	2.489	2.489	2.49
dwDailyKWH	0.357	0.377	0.369	0.427	0.756	0.732	0.723	0.899
washerDailyKWH	0.08	0.099	0.106	0.115	0.144	0.139	0.131	0.174
dryerDailyKWH	0.458	0.584	0.637	0.691	0.901	0.878	0.834	1.103
hhSize	2.927	2.74	2.715	2.61	3.019	3.075	2.901	3.293
income	6366.96	22541.617	38253.898	52843.805	264532.432	151225.389	88375.719	512357.68
houseAge	47.525	48.116	49.235	48.779	12.754	12.816	13.156	12.859
sqft	1181.878	1449.487	1639.064	1743.115	3505.091	3428.464	2661.76	3638.34

C. Statistical Analysis

1) *ANOVAs: Four one-way ANOVAs were conducted to test for significant difference between race groups in terms of 4 numeric variables: (1) energy consumption, (2) income, (3) home age, and (4) area. Results* show significant differences between race and each of these variables with a p-value<0.001for all variables.

Tukey's Post-Hoc test was conducted after every ANOVA to test between which races, does the significance lie for a particular variable. Table IV summarizes the results of thefour post-hoc tests conducted. Results show that there is a positive significant difference in all variables 1, 2, 3, and 4 when compared to all other racial groups. This means that people who were identified as white have a high amount of income, larger houses, higher house ages, and they consume more energy than all other races.

2) *Chi-Squared Tests:* Six Chi-Squared Tests were con- ducted to test for significant differences between racial groupsin terms of 6 categorical variables which are (5) Own/rent, (6) Dwelling Type, (7) Old Dishwasher, (8) Old CD, (9) Old CW, and (10) Old Insulation. Results show significant differences between race and each of these variables with a p-value<0.001for all variables.

Post-Hoc tests were also conducted after every Chi-Squaredtest to test between which races, does the significance lie for aparticular variable. Table V summarizes the results of the six post-hoc tests conducted. Results show that White and Native Hawaiian or Other Pacific Islander have the highest percentageof homeownership (70.5% and 68.3% respectively), while Black/African American and Asian have the lowest (54.8% and 53.4% respectively). Moreover, Black/African American have the highest percentage of renters (45.8%), while White have the highest percentage of type 2 dwellings (68%). White have the highest percentage of no old insulation (83.%), while Black/African American have the lowest (77.1%).

Table IV. TUKEY'S POST-HOC TEST RESULTS.

Race	White	Black or African American	American Indian or Alaska Native	Asian	Native Hawaiian or Other Pacific Is-lander	2 or More Races Selected
White		+ sig. difference in variables 1,2,3, and 4	+ sig. difference in variables 1,2,3, and 4	+ sig. difference in variables 1,3,4	+ sig. difference in variables 1,2,3	+ sig. difference in variables 1,2,3
Black or African American			+ sig. difference in variable 3 only	+ sig. difference in variable 4 - sig. difference in variables 2,3	- sig. difference in variable 3	+ sig. difference in variable 4 - sig. difference in variables
American Indian or Alaska Native				+ sig. difference in variable 4 - sig. difference in variable 2	No Significant Differences	+ sig. difference in variable 3
Asian					+ sig. difference in variable 2 - sig. difference in variable 4	+ sig. difference in variable 2,3 - sig. difference in variable 4
Native Hawaiian or Other Pacific Is-lander						- sig. difference in variable 1

Table V. RESULTS OF THE CATEGORICAL VARIABLES.

