Leveraging Machine Learning for Equitable Transition of Energy Systems

Enea Dodi [*1] **Anupama Sitaraman** [*1] **Mohammad Hajiesmaili** [1] **Prashant Shenoy** [1]

Abstract

Our society is facing overlapping crises of climate change and systemic inequality. To respond to climate change, the energy system is in the midst of its most foundational transition since its inception, from traditional fuel-based energy sources to clean renewable sources. While the transition to a low-carbon energy system is ongoing, there is an opportunity to make the new system more just and equitable than the current one which is inequitable in many forms. Measuring inequity in the energy system is a formidable task since it is large scale and the data is coming from abundant data sources. In this work, we lay out a plan to leverage and develop scalable machine learning (ML) tools to measure the equity of the current energy system and to facilitate a just transition to a clean energy system.

We focus on two concrete examples. First, we explore how ML can help to measure the inequity in the energy inefficiencies of residential houses at the scale of a town or a country. Second, we explore how deep learning techniques can help to estimate the solar potential of residential buildings to facilitate a just installation and incentive allocation of solar panels. The application of ML for energy equity is much broader than the above two examples and we highlight some others as well. The result of this research could be used by policymakers to efficiently allocate energy assistance subsidies in the current energy systems and to ensure justice in their energy transition plans.

[*]Equal contribution [1]College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA, USA. Correspondence to: Mohammad Hajiesmaili <hajiesmaili@cs.umass.edu>.

Tackling Climate Change with Machine Learning Workshop at ICML 2021.

1. Introduction

The current energy system is steeped in inequities. People of color and people with low incomes are especially vulnerable to serious negative impacts from the way we generate, store, transmit, use, and invest in energy (Jessel et al., 2019). Polluting power plants are more likely to be located in low-income and under-represented communities. In a crisis, power is restored last for minority neighborhoods, e.g., the recent blackout in Texas (Dobbins & Tabuchi, Feb 2021). High electricity and heating fuel bills due to poor energy efficiency are a major burden for low-income minority households, and are becoming even more onerous during the COVID-19 crisis (Graff & Carley, 2020).

As energy system transitions to a low carbon one to respond to climate change, we must intentionally build an equitable system that not only reduces harm but enables wealth creation in the communities that most need it. Planning for a just transition requires negotiating between many competing demands and tradeoffs. Understanding the aggregate economic, health and environmental impacts of energy decisions is already difficult; when equity is a priority, this is an even greater challenge.

There are signs that the underserved communities vulnerable to adverse energy impacts today may be cut out of benefits of the energy transition, or even further harmed. For example, subsidies for rooftop solar installations disproportionately benefit higher-income suburban homeowners; yet the surcharges to fund them raise everyone's energy bills, further burdening people who already spend a large portion of their income on energy costs (Roth, June 2020). To mitigate this inequity, there are nation-wide programs around the world, such as Weatherization Assistance Program (WAP) (Department of Energy, 2021b) or Low Income Home Energy Assistance Program (LIHEAP) (Department of Energy, Accessed January 2021) both in the U.S., with fiscal budgets above $3 billion (Department of Energy, 2021a).

However, according to DataKind (DataKind, 2021), a major challenge in these programs is on how to efficiently allocate these funds to the communities that most need it. To respond this challenge, some organizations have begun to identify measures of interest; the Energy Justice Scorecard from the Institute for Energy Justice (Institute for Energy Justice, 2021), for example, allows for the evaluation of individual policies for their equitable process and impacts. In another initiative, the policymakers in the New York State are insuring benefits of clean energy initiatives reach disadvantaged communities (NYSERDA, 2021). These solutions, however, are labor-intensive and limited in scale. In this work, we lay out initial ideas on how one can develop scalable, replicable, and practical ML tools to improve the justice in the current energy system and more importantly, ensure equitable transition of energy systems. Our vision is that this research will inform policies to induce green growth, reduce pollution, and encourage decentralized energy ownership and community representation. The high availability of the public data used in our approach allows the usage of machine learning tools in energy equity in an expansive set of locations.

To highlight the potential of ML for scalable and replicable enhancement of energy equity, in this work, we focus on two concrete applications and explain the challenges of developing ML tools for both. First, as an application for measuring energy equity in the current energy system, we present how ML could be leveraged to detect the equity in energy inefficient residential houses at scale (Section 2). Secondly, we focus on how ML (and especially deep learning) can facilitate an equitable installation of residential solar panels (Section 3). Finally, in Section 4, we present the impact of this research for different stakeholders, and shed light on the other potential impact of ML for energy equity in broader domains beyond our target applications.

2. ML for Equity Analysis of Energy Inefficient Buildings

Building infrastructure constitutes around 40% of total energy and 70% of the overall electricity usage in the United States (US Energy Information Administration, 2020). Consequently, residential energy-efficiency is a critical area that can have a significant impact on how the energy system is equitable. Our goal is to develop a ML model to automatically characterize the least energy efficient residential homes in a town and then classify these inefficient homes into different demographic groups, based on the publicly available data. We particularly focus to make our ML approach to be scalable and replicable such that it can be used beyond our pilot town.

