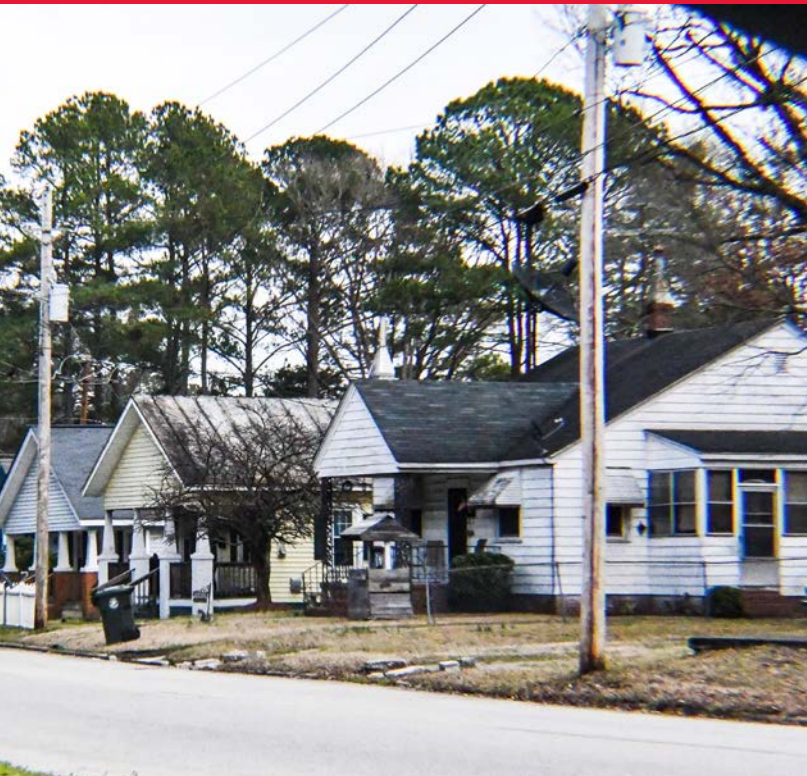


POWERING ENERGY EFFICIENCY AND IMPACTS

A Data-Driven Project Supporting Low-Income Households in Northeastern North Carolina

March 2019



POWERING ENERGY EFFICIENCY AND IMPACTS

A Data-Driven Project Supporting Low-Income Households in Northeastern North Carolina

REPORT CONTRIBUTORS

Anne Tazewell, North Carolina Clean Energy Technology Center

Ron Townley, Upper Coastal Plain Council of Governments

Tirence Horne, Upper Coastal Plain Council of Governments

Laura Langham, North Carolina Sustainable Energy Association

Daniel Pate, North Carolina Sustainable Energy Association

Mark James, Vermont Law School

Joshua Randall, North Carolina State University Center for Geospatial Analytics

William S. Slocumb, North Carolina State University Center for Geospatial Analytics

Mark Griffin, North Carolina State University Center for Geospatial Analytics

Scott Ferguson, North Carolina State University System Design Optimization Lab

Daniel Kauffman, Enpira (formerly ResiSpeak)

Alfred Ripley, North Carolina Justice Center

Ookie Ma, U.S. Department of Energy

PARTNERS



Center for
Geospatial Analytics



NC SUSTAINABLE ENERGY ASSOCIATION



VERMONT LAW SCHOOL



UNIVERSITY OF SOUTH CAROLINA

FOREWORD

The project partners wish to thank our data-providing partners for their involvement in and contributions to the Powering Energy Efficiency & Impacts Framework (PEEIF) project: the North Carolina Department of Environmental Quality (NCDEQ) Weatherization Assistance Program, the North Carolina Department of Health and Human Services (NCDHHS) Energy Services Program, Roanoke Electric Cooperative (REC), Wilson Energy, and the Town of Enfield. Without these partners, our efforts on behalf of income-challenged households in the Upper Coastal Plain Council of Governments (UCPCOG) region would be purely theoretical. Thanks to their contributions, the project team was able to demonstrate, in a proof of concept, both the challenges and opportunities of using "big data" and Geographic Information Services (GIS) mapping to address the nexus of energy and poverty. The partners also appreciate the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Cities Leading through Energy Analysis and Planning (Cities-LEAP) project for selecting this endeavor as one of three 2016 financial assistance awards.

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government, or any agency thereof, or any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

CONTENTS

- SUMMARY** 1
 - Need 1
 - Challenges..... 2
 - Approach and Results 2
 - Learnings and Recommendations 3
 - Key accomplishments 4
- REGIONAL CONTEXT** 5
 - Upper Coastal Plain Region..... 5
 - Energy and Poverty 6
 - Energy Service Provider Partners 7
 - Data-Sharing Partners 8
- ENERGY SERVICE DATA AND PROGRAM EVALUATION FRAMEWORK** 9
 - Framework Principles 10
 - Data Confidentiality..... 11
 - Data Types and Attributes 11
 - Data Analysis..... 13
- ANALYSIS PARTNERS** 16
- CREATED APPLICATIONS** 17
 - North Carolina Department of Environmental Quality 17
 - North Carolina Department of Health and Human Services 18
 - Energy Utilities 20
 - REC Map Analysis 22
 - Utility Attributes Lessons Learned..... 23
 - Legal Process and Steps 23
- RECOMMENDATIONS** 25
 - Spatial Analysis 25
 - Future Legal Agreements 25
 - Multi-Party Non-Disclosure Agreements..... 25
 - Legal Considerations..... 27
 - Future Possibilities 28
- APPENDIX** 29

SUMMARY

The Powering Energy Efficiency and Impacts Framework (PEEIF) project, a two-year U.S. Department of Energy (DOE) sponsored initiative, seeks to develop a data-driven framework to increase energy-related program effectiveness in low-income households. As part of this project, a diverse, multi-sector partnership developed a database and a geospatial-mapping tool of homes located in five eastern North Carolina counties served by the Upper Coastal Plain Council of Governments (UCPCOG), the lead administrator for the effort. This report outlines the needs, challenges, and solutions identified through the project and presents opportunities for state agencies, utilities, and organizations that are currently providing, or interested in supporting, low-income energy-efficiency services. It details the project’s methods and results and identifies strategies for future iterations of the mapping tool.

Need

Despite overall growth and economic vitality in North Carolina, many communities throughout the state struggle with chronic and persistent poverty due to industrial downsizing, loss of main street businesses, educational challenges, and the lack of availability of living wage jobs. Low-income households often face high energy costs as a

percentage of income (i.e., they incur a high energy burden), resulting in less disposable income that could otherwise be used to make homes more energy efficient and used in the local economy. In response to the challenge of reducing energy burden, a number of federal, state, utility, nonprofit, and local energy programs focus on low-income households. These range from funds to pay for inordinately high utility bills to improving energy efficiency in homes through physical improvements of the building and systems.

Low-income communities are desirable focus areas for these programs because they are more likely to contain older homes in higher need of maintenance and improvements. These communities also often contain the most vulnerable populations, such as at-risk children, the elderly, and the disabled. If not addressed, home maintenance challenges for low-wealth families can compound over time. For example, without regular monitoring or upgrades, low-income homes are more likely to contain pollutants such as mold and pests that can trigger respiratory and other illnesses.

These needs can be addressed by services such as home weatherization, energy efficient appliance



upgrades, and other improvements provided by state and local agencies. Energy-efficiency improvements can improve the quality of life for low-income residents through lower bills, higher comfort levels, and reduced negative health effects. Moreover, energy bill assistance for income-challenged households helps keep their lights on.

Challenges

Overlap of Services

State energy efficiency and utility assistance programs are often run independently, with little coordination. The federal Weatherization Assistance Program (WAP), Low Income Home Energy Assistance Program (LIHEAP), and electric utilities' energy-efficiency rebate programs are all designed to reduce energy burden. While some homes receive assistance from more than one program provider, the providers are often unaware of each other's efforts and are therefore not collectively optimizing their services. For example, WAP currently has a long list of applicants who are not being served because either the residence needs repairs that disqualify them from WAP or funding is not available to upgrade the home. In Roanoke Electric Cooperative service area, the utility has a tariffed, on-bill energy efficiency program called "Upgrade to \$ave"¹ that could retrofit these homes.

Coordination between Organizations

In North Carolina and other states, state agencies serve, in part, as a "pass through" for federal funding to other entities for energy bill assistance and home energy-efficiency upgrades. There is often limited coordination on their mutual common goals of serving low-income households. Furthermore, multiple parties have data that could aid with other organizations' applicant selection processes. For example, utilities have energy data that could be used to document the need for bill assistance and to assist in weatherization applicant selection.

Confidentiality and Data Sharing

The individualized nature of agency data management and data access practices presents a challenge to the aggregation and integration of



different data sets. Agencies have different confidentiality regulations and statutes—some of which are derived from federal sources, and some from the state level. The classification of the service provider—contractor, agent, or local government—providing service to the client varies from agency to agency, which creates additional obstacles to address. Each agency also maintains its own database of client information, which requires outreach and communication with multiple IT departments, which follow their own organization's internal data management guidelines.

Reach and Effectiveness

Organizations serving income-challenged residents often do not know if they have reached the communities and households that are most in need of energy assistance. With a robust data management tool, agencies could identify in-need service areas and market programs using information provided from a variety of parties providing complementary services to low-income residents. For homes that have received prior assistance from other organizations, agencies could know the service history for each home, which would help evaluate the current needs of the resident. Furthermore, many organizations providing energy-efficiency upgrades do not know to what extent their assistance has reduced energy burdens in the households they have served.

Approach and Results

The project partners established a legal and technical framework to enable government, utility,

¹ <https://roanokeelectric.com/upgradetosave>

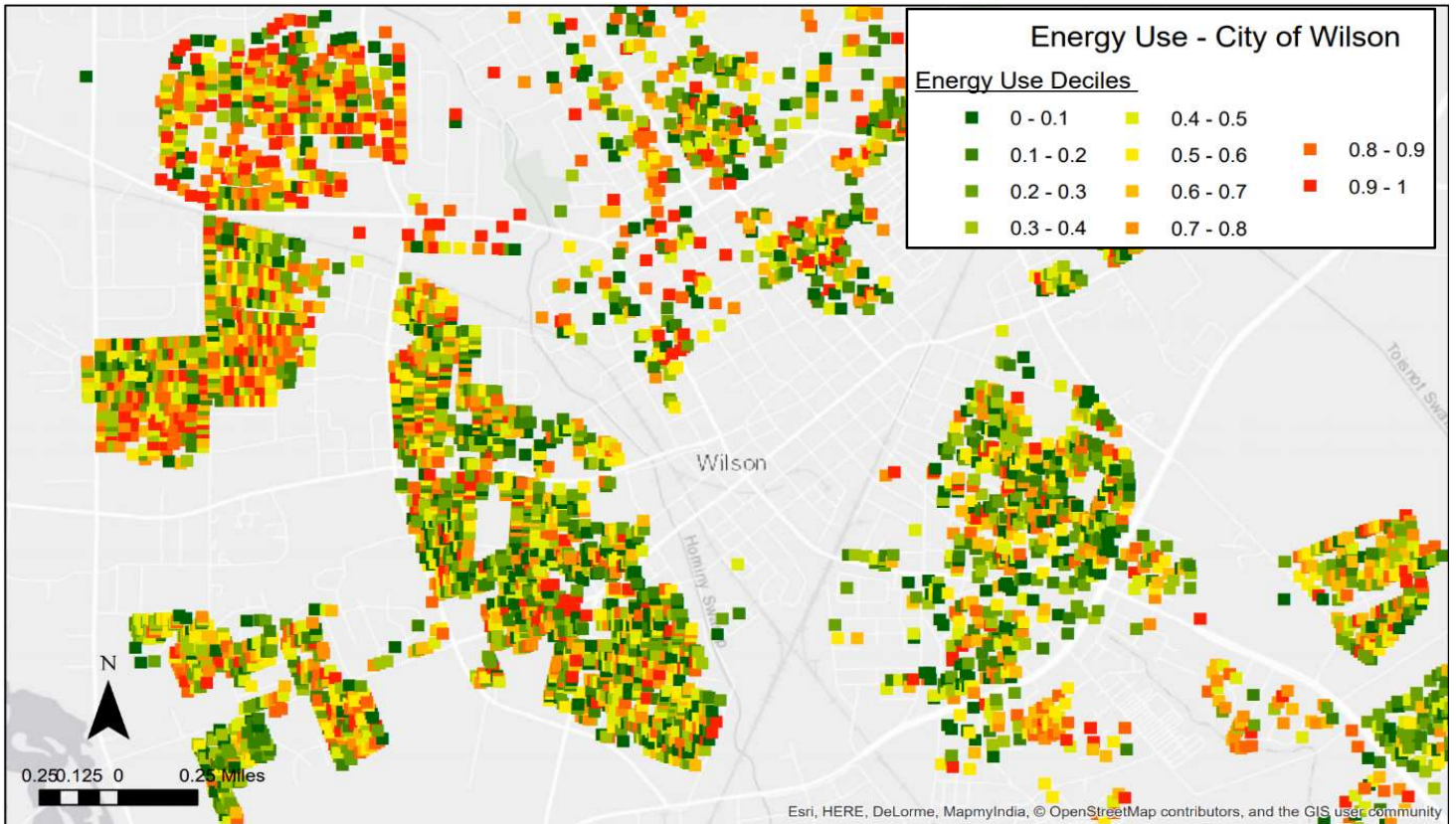


Figure 1: Map of Wilson homes after categorization into deciles for analysis.

and nonprofit service providers to view their own confidential information alongside aggregated, blended, and/or publicly available information. In-person meetings were conducted to gather input into the construction and optimization of a geospatial database where users could access both their own data and other relevant data and analytics. The use of Esri's ArcGIS Online was utilized for both data visualization and data manipulation. Using both information and feedback from data-providing partners, the PEEIF project has demonstrated various value propositions for organizations focused on enhancing energy services for low-income households (see the "Key Accomplishments" text box on page 4).

If the program continues, there is a potential for even greater improvements, including:

- An analysis of the cost effectiveness of various energy efficiency measures
- A comparison of contractor effectiveness in home weatherization measures.

Learnings and Recommendations

The PEEIF project combines data and mapping to better identify and target the nexus of energy and poverty. Its applied experience yielded several insights that could inform future efforts. For example, Figure 1 compares the energy usage of households across the City of Wilson, which could help identify candidates for energy-saving retrofits. These recommendations could also be used by other organizations that are interested in improving the energy landscape of low-income households in their community, region, state, or nation.

For future efforts, the PEEIF team suggests the following:

- Create an online GIS application for a specific region that includes poverty- and housing related Census Tract (an area with 2,500-8000 people) data with aggregated data from utilities and state and local agencies that provide energy services to low-income households. For organizations that want to track visitors to the site, create a short login form.

- Use information from county tax files, such as square footage, year built, and type of housing, when available, to help identify energy use intensity and energy efficiency potential.
 - When developing a public facing GIS application, include a section on available services in each county and region, as well as points of contact for services. Update these annually.
 - When recruiting data-providing partners, understand the value proposition for their involvement and then have a standardized process for two critical issues: legal agreements and data transfer. Have specific points of contact to answer questions about these issues and a predetermined, agreed-upon process for tracking progress in completing these first two critical steps of the partnership.
- As staff changes, having a standard process for orienting new partners is important.
- Create a legal data-sharing framework that facilitates data sharing among partners as they join the project. Refer to the *Multiparty Data Sharing Agreement Template* in Appendix A.
 - Do not create separate login portals for data-providing partners that contribute confidential information, unless requested. Rather, understand client needs and develop and deliver a suite of services, including analysis and mapping, upon request and consultation with data partners.
 - Bring stakeholders from a specific region together to discuss ways to increase collaborative efforts at reducing the energy burden for low-income households, then prioritize these needs, track your actions, and celebrate success.

Key accomplishments

- Providing publicly available data grouped by Census Tract, such as the population living at or below 150% or 200% of the federal poverty rate, overlaid with confidential information informs county social service and state agencies if they have adequately reached all areas of need in their communities.
- Breaking out electric utility data into deciles allows partners to identify households that can most benefit from increased energy efficiency, as well as identify homes that have decreased energy use after receiving weatherization and other energy saving services.
- Utilizing services provided by ResiSpeak, energy efficiency programs could better monitor and verify energy savings of specific households that have received assistance.
- Facilitating connections among PEEIF project partners provided for nine additional energy efficiency upgrade applications through Roanoke Electric Cooperative's (REC) Upgrade to \$ave program when the Choanoke Area Development Association (CADA), the Community Action Agency providing federally-supported weatherization assistance, mailed letters provided by REC to their waitlisted clients.
- Calculating energy use intensity (energy use/square footage) to compare energy use between similar sized homes assisted with the identification of homes most in need of retrofit services and/or energy efficiency education.
- Providing a platform to use aggregated information from data-providing partners overlaid with Census tract information helps researchers and grant writers to better understand the effectiveness of, and illustrate the need for, services provided in the project region.
- Catalyzing the process of removing legal barriers so NCDHHS and NCDEQ can share confidential data to inform and document potential savings from energy efficiency upgrades through WAP & HARRP programs, serving to reduce further need of CIP and LIEAP energy bill assistance.

REGIONAL CONTEXT

The Powering Energy Efficiency & Impacts Framework (PEEIF) project developed a comprehensive, data-driven tool that allows users to compare data from independent, energy-related programs that serve low-income households alongside other relevant inputs. This helps support integrated planning, deployment, and evaluation of program effectiveness. This work is a template, both from an organizational and data-management perspective, that can be replicated and/or modified to serve other regions.

The goal of the PEEIF project is to foster regional collaborations between multiple government and non-government entities and use data-driven analysis to facilitate local government, utility, and community decision-making with respect to energy services for low-income households. Increased collaboration and targeted analysis will improve the effectiveness of organizations serving low-income households and reduce energy use in energy-intensive homes. Critical success factors in achieving this goal include the resolution of data security and privacy issues, adoption of data-sharing agreements, and the development of an informative, accurate, and accessible database to consider policy, energy intensity reductions, and tailored program implementation efforts.

Led by the Upper Coastal Plain Council of Governments (UCPCOG), the PEEIF project has been a collaborative effort of three North Carolina State University partners: the North Carolina Clean Energy Technology Center (NCCETC), the Center for Geospatial Analytics (CGA), and the System Design Optimization Lab (SDOL), as well as the North Carolina Sustainable Energy Association (NCSEA), ResiSpeak, Vermont Law School (VLS), the NC Justice Center (NCJC), and the University of South Carolina (USC). The geographical focus of the project is a five-county region in northeastern North Carolina served by UCPCOG.

Upper Coastal Plain Region

The Upper Coastal Plain Council of Governments represents forty-six member governments and allies in a strategically important region in eastern



North Carolina. Consisting of five counties—Edgecombe, Halifax, Nash, Northampton, and Wilson—the region covers 2,707 square miles, totaling 1.7 million acres of valuable working lands and communities dating back to colonial times.

The Upper Coastal Plain borders Virginia to the north, where Northampton County shares a 40-mile border in close proximity to Norfolk’s International Port Terminals. Halifax and Edgecombe counties are within 50 miles of the nearest inlets to the Atlantic Ocean and Outer Banks. Nash and Wilson Counties are less than an hour’s drive from the state’s capital and the Research Triangle area.

However, like much of rural America, drastic declines in manufacturing and the consolidation of agricultural farming practices, along with urban migration, has resulted in a slow-growing and aging population and put a severe strain on the largely rural regional economy. While the region has numerous laudable assets, like a strategic location, affordable lands and homes, vibrant and diverse cultural resources and communities, and a ready and willing workforce, many of its citizens have been impacted by the changing economy and other stress factors. These individuals are the focus of this project.