Race/Variable	Own/Rent	Dwelling Type	Old DW	Old CD	Old CW	Old Insulation
White	Own: 70.5%	68% type 2	No: 76.6%	No: 75.7%	No: 79.3%	No: 83.7%
Black/African American	Rent: 54.8%	45.8% type 2	No: 78.8%	No: 78.3%	No: 81.4%	No: 77.1%
American Indian or Alaska Native	Own: 57.7%	64.1% type 2	No: 77.6%	No: 74.8%	No: 78.2%	No: 80.1%
Asian	Own: 53.4%	44.5% type 2	No: 77.1%	No: 77.6%	No: 80.9%	No: 83.2%
Native Hawaiian or Other Pacific Is-lander	Own: 68.3%	50.4% type 2	No: 81.3%	No: 74.8%	No: 74.8%	No: 78%
2 or More Races Selected	Own: 66.2%	47.7% type 2	No: 79.4%	No: 77.4%	No: 80.6%	No: 80.3%

IV. LIMITATIONS & FUTURE WORK

One limitation to this study is that the data analyzed in this study is limited to 2019, which may not reflect changes that have occurred since the COVID-19 pandemic. Moreover, the COVID-19 pandemic has caused significant disruptions to energy consumption patterns, making the data analyzed during this time period unstable. Future work should consider these changes and analyze post-COVID data when it becomes available.

Additionally, the analysis in this study was based on per-day average, which may obscure important patterns that emerge over longer time periods. Therefore, future work should focus on collecting and analyzing data from all days of each month to better understand social equity trends. Moreover, future work should include annual data to identify longer-term trends in social equity and energy consumption patterns.

This study focused on time series analysis of energy data to understand trends of social inequity. While this approach is useful, future work should also consider qualitative data collection methods, such as interviews or surveys, to gain a more nuanced understanding of social equity issues in relation to energy consumption.

Finally, policymakers should develop policies with specifications that follow equity constraints. This approach would ensure that all households have access to high-quality, energy-efficient housing regardless of their race, income, or age of the house. Such policies could include programs to help low-income households upgrade their homes, incentives for developers to build more energy-efficient housing, and regulations to ensure that all new buildings meet energy efficiency standards. By addressing the impact of race, income, and age of the house on residential attributes and energy consumption, we can work towards a more equitable and sustainable future.

V. DISCUSSION & CONCLUSION

The focus of this research is to investigate the influence of race on residential features and energy consumption. The impact of race and income on residential attributes and energy consumption is a critical issue that requires attention. Studies have shown that race plays a significant role in determining the location and quality of housing, with people of color often living in areas with higher pollution and less access to green spaces. The statistical analyses presented examine the relationship between race and various variables, including energy consumption, income, homeownership, dwelling type, and the presence of old infrastructure in the home. The results suggest that significant differences exist between different racial groups with respect to these variables. The analysis finds that White and Native Hawaiian or Other Pacific Islander individuals have the highest rates of homeownership, while Black/African American and Asian individuals have the lowest. This finding may reflect differences in income, wealth, or access to credit, among other factors, which could affect the ability of individuals from different racial groups to purchase a home. The analysis also examines the presence of old infrastructure in the home, including old dwelling and insulation. The results indicate that White individuals have the highest percentage of homes without old infrastructure in all categories, where Black/African American individuals have the lowest percentage. This finding suggests that individuals from different racial groups may have differing levels of access to and quality of housing infrastructure. Furthermore, income levels play a crucial role in deciding the type of dwelling and ownership of the house, which can have implications for energy consumption. Lower-income households are more likely to live in older, less energy-efficient homes, which can result in higher energy bills and increased carbon emissions. Moreover, the age of the house is also an important factor in determining energy efficiency. Older appliances and insulation are more common in houses that are around 15 years old, which can lead to higher energy usage. However, it is worth noting that many older buildings are refurbished to meet building standards, which can help to improve energy efficiency. Taken together, these results highlight the ongoing importance of addressing disparities in housing access and quality among individuals from different racial groups. Policies aimed at reducing income and wealth inequality, increasing access to credit, and promoting equitable distribution of housing re-sources