An essential first step for improving the energy-efficiency of buildings in an equitable manner is to identify the least efficient ones that have the greatest need for improvements. However, naive approaches such as using the age of the building or its monthly energy bill to identify inefficiencies do not work well. While older buildings are usually less efficient than newer ones, age alone is not an accurate indicator since older buildings may have undergone energy improvements. Similarly, the total energy usage is not well correlated to energy inefficiency since larger buildings will consume more energy than smaller ones. Even normalizing for size, the energy usage does not necessarily point to inefficiencies. For example, two identical size house with a different number of residents merely reflects a higher actual demand rather than inefficiency. Thus, finding truly inefficient buildings from a large number of buildings is a challenging problem.

In this work, we will develop a model-driven ML approach for measuring the inequity of the inefficient residential homes from a large population of buildings in a pilot town in Massachusetts, U.S., and then use other publicly available data to expand this analysis.

Datasets. To implement ML-based approach for measuring inequity of inefficient residentials, We will use the following three datasets:

1) *Town-level energy dataset:* We use a dataset of energy usage from smart meters of 10K+ homes in Massachusetts. More than 70% of residential buildings in the United States have smart meters, and hence, our prototype could be extended for other towns as well.

2) *Nation-wide energy dataset:* We use several nation-wide publicly available data for the energy usage of buildings from the U.S. Dept. of Energy provides the Building Performance Database (Building Performance Database, 2021), which is the nation's largest information source for energy efficiency of all types of buildings across various sectors.

3) *The American Community Survey data:* To measure the energy inefficiency in low income and other different demographic groups, we use the publicly available census data. This data is part of a large dataset that is publicly available from multiple resources such as The American Community Survey (ACS) (acs) or the American Housing Survey (AHS) (ahs). This date is available in different categories in tract level.

Why machine learning? While building energy models have been extensively studied in the building science, and practitioners such as energy auditors use them to analyze a building's energy performance, we believe that developing ML tools can bring unique advantages as compared to traditional approaches. First, the current energy models assume simple (manual) regression analysis over observed energy data. In contrast, state-of-the-art ML tools will allow using novel probabilistic estimates of energy use, leading to more accurate energy efficiency analysis. Second, current models incur several parameters that are often chosen manually, based on intuition. With ML tools one can use advanced optimization approaches to learn the "optimal" values of these parameters so that they best explain the ob-

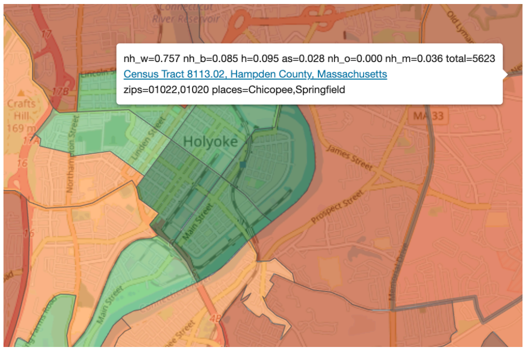

Figure 1. A Map of racial demographics in a town in MA

served data. Third, current approaches typically tend to be manual, which does not scale to modeling thousands of nation-wide buildings to identify the most inefficient ones. ML-based approaches are scalable and replicable, and can fully automate this process.

The implementation plan. We will build our proposed research based on the recent results WattHome (Iyengar et al., 2018) (at town-scale) and WattScale (Iyengar et al., 2020) (at nation-scale). While WattHome and WattScale are designed for residential buildings, in our proposed research, we will focus on extracting inequity patterns from the least efficient buildings. In this previous work, it is found that inefficient homes are geographically co-located. This is an initial observation that motivates us to systematically focus on how inequitable are the inefficient homes. Towards this, we need additional datasets that enable this extension. This is possible by using the dataset listed above. We did an initial investigation on extracting the demography information of Holyoke, MA, using the ACS data, with a representative sample shown in Fig 1 that shows the percentage of racial demography in a particular regions. Based on the available data, we will expand our solution to address racial justice directly by estimating the distribution of residential energy inefficiency by race and ethnicity.

3. ML For Equity Analysis of Residential Solar Potential Estimation

In this section, we briefly explain another potential application. Our goal is to develop a deep learning tool that is fed by satellite images that can help to characterize the residential houses with the most solar potential in certain minority neighborhoods. The result of this research proposal is a scalable response to the challenges that are already notified by social scientists on the inequitable allocation of clean energy incentives. For example, an initial analysis shows that that electric vehicles and rooftop solar installations in California are way more common in wealthier neighborhoods (Fournier et al., 2020). This is in conflict with the fact that California tries to deploy climate policies that not only green, but also spread the benefits of clean energy equitably.

This research extends DeepRoof (Lee et al., 2019), a convolutional neural network (CNN) approach for solar potential estimation, uses real estate data, solar irradiance data, and satellite imaging to estimate size, orientation, and geometry of rooftops, classify shadow casting buildings and trees, and provide a per-pixel generation potential of planar roof segments. The accuracy result of DeepRoof substantially outperforms alternative traditional approaches. The only dataset needed for this work are satellite images, which are widely available. Last, the socioeconomic data is available through the sources that are mentioned in Section 2.