Of the estimated 303,922 residents, the region has an average per capita income of \$35,000, with 23.2% of the residents earning an income below the poverty level. The region continues to struggle with unemployment when compared to the rest of North Carolina and the nation. As a result:

- 65% of the total housing stock was built before 1990
- The 2010 Vacant Housing rate was 14.3%
- The 2015 Estimated owner housing vacancy rate was 2.0%

- The 2015 estimated renter housing vacancy rate was 6.2%.

Energy and Poverty

The southern region of the United States has the highest rate of poverty in the country, especially in rural areas. According to the U.S. Department of Agriculture (USDA), 21.7% of non-metro residents live in poverty, compared to 14% to 17.4% in other regions of the country. Moreover, residents of the South suffer from higher persistent poverty than their counterparts in other U.S. regions. In Figure 2, the USDA's Economic Research Center illustrates counties identified as persistently poor. Persistently poor counties are defined as those where 20% or more of their populations have lived in poverty over the last 30 years (as measured by the 1980, 1990, and 2000 decennial censuses and 2007–2011 American Community Survey five-year estimates).

Federal guidelines define poverty for a household of four when their collective annual income is \$24,300 or less. In 2016, more than 1.5 million North Carolina residents lived in poverty. Individuals living in poverty spend a significant amount of their incomes on utilities, including electricity and heating. In North Carolina, households living below 50% of the federal poverty level spent 29% of their income on utilities, while those living up to the

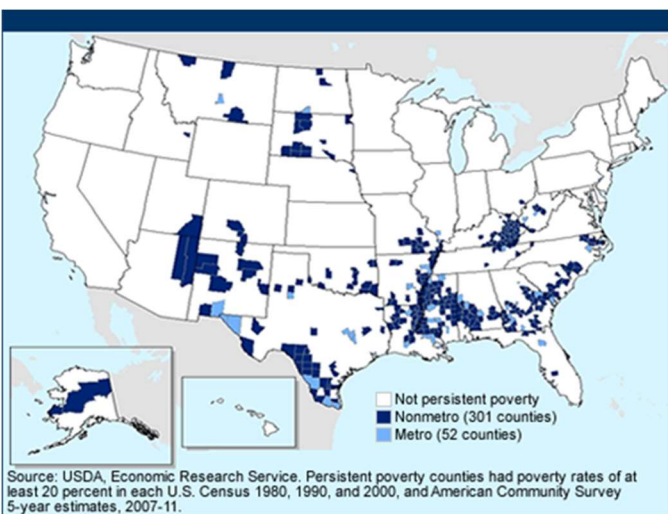


Figure 2: Persistently Poor Counties in the United States, 2007–2011.

federal poverty level had a 16% home energy burden rate.²

Energy burden refers to the percentage of household income spent on home energy bills. For example, a household of four earning an income of less than \$12,500 is spending \$3,524—29% of their income—on utilities, while a family earning \$24,300 is spending 16%, or \$3,888, on utilities. Lower income households typically have higher energy burdens, highlighting the importance of opportunities for these customers to reduce their electric bills.

To reduce the burden, a number of federal, state, utility, and local energy programs focus on low-income households. The federal Weatherization Assistance Program (WAP), Low Income Home Energy Assistance Program (LIHEAP), and energy efficiency rebate programs offered by electric utilities are designed to reduce energy burden. The North Carolina Low Income Energy Assistance Program (LIEAP) and Crisis Intervention Programs (CIP)³ provide energy bill assistance payments directly to utilities to help low income households pay their energy bills and avoid costly shut-off and turn-on fees when bills are not paid. Traditionally,

these programs are run independently and with little coordination, which reduces the potential benefits for low-income households.

Energy Service Provider Partners

The PEEIF project originally planned to work with a minimum of four data-providing partners serving low-income households in the UCPCOG region, including the NC Dept of Health and Human Services Low Income Energy Assistance Program (LIEAP), the North Carolina Department of Environmental Quality (NCDEQ) Weatherization Assistance Program, the North Carolina Housing Finance Agency (NCHFA), Wilson Energy, and Roanoke Electric Cooperative. However, as legal concerns were more fully explored, the NCHFA was unable to participate because federal law (Gramm-Leach-Bliley) restricted it from disclosing required data, such as client addresses. Unlike the other participating government agencies, NCHFA is considered a financial institution under the law and is subject to its privacy and disclosure obligations. Consequently, the town of Enfield, which serves as a municipal utility, joined the project in its development phase and served as the fifth service provider participating in this pilot project.

² Ripley, A. (2018). *Community Solar in the Southeast for Low Income Customers*. Retrieved from North Carolina Justice Center: <https://nccleantech.ncsu.edu/wp->

<content/uploads/2018/05/Ripley-Alfred-NCJC-Community-Solar-Presentation.pdf>

³ <https://www.ncdhhs.gov/crisis-intervention-program>

Data-Sharing Partners



The North Carolina Department of Health and Human Services (NCDHHS) manages federal and state funds for a variety of programs that serve low-income households across the state. Federal funds are distributed through an application process administered by county social service departments. The Low Income Energy Assistance Program (LIEAP) and Crisis Intervention Program (CIP) beneficiaries spanning July 2014 through February 2017 are included in the PEEIF tool, totaling close to 40,000 records with 6,588 unique addresses and about 14,463 assistance instances (some addresses received more than one assistance due to multiple years in the data set and multiple people in the house).



The North Carolina Department of Environmental Quality (NCDEQ) administers two low-income energy related programs: the Weatherization Assistance Program (WAP) and Heating Appliance Repair and Replacement Program (HARRP). WAP focuses on increasing energy efficiency with client education, air sealing, and insulation, while HARRP provides efficient heating and air systems for qualifying applicants. Federal funds are

distributed through twenty-three community agencies across North Carolina, including three that serve the UCPCOG region: the Choanoke Area Development Association (CADA), the Wayne Action Group for Economic Solvency, Inc. (WAGES), and Franklin-Vance-Warren Opportunity, Inc. (FWWOP). Both programs require beneficiaries to earn no more than 150% of the federal poverty rate. The PEEIF project incorporates 248 homes from February 2014 through August 2017 that received WAP or HARRP services in the five-county project area.



Roanoke Electric Cooperative (REC) is a rural electric cooperative serving more than 14,500 member-owners in Bertie, Halifax, Hertford, Northampton, Gates, Perquimans, and Chowan Counties. Two

counties overlap with the PEEIF project area: Halifax and Northampton. This area represents about 2,000 member-owner households. REC offers a number of energy savings opportunities for member-owners; its flagship is Upgrade to \$ave (U2\$), which provides on-bill financing for insulation, duct, and air sealing; heat pump improvements; water heater wraps; and LED lighting that generate immediate monthly savings to the member-owners who enroll (Roanoke Electric Cooperative, 2018). REC provided utility data on meters in Halifax and Northampton counties from May 2014 through to June 2018, and more 1,900 REC meters were geolocated. The median energy use for the most recent year was 15,700 kWh. Square footage data was available for about 1,400 of the homes, and the median electrical usage intensity was 11.5 kWh/sq. ft. REC's U2\$ has upgraded 90 homes in the PEEIF project area.



Wilson Energy is responsible for the operation and maintenance of the city's electric and gas facilities. The City of Wilson has more than 49,000 residents (U.S. Census Bureau, 2017). The municipal utility offers free in-home energy audits and an online energy savings calculator. PEEIF project data includes electricity and natural gas use for two years spanning from January 2016 to May 2018. Wilson Energy provided electrical and gas usage information for homes, of which about 7,000 electric meters and about 4,500 natural gas meters were geolocated. The median electricity usage for homes in Wilson was about 12,000 kWh, and the median natural gas usage was about 420 therms. Square footage data was available for nearly all of the homes, median electricity usage intensity was about 7.7 kWh/sq. ft, and median natural gas usage was about 0.25 therms/sq. ft.



The Town of Enfield operates a municipal electric utility that serves its more than 2,300 residents. The town provided electrical usage information from August 2015 to May 2018 for about 800 residential accounts. Their median energy usage was about 9,700 kWh/year.

ENERGY SERVICE DATA AND PROGRAM EVALUATION FRAMEWORK

The PEEIF energy service data and program evaluation framework has a number of components, which are depicted in Figure 3 below. Users interact with the framework primarily through GIS-based online web applications. Confidential portals were created for individual data-providing partners to view their data overlaid with aggregated and blended data contributed by other data providing partners, as well as publicly available Census tract data, such as the percentage of households living at various poverty levels, number of elderly persons, and number of rental households.

In order to display the tool to a wider, public group of beneficiaries—including potential project duplicators, elected officials, future partners, and others—the CGA team created a project story map⁴. This portal gives visitors a limited but critical

example of the PEEIF project and GIS tool. The images in the story map overlay Census tract data for the UCPCOG region with aggregated and scrambled confidential information from data-providing partners that protects the confidentiality of the information. A Google login form was created to register visitors and a link to a short, five-question survey was created to gather feedback on the presented information. This feedback helps project partners understand the project’s value to a variety of stakeholders and determine future potential improvements and partner interests.

Behind the scenes, PEEIF uses a hub and spoke non-disclosure agreement (NDA) structure to facilitate confidential data flows in and out of the database. NDAs were completed between UCPCOG and each of the data-providing parties and between UCPCOG and each of the cooperative

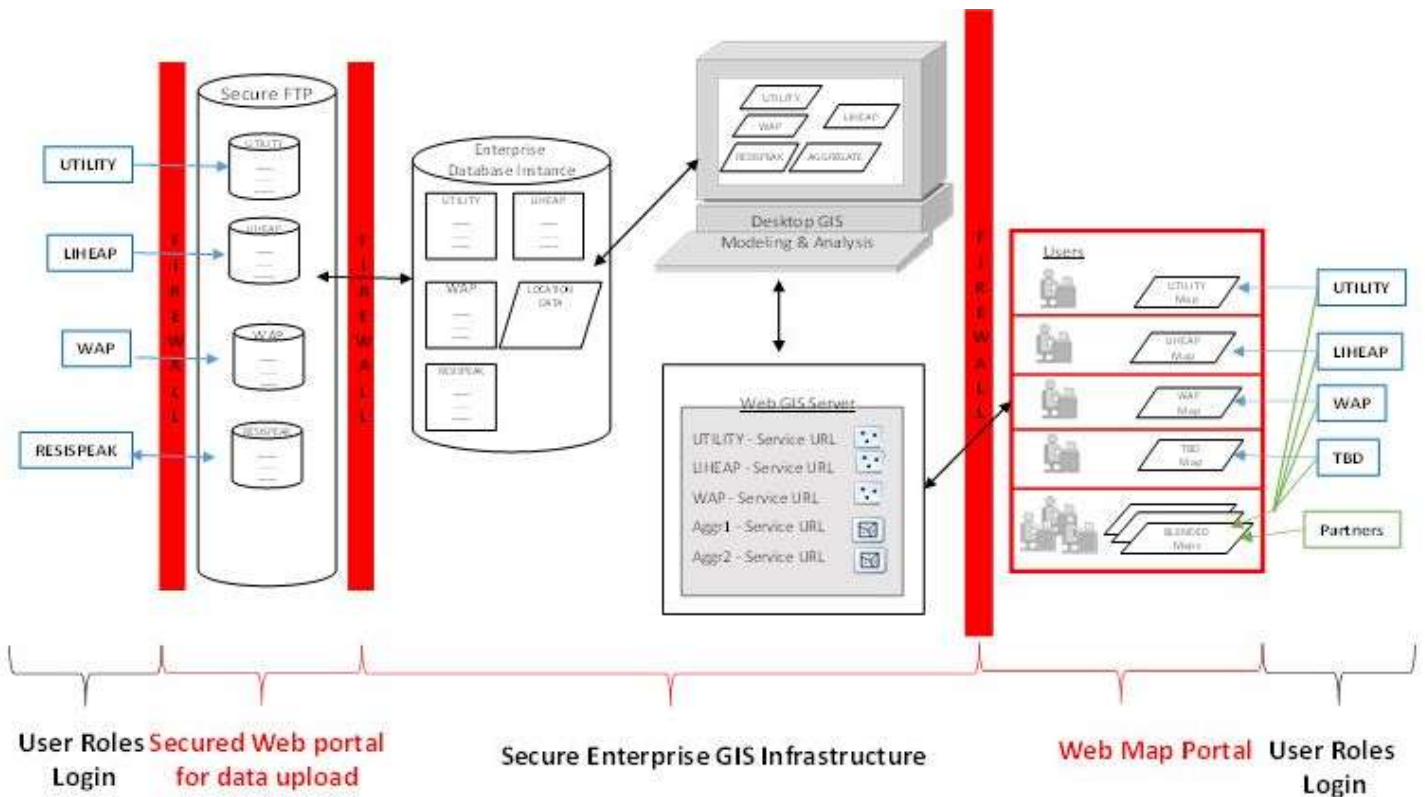


Figure 3: PEEIF data flowchart demonstrating processing and security steps from provider to end user.

⁴ PEEIF Story Map PDF, <https://www.energy.gov/eere/analysis/powering-energy-efficiency-and-impacts-project-framework-peeif>

Framework Principles

A collaborative, residential energy services data and program evaluation framework can be organized many ways as long as some key services and individuals are included in the process, including:

- A secure web portal and protocols for hosting, providing, and sharing confidential data. For example, GIS data could also be hosted on Google Cloud Platform's (GCP) technology, which can provide a secure environment to scale storage and analytics capabilities without affecting speed or security.
- Geographic information system (GIS) expertise to analyze and display information in a visual, map-based platform. Esri ArcGIS was selected because North Carolina's sixteen regional Council of Governments have subscriptions to this widely utilized software. There are also open source GIS platforms such as QGIS Server (<https://qgisenterprise.com/en/qgis-suite/server/>), which can be used to create a Web Mapping Service (WMS) and OpenLayers and that can support the web mapping aspect of the project with equal success.
- Legal counsel to craft and review non-disclosure and data sharing agreements.
- Energy analysis services and software such as ResiSpeak (<https://www.resispeak.com/>) that can weather normalize utility energy use for comparison and evaluation over time.
- Project managers to solicit, interact, and manage data-providing partners. Project managers can also serve as quality control agents by reviewing best practices for data formatting to minimize staff time spent "cleaning" data.
- Subject matter experts that understand the regulatory, policy, energy efficiency, low income, and/or data analytics and big data landscapes.

agreement subcontractors accessing confidential information. The combination of NDAs allows for UCPCOG to flow appropriate data between the sub-awardees building the data management and mapping framework, to provide data analytics services, and to return those analytics to the data-providing parties through the framework and maps. Each NDA restricts data access to approved individuals in data-providing and data-receiving parties, a restriction that is reinforced by the physical process for accessing and receiving data. By knitting together the data collection and data access, confidential information is protected while making it accessible to approved individuals.

Data transfer and storage were achieved through a multi-step process. First, data were transferred between parties via a secure file transfer protocol (FTP). Velocity, an FTP system operated and managed by the North Carolina State University Office of Information Technology, was used for data transfer. Velocity stores data in "cabinets," where each data provider was given access to an individual cabinet and only data providers and selected project team personnel had the ability to

read, write, and update data. In addition, using an automated Velocity procedure, cabinets were regularly purged to ensure data security and confidentiality.

Data was transferred to an internal enterprise geodatabase and published as web GIS services in the representational state transfer (REST) protocol. The enterprise geodatabase server and web GIS server were administered by the Center for Geospatial Analytics and provided for data cleaning, storage, analysis, and visualization of confidential data. The web GIS server hosted the web GIS services and allowed for access by ArcGIS Online for Organizations (AGOL). Refer to Appendix B for the PEEIF virtual machine specifications. Data from the enterprise geodatabase servers were registered with the web GIS server, which allowed data to be maintained internally by NCSU and provided an additional layer of protection. Further protection was added through the use of groups in AGOL, where only specific users were granted access to view web applications specific to their organization, ensuring no data was shared improperly.

Data Confidentiality

Confidential data is protected data that a specific data-providing partner is only able to see under the terms of an NDA. This data is not available to other data-providing partners or the public, for whom it is either aggregated or blurred based on the need and the privacy requirements of the original data set.

Blurring data is an integral part of creating datasets that can be accessed by multiple users. The process of blurring data creates a data set where individual locations are retained, but the data assigned to those points are generalized. For example, households have individual annual energy data that cannot be shared with other users other than the original data provider. Blurred data is created through "binning" the data into groups of use based on amount. In the PEEIF project, ten groups or deciles were created of even amounts of houses in each. These groupings effectively retain relative information on kWh usage without showing actual usage. Comparative analysis can still be performed (e.g. did this house move from one decile to another?) without compromising privacy.

Aggregation of data combines multiple observations into one. Aggregated data is another technique to protect data privacy while still providing useful analysis for the project. Aggregation is used when specific locations cannot be used. Data are combined within a specified geographic location and are then calculated to represent the areas. For example, the energy usage for all of the houses in a census block group are used to calculate an average use for the area. The census block groups then represent that average

across the study area. While aggregation lowers the specificity of the data, it does allow for some analysis and context for use.

De-identification reveals confidential information about an individual house, but removes identifiable information about it, such as its address or county parcel identifier. De-identification is used to produce demonstration maps where actual data are randomly assigned to nearby homes, and is used for presentation purposes to display examples of individual homes without revealing their identity.

Data Types and Attributes

The attributes used for visualization, query, and display in the final database were selected based on their usefulness. These attributes included energy uses, household characteristics, and characteristics from programs that provide energy-efficiency and energy services to low-income households. Data came from three sources: publicly available housing data through county web sites and North Carolina OneMap, utility data from project utility partners (e.g., Roanoke Electric Cooperative), and program assistance data from North Carolina state agencies (e.g., the North Carolina Department of Environmental Quality). Attributes were also created as part of the project based on analysis from partner-provided data. Data sets at the Census-block-group level were included in the final versions of the application to provide context for areas of study. All of the data that was provided was retained in the event that it should be needed for analysis.

Housing data	Occupant data
Location (mapped)	Name
Size of home (sq. ft.)	Census block
Type of home	Program enrollment and completion
Age of home (years)	
If retrofitted: Date	
Retrofit measures	
Crew	
Cost	
Agency and Program	
Pre and post retrofit electric utility usage	

Table 1: Table with schema for standardizing housing data

An internal "PEEIF ID" was created to systematically match parcel data with multiple data sets provided by project partners. The PEEIF ID was assigned to parcel data and was assigned to new datasets, as they were geolocated. This unique ID improved reproducibility in data matching as the project progressed and additional data were acquired from partners.

The PEEIF tool provides energy data, square footage, and other critical information required to perform energy analysis on program homes.

Housing Data

Housing information can be shared among data-providing partners and contains all the variables available in the parcel information. Housing information is used to provide context for individual residences and identify the types of houses that may be experiencing symptoms of inefficiency. Building characteristics can result in differences in performance and including them allows for pattern analysis of the house. The most important aspect of these attributes is square footage, as it is a critical variable in comparing household energy use and determining energy-efficiency intensity.

Housing data can often be found in state-run geospatial repositories. Other places include state planning or county planning offices, although the information may be in a paper or non-spatial form, and thus may be difficult to use. Data can also be retrieved from real estate data service providers such as Zillow, and from consumer data service providers such as USADATA. Refer to Appendix C for the data attributes of PEEIF housing. Table 1 provides a schema for standardizing housing data.

Utility Data

Electricity data was acquired from three utility providers: Roanoke Electric Cooperative, the City of Wilson Electric (Wilson Energy), and the Town of Enfield. The data acquired were at the individual household level at many temporal resolutions. Data were acquired in Microsoft Excel spreadsheets or comma-separated values (CSV) ASCII text format. Refer to Appendix D for PEEIF project utility data attributes.

Multiple aspects should be considered when using utility information. The most important is the

temporal aspect. Data are provided anywhere from 15-minute intervals to annual intervals. Depending on the goal of the study, the lowest common denominator of all energy usage values will have to be used in order to provide comparable information across utility providers. For example, the PEEIF project was provided information in 15 minute, monthly, and annual intervals. In order to provide comparable statistics across providers and the study area, all energy use was compiled at in annual intervals.