may be important steps in addressing these disparities.Through our analysis, we have demonstrated the necessityof considering social equity constraints in developing federal initiatives and policies to support households in energypoverty. These policies not only help the community to havea safer and healthier place to live but also has a greatergood of reducing CO_2 emissions. We have provided insightson race's significant impact on residential attributes, energyconsumption, and income level. The current policies andinitiatives by Federal and state governments to support clean and sustainable green energy are focused on energy-burdened low-income households. One example is by the governmentto advance building codes and standards [14] for climate resilience and lower utility bills. These initiatives are intendedto support building projects through federal funding opportunities, provide incentives for communities for retrofitting,and deploy bipartisan infrastructure law to implement building standards. Another one is the Weatherization Assistance Pro- gram (WAP) initiated by the Department of Energy (DOE), where households below the poverty of 200% are eligible to participate in retrofitting programs. While these policies focuson one of the attributes, the income of households, we need more policies that take care of diverged attributes, including race, income level, dwelling type, house age and ownershipof houses.

In conclusion, while this study provides valuable insights into the relationship between social equity and energy con- sumption patterns, there are limitations that must be addressedin future work. By addressing these limitations, future studies can provide a more comprehensive understanding of social equity issues and develop more effective policies to address them.

References

1. Wang, Qiang, et al. "Racial disparities in energy poverty in the
2. United States." Renewable and Sustainable Energy Reviews 137 (2021):110620.
3. Goldstein, Benjamin, Dimitrios Gounaridis, and Joshua P. Newell. "The carbon footprint of household energy use in the United States." Proceedings of the National Academy of Sciences 117.32 (2020): 19122-19130.
4. Goldstein, Benjamin, Tony G. Reames, and Joshua P. Newell. "Racial inequity in household energy efficiency and carbon emissions in the United States: An emissions paradox." Energy Research & Social Science 84 (2022): 102365.
5. Thorve, Swapna, et al. "High resolution synthetic residential energy use profiles for the United States." Scientific Data 10.1 (2023): 76.
6. Lum, Kristian, et al. "A two-stage, fitted values approach to activity matching." International Journal of Transportation 4.1 (2016).
7. Gershuny, Jonathan. "Multinational time use study." (2013).
8. Mayer, Ines, et al. "The two faces of energy poverty: a case study of households' energy burden in the residential and mobility sectors at the city level." Transportation Research Procedia 4 (2014): 228-240.
9. Swarup, S., and M. V. Marathe. "Generating synthetic populations for social modeling." (2017).
10. U.S. Energy Information Administration. "2015 RECS Survey Data." RECS2015. https://www.eia.gov/consumption/residential/data/2015/ (ac-cessed Mar, 2023).
11. U.S. ATUS Survey."U.S. Bureau of Labor Statistics: American Time Use Survey." ATUS2019. https://www.bls.gov/tus/data/datafiles- 2019.htm (accessed Mar, 2023).
12. U.S. PUMS."PUMS 2017." PUMS. https://www2.census.gov/programs- surveys/acs/data/pums/2017/5-Year/ (accessed Mar, 2023).
13. NHTS."NHTS." NHTS2017. https://nhts.ornl.gov (accessed Mar, 2023).
14. Hothorn, Torsten, Kurt Hornik, and Achim Zeileis. "ctree: Conditional inference trees." The comprehensive R archive network 8 (2015).
15. White house breifing room statement. "Government initiatives". WhiteHouse. https://www.whitehouse.gov/briefing-room/statements- releases/2022/06/01/fact- sheet-biden-harris-administration-launches- initiative-to-modernize-building-codes- improve-climate-resilience-and-reduce-energy-costs/ (accessed Feb, 2023).
16. Wade, Faye, and Henk Visscher. "Retrofit at scale: accelerating capabil-ities for domestic building stocks." Buildings and cities 2.1 (2021).
17. Atweh J. A., Moacdieh N. M., Riggs S. L. (2022). Identifying individual-, team-, and organizational-level factors that affect team performance in complex domains based on recent literature. InProceedings of the Human Factors and Ergonomics Society An- nual Meeting (Vol. 66, No. 1, pp. 1795–1799). Sage Publications. https://doi.org/10.1177/1071181322661213