There are additional challenges to be resolved for the purpose of equity analysis of solar potential estimation. More specifically, the current DeepRoof project is location-based, which means that one has to enter the exact location of a building to return its solar potential. However, characterizing the the buildings with the most solar potential in a neighborhood needs a learning process to detect the buildings in a neighborhood and then estimate the solar potential.

4. Expected Impact and Future Directions

Beyond the research potential of this work for ML researchers, and by integrating the developed ML tools for energy equity into an interactive online visualization software, the potential outcomes of our proposal could benefit the following groups:

1) *Government policymakers*: Visualization of present inequalities in their community, leading to motivation for new policies focused on fixing disparities in opportunity present in the community. Some potential questions that we expect our research could answer: What are common demographics of energy inefficient residential houses? or what communities have under-utilized solar energy?

2) *Residents*: An opportunity to contribute their own data to make the analysis more accurate, thereby helping their policymaker. Additionally, it will act as an insight tool for residents to better understand their community.

3) *Social and environmental scientists*: To observe long-standing trends and patterns in energy equity (or lack thereof). Researchers will be able to use this tool to answer important research questions in the fields of public policy, environmental science, and social justice.

Last, we highlighted two concrete research topics at the intersection of ML, energy, and equity. We believe that this research is broad and can be expanded into other broader topics as well. An example is measuring the inequity of the carbon intensity of different geographical neighborhoods in cities and towns. The carbon intensity of different geographical neighborhoods is not publicly known, however, there are plenty of publicly available data sources, such as time-varying energy supply and demand, that could be fed into ML tools to estimate the carbon intensity. A closely related ML research is the emerging topic of carbon-intelligent computing, which is a paradigm shift in the way that technology companies operates their datacenters (Radovanovic, 2020; Koningstein, 2021). In this approach, some ML tools have been developed to shift the computing workload across temporal and spatial domains such that the overall carbon footprint of the digital infrastructure is minimized. In doing so, one need to estimate the time-varying carbon intensity of the electric grids that supply datacenters. The carbon-intensity estimation output, however, could be used to measure the carbon intensity of different neighborhoods in a town from an equity perspective.

Acknowledgment

This research is supported by NSF grants CAREER 2045641 and REU supplements of CNS 1908298.

References

American community survey. *Census Reporter*. URL https://censusreporter.org/.

American housing survey. *https://census.gov/programs-surveys/ahs.html*.

Building Performance Database, 2021. https://bpd.lbl.gov/.

DataKind. DataKind partners with data.org on \$10m inclusive growth & recovery challenge. *https://bit.ly/3wHFFmX*, 2021.

Department of Energy. LIHEAP and WAP funding. *https://liheapch.acf.hhs.gov/Funding/funding.htm*, 2021a.

Department of Energy. Weatherization assistance program for low-income persons. *https://benefits.gov/benefit/580*, 2021b.

Department of Energy. Low income home energy assistance program (liheap). *https://www.benefits.gov/benefit/623*, Accessed January 2021.

Dobbins, J. and Tabuchi, H. Texas blackouts hit minority neighborhoods especially hard. *https://nyti.ms/2Tl1rOP*, Feb 2021.

Fournier, E. D., Cudd, R., Federico, F., Pincetl, S., Iles, A., and Mulvaney, D. On energy sufficiency and the need for new policies to combat growing inequities in the residential energy sector. *Elementa: Science of the Anthropocene*, 8, 2020.

Graff, M. and Carley, S. COVID-19 assistance needs to target energy insecurity. *Nature Energy*, 5(5):352–354, 2020.

Institute for Energy Justice. Energy justice scorecard. *https://iejusa.org/*, 2021.

Iyengar, S., Lee, S., Irwin, D., Shenoy, P., and Weil, B. Watthome: A data-driven approach for energy efficiency analytics at city-scale. In *Proc. ACM SIGKDD*, pp. 396–405, 2018.

Iyengar, S., Lee, S., Irwin, D., Shenoy, P., and Weil, B. Wattscale: A data-driven approach for energy efficiency analytics of buildings at scale. *ACM Transactions on Data Science, to appear*, 2020.

Jessel, S., Sawyer, S., and Hernández, D. Energy, poverty, and health in climate change: a comprehensive review of an emerging literature. *Frontiers in public health*, 7:357, 2019.

Koningstein, R. We now do more computing where there's cleaner energy. available at https://bit.ly/3c9o27w/, 2021.

Lee, S., Iyengar, S., Feng, M., Shenoy, P., and Maji, S. Deeproof: A data-driven approach for solar potential estimation using rooftop imagery. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2105–2113, 2019.

NYSERDA. Climate leadership and community protection act (climate act), 2021. URL www.nyserda.ny.gov/All-Programs/Programs/CLCPA.

Radovanovic, A. Our data centers now work harder when the sun shines and wind blows. available at https://bit.ly/3wICF9O, 2020.

Roth, S. California's clean energy programs are mainly benefiting the rich, study finds. *https://lat.ms/3vCiavi*, June 2020.

US Energy Information Administration, 2020. https://www.eia.gov/, accessed September 2020.