The second consideration is the creation of blurred data so that residential energy use may be shared without divulging personally identifiable information (PII), such as specific kWh electricity usage. PII is any data that may identify a specific individual. Blurred data retains the spatial attributes of the data points while blurring specific information (such as energy use) about the household. In this case, a set of deciles were created for each set of energy use data. This allows an approximation of energy use without providing specific energy information.

Energy data is available in multiple formats for both electric and gas. Some residences also have designators on the type of heating fuel used, such as propane and wood, which were provided by the Community Action Agencies (CAA) that distribute WAP program funds. While this information is collected by the CAAs as part of their application process, the data were more limited than the electricity data within the PEEIF project area. In the Wilson area, gas combined with electric data provided a multi-season view, especially in homes that use gas for heat. It is important to note that multiple energy attributes increase the ability to characterize household energy use in a more complete manner.

Acquiring utility data is challenging due to the confidentiality of the energy data and the number of energy sources that may be used by individual households. In addition to electricity and natural gas, household energy use may include solar, wood, propane, kerosene, and fuel oil. While energy providers have localized data, home confidentiality requirements are a concern.

Energy Efficiency and Assistance Data

Energy-related service provider data were acquired from two main sources, the NCDEQ and NCDHHS. Refer to Appendix E for NCDEQ and NCDHHS PEEIF data attributes.

The Low-Income Energy Assistance Program and Crisis Intervention Program is administered by NCDHHS on the county level and is one of the two main types of federal assistance used to address energy poverty. Both programs provide utility bill payments to qualifying low-income households. When mapping these energy services, the information should contain the program type and amount of assistance provided. Another consideration may be how many times assistance has been provided to a specific house. One of the challenges to using this information, or any client information that has to be manually entered, is that the program data may be entered incorrectly. Moreover, low-income households are more likely to rent and thus may move more frequently. They may use P.O. boxes when applying for assistance. Therefore, extra care should be taken when mapping this data to make sure that the information is accurate and, if it is not complete, to provide clear notation of such in each service area.

Weatherization Assistance Program (WAP) funds are provided through NCDEQ and distributed by CAAs and are used to provide home upgrades that improve energy efficiency. In addition to the allocated funds, data associated with weatherization can include pre- and post-retrofit blower door testing, installation costs, material costs, contractors, and resident application and assistance dates. Through the PEEIF project data acquisition process, weatherization data was sometimes provided in multiple tables, which made visualization difficult for specific data. For example, multiple types of weatherization measures—such as weather stripping, insulation and/or a new, more efficient heating and air conditioning system—may take place in one home, which makes the visualization of multiple assistance types difficult in a spatial format due to existing limitations of software and the ability to query and visualize related tables.

Data can be found by contacting assistance program data managers. As was the case with utility information, this data may be difficult to acquire without proper legal agreements.

Data Analysis

A regression-based model that uses home information can be used to identify homes using more energy than expected when corrected for age, size, and the contribution of heating and cooling loads. Ranking homes by their performance on these two criteria provides a way to screen for homes that may be good candidates for energy-assistance measures. The model assumes that housing-related variables, such as the age and size of the home and total degree days, are correlated with energy use. The results are then compared with occupant income or income proxy data such as the home tax assessed value to select candidates for follow-up. The candidate lists' location data provide geographical clusters for additional study.

The following data resulted from a case study on homes in Wilson County, North Carolina. A sample of homes is given in Table 2.

Tracking number	Age (years)	Size (sq. ft.)	Use (kWh/year)
1	37	1044	14,074
2	12	1466	13,517
3	47	1058	5,615
4	6	1156	5,604
5	61	810	2,971
6	59	1131	16,209
7	32	1884	10,430

Table 2: Characteristics of houses used in the study

Step 1: Perform a Regression Using Size and Age to Predict Annual Power Use.

Better predictions can be made by dividing homes into size and age clusters with similar characteristics. For example, in North Carolina, energy-efficiency-related building codes changed in 1975, 1982, 2002, and 2005. Clusters were found with breaks at 2005 and 1974.

The results from regressions for a subset of homes is seen in Table 3. In the case study, predicted use was used to correct for the age and size of the

Tracking number	Actual use (kWh/year)	Model use (kWh/year)	Error
1	14,074	10,749	(0.24)
2	13,517	12,192	(0.10)
3	5,615	10,776	0.92
4	5,604	10,071	0.80
5	2,971	9,867	2.32
6	16,209	11,007	(0.32)
7	10,430	13,732	0.32

Table 3: Comparison of model predictions with actual energy use values

home. The more the model under-predicts the annual energy usage (as it did for homes one, two, and six), the more likely the home is suitable for weatherization.

Step 2: Calculate the Correlation Between Total Degree Days (TDD) and Daily Energy Use.

Homes with strong weather correlations use more energy to maintain indoor temperature. This can be because the occupants choose to use more energy by setting the thermostat farther from the outside temperature or because the home is less energy efficient. A strong correlation may demonstrate the potential need for weatherization assistance, but is not a perfect indicator.

Table 4 gives sample energy use correlation with weather data from Wilson, North Carolina, from January–May 2017. A perfect correlation would have a score of 1, no correlation would have a score of 0, and an inverse correlation would have a score of -1. A home with a score near 1 would use more energy to heat or cool when the outside temperature is colder or hotter.

Tracking number	Energy use (kWh/year)	Weather correlation
1	10,430	0.16
2	2,971	0.28
3	14,074	0.83
4	13,517	0.61
5	5,604	(0.18)
6	5,604	0.81
7	16,209	0.04

Table 4: Correlation between energy use and weather.

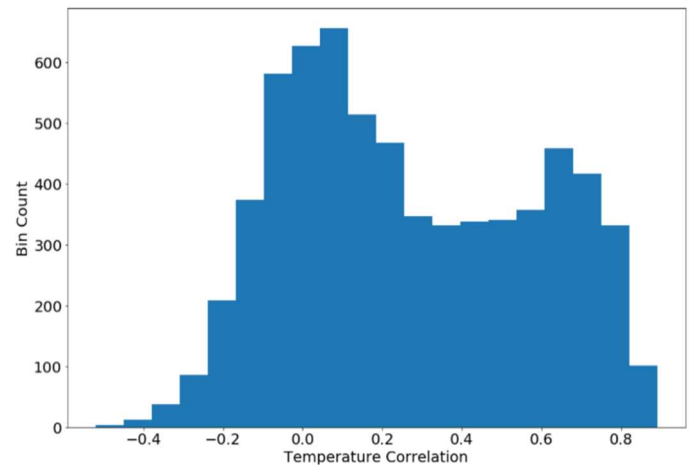


Figure 4: Histogram of Correlated Temperatures, Wilson, North Carolina, January–May 2017

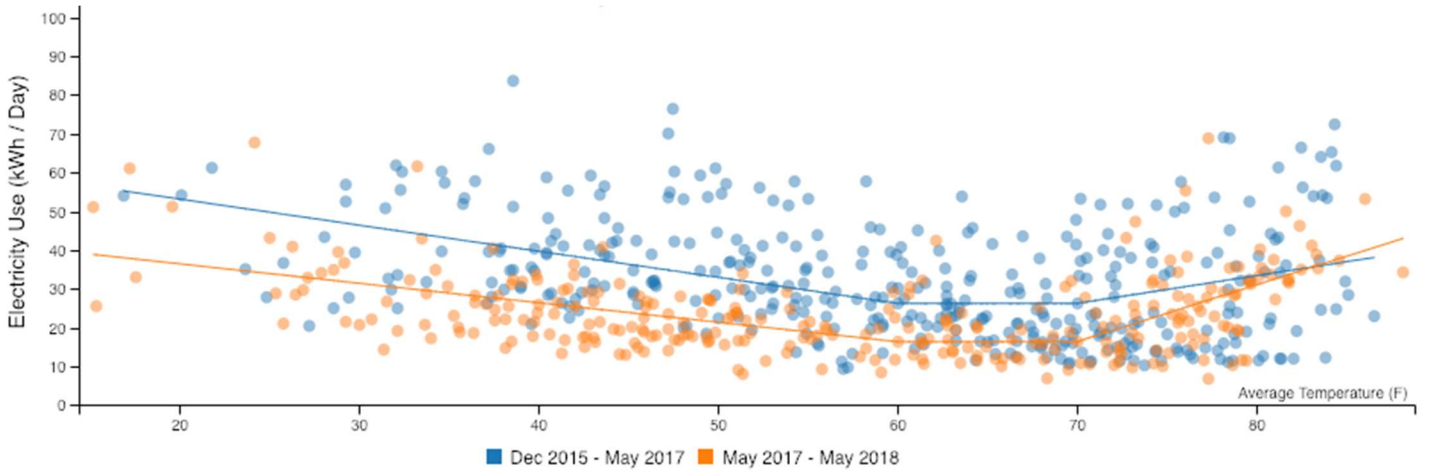
A histogram of the correlation scores (Figure 4) for the entire dataset reveals a bimodal distribution, with one set of homes clustered around 0.7 and another around 0. This dataset was captured during winter months, so homes with alternative heating sources fell into the distribution near 0, while homes that used electric heaters or heat pumps (since the data was average annual electricity use) showed a stronger correlation.

Step 3: Rank and Sort Based on Regression and Weather Correlation Analysis Results.

The two analyses in steps 1 and 2 were then used to rank-order the homes and identify the ones that were potentially in need of greatest assistance. Taking the sample of homes in Table 5, home number six consumes significantly more energy than expected (error rank of 1) and has a strong weather correlation (correlation rank of 2).

Tracking number	Model error	Weather correlation	Error rank	Correlation rank
6	(0.32)	0.81	1	2
3	0.92	0.83	6	1
2	(0.10)	0.28	3	4
1	(0.24)	0.16	2	5
4	0.80	0.61	5	3
7	0.32	0.04	4	6
5	2.32	(0.18)	7	7

Table 5: Ranking of homes based on the correlation between energy use and weather and model under prediction of energy use, indicating a possible need for weatherization.



House Number	Home Type	Year Built	Square Footage	Pre Improvement Consumption	Post-Improvement Consumption	Annual Savings (kBtu/year)	Annual Energy Savings
House 1	SINGLE FAMILY	1955	950	79426	66657	12769	16.1%
House 2	SINGLE FAMILY	1975	1050	187063	143638	43425	23.2%
House 3	SINGLE FAMILY	1950	2600	94061	95332	-1271	-1.4%
House 4	SINGLE FAMILY	1955	850	25488	23021	2467	9.7%
House 5	SINGLE FAMILY	1950	1400	36270	27873	8397	23.2%
House 6	SINGLE FAMILY	1955	1200	21547	21103	444	2.1%
House 7	SINGLE FAMILY	1960	1750	15931	15490	441	2.8%

Figure 5: Scatter plot of the pre-improvement and post-improvement daily electricity usage compared to the daily average outdoor temperature for a home. The data table illustrates the total normalized annual energy consumption and energy savings calculated for seven weatherized homes.

The top-ranked homes are more likely to need weatherization assistance. Users could define weights for each attribute and select as many homes as desired. Table 5 shows a ranking from a subset of homes that illustrates how the top homes make the best candidates for weatherization assistance based on the information available.

This method can also be extended with additional information. Agency-dependent criteria like energy burden or household income can be added to the table and ranked to adjust the search criteria. Furthermore, top-ranked homes could be placed on a map using GIS data to identify neighborhoods with clusters of identified homes. Refer to Appendix F for a Python script used in the PEEIF project to determine the energy-efficiency potential of various homes.

Step 4: Evaluate Energy-Saving Measure Retrofit Outcomes.

The process of measuring savings from a whole-home retrofit involves comparing the pre-improvement and post-improvement periods using temperature-based regression models. Each regression model comprises five coefficients for electricity usage and three coefficients for natural gas usage:

- Marginal heating use in kWh or therms per degree day
- Heating balance point temperature (i.e., the outdoor temperature at which the home begins to heat)
- Non-temperature related average energy usage in kWh or therms per day
- Marginal cooling use in kWh per degree day

- Cooling balance point temperature (i.e., the outdoor temperature at which the home begins to cool).

The computational steps for measuring energy savings are:

1. Determine a best-fit degree day regression model for the pre-improvement and post-improvement periods for each home.
2. Calculate normalized annual consumption separately for the pre-improvement and post-improvement periods by fitting the coefficients of the regression models with daily climate normal temperatures.
3. Subtract the post-improvement from the pre-improvement normalized annual consumption,

with the difference representing the energy savings.

As shown in Figure 5, the blue line represents the best-fit, pre-improvement weather usage model, and the orange line represents a commensurate post-improvement model. For this home, a normalized annual electricity consumption savings of 23% is calculated. Annualized electricity and natural gas usage are then converted to kBtu and added together to determine total pre-improvement and post-improvement normalized annual home energy consumption values. The seven weatherized homes in Wilson, North Carolina, realized an average combined electricity and natural gas savings of about 9500 kBtu, or 11% of annualized total energy consumption.

ANALYSIS PARTNERS

Center for GeoSpatial Analytics

The PEEIF tools created by NCSU's Center for GeoSpatial Analyses enables users to focus on areas of interest and ask questions by using combinations of filters. The filters refine the map display to reveal data that help the user understand the combination of assistance, housing and electricity use. Specific data is revealed through pop up boxes when the user clicks on a house or area.

ResiSpeak:

ResiSpeak is a utility data analysis software and services provider servicing utilities, building portfolios, and local government. ResiSpeak operates a weather database with data provided by NOAA in order to compute whole building energy use models for compliance with industry accepted guidelines for energy savings measurement & verification as well as for inefficient home identification. ResiSpeak collected and formatted usage data sets from the data providing utilities and computed metrics for upload to the PEEIF database, as well as provided data visualization services through its online portal. Energy usage metrics provided by ResiSpeak included annualized energy usage, energy usage intensity, usage percentile within a peer group, usage decile within a peer group, percentile and decile range values for the peer groups, marginal energy usage for heating and cooling, and non-temperature related energy usage. In addition, ResiSpeak contributed to the analytics developed specifically for the PEEIF grant by the System Design Optimization Lab. Note that as of 2019, ResiSpeak is known as Enpira.

System Design Optimization Lab:

As a part of NC State University's Mechanical & Aerospace Engineering department, the System Design Optimization Lab (SDOL) explores how design decisions can be made to maximize system value by creating tools and methodologies that advance the state-of-the-art in design theory and design automation. SDOL products help navigate the interdisciplinary challenges associated with designing consumer products and complex systems. For the PEEIF project, the SDOL group explored regression analysis of energy use data and use data correlations. The aim of this analysis was determining the energy efficiency potential of homes in the PEEIF data such that those structures that need assistance can be identified.

CREATED APPLICATIONS

GIS-based web mapping applications were created for the data-providing partners and tailored to their specific needs. The selected platform, ArcGIS Online for Organizations, uses user accounts and groups to compartmentalize and share online content. Users are assigned to groups based on their application access requirements. To provide application access, a group was created for each data-providing partner and web-mapping applications were shared within that group.

To provide web-mapping application access to each data-providing and stakeholder organization, each organization provided contact information to PEEIF project personnel, who created the requested account and assigned the user to the appropriate ArcGIS Online for Organizations group. This workflow ensured that all users gained access to the confidential data specific to their organization, publicly available data, and aggregated or blurred information from other project stakeholders. Appendix G provides more detailed description of

the PEEIF technical process workflow, challenges, and future considerations.

Below are descriptions of the applications that were created and examples of data analysis functionalities.

North Carolina Department of Environmental Quality

The NCDEQ application portal was created to provide visualization and analysis of weatherization assistance implementation data for NCDEQ personnel and contractors. In particular, the application displayed before and after energy usage data for houses that received weatherization assistance funds through WAP and HARRP. Note that additional application portals were created for the three Community Action Agencies that distribute NCDEQ assistance funds in the UCPCOG region: CADA, WAGES, and Franklin-Vance-Warren Opportunities, Inc.

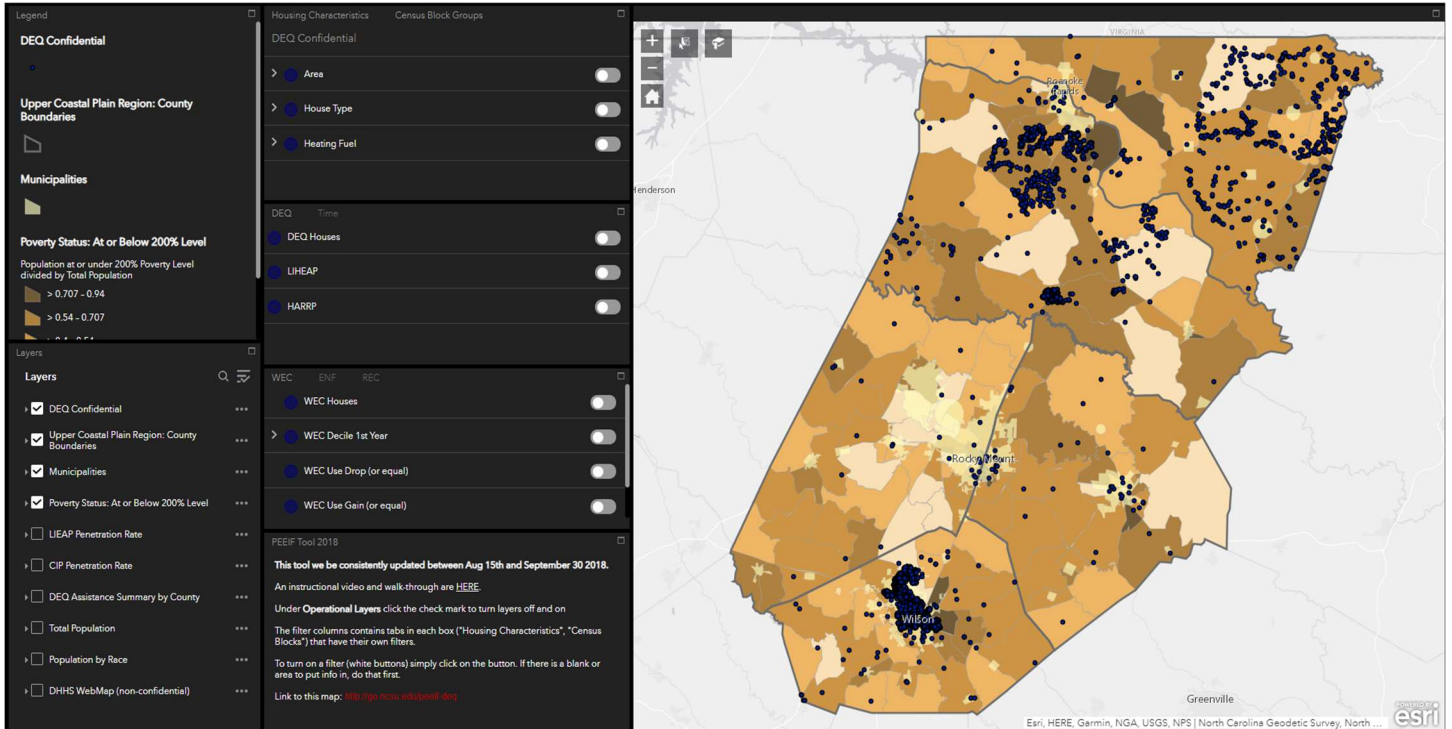


Figure 6: Screenshot of the NCDEQ application displaying houses that have received DES assistance overlaid on census tracts and categorized by the concentration of population that is more than 200% of poverty level

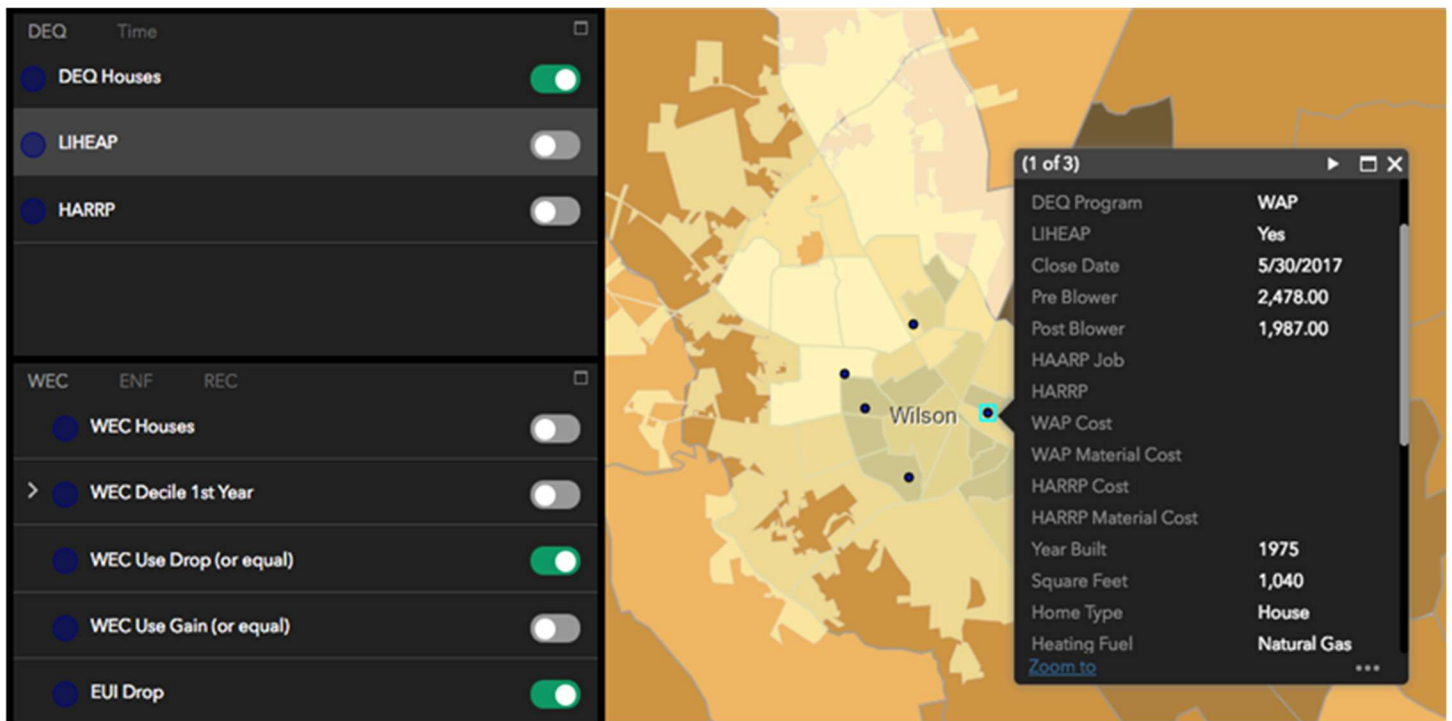


Figure 7: Screenshot visualization of NCDEQ homes in the Wilson, North Carolina, area that have shown a decrease in energy use intensity after receiving weatherization assistance

As shown in Figure 6, the application contains the following layers:

- NCDEQ confidential house information, containing all NCDEQ attributes and non-confidential data from the remainder of the PEEIF data set
- The Upper Coastal Plain region boundaries (within the five-county region)
- Non-confidential NCDHHS data showing the amount of the population at 200% or greater of the poverty level per household (Census Block Groups)
- The LIEAP Penetration Rate (Census Block Groups)
- The CIP Penetration Rate (Census Block Groups)
- NCDEQ's Assistance Summary by County (County)
- Total Population (Census Block Groups)
- Population by Race (Census Block Groups)
- DHHS assistance statistics (Census Block Groups).

The final application contains functionality to:

- View active layers

- Toggle layers on and off
- Display selected features based on attribute filters using existing layers.

Specific data is revealed through pop-up boxes when the user clicks on a house or area. Figure 7 shows houses in the Wilson area that have received NCDEQ assistance and consequently had their overall electricity usage and Energy Use Intensity (EUI) decrease. The pop-up box reveals some specific information about the assistance received, housing characteristics, and electricity use (not pictured). Note that the combination of filters that have been turned on: "DEQ Houses," "WEC Use Drop," and "EUI Drop" have been used for this output. The map can be zoomed and panned to see other homes that meet these criteria.

North Carolina Department of Health and Human Services

The NCDHHS application was created to provide visualization and analysis of income assistance throughout the five-county region. The application displays the variability of income assistance across the area. In particular, the application provides the ability to visually analyze Census tract areas that

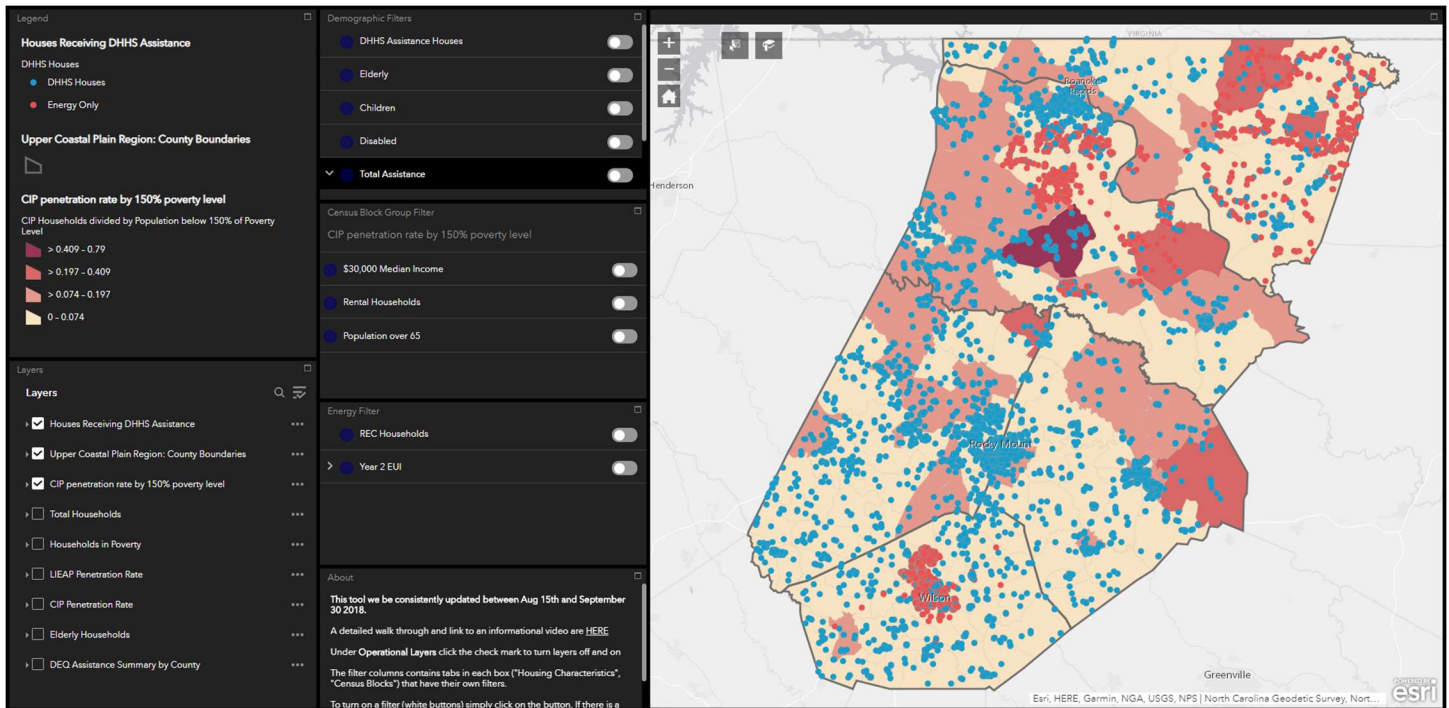


Figure 8: The NCDHHS application displaying NCDHHS houses (blue dots) and energy data houses (red dots) overlaid on a map showing CIP penetration rates by Census tract

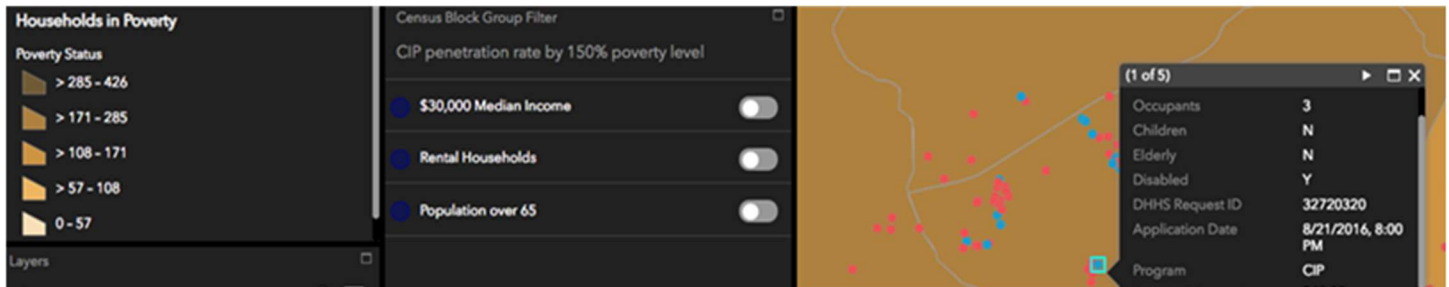


Figure 9: Roanoke Rapids area in the NCDHHS web application

have received significant assistance and those that have not.

As shown in Figure 8, the application contains the following layers:

- NCDHHS confidential information, including all NCDHHS attributes and non-confidential data from the remainder of the PEEIF data set
- The Upper Coastal Plain region boundaries (for the five-county region)
- The LIEAP Penetration Rate (Census Block Groups)
- The CIP Penetration Rate (Census Block Groups)
- The DEQ Assistance Summary by County (County)
- Total Population (Census Block Groups)

- Population by Race (Census Block Groups)
- Elderly Households (Census Block Groups).

The final application contains functionality to:

- View active layers
- Toggle layers on and off
- Display selected features based on attribute filters using existing layers.

The PEEIF tool has multiple data layers that allow census tract data to be visualized in a variety of ways. Overlaying an individual house's electric and assistance data on this layer data and using filters to further refine the display can help identify areas and/or individuals that may be in need of assistance. Appendix H contains U.S. Census and

data map outputs related to NCDHHS penetration and visualization by ZIP code.

Figures 8 and 9 show houses that have received NCDHHS Assistance (shown in blue) and houses with electric data (shown in red) overlaid on the poverty layer. The dark brown color indicates high poverty in the Census tracts. Additionally, a filter has been used to display homes with a high EUI. The specific house selected in Figure 9 has received NCDHHS assistance and still has a high EUI (.90 decile, not pictured). This home could be a good candidate for weatherization to prevent further need for NCDHHS assistance. There are a number of high-EUI homes in this relatively small area, which may indicate there is a greater need for assistance here.

Energy Utilities

Utility applications were created to provide visualization and analysis of residential energy usage for the specific utility service area in UCPCOG region:

- Roanoke Electric Cooperative (REC)

- Wilson Energy—City of Wilson, North Carolina (WEC)
- Town of Enfield—North Carolina Electric Service (ENF).

These applications display home energy usage in relation to aggregated information about low-income homes and assistance programs. In Figures 10, 11, and 12 are example screenshot images from Wilson Energy, the Town of Enfield utility, and the Roanoke Electric Cooperative applications that provide the utilities with the ability to visualize homes by their annual energy usage and square footage. For Wilson and Roanoke Electric, where two years of energy use data were provided, users are able to identify homes that have increased or decreased energy use by percentiles over the two years.

The applications contain the following layers:

- Utility-confidential house information containing all energy attributes and non-confidential data from the remainder of the PEEIF data set

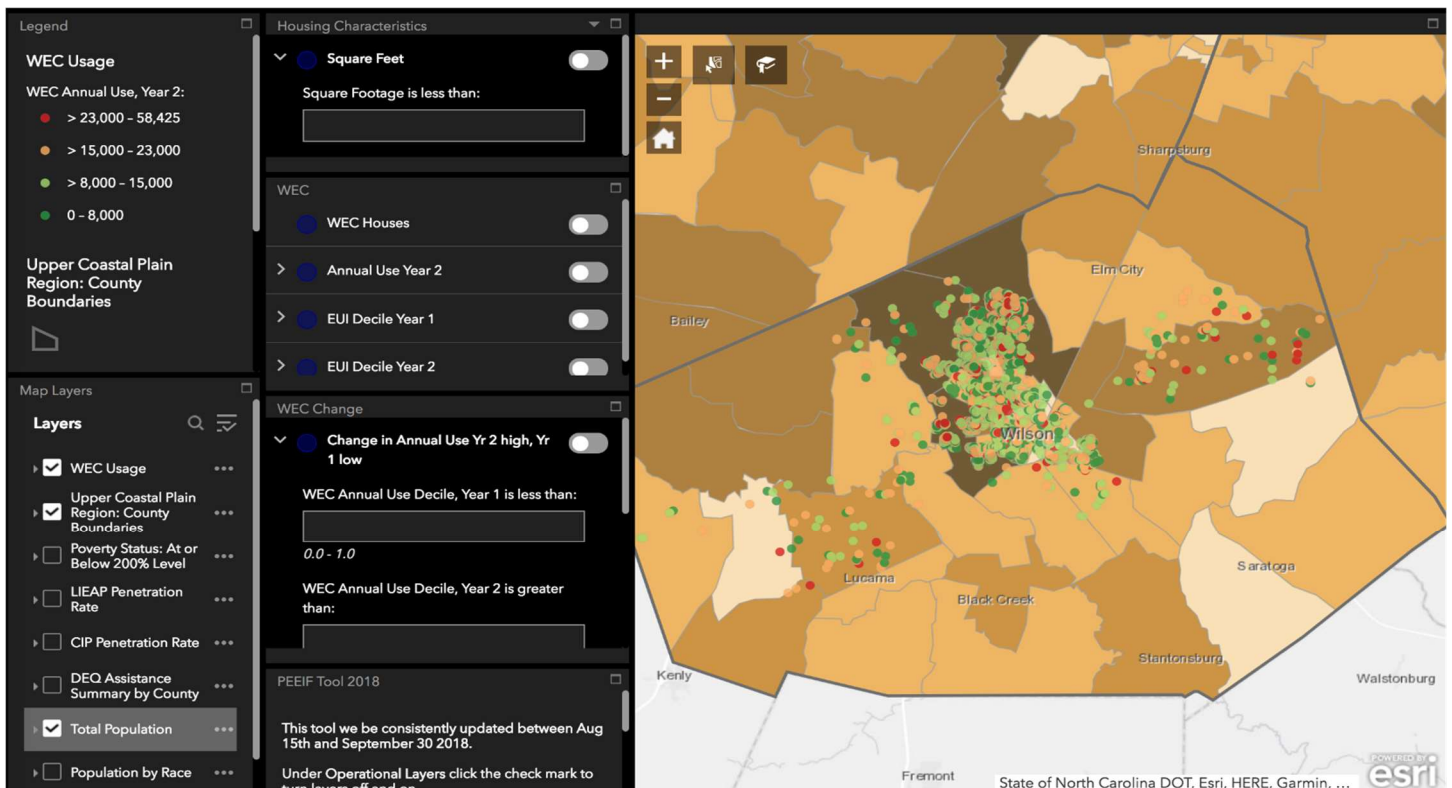


Figure 10: Screenshot of the Wilson Electric web application displaying the annual energy use of individual and the total population from the Census tracts in the Wilson Electric region. To the left of the map are the legend, additional data layers, and filters that can be used to refine the data displayed.

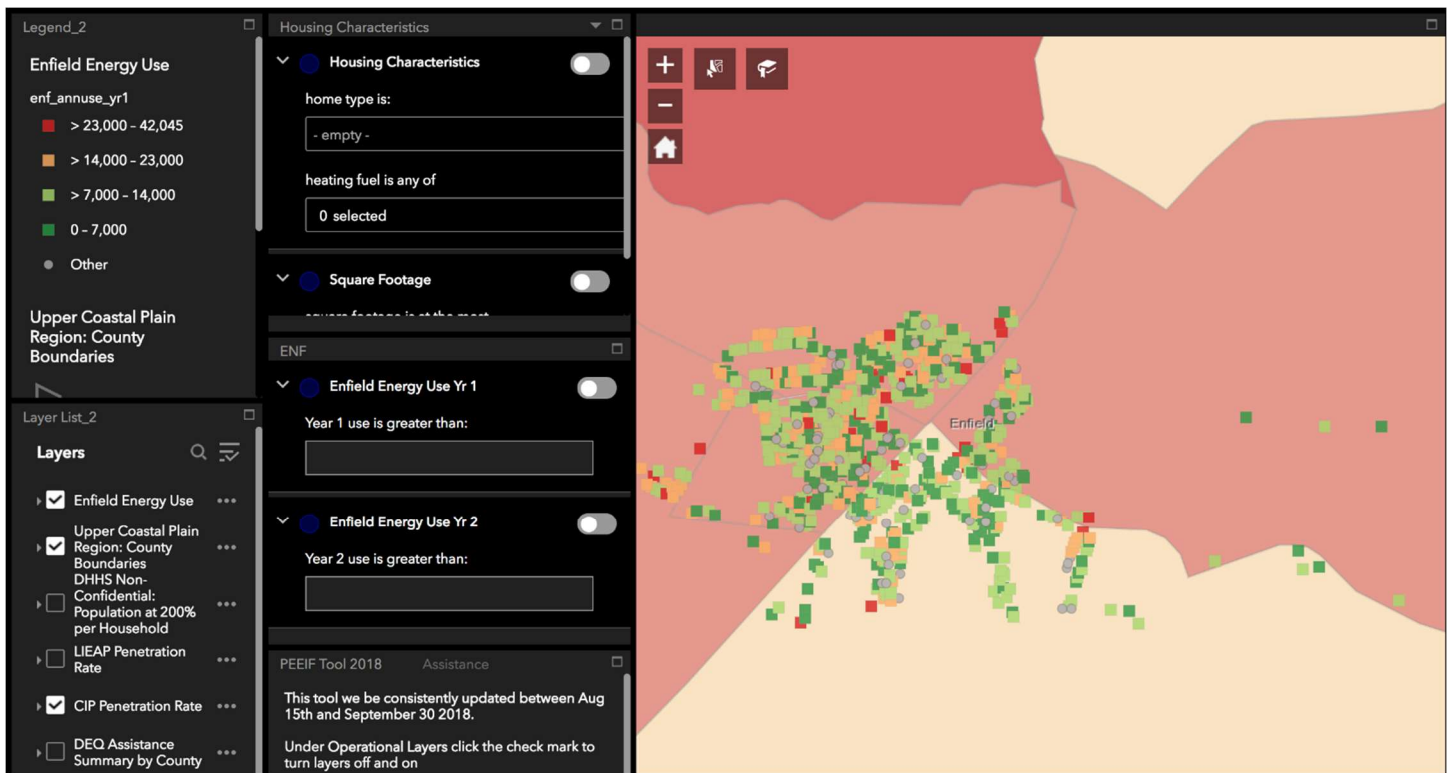


Figure 11: Screenshot of the Enfield Electric web application displaying the annual energy use of individual homes and the CIP penetration rate of the census tracts in the Enfield region. To the left of the map are the legend, additional data layers, and filters that can be used to refine the data displayed.

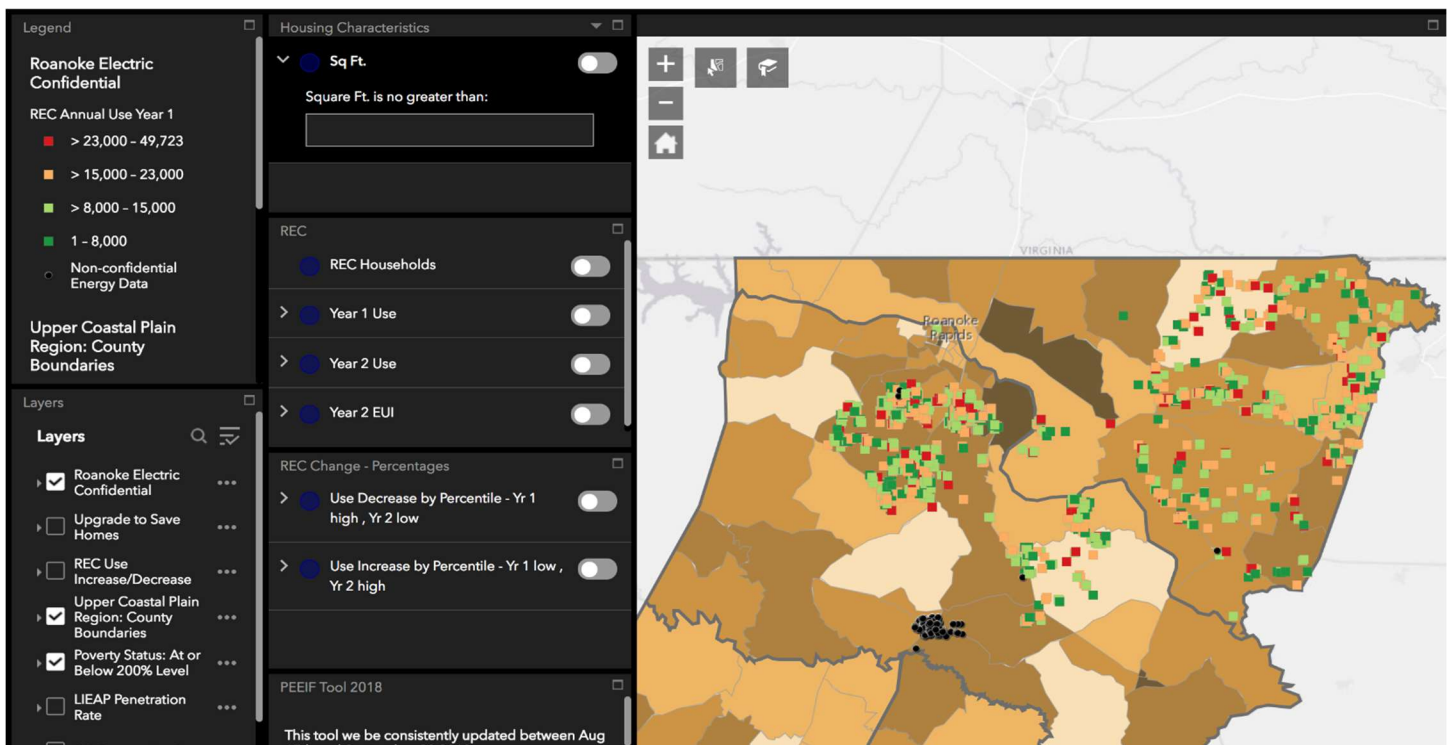


Figure 12: Screenshot of the Roanoke Electric Co-op web application displaying home annual energy and the REC region's Census tract poverty status percentage. To the left are the legend, additional data layer, and filters that can be used to refine the displayed data.

- Utility-confidential data on increased or decrease energy usage between year one and year two
- Upgrade to Save Homes data (for REC only), which includes data related to REC’s Upgrade to Save program
- Upper Coastal Plain region boundaries (for the five-county region)
- The LIEAP Penetration Rate (Census Block Groups)
- The CIP Penetration Rate (Census Block Groups)
- NCDEQ Assistance Summary by County (County)
- Total Population (Census Block Groups)
- Population by Race (Census Block Groups)
- DHHS statistics (Census Block Groups)
- Poverty Status for homes below 200% of the poverty level (Census Block Groups).

The final application contains functionality to:

- View active layers
- Toggle layers on and off
- Display selected features based on attribute filters using existing layers.

REC Map Analysis

As shown in Figure 12, the REC application contains a layer on the Upgrade 2 \$ave program that contains weatherization information for REC houses and allows for visualization of efficiency programs in addition to energy use.

An electric utility has a large amount of data available on energy usage. The PEEIF tool, however, enables the utility to view that data spatially against a backdrop of demographic data. This may help an interested utility in decision making, program targeting, and marketing.

Figure 13 uses layers and filters to discover what areas are being served and where future efforts may need to be made. The circles represent homes that have participated in Roanoke Electric Coop’s Upgrade to \$ave program. The squares are REC homes with a high EUI. The background layer is the "Poverty Status: At or Below 200%" layer. This layer also includes the "Median Income Less than \$30,000" filter. The brown Census tracts indicate a high concentration of poverty and low income. The gray census tracts have a median income greater than \$30,000 and less concentrated poverty. There are several Upgrade to \$ave homes in this gray



Figure 13: Weldon area displayed in the REC application

area. Although there are more homes overall in this gray area, REC may be interested in targeting more of the homes in the higher poverty areas of its territory. Since the homes have been filtered for a high EUI, the homes represented by orange and red squares (indicating high overall energy use) may be of specific interest.

Utility Attributes Lessons Learned

Utilities contributed data in multiple resolutions, units, and formats. PEEIF encountered files in XLS, CSV, XML, TXT, and ACCDB formats. The electricity data time resolutions ranged from 15 minutes to non-timestamped monthly. A natural gas utility contributed raw natural gas meter reads in units of CCF. For conversion to therms, kBtu and other standard energy metrics, heat content conversions had to be applied based on the time of usage. It was important to standardize a set of metrics and tools that could be used regardless of the resolution of the input data and could output additional metrics and graphics if a given data set lent itself to such. This ensured that the toolset could be scalable, regardless of the quality of the inputted data.

In addition, the utilities had varying degrees of geospatial identification for their customer accounts. In the best-case scenario, a utility would have a master account list that included a location identifier that was also present in the county parcel information. In other circumstances, no matching fields existed between the utility's master account list and the county parcel information. If the utility provided a latitude and longitude for the location, the coordinates could be used to geospatially locate the residence, though the precise latitude and longitude coordinates may have been for the physical utility meter rather than, for example, the driveway or mailbox, which meant that matching utility data to geospatial data based on latitude and longitude required careful oversight. In the worst-case scenario, no overlapping information was available and geolocation had to be performed through address text matching.

Extracting useful utility data for use in the PEEIF application presented several notable challenges, including the following.

Data Access: The utilities may be unfamiliar with the processes required to extract usage data in a useable form from their meter data systems, and may require technical meetings to establish procedures.

Data Format: The utilities all had unique data formats in their outputs. Multiple input formats were required, even within individual utilities.

Geolocation: The utilities may not have a way to map their meters or customer premises to county parcel IDs. In such cases, geolocation was performed by text matching physical addresses or by applying latitude and longitude data provided by the utility. Both techniques were problematic because:

- Physical addresses may not match parcel addresses due to the use of abbreviations, hyphens, spaces, and other nuances. For example, Martin Luther King Junior Boulevard could be spelled "ML KING JR BLVD" or "MARTIN LUTHER KING JR BLVD".
- Latitude and longitude data from the utility may correspond with the physical location of the electric meter, whereas the latitude and longitude data from the parcel ID may correspond with the primary street entrance to the property.

Time Range: Generally, the utilities provided usage data that only went back 1–3 years. Although a single year of continuous data was sufficient for calculating usage metrics for homes, other award partners such as NCDHHS and NCDEQ had records going back many years (5–8), so there was only a very small number of homes where it was possible to compare energy use before and after NCDHHS or NCDEQ intervention. This challenge could be remedied by knowing the years of data that state agencies and/or non-profits are providing in advance of facilitating utility data contributions in order to have the time frame of data providers line up with conduct-related analysis.

Legal Process and Steps

The core functions of the PEEIF database are data in, data out; and data in, analytics out. Functions must abide by the privacy protections imposed on

the movement and use of the data. The movement of data is accomplished by a legal structure and non-disclosure agreements (NDAs) that provide access to data while protecting it. The NDAs allowed UCPCOG to receive data from state agencies and electric utilities and pass it through the secure NCSU Velocity server, where it was cleaned, formatted, and assessed by the subrecipients. Subrecipient access was managed through a secondary NDA that required the subrecipient to maintain the same level of data security as the data-providing party and to vet individuals who were granted access to the data and preserve all confidential information within a protected database.

This combination of NDAs allowed for data to flow into PEEIF while allowing data analytics to flow back out of PEEIF. The privacy protections imposed on the data arose from its confidential nature. The program data was used to indicate which households may qualify for or had already received assistance. However, program eligibility is determined by several sensitive criteria, including income level, age, disability status, and the presence of children. To protect the privacy of applicants and recipients, this confidential information can only be shared in limited circumstances. Data can only be shared with the original data-providing party or with other parties who have completed secondary agreements. These analytics are helpful, but they do not allow the full capture of all the value in the database. To expand the potential applications of PEEIF analytics, the NDAs also allowed for non-confidential information to be disclosed to parties outside of the NDAs. The PEEIF data analytics team employed blurring and aggregation techniques to remove identifying factors from the data records while still maintaining the value of the data analytics.

The NDAs also integrate how data is collected from the different programs. While federal funding flows to state agencies, it is the county-level Divisions of Social Services and regional Community Action Agencies that are the point of contact with program applicants and contractors. Consequently, the

NDAs needed to be constructed to allow for access to data and data analytics at the state, county, or regional levels. This was accomplished by expanding the definition of a "data-providing party" to include contractors and agents and by assessing existing state regulations to determine when information can move between the state agency and the county Divisions of Social Services.

The construction of the NDAs reflects the legal restraints on data sharing and the practical realities of program delivery. Pairing these conditions with the cooperative agreement structure allowed for data to move into the PEEIF database, to be cleaned, to be evaluated, and to flow back out. This allowed for data-providing parties to see their own data analytics with the highest level of granularity and other program information at an aggregated level. This structure drove internal program development while also incentivizing future collaboration between agencies. Refer to Appendix I for a sample NDA.

There are a number of steps in the legal process:

- The first step must be to identify potential sources of program data, the agency collecting the data, and the source of funding for the program.
- Develop your business case before your legal arguments.
- Plot data flows into and out of agency and utility systems. Use the data flow map to guide the development of your legal documents.
- Identify the highest standard for data security among all potential data providers. Build your data protection to that standard.
- Locate any existing utility customer data protection policies. Start by identifying the applicable regulatory body for each type of utility that may upload data into the database and search for existing utility customer data protection policies.
- Leverage the database to build a system and relationships that can address and evaluate current program delivery and enhance future program delivery.

RECOMMENDATIONS

Stakeholders (e.g., data providing partners and database and application developers) must coordinate and discuss the requirements for and strategies to implement data provider needs at the earliest stages. This coordinated involvement will provide an iterative application development process that ensures that all stakeholders are provided applications that are functional and resilient enough to help identify and assist with the needs of their low-income energy service program. In addition, stakeholders must work together to ensure data is provided in a uniform manner and to expedite database and application development and accuracy. The data must be provided in a consistent and repeatable format that allows subsequent data at differing temporal and geographic scales to be quickly integrated into the system and deployed through web-mapping applications. This assists in alleviating data irregularities and anomalies that require data manipulation before entering into the database and web-mapping systems. Finally, involving stakeholders at all points in the application development process may identify the need to develop a suite of focused applications for each agency instead of one application interface that may not be optimum for all application users.

Spatial Analysis

Several distinct spatial topics could be explored in this project. In some cases, data were aggregated to different spatial scales to examine spatial patterns at a broader scale. However, NCDHHS's data were unable to accurately be located at a high rate due to data quality, which meant that fine spatial analysis was difficult to perform. Certain types of analyses, such as penetration rate, were inaccurate due to the total houses being lower than normal. In other cases, the data were only available by P.O. Box address, which meant that no house could be geolocated. However, ZIP code-level data was provided as part of the dataset, and these could be linked to ZIP code geographies, essentially allowing accurate statistics at these levels. ZIP code areas are broad, which means that spatial patterns cannot be fully recognized, but they can still provide finer resolution results than to county geographies.

There are multiple opportunities for using this data in explicit spatial analysis. Tools that perform clustering analysis on multiple variables could provide insight on the ways that the variables exist, and may discern patterns that weren't available before the data were collected. This is especially true of energy data.

Future Legal Agreements

The PEEIF platform opens new opportunities to facilitate the development, implementation, and targeting of low-income, energy-related services. While the established legal and technical frameworks work in concert to move data into the database and flow analysis back out of it, the structure and organization could be improved to create a sustainable and adaptable platform that could facilitate the addition of new data-sharing parties and other parties granted access to data. Ideally, this system could allow for the development of new data uses.

The PEEIF legal framework is based on individual non-disclosure agreements signed with a centralized party. Within PEEIF, the data-sharing parties and other parties that were granted access to the data signed NDAs with the UCPCOG, and the terms and conditions were restricted to sharing data with and through UCPCOG. While this template links individual data-sharing agreements, it did not encourage data sharing between parties, and it constrained the ability to evolve the database's functions in response to user needs. A multi-party non-disclosure could address these and other issues.

Multi-Party Non-Disclosure Agreements

PEEIF is built on primary and secondary non-disclosure agreements where UCPCOG serves as the central hub for the collection of and movement of data. NDAs were constructed for data-sharing parties and data-accessing parties. This structure was used because it mimics the DOE grant structure, but it could be improved to focus on the location of the database and the organization that controls it.

A multi-party NDA between all the parties sharing and accessing data should be built around where the database is housed and include all of the parties' agreed-upon security protocols and conditions. Below are key considerations in the development of a multi-party NDA.

Unified Data Security Protocols

A single-data security protocol will reduce the technical limitations on enhanced data sharing. State agencies with the authority to share data must often demonstrate that the recipient of the data can provide a similar level of data protection as the data-sharing agency. Each participating North Carolina agency follows the universal compliance obligations contained in the *North Carolina Statewide Information Security Manual*, as well as their own data-management protocols. Building a system that addresses individual security needs and shared security obligations is critical to creating a multi-party platform.

Acknowledgement and Incorporation of Data Protection and Data-Sharing Regulations and Laws

The NDA must be able to acknowledge and respond to the different sources of data-privacy regulations and the mechanisms by which those regulations and statutes are enforced. A multi-party NDA may include state agencies, investor owned utilities, electric membership cooperatives, municipal owned utilities, and third-party service providers, each of which may be responding to a different combination of guidelines, policies, ordinances, by-laws, regulations, and statutes. Protecting confidential data from unauthorized use is a core responsibility of every signatory to the NDA. Unlocking the potential of data sharing is core goal of every signatory to the NDA. Consequently, the NDA should include terms that recognize the current legal environment while having the flexibility to adapt to the needs of new parties without affecting data security.

Flexibility of Purpose

Unlocking the multiple uses for and values of each piece of data stored in the database requires a flexible NDA capable of evolving to the needs of its signatories. A PEEIF-style database can produce

value by allowing for single-party data analysis or by facilitating cross-party data sharing and analysis. In single-party data analysis, the data provided by a party is uploaded, analyzed, and returned to the uploading party. The application of third-party analytics can create new methods of viewing data, synthesizing data for reports and audits, and compiling historic trends. The single source of data and single destination of analytics is easy to capture in an NDA. The difficulty lays in formulating a system that permits for cross-party data sharing. Cross-party data sharing opens the possibility of enhanced synergies between different programs and program providers to assess quality of services provided, identify unserved and underserved populations, and reduce administrative burdens.

Cross-party data sharing requires an NDA with a flexible purpose, an open-ended period of operation, and the ability to capture individual agreements between parties that can modify the purpose and the length of a specific data-sharing arrangement.

The purpose of the NDA must be flexible enough to capture a variety of known and unknown uses. The known uses are present when the NDA is created, and they help frame out the range of potential uses served by allowing for data sharing. The unknown uses arise as parties develop their knowledge of the potential uses and benefits of the database. Having the flexibility to adapt to unknown, beneficial uses requires the creation of an NDA that is expansive, yet focused. It must be expansive enough to accommodate a variety uses and be focused on the delivery of program services tied to agency or entity goals. In combination, the two elements produce a range of uses that are flexible yet constrained.

The period of operation for the NDA should be open-ended, with clear rules for how parties can enter into and withdraw from the agreement and provisions that address how data will be managed after a decision to terminate participation is made.

A narrow time period does not account for how agencies function and how resources are tasked to implement data sharing processes.

Legal Considerations

- Early knowledge will streamline the process of completing data-sharing agreements. Identify the desired types of data, which entity holds the data, and if there is federal or state funding for the program. This will allow project proponents to recognize potential conflicts or barriers before engaging in efforts to secure data. Federal program funds require the receiving agency to comply with federal statutes and regulations. The availability of the data is affected by the privacy and consent obligations imposed as a condition of receiving federal funds. For example, the LIHEAP statute and regulations encourage coordination and data sharing between low-income energy service providers and sources of program funding. On the opposite end of the spectrum, federal banking regulations may impair the sharing of energy efficiency mortgage program data. In the PEEIF project, the North Carolina Housing Finance Agency was prevented from sharing historic program data without securing the consent of each client.
- Develop the business case first. Before drafting any legal documents, present the project's business case to the policy and technical teams. This step builds consensus with potential data providers and gains access to insider knowledge of data security compliance obligations while mapping data flows into and out of system. Both items must be addressed in the legal documents supporting the data sharing.
- Use the map of data flows into and out of agency and utility systems to guide the development of your legal documents. State agencies work with local program partners to evaluate applicants and service clients. As such, the agency funding the program is often not the party collecting or maintaining the data or the only party that would benefit from access to data analytics. For example, CAAs deliver most of the weatherization programming in North Carolina. CAAs subcontract with the NCDEQ, receive and evaluate applications, and perform project audits. Emergency assistance and crisis intervention programs are offered through county Divisions of Social Service. None of the employees in these organizations were state employees, a fact that needed to be considered in legal agreements.
- The data protection standard for the database must meet or exceed the highest individual standard of all data-providing parties. Working with multiple data-providing parties will mean that there will be multiple data-protection standards. Identify the most stringent standard and adopt it. In some jurisdictions, a statewide information security protocol might apply in addition to agency-specific procedures. For North Carolina, all state agencies must comply with the *Statewide Information Security Manual*. Additionally, the PEEIF database was housed on the North Carolina State University computer system, which had its own database security protocols. Team members evaluated the state requirements and determined that the university data management procedures exceeded them. This allowed for the application of a universal set of data security protocols that met all compliance requirements.
- Identify the regulatory body for each type of utility uploading data into the database. Within a state, investor-owned utilities (IOUs), electric membership cooperatives (EMCs), and municipal-owned utilities might be subject to the same regulatory oversight or governed by three separate organizations. In North Carolina, the North Carolina Utility Commission is the regulator for IOUs, but it does not oversee the activities and operations of EMCs or municipal utilities. Those entities report to local government officials or a member-elected board of directors. The commission has developed rudimentary data privacy guidance for IOUs; the EMCs and municipal-owned utilities are not subject to the same rules.
- Leverage the database to build a system and relationships that can evaluate current program delivery and enhance future program delivery. The database is an opportunity to retrospectively evaluate program performance and streamline current administrative and reporting processes. Those are critical to making a successful business case. However, the database is also an opportunity to improve the client application process and inter-agency and third-party data sharing practices. Leverage engagement with key policy, legal, and technical personnel to identify and pursue opportunities to improve application processes and agency data collection, sharing, and reporting, and to synchronize data-protection processes in a manner that will enhance the delivery and receipt of low-income energy efficiency services. Think big and use this opportunity as a first step, not a final stop.

Individual agreements between parties can spring from negotiations between parties or from pre-existing regulatory pathways. The individualized nature of the agreements can be captured in the NDA appendices. Memorandums and letters of understanding appended to the NDA will provide additional details on how data will be shared between parties. Details that are necessary to meet statutory or regulatory duties (e.g., identifying how the data will assist an agency in discharging a statutory obligation) should build on the overall terms agreed to by all parties who are signatories to the NDA.

Building a multi-party NDA requires knowledge of the applicable regulations and statutes, an understanding of what technologies parties are using, familiarity with program purposes, and a willingness to collaborate on complicated issues. The task is not easy, but can result in a flexible, manageable, and valuable system of collaborative data sharing that transforms program delivery to the benefit of all residents of the state. Appendix A contains a sample multi-party NDA.

Future Possibilities

In the future, there are at least three significant ways that a PEEIF-like database and construct could be utilized. One option would be for a state agency to administer a PEEIF-like database utilizing data already available to the agency in concert with other public and private partners.

For example, in North Carolina, the NCDEQ and/or the NCDHHS could host a database using the data already available to each respective agency while partnering with utility providers to achieve comparative analysis of utility data, WAP-related data, and CIP and LIEAP data.

A second—or additional—option would be for a federal agency to host a PEEIF like database for the purpose of administering federal energy-related programs (e.g., WAP, CIP, LIEAP and others). A possible advantage of this structure is that the federal agency has the ability to create uniform data standards across all states and territories.

A third option would be a PEEIF like database administered by a state or federal agency that is designed to perform a sample analysis of subject program performance in concert with utility data. This structure could take advantage of smaller data sets that are easier to obtain and administer.

In any of these three options, there is a profound opportunity to analyze and improve the benefit and performance of programs serving low-income communities. By comparing WAP, CIP, LIEAP, and other data points with utility data, the ability to better deploy funds and monitor program performance could deliver significant benefits to low-income communities while also improving the cost-benefit performance of deployed funds.

APPENDIX

Appendix A: Multi-Party Data Sharing Agreement Template

Appendix B: PEEIF Virtual Machine Specifications

Appendix C: Data Attributes of PEEIF Housing

Appendix D: PEEIF Project Utility Data Attributes

Appendix E: NCDEQ and NCDHHS PEEIF Data Attributes

Appendix F: Household Energy Efficiency Potential Example Python Script

Appendix G: PEEIF Technical Process Workflow, Challenges, and Future Considerations

Appendix H: U.S. Census and Data Map Outputs

Appendix I: Sample Non-Disclosure Agreement

Appendix A: Multi-Party Data Sharing Agreement Template

Introduction

This section discusses the creation of a multi-party data sharing agreement that facilitates the inflow and outflow of data plus the cross-sharing of data between data sharing entities and other vetted third parties. It does so in two parts: first, by discussing the current PEEIF framework and its limitation, and second, by detailing the elements that would make a more flexible and inclusive multi-party framework for future collaboration.

PEEIF Framework

The PEEIF platform opens new opportunities to facilitate the development, implementation, and targeting of low-income energy efficiency services. The legal and technical frameworks work in concert to move data into the database and to flow analysis back out of the database. However, the database structure and organization could be improved to create a sustainable and adaptable platform that would facilitate the addition of new data sharing parties and parties granted access to data and would allow for the development of new data uses. The PEEIF legal framework is based on individual non-disclosure agreements signed with a centralized party. Data sharing parties and parties granted access to data sign NDAs with the Upper Coastal Plains Council of Governments. Terms and conditions are restricted to a sharing of data with and through UCPCOG. While this template links together individual data sharing agreements, it did not create conditions for enhanced data sharing between parties and it constrains the ability to evolve database functions in response to user needs. A multi-party non-disclosure could address these and other issues.

Multi-Party Non-Disclosure Agreements

PEEIF is built on primary and secondary non-disclosure agreements in which UCPCOG serves as the central hub for the collection of and movement of data. NDAs were constructed for data sharing parties and for data accessing parties. This structure was selected as it mimics the DOE grant structure, but it could be improved to focus on the location of the database and the organization that controls the database. A multi-party NDA constructed between all the parties sharing and accessing data should be built around where the database is housed and agreed upon security protocols and conditions. Below are key considerations in the development of a multi-party NDA.

1. Unified Data Security Protocols – A single data security protocol will reduce technical limitations to enhanced data sharing. State agencies with the authority to share data must often demonstrate that the recipient of the data provides a similar level of data protection as the data sharing agency. Each participating North Carolina agency has its own data management protocols as well as universal compliance obligations contained in the North Carolina Statewide Information Security Manual. Building a system that addresses individual security needs and shared security obligations is critical to creating a multi-party platform.
2. Acknowledgement and Incorporation of Data Protection and Data Sharing Regulations and Laws – The NDA must be able to acknowledge and respond to the different sources of data privacy regulations and the mechanisms by which the regulations and statutes are enforced. A multiparty NDA may include state agencies, investor owned utilities, electric membership cooperatives, municipal owned utilities, and third-party service providers, each of which may be responding to a different combination of guidelines, policies, ordinances, by-laws, regulations, and statutes. Protecting confidential data from unauthorized use is a core responsibility of every signatory to the NDA. Unlocking the potential of data sharing is core goal of every signatory to the NDA. Thus, the NDA should include terms that recognized the current legal environment while having the flexibility to adapt to the needs of new parties without affecting data security.

3. Flexibility of Purpose – Unlocking the multiple uses for and values of each piece of data stored in the database requires a flexible NDA capable of evolving to the needs of its signatories. A PEEIF-style database can produce value by allowing for single party data analysis or by facilitating cross-party data sharing and analysis. In single party data analysis, the data provided by a party is uploaded, analyzed, and returned to uploading party. The application of third-party analytics can create new methods of viewing data, synthesizing data for reports and audits, and compiling historic trends. The single source of data and single destination of analytics is easy to capture in an NDA. The difficulty lays in formulating a system that permits for cross-party data sharing.

Cross-party data sharing opens the possibility of enhanced synergies between different programs and program providers to assess quality of services provided, identify unserved and underserved populations, and to reduce administrative burdens. Cross-party data sharing requires an NDA with a flexible purpose, an open-ended period of operation, and with the ability to capture individual agreements between parties that can modify the purpose and the length of a specific data sharing arrangement.

- a. The purpose of the NDA must be flexible enough to capture a variety of known and unknown uses. The known uses are present when the NDA is created, and they help frame out the range of potential uses served by allowing for data sharing. The unknown uses arise as parties develop their knowledge of the potential uses and benefits of the database. Having the flexibility to adapt to unknown beneficial uses requires the creation of a purpose that is expansive yet focused. Expansive enough to accommodate a variety of uses; focused on the delivery of program services tied to agency or entity goals. In combination, the two elements produce a range of uses that are flexible yet constrained.
- b. The period of operation for the NDA should be open-ended with clear rules for how parties can enter into and withdraw from the agreement along with provisions that address how data will be managed after a decision to terminate participation is made. A narrow time period does not account for how agencies function and how resources are tasked to implement data sharing processes.
- c. Individual agreements between parties can spring from negotiations between parties or from pre-existing regulatory pathways. The individualized nature of the agreements can be captured in the NDA appendices. Memorandums and letters of understanding appended to the NDA will provide additional details on how data will be shared between parties. Details that are necessary to meet statutory or regulatory duties, e.g. identifying how the data will assist an agency in discharging a statutory obligation. These details will build upon the overall terms agreed to by all parties who are signatories to the NDA.

Summary

Building a multi-party NDA requires knowledge of the applicable regulations and statutes, an understanding of what technologies parties are using, familiarity with program purposes, and a willingness to collaborate on complicated issues. The task is not easy but the results of the work can be a flexible, manageable, and valuable system of collaborative data sharing that transforms program delivery to the benefit of all residents of the state.

Prepared by Mark James, Vermont Law School Dec, 2018

Appendix B: PEEIF Virtual Machine Specifications

Powering Energy Efficiency & Impacts Framework Virtual Machine Specifications

System Information:

Item	Value
OS Name	Microsoft Windows Server 2012 R2 Standard
Version	6.3.9600 Build 9600
Other OS Description	Not Available
OS Manufacturer	Microsoft Corporation
System Name	CASCADES
System Manufacturer	VMware, Inc.
System Model	VMware Virtual Platform
System Type	x64-based PC
System SKU	
Processor	Intel(R) Xeon(R) CPU E7-L8867 @ 2.13GHz, 2128 Mhz, 1 Core(s), 1 Logical Processor(s)
BIOS Version/Date	Phoenix Technologies LTD 6.00, 7/30/2013
SMBIOS Version	2.4
Embedded Controller Version	0.00
BIOS Mode	Legacy
BaseBoard Manufacturer	Intel Corporation
BaseBoard Model	Not Available
BaseBoard Name	Base Board
Platform Role	Desktop
Secure Boot State	Unsupported
PCR7 Configuration	Not Available
Windows Directory	C:\Windows
System Directory	C:\Windows\system32
Boot Device	\Device\HarddiskVolume1
Locale	United States
Hardware Abstraction Layer	Version = "6.3.9600.17196"
User Name	Not Available
Time Zone	Eastern Standard Time
Installed Physical Memory (RAM)	5.00 GB
Total Physical Memory	5.00 GB
Available Physical Memory	2.41 GB
Total Virtual Memory	6.94 GB
Available Virtual Memory	4.15 GB
Page File Space	1.94 GB
Page File	C:\pagefile.sys
A hypervisor has been detecte...	

Database Management System:

Microsoft

Microsoft SQL Server 2012

Component Name	Versions
Microsoft SQL Server Management Studio	11.0.6020.0
Microsoft Analysis Services Client Tools	11.0.6020.0
Microsoft Data Access Components (MDAC)	6.3.9600.17415
Microsoft MSXML	3.0 5.0 6.0
Microsoft Internet Explorer	9.11.9600.18538
Microsoft .NET Framework	4.0.30319.42000
Operating System	6.3.9600

To copy component name and version information, click Copy Info.

Warning: This computer program is protected by copyright law and international treaties. Unauthorized reproduction or distribution of this program, or any portion of it, may result in severe civil and criminal penalties, and will be prosecuted to the maximum extent possible under the law.

© 2012 Microsoft. All rights reserved.

Microsoft

Appendix C: Data Attributes of PEEIF Housing

Source: House data was acquired from NC OneMap, the North Carolina spatial data repository. County zoning maps are all available, however not all attributes listed below were available for all counties. For example some counties have square footage available to NCONE while others do not.

Housing Attributes:

ORIGINAL	Alias	Data Type	Data Length
OBJECTID_1 *			
Shape *			
Join_Count			
TARGET_FID			
OBJECTID			
PEEIF_ID	PEEIF ID	string	254
ZIPCODE	Zip Code	number	
CENSUS_TRA	Census Tract	number	
CENSUS_BG	Census Block Group	number	
PARCEL_NUM	Parcel Number	string	254
STRUCTYEAR	Year Built	number	
HOME_TYPE	Home Type	string	254
MH_TYPE	Mobile Home Type	string	254
STORIES	Stories	string	254
TOT_RMS	Total Rooms	string	254
FULL_BATH	Number Full Baths	string	254
HALF_BATH	Number Half Baths	string	254
SF	Square Foot	string	254
HEATINGFUE	Heating Fuel Type	string	254
HW_UTILITY	Hot Water Fuel	string	254
STOVE_UTIL	Stove Fuel	string	254
SF_2	Square Foot	number	

Appendix D: PEEIF Project Utility Data Attributes

Original	Alias	Data Type	Data Length
ENF_ACCT_NUM	Enfield Account Number	number	
ENF_ANNUSE_YR1	Enfield Year 1 Usage	number	
ENF_ANNUSEPER_YR1	Enfield Year 1 Usage Percentile	number	
ENF_ANNUSEDEC_YR1	Enfield Year 1 Usage Decile	number	
ENF_ANNUSEDECMIN_YR1	Enfield Year 1 Usage Decile Min Usage	number	
ENF_ANNUSEDECMAX_YR1	Enfield Year 1 Usage Decile Max Usage	number	
ENF_EUI_YR1	Enfield Year 1 Energy Use Intensity	number	
ENF_EUIPERC_YR1	Enfield Year 1 EUI Percentile	number	
ENF_EUIDEC_YR1	Enfield Year 1 EUI Decile	number	
ENF_EUIDECMIN_YR1	Enfield Year 1 EUI Decile Min	number	
ENF_EUIDECMAX_YR1	Enfield Year 1 EUI Decile Max	number	
ENF_ANNUSE_YR2	Enfield Year 2 Usage	number	
ENF_ANNUSEPER_YR2	Enfield Year 2 Usage Percentile	number	
ENF_ANNUSEDEC_YR2	Enfield Year 2 Usage Decile	number	
ENF_ANNUSEDECMIN_YR2	Enfield Year 2 Usage Decile Min Usage	number	
ENF_ANNUSEDECMAX_YR2	Enfield Year 2 Usage Decile Max Usage	number	
ENF_EUI_YR2	Enfield Year 2 Energy Use Intensity	number	
ENF_EUIPERC_YR2	Enfield Year 2 EUI Percentile	number	
ENF_EUIDEC_YR2	Enfield Year 2 EUI Decile	number	
ENF_EUIDECMIN_YR2	Enfield Year 2 EUI Decile Min	number	
ENF_EUIDECMAX_YR2	Enfield Year 2 EUI Decile Max	number	
REC_MTR_NUM	Roanoke Meter Number	string	
REC_ANNUSE_YR1	Roanoke Year 1 Annual Usage	number	
REC_ANNUSEPER_YR1	Roanoke Year 1 Usage Percentile	number	
REC_ANNUSEDEC_YR1	Roanoke Year 1 Usage Decile	number	
REC_ANNUSEDECMIN_YR1	Roanoke Year 1 Usage Decile Min Usage	number	
REC_ANNUSEDECMAX_YR1	Roanoke Year 1 Usage Decile Max Usage	number	
REC_EUI_YR1	Roanoke Year 1 Energy Use Intensity	number	
REC_OWN_RENT_YR2	Roanoke Year 2 Owner Renter	string	254
REC_ANNUSE_YR2	Roanoke Year 2 Annual Usage	number	
REC_ANNUSEPER_YR2	Roanoke Year 2 Usage Percentile	number	
REC_ANNUSEDEC_YR2	Roanoke Year 2 Usage Decile	number	
REC_ANNUSEMIN_YR2	Roanoke Year 2 Usage Decile Min Usage	number	
REC_ANNUSEMAX_YR2	Roanoke Year 2 Usage Decile Max Usage	number	
REC_EUI_YR2	Roanoke Year 2 Energy Use Intensity	number	
REC_EUIPERC_YR2	Roanoke Year 2 EUI Percentile	number	
REC_EUIDEC_YR2	Roanoke Year 2 EUI Decile	number	
REC_EUIDECMIN_YR2	Roanoke Year 2 EUI Decile Min	number	
REC_EUIDECMAX_YR2	Roanoke Year 2 EUI Decile Max	number	

WEC_MTR_NUM	Wilson Meter Number	string	254
WEC_PARCEL_PIN	Wilson Parcel PIN	string	254
WEC_ANNUSE_YR1	Wilson Year 1 Energy Use	number	
WEC_ANNUSEPERC_YR1	Wilson Year 1 Usage Percentile	number	
WEC_ANNUSEDEC_YR1	Wilson Year 1 Usage Decile	number	
WEC_ANNUSEDECMIN_YR1	Wilson Year 1 Usage Decile Min Usage	number	
WEC_ANNUSEDECMAX_YR1	Wilson Year 1 Usage Decile Max Usage	number	
WEC_EUI_YR1	Wilson Year 1 Energy Use Intensity	number	
WEC_EUIPERC_YR1	Wilson Year 1 EUI Percentile	number	
WEC_EUIDEC_YR1	Wilson Year 1 EUI Decile	number	
WEC_EUIDECMIN_YR1	Wilson Year 1 EUI Decile Min	number	
WEC_EUIDECMAX_YR1	Wilson Year 1 EUI Decile Max	number	
WEC_ELEC_COST	Wilson Electricity Cost	number	
WEC_ANNUSE_YR2	Wilson Year 2 Energy Use	number	
WEC_ANNUSEPERC_YR2	Wilson Year 2 Usage Percentile	number	
WEC_ANNUSEDEC_YR2	Wilson Year 2 Usage Decile	number	
WEC_ANNUSEDECMIN_YR2	Wilson Year 2 Usage Decile Min Usage	number	
WEC_ANNUSEDECMAX_YR2	Wilson Year 2 Usage Decile Max Usage	number	
WEC_EUI_YR2	Wilson Year 2 Energy Use Intensity	number	
WEC_EUIPERC_YR2	Wilson Year 2 EUI Percentile	number	
WEC_EUIDEC_YR2	Wilson Year 2 EUI Decile	number	
WEC_EUIDECMIN_YR2	Wilson Year 2 EUI Decile Min	number	
WEC_EUIDECMAX_YR2	Wilson Year 2 EUI Decile Max	number	
WEC_GASUSAGE_YR2	Wilson Year 2 Energy Use	number	
WEC_GASCOST_YR2	Wilson Gas Cost	number	
WEC_GASPERC_YR2	Wilson Gas Year 2 Usage Percentile	number	
WEC_GASDEC_YR2	Wilson Gas Year 2 Usage Decile	number	
WEC_GASDECMIN_YR2	Wilson Gas Year 2 Usage Decile Min Usage	number	
WEC_GASDECMAX_YR2	Wilson Gas Year 2 Usage Decile Max Usage	number	
WEC_GASEUI_YR2	Wilson Gas Year 2 Use Intensity	number	
WEC_GASEUIPERC_YR2	Wilson Gas Year 2 EUI Percentile	number	
WEC_GASEUIDEC_YR2	Wilson Gas Year 2 EUI Decile	number	
WEC_GASEUIDECMIN_YR2	Wilson Gas Year 2 EUI Decile Min	number	
WEC_GASEUIDECMAX_YR2	Wilson Gas Year 2 EUI Decile Max	number	
WEC_GAS_ELEC_YR2_csv_GasMeterNumber	Wilson Gas Meter	string	254
WEC_GAS_ELECANNUSE_YR2	Wilson Gas Year 2 Gas & Elec Annual Use	number	
WEC_GAS_ELEC_THERM_YR2	Wilson Gas Year 2 Gas & Elec Therms	number	
WEC_GAS_ELEC_USEBKU_YR2	Wilson Gas Year 2 Gas & Elec BTUS	number	
WEC_GAS_ELEC_COSTKWH_YR2	Wilson Gas Year 2 Gas & Elec cost per KWH	number	
WEC_GAS_ELEC_COSTTHERM_YR2	Wilson Gas Year 2 Gas & Elec cost per Therm	number	

WEC_GAS_ELEC_FULLCOST_YR2	Wilson Gas Year 2 Gas & Elec Full Cost	number	
WEC_GAS_ELEC_PERC_COST_YR2	Wilson Gas Year 2 Gas & Elec Percentile Cost	number	
WEC_GAS_ELEC_DEC_COST_YR2	Wilson Gas Year 2 Gas & Elec Decile Cost	number	
WEC_GAS_ELEC_DEC_MIN_COST_YR2	Wilson Gas Year 2 Gas & Elec Decile Min Cost	number	
WEC_GAS_ELEC_DEC_MAX_COST_YR2	Wilson Gas Year 2 Gas & Elec Decile Max Cost	number	
WEC_GAS_ELEC_DEC_COSTINTENSIT_Y_YR2	Wilson Gas Year 2 Gas & Elec Decile Cost Intensity	number	
WEC_GAS_ELEC_DEC_COSTINTENSIT_Y_PERC_YR2	Wilson Gas Year 2 Usage Percentile	number	
WEC_GAS_ELEC_DEC_COSTINTENSIT_Y_DEC_YR2	Wilson Gas Year 2 Usage Decile	number	
WEC_GAS_ELEC_DEC_COSTINTENSIT_Y_DECMIN_YR2	Wilson Gas Year 2 Usage Decile Min Usage	number	
WEC_GAS_ELEC_DEC_COSTINTENSIT_Y_DECMAX_YR2	Wilson Gas Year 2 Usage Decile Max Usage	number	

Appendix E: NCDEQ and NCDHHS PEEIF Data Attributes

Original	Alias	Data Type	Data Length
DEQ_JOB	DEQ Job Type	string	254
DEQ_PROGRAM	DEQ Program	string	254
DEQ_LIHEAP	DEQ Liheap	string	254
DEQ_APPR_DATE	DEQ Job Approval Date	date	
DEQ_SELECT_DATE	DEQ Job Selection Date	date	
DEQ_CLOSE_DATE	DEQ Job Close Date	date	
DEQ_PRE_BLWR	DEQ Job Pre Blower Reading	number	
DEQ_POST_BLWR	DEQ Job Post Blower Reading	number	
HAARP_JOB	HAARP Job	string	254
HARRP	HAARP Job	string	254
DHHS_CASEID	DHHS Case ID	string	254
DHHS_CASEID_NUM	DHHS Case ID	number	
DHHS_COUNT	DHHS Assistance counts at home	number	
DHHS_CASE_ID_INT	DHHS Case ID	number	
DHHS_CHILD	DHHS Child Present	number	
DHHS_ELDERLY	DHHS Elderly Present	number	
DHHS_DISABLED	DHHS Disabled Present	number	
DHHS_HH_ELLIGIBLE	DHHS Number in HH eligible	number	
DHHS_FUEL_TYPE	DHHS home utility fuel type	string	254
DHHS_REQUEST_ID	DHHS request ID	string	254
DHHS_APPLIED_DATE	DHHS Application Date	date	
DHHS_APPRV_DATE	DHHS Approval date	date	
DHHS_APPRV_TYPE	DHHS approval type	string	254
DHHS_APPRV_AMT_SUM	DHHS total assistance sum	number	

Appendix F: Household Energy Efficiency Potential Example Python Script

The Python script used to complete the analysis described in section Energy Service Data and Program Evaluation Framework, subsection Data Analysis requires 4 data input files to be created. In general, the software is robust to missing data. Entries with missing information are dropped from the calculations.

The input files are:

input_options.txt

This is a text file with user specified parameters to run. An example of the input file is found below.

```
"This input file contains input options for the Energy Use Regression Correlation Tool
"This section is for data scrubbing. The entries below trim age,"
" use, and size for entries which exceed certain thresholds"
max_size 30000
min_size 100
max_age 250
max_use 50000
min_use 1000
"This section contains the bin edges for the segmented regression."
"The format is [0,first edge, second edge,...,last edge]"
age_bins 0 15 50
size_bins 0 1500 4000
Input File Example
```

home_data.csv

This is a comma separated value file with information required to perform the regression analysis. This includes home information (age, size, and a unique identifier (like meter number)) and the annual energy use. Below is an example input. The software is not unit dependent, so values are at the user’s discretion (size can be square feet or square meters for example). The column names must be exactly as shown in the example below but can be in any order.

Size	Name	Use	Age
1166	Meter 1	22359	58
3614	Meter 2	35090	5
6211	Meter 3	20508	4
4030	Meter 4	42763	9
2386	Meter 5	14637	3
5284	Meter 6	57709	41
4390	Meter 7	13357	41
4534	Meter 8	597	49
3521	Meter 9	886	54
3552	Meter 10	695	53

temperature_data.csv

This file contains dates associated with the period of interest and the TDD (temperature degree days) for each day. The software is not unit dependent, so values are at the user’s discretion (size can be square feet or square

meters for example). The column names must be exactly as shown in the below example but can be in any order.

Date	TDD
12/1/2016	10
12/2/2016	21
12/3/2016	22
12/4/2016	21
12/5/2016	13
12/6/2016	18
12/7/2016	16
12/8/2016	21

daily_use.txt

This file contains a JSON containing energy use data. The file should be formatted as follows:

```
{"Home Identifier": {"Date": Energy used on Date}}
```

An example would be:

```
{"Meter1": {"2017-03-31": 6.3198, "2017-04-01": 6.8262, "2017-04-02": 6.9777}, "Meter2": {"2017-03-31": 1.644525, "2017-04-01": 1.8162, "2017-04-02": 1.95285}, "Meter3": {"2017-03-31": 2.493225, "2017-04-01": 1.798425, "2017-04-02": 1.798425}}
```

The home identifiers should be exact matches to the identifiers in the home_data.csv file. The dates match those found in the temperature_data.csv file.

The Python script follows. The input file paths should be updated to reflect the location of the files above.

energyAnalysis.py

```
import datetime as dt
import json
import numpy as np
import pandas as pd
import os
from sklearn import linear_model
from sklearn.metrics import mean_squared_error, r2_score
from mpl_toolkits.mplot3d import Axes3D
import matplotlib.pyplot as plt
import scipy.stats as st
import matplotlib as mpl

"Load input options"
input_options = {}
with open("input_options.txt") as f:
    for line in f:
        dline = line.split()
        if line[0][0] != ' ':
            input_options[dline[0]] = list(map(float, dline[1:]))
```

```

"Load Data from Larger Database"
"read in input"
# data = pd.read_csv("/Users/Akodo/OneDrive/Academic Stuff/Current/CES
Research/Energy Project/Turnover Script/output_summary_conf.txt",delimiter="\t",)
#data.set_index('PEEIF_ID',inplace = True)
data =
pd.read_csv("/Users/mark/Desktop/peeifPythonData/ScriptAndInputs/home_data.csv",d
elimiter=',', encoding="utf-8-sig")
df = data.copy()
df.set_index('Meter', inplace = True)
df.dropna(inplace = True)
df.drop_duplicates(subset = ['Meter'])

"Will need to import these numbers from the input file"
try:
    df = df.drop(df[df['Size']<input_options['min_size']].index)
    df = df.drop(df[df['Size']>input_options['max_size']].index)
    df = df.drop(df[df["Age"]>2018-1850].index)
    df = df.drop(df[df['Use']<1500].index) # < 2 100W lightbulbs running 24x7
    df = df.drop(df[df['Use']>50000].index)
except:
    pass

"This block calculates the correlation between meter number usage and today
degree days"
"This block will be cut because daily energy use is expected to be loaded as a
json as {Name {Date:Daily Energy Use},}"

"Load use data"
with
open('/Users/mark/Desktop/peeifPythonData/ScriptAndInputs/daily_use.txt','r') as
outfile:
    daily_use_json = json.load(outfile)
use = pd.DataFrame(daily_use_json)

"drop all entries not included in the home dataset"
use=use[df['Meter']]

"Load Temperature Data as Panda"
temperature_data =
pd.read_csv(os.getcwd()+'\temperature_data.csv',delimiter=',')
temperature_data.set_index('Date',inplace = True)

'remove indices from temperature df with no corresponding meter data'
diff = set(use.index) ^ set(temperature_data.index)
try:
    temperature_data = temperature_data.drop(list(diff))
    use = use.drop(list(diff))
except:
    pass

"sort the index for each array so they are exactly the same"
temperature_data.sort_index(inplace = True)
use.sort_index(inplace = True)

```

```

cor = use.corrwith(temperature_data['TDD']).dropna()
cor = pd.Series.to_frame(cor)

"Combined the Correlation with the original data"
df_indexed = df.set_index('Name')
final = df_indexed.join(cor)
final.rename(columns = {0:"Temp_Correlation"},inplace = True)
final.dropna(inplace = True)

'Functions for Binning and Regression'
def regress(data,yvar,xvars):
    Y = data[yvar]
    X = data[xvars]
    clf = linear_model.LinearRegression()
    return clf.fit(X,Y)

class predict:
    def __init__(self,fit):
        self.fit = fit
    def pred(self,row):
        return
self.fit[(row['age_bracket'],row['size_bracket'])].predict(np.array([[row['Age']]
,[row['Size']]])).T)
    def prnt(self,data):
        print(data['age_bracket'],data['size_bracket'],data['Age'],data['Size'])

class fiteval:
    def __init__(self,df,bins):
        self.df = df
        "Optimum bins found at size = 3400, age = [14,44]"
        self.age_bins = np.array(bins['age_bins'])
        self.size_bins = np.array(bins['size_bins'])

    def fitfun(self):
        "copy the dataframe and create columns with binned ages/sizes"
        tmp = self.df.copy()
        tmp['age_bracket'] = np.digitize(tmp['Age'],self.age_bins)
        tmp['size_bracket'] = np.digitize(tmp['Size'],self.size_bins)
        tmp_grouped = tmp.groupby(['age_bracket','size_bracket'])

        "find regressions for each group"
        fitmodels = dict(tmp_grouped.apply(regress,'Use',['Age','Size']))

        "use the regression variables to predict what the error for each home is"
        pre_model = predict(fitmodels)
        tmp_use_df = tmp.apply(pre_model.pred,axis = 1)
        tmp['Puse'] = tmp_use_df['Use']

        return tmp.copy()

bins = {}
try:
    bins['age_bins'] = input_options['age_bins']
except:
    bins['age_bins'] = [0,float('inf')]

```

```

try:
    bins['size_bins'] = input_options['size_bins']
except:
    bins['size_bins'] = [0,float('inf')]

fe = fiteval(final,bins)
regressed = fe.fitfun()
regressed['error'] = (regressed['Puse']-regressed['Use']/regressed['Use']
regressed.sort_values(by=['error'],inplace = True)
regressed['Regression Error Rank']= [x for x in range(len(regressed['error']))]

regressed.sort_values(by=['Temp_Correlation'],ascending = False, inplace = True)
regressed['Temp Correlation Rank']= [x for x in
range(len(regressed['Temp_Correlation']))]

regressed.to_csv('out.csv')
print ("Program terminated normally")

```

Appendix G: PEEIF Technical Process Workflow, Challenges, and Future Considerations

The following document, provided by NCSU Center for Geo Spatial Analytics, includes the process workflow utilized by PEEIF project partners for data preparation, as well as some of the challenges encountered and considerations for future similar projects.

Source and Spatialization Workflow

- Acquire raw data from data providing partner
- Clean data so that attributes and all data points are standardized
- Geolocate all data to house location
 - PEEIF project used a parcel shapefile for all 5 counties to provide “bins” for points that all data could be matched to.
- Add PEEIF ID to new data for each parcel and then join to each cleaned data set to standardize across all sets
- Join data to main data sheet which collects all geolocated data into one shapefile
- Incorporate into registered geodatabases to be used in user portals.

Gaps and Challenges

Data cleaning requires substantive efforts to ensure quality data is used. Many new data sets will have a significant number of data points that are also difficult to manage directly through a GIS interface. The most pressing issue that required substantial effort was geolocation and matching houses from multiple datasets. In many cases, particularly with utilities data, meter locational data were the only locational data included. Houses in other datasets, including assistance data, can only be located through address geocoding.

In future projects, geolocation technologies will need to be assessed to determine the optimal geolocation outputs. In addition, associating one home with multiple stakeholders’ datasets for query and display in the ArcGIS 10.5.1 ecosystem does not support web application needs. Examples of technological constraints include, but are not limited to:

- Address geolocation issues: certain geolocation application programming interfaces (APIs) do not locate the address on the house, but instead the location is positioned on the street. The point location should be located on the parcel to allow joining parcel information, e.g. home square footage, with a geolocated home. Associating the geolocated point with the actual parcel is time intensive, thus costly. Google API was chosen because it does geo locate the house on the parcel through rooftop locations, however Google limits the number of daily geocoding requests. This also does not solve some issues such as poor address quality (multiple mistakes in names, variation in attributing the same street name) or meter location (often not on a parcel).
- Incorporating more than one data set at a time becomes difficult due to the extensive amount of data, including the inconsistencies listed above. Scripting workflows would be beneficial and is recommended moving forward. It is recommended to automate as many cleaning processes of the data tables to minimize data processing time.

Storage and Incorporation into a Geodatabase

Data is stored in two locations in the PEEIF data organization schema. Raw data and processed data are being stored on a virtual server, but can be stored in any typical way (as long as data is backed up). However, production data layers must be imported and stored in a geodatabase registered with the web GIS server, e.g. ArcGIS Server. This configuration allows the final data to be shared as a web GIS service, e.g. Representational State Transfer (REST) service for consumption into web mapping applications.

Raw and processed data are located on a server in order to preserve data continuity and take advantage of security features. Data is organized in a folder structure by data provider. Raw data is stored in a separate folder as it is downloaded as a backup in case the production must be recreated due to system failures. There was a choice made in the PEEIF project to create a singular point for all data. A single shape file was created that was linked using a base file of all parcels in the area. This was done due to the limitations in ArcGIS Online to incorporate a joined table for query and display and when visualization and querying data. Data was exported from this format on an as-needed basis for each new data set based on data security.

Portal Configuration

Creation of Services

In the ArcGIS platform, a registered geodatabase is used to store data and a Representational State Transfer (REST) web service is created from the source data using ArcGIS Server. ArcGIS Online for Organizations is then able to consume those services for use in web mapping applications. The registered geodatabases contained final, production data. Intermediate data, to create the production data set, is not stored. This reduces the amount of managed data and the possibility of data compromise. ArcGIS Server essentially becomes a pointer to data located in the geodatabase and accessible through a URL for web access. These URLs then were used to create web maps and web applications as requested by project partners. As a best practice for data security, data should not be copied to ArcGIS Online for Organizations. Not all entities however will have the ability to configure, manage and host a web GIS server and register a geodatabase. In this situation, loading data directly to ArcGIS may be an option but the security limitations should be made known to stakeholders.

Creation of Web Applications

Using REST services, applications were developed with project data, e.g. US Census data, stakeholder confidential data and stakeholder aggregated and blurred confidential data. These layers were created, converted to REST services and incorporated into each web application based on stakeholder and user program requirements. For example, the DEQ application contains Census data at the block group level, confidential DEQ program data and aggregated information for other data providers.

ArcGIS Online for Organizations applications were developed in two steps. First, online maps are created. In this phase, all layers are incorporated into the ArcGIS map document. The ArcGIS map document is then configured to visualize the data to communicate the desired message and allow for query and display of data in a meaningful manner. Visualization may include symbolizing energy usage decile data, ordering of data layers in a stacked order to allow visualization of multiple layers simultaneously and basic cartographic principles including a legend and title. Additional configurations included the development of informational pop-ups for each layer and meaningful layer names. Following the development of the map document, the layers are published as REST services using ArcGIS Server. The REST services are then used to develop ArcGIS Online web maps and a corresponding web mapping application. Finally, the web mapping application is configured with tools to allow for data exploration, query and display. Custom tools may include the following ability:

- to toggle layers on and off,
- switches and dialog boxes to filter and refine data, and
- zoom and pan to focus the display on critical areas.

Gaps and Challenges

Combining the disparate stakeholder data into a single, flattened data layer is not the optimal method to organize the data. However, due to documented limitations with the ArcGIS Online for Organizations platform presented earlier, this data format was chosen. Further, this data format is not a scalable data model as additional data may be identified and incorporated in the database causing the need to re-configure the database schema when additional data is discovered and incorporated. This limitation may be solved by creating a custom web mapping application. Although this approach was considered, it was not adopted due

to the potential lack of programming expertise and the potential for applications to become inoperable due to ArcGIS Online for Organizations programming library updates and other modification developed by Esri and pushed out without notification.

Moving Forward

The scalability of one data layer storing all of the data needs to be addressed in some capacity if ArcGIS Online for Organizations hosted applications will be used moving forward. Generally using related tables is more efficient, however ArcGIS does not, as of December 2018 provide a method for visualizing related data tables. This missing feature limits the spatial analysis addressing multi to single point data, such as multiple weatherization measures on one house. Data providing partner interests, priorities and options must be thoroughly reviewed and discussed in order to insure data is well organized before being placed in the online environment through any future iteration.

Appendix H: U.S. Census and Data Map Outputs

I. Overview

The intention of the production of these maps is to present both census data and proprietary DHHS data in an easy to read format, where PEEIF study outputs can be presented in tandem with general demographic data of the same area.

The map outputs consist of combination of layouts, presenting a variety of census or DHHS data. The layout is a large map of all 5 counties + plus insets of the 3 densest populated areas in the study area: Wilson, Rocky Mount and Roanoke Rapids.

These maps are meant to represent general trends for comparison. The next section will discuss the maps and the last section the limitations of the maps and data.

II. Data sources

Data was compiled by the PEEIF Technical Team from the following sources:

1. NC DHHS
2. US CENSUS
3. NC DOT

III. Spatial extent

All Census block groups in 5 counties in NC

1. Northampton
2. Wilson
3. Edgecombes
4. Nash
5. Halifax

IV. Maps

1. Data

- i. **Quantiles:** Data is presented in quantiles. Quantiles are 5 equal sections of any set of data. For example, if there are 20 homes with income, the 4 of highest income will be one groups, the next 4 another group etc. This means that data ranges for each group will vary, especially at the lowest and highest groups where outliers may be present

2. Map outputs – These are the map current map outputs

- i. **% Rental Households:** Rental households as a percentage of total households total households
- ii. **Median Income:** Median Household Income of all households in census block group
- iii. **CIP penetration rate:** Number of houses receiving CIP assistance as a percentage of total households in the census block group
- iv. **LIEAP penetration rate:** Number of houses receiving LIEAP assistance as a percentage of total households in the census block group
- v. **Households at 200% poverty:** Population at or under 200% of the federal poverty level in the census block group
- vi. **150% poverty:** Population at or under 150% of the federal poverty level in the census block group

V. Definitions

1. Census Block Group⁵

- i. "Block Groups (BGs) are statistical divisions of census tracts, are generally defined to contain between 600 and 3,000 people, and are used to present data and control block numbering. A block group consists of clusters of blocks within the same census tract that have the same first digit of their four-digit census block number. For example, blocks 3001, 3002, 3003,..., 3999 in census tract 1210.02 belong to BG 3 in that census tract. Most BGs were delineated by local participants in the Census Bureau's Participant Statistical Areas Program. The Census Bureau delineated BGs only where a local or tribal government declined to participate, and a regional organization or State Data Center was not available to participate. A BG usually covers a contiguous area. Each census tract contains at least one BG, and BGs are uniquely numbered within the census tract. Within the standard census geographic hierarchy, BGs never cross state, county, or census tract boundaries but may cross the boundaries of any other geographic entity."

2. Households⁶

- i. "A housing unit is a house, an apartment, a group of rooms, or a single room occupied or intended for occupancy as separate living quarters. Separate living quarters are those in which the occupants do not live and eat with other persons in the structure and which have direct access from the outside of the building or through a common hall. For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible. If the information cannot be obtained, the criteria are applied to the previous occupants."

"A unit is owner occupied if the owner or co-owner lives in the unit, even if it is mortgaged or not fully paid for. A cooperative or condominium unit is "owner occupied" only if the owner or co-owner lives in it. All other occupied units are classified as "renter occupied," including units rented for cash rent and those occupied without payment of cash rent."

3. Poverty⁷

- i. "People and families are classified as being in poverty if their income is less than their poverty threshold. If their income is less than half their poverty threshold, they are below 50% of poverty; less than the threshold itself, they are in poverty (below 100% of poverty); less than 1.25 times the threshold, below 125% of poverty, and so on."

VI. Limitations

1. Data Location

- i. Geolocation of assistance programs is the main limitation. It is important to note, specifically with the penetration rates, that lack of number of assistance not geolocated shouldn't be conflated with lack of assistance generally. One is a data issue and another is a pattern of assistance that should be addressed. We can provide a map by zipcode of houses that have received assistance but have not been geolocated in order to qualify some penetration rates.

VII. Statistics

1. Census Block Groups – 240 total CBG

- i. Household Count by CBG
 1. Range – 165 to 2353
 2. Average – 583

⁵ <https://www.census.gov/prod/cen2010/doc/sf1.pdf#page=605>

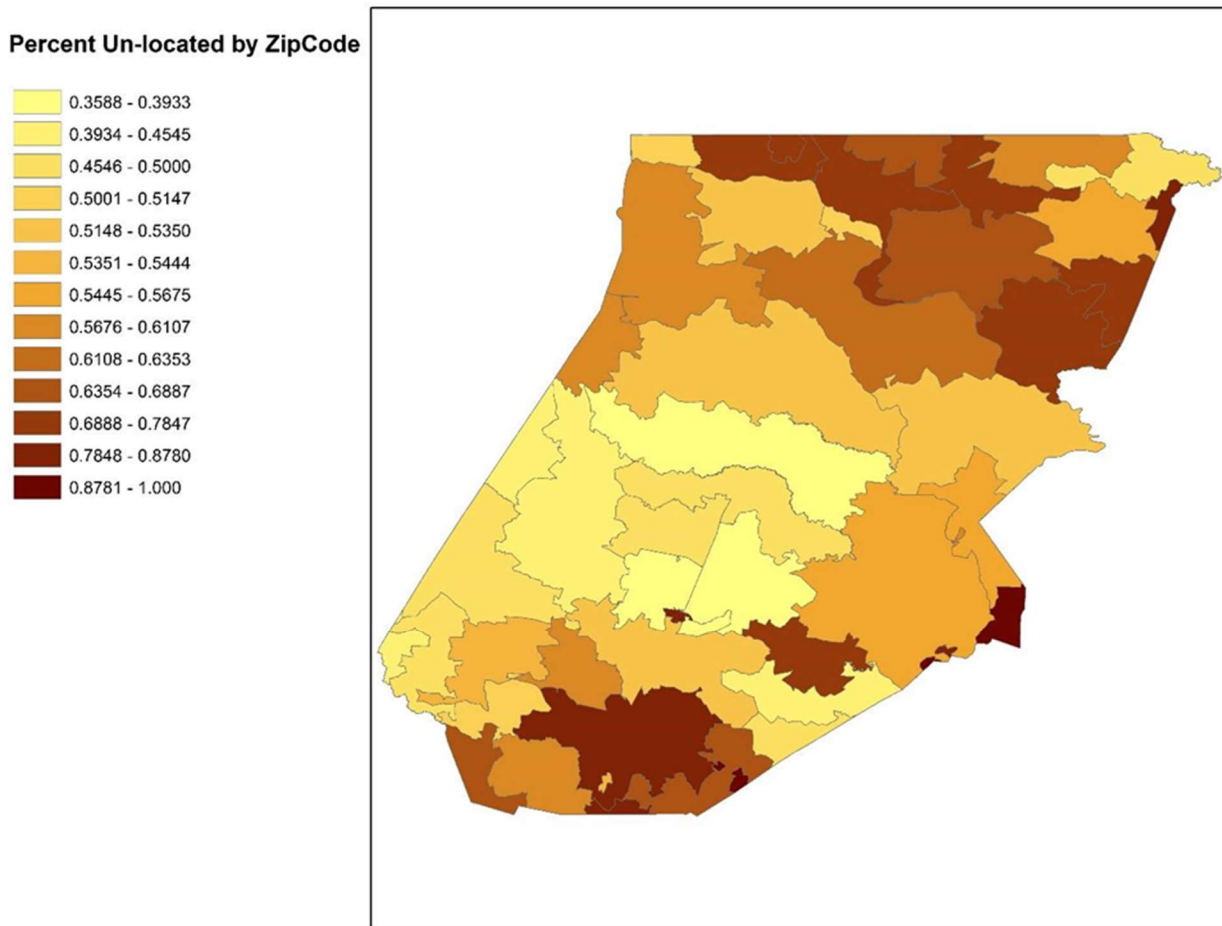
⁶ <https://www.census.gov/housing/hvs/definitions.pdf>

⁷ <https://www.census.gov/topics/income-poverty/poverty/about/glossary.html>

- ii. Population by CBG
 - 1. Range – 183 to 6200
 - 2. Average – 1267
- 2. Assistance
 - i. Total Assistance Instances - 39376
 - ii. Located CIP Assistance Instances – 3845
 - iii. Average – 16 houses per CBG
 - iv. Located LIEAP Assistance Instances - 5487
 - v. Average – about 23 houses per CBG
 - vi. Un-located homes – 22977
 - vii. Avg- around – 441

VIII. Example

Percent Un-located Assistance Instances by Zip Code



Appendix I: Sample Non-Disclosure Agreement

NON-DISCLOSURE AGREEMENT – Template #1

This Agreement (“**Agreement**”) is made and entered into as of this [REDACTED] day of [REDACTED], (“**Effective Date**”), by and among the [REDACTED] (“**Disclosing Party**”), and [REDACTED] (“**Receiving Party**”) (each of the Disclosing Party and the Receiving Party being referred herein as a “**Party**” and collectively as the “**Parties**”).

WHEREAS, the Receiving Party is [REDACTED].

WHEREAS, the Disclosing Party, [REDACTED] is a data providing partner assisting in the creation and functionality of the replicable and scalable geospatial database management system and data analytics framework

WHEREAS, the Receiving Party is required to work collaboratively under Subawards (“**Subawards**”) with **Subrecipients** to develop, test, implement and promote a replicable and scalable geospatial database management system and data analytics framework to facilitate the creation, enhancement, expansion, understanding, and/or deployment of energy efficiency programs to reduce energy use within the [REDACTED] region.

WHEREAS, the Receiving Party has concluded Subawards with Subrecipients for the performance of a portion of the work. Under the Subawards, the Subrecipients will supply all personnel, equipment, and materials necessary to accomplish the Purpose of this Agreement.

WHEREAS, the Parties are entering into this Agreement to set out the practices, protocols, and procedures for the transfer of Confidential Information held by the Disclosing Party to the Receiving Party and how such Confidential Information will be secured, managed, and treated, as well as when such information may be disclosed.

WHEREAS, the Receiving Party will serve as project administrator and the Subrecipients will be the immediate recipients of the Confidential Information. The Receiving Party will grant an immediate bypass to the Confidential Information disclosed by the Disclosing Party to enable the Confidential Information to flow directly to the Subrecipients provided that similar NDAs have concluded between the Receiving Party and the Subrecipients prior to or on the Effective Date of this Agreement.

NOW, THEREFORE, in consideration of the foregoing premises and for other good and valuable consideration, the receipt and sufficiency of which is hereby acknowledged, the Parties hereto agree as follows:

ARTICLE I DEFINITIONS

In this Agreement, unless the context requires otherwise, the following words shall have the following meanings:

“**Affected Party**” has the meaning set forth in Article V, section 1 of this Agreement.

“**Agreement**” means this Non-Disclosure Agreement as it may be amended, modified or otherwise supplemented, as in effect from time to time.

“**Confidential Information**” shall have the meaning set forth in Article III, section 1 of this Agreement. Confidential Information may include, but is not limited to, all information from data fields listed in Appendix A.

“**De-identification Mechanisms**” refers to data management operations performed on Confidential Information such that the operation renders the information non-confidential. Data De-identification Mechanisms may include, but not limited at: deleting or masking personal identifiers, data anonymization, data aggregation, data blurring, and computing non-confidential derivatives.

“**Disclosing Party**” means the Party disclosing Confidential Information to another Party. For the purposes of this Agreement the Disclosing Party is [REDACTED] and includes affiliates, subsidiaries, agents, directors, officers, employees and consultants.

“**Effective Date**” has the meaning set forth in the Preamble to this Agreement and means the date of execution of this Agreement.

“**Indemnified Parties**” has the meaning set forth in Article VII of this Agreement.

“**Law**” means any applicable constitutional provision, statute, act, code, law, regulation, rule, ordinance, order, decree, ruling, proclamation, resolution, judgment, decision, declaration or interpretive or advisory opinion of an Authority.

“**Non-Confidential Information**” shall have the meaning set forth in Article III, section 2 of this Agreement.

“**Party**” or “**Parties**” has the meaning set forth in the Preamble to this Agreement and Article IX, section 5.

“**Public Information**” means any information that is available to the Public, including but not limited to Non-Confidential Information.

“**Public**” refers to individuals, groups, or entities, including but not limited to a Party or Subrecipient as defined in the Agreement.

“**Purpose**” has the meaning set forth in Article II of this Agreement.

“**Receiving Party**” means the Party, and its respective affiliates, subsidiaries, agents, directors, officers, employees and consultants, receiving Confidential Information from a Disclosing Party. For the purposes of this Agreement, the Receiving Party is the [REDACTED].

“**Residuals**” means any information that is retained in the unaided memory of the Receiving Party's, its Subrecipients, and the Subrecipients employees, agents, representatives who have had access to the Disclosing Party's Confidential Information in compliance with the terms and conditions of this Agreement. The Receiving Party's, its Subrecipients, and the Subrecipients employees, agents, representatives' memory is unaided if they have not intentionally memorized the Confidential Information for the purpose of retaining and subsequently using or disclosing it and has not intentionally referred to the Confidential Information to refresh his or her memory for any such purpose.

“**Subaward**” means a legal instrument to provide support for the performance of any portion of the substantive project or program for which the Receiving Party subcontracted to an eligible Subrecipient.

“**Subrecipients**” means an entity that receives a Subaward from the Receiving Party

- **INSERT NAMES OF SUBRECIPIENTS**

ARTICLE II

PURPOSE

1. The Purpose of this Agreement is to facilitate the development and the accomplishment of the ...
2. Each Party hereby agrees that all Confidential Information will be used by the Receiving Party solely for the Purpose of this Agreement and will not be used by the Receiving Party for any other purpose unless agreed to in writing by both Receiving Party and Disclosing Party.

ARTICLE III

CONFIDENTIAL INFORMATION

1. Confidential Information, as used in this Agreement, means all information, data and knowledge, or material of a proprietary or confidential nature whether or not owned or developed by the Disclosing Party, which the Receiving Party may obtain through any direct or indirect contact with the Disclosing Party that is delivered or disclosed by or on behalf of the Disclosing Party to the Receiving Party, whether written or oral, not generally known to the Public in the industry in which the Disclosing Party is or may become engaged in the future. More specifically, for the purposes of this Agreement, Confidential Information shall include, but not limited to:
 - a) Information, as listed in Appendix A, that is furnished by the Disclosing Party to the Receiving Party;
 - b) Information that has been expressly identified in writing by the Disclosing Party as Confidential Information or proprietary at the time of disclosure either in the document which provided such information, in the transmittal materials accompanying such Confidential Information, or in a separate document, which identifies the information with sufficient specificity and clarity so that the entity receiving such information has been made aware that Disclosing Party seeks

confidential treatment for such information or within 30 days after any oral disclosure.

- c) All records, documents, financial information, function, system, design developments, ideas, concepts, business methods or other information, whether oral, written, electronic or otherwise, furnished by the Disclosing Party to the Receiving Party, either before or after the date of this Agreement relating to the Disclosing Party or its business activities, customers, customer lists, addresses, phone numbers or similar identifying indicia, or clients, client lists, or analyses, compilations, studies, summaries, extracts or other documents or records prepared by the Disclosing Party that contain, otherwise reflect, or are generated from such records, documents, financial information, ideas, concepts, business methods or other information, provided by the Disclosing Party, but only to the extent that Confidential Information is specifically contained therein; or
- d) Software, hardware, prototypes, products in development, computer code or data, equipment, notes, memoranda, manuals, disks, know-how, procedures, materials, reports, interpretations, projections and other information, whether written or oral, to the extent it relates to the Disclosing Party, its employees, its loan agents, its affiliates or subsidiaries, which has been previously provided to the Receiving Party, or which will be provided to the Receiving Party pursuant to this Agreement.

2. For the purposes of this Agreement, Confidential Information shall not include:

- a) Information that is or becomes generally available to the Public other than as a result of unauthorized disclosure by the Receiving Party, its agents or employees;
- b) Information that becomes available to the Receiving Party on a non-confidential basis from a source other than the Disclosing Party, provided that (i) such source is not prohibited from transmitting such information by contractual, legal or other obligation; or (ii) was in Receiving Party's possession prior to disclosure of the same information by Disclosing Party;
- c) Information that is defined as "Public Information";
- d) Information that can be shown by the Receiving Party's prior records to have been already known to the Receiving Party other than through disclosure by a third party which would not be subject to exclusion based on Article III section 2.b)(i) above.;
- e) Information independently developed by the Receiving Party's employees without use, knowledge or reference of such Confidential Information;
- f) Confidential Information shared by the Disclosing Party to the Receiving Party to which De-identification Mechanisms have been applied after disclosure; and
- g) Subject to Article IV section 3.c), information that is required to be disclosed or made available by the Receiving Party pursuant to any applicable Law, or decision of any court or tribunal of competent jurisdiction or any government body, agency or regulatory body, or by subpoena or other legal process.

3. All Confidential Information, as defined in this Agreement, will be distributed only in accordance with this Agreement. All other information, which is not defined as Confidential Information may be released and used at the discretion of the Parties.

ARTICLE IV
SHARING, USE, AND TREATMENT OF INFORMATION

1. The Disclosing Party will provide written certification to the Receiving Party that it has the requisite authority to share the Confidential Information with the Receiving Party.
2. Subject to the non-disclosure and data security and management requirements provided herein, the Disclosing Party agrees to share information with the Receiving Party as reasonably necessary to achieve the Purpose of the Project.
3. Disclosure of Confidential Information
 - a) Notwithstanding the foregoing, the Receiving Party may disclose Confidential Information of the Disclosing Party to a listed Subrecipient only if such disclosure is authorized in writing by the Disclosing Party and the Receiving Party, as provided in this Agreement and under terms consistent with and similar to those contained herein.
 - b) The Receiving Party will serve as project administrator and the Subrecipients will be the immediate recipients of the Confidential Information. The Receiving Party will grant an immediate bypass to the Subrecipients to enable the Confidential Information disclosed by the Disclosing Party to flow directly to the Subrecipients provided that similar NDAs have concluded between the Receiving Party and the Subrecipients prior to or on the Effective Date of this Agreement.
 - c) The Receiving Party shall promptly advise the Disclosing Party if the Receiving Party becomes aware of any possible unauthorized disclosure or use of the Confidential Information.
4. For information to receive protection under this Agreement as Confidential Information, the information must meet the provisions listed in Article III section 1 and the information is:
 - a) disclosed in writing or other tangible form. The information must be plainly marked as Confidential Information and shall bear the date of disclosure; or
 - b) orally disclosed. The information shall be identified as Confidential Information, orally, at the time of disclosure and confirmed in a written summary thereof, appropriately marked by the Disclosing Party, within thirty (30) days as being Confidential Information and shall bear the date of disclosure. Inadvertent failure to mark Confidential Information as “Confidential” at the time it is disclosed shall not be deemed a waiver by the Disclosing Party of the protections of this Agreement provided that such Confidential Information is, no more than 30 days from disclosure, identified and marked "Confidential" promptly upon the discovery of its inadvertent disclosure and the Receiving Party shall have no liability for its use of Confidential Information during the period of disclosure to the time of identification.
5. The Receiving Party will ensure that the Subrecipients, as primary recipients of the data,

protect all such Confidential Information from disclosure to others using the same degree of care used to protect their own similar confidential information, but in any case, using no less than a reasonable degree of care. The Receiving Party shall ensure that the Subrecipients will use all reasonable security measures to protect the Confidential Information from unauthorized access, destruction, use, modification, or disclosure. Upon request, the Receiving Party will provide certification that the Subrecipients have adopted all reasonable security measures. These measures shall include but are not limited to:

- a) Written policies regarding information security;
 - b) Password protected workstations at the Subrecipients' premises, any premises where work or services are being performed, and any premises of any person who has access to such data;
 - c) Applicable data security standards and procedures that comply with all federal and state laws, regulations, rules, and policies including **(LIST SPECIFIC STATE POLICIES OR REGULATIONS)**. Such standards and procedures include, but are not limited to access to systems and data; transmission and storage of data; access to, disclosure, sharing and confidentiality of client records; use, retention and disposal of data; and reporting of security incidents and breaches;
 - d) Applicable Subrecipient policies and standards; and
 - e) Measures to safeguard against the unauthorized access, destruction, use, alteration or disclosure of any such data including, but not limited to, restriction of physical access to such data and information, including hard drives, and establishment of an information security program that at all times is in compliance with reasonable security requirements as agreed to between recipient and disclosing party.
6. The Receiving Party shall with regard to any Confidential Information disclosed pursuant to this Agreement by or on behalf of a Disclosing Party on or after the Effective Date hold the Confidential Information in confidence and, except as is otherwise stated herein or agreed in writing by the Disclosing Party, and shall use any Confidential Information disclosed to it pursuant to this Agreement only for carrying out the Purpose of the Agreement.
7. Notwithstanding any provision herein to the contrary, the Receiving Party may share Confidential Information with their respective Subrecipients as the Confidential Information will flow directly from the Disclosing Party to the Subrecipients, provided, however, that the Receiving Party shall take all reasonable measures to ensure that its respective Subrecipients may not disclose such Confidential Information to any other employee, representative, or agent of the Subrecipients or any other person except as permitted under this Agreement or similar agreements concluded between the Receiving Party and Subrecipients. Thus, the Subrecipients, in order to be granted access to the Confidential Information, must agree to be bound by terms consistent with and similar to this Agreement and to maintain the confidentiality of the information, concluded prior or on the Effective Date of this Agreement. The Receiving Party agrees to maintain, update and share a list of non-disclosure agreements/memorandums of understanding, and Subaward terms of those Subrecipients that have been provided access to and have access to the Confidential Information that can be shared upon request with the Disclosing Party. The Receiving Party shall not use the

Confidential Information or any portion thereof to give any Party or a competitor of any Party a competitive or commercial advantage. In the event a Subrecipient ceases to be employed or engaged by the Receiving Party, or is employed, retained, or given duties that include competitive duties, (i) the Receiving Party shall terminate the Subrecipients' access to Confidential Information, and (ii) the Receiving Party shall cause the Subrecipients to return or dispose of the Confidential Information, or transfer the information to another Subrecipients of the Receiving Party.

8. The Receiving Party shall not disclose any Confidential Information to any employees, representatives, or agents of the Receiving Party, except those employees, representatives or agents who are required to have access to the Confidential Information to perform job duties connected to the Purpose of the Agreement. The Receiving Party shall require those persons to comply with the provisions of this Agreement. The Receiving Party will maintain a list of the employees, representatives or agents that have been given access to the Confidential Information in order to fulfil the Purpose of this Agreement.
9. Confidential Information shall not be used by the Receiving Party other than for fulfilling the Purpose of the Project.
10. In using the De-identification Mechanisms on the Confidential Information disclosed by the Disclosing Party, the Receiving Party will take all reasonable steps to ensure that the individual data points cannot be re-identified.
11. Notwithstanding any provision herein to the contrary, Public Information may be made available to the Public. Fees may apply to cover process and handling expenses.
12. The Receiving Party shall report a suspected or confirmed security breach to the state Division of Social Services within twenty-four (24) hours after the breach is first discovered, provided that the Receiving Party shall report a breach involving Social Security Administration data or Internal Revenue Service data within one (1) hour after the breach is first discovered. The Receiving Party agrees that, if the State/County determines that some or all of the activities within the scope of this Agreement are subject to the Health Insurance Portability and Accountability Act of 1996, P.L. 104-91, as amended (“HIPAA”), or its implementing regulations, it will comply with the HIPAA requirements.

ARTICLE V REMEDIES

1. Each Party acknowledges that if a Party to this Agreement brings an action to enforce the provisions of this Agreement (the “**Affected Party**”), the damages may be irreparable and difficult to measure, and the Affected Party shall be entitled to seek equitable relief, including injunction and specific performance, as a remedy for such breach. Such remedies shall not be

deemed to be the exclusive remedies for a breach of this Agreement but shall be in addition to any and all other remedies available at Law or equity. Should litigation arise concerning this Agreement, the prevailing party, as determined by the court in an express opinion relating thereto, shall be entitled to seek its attorney's fees and court costs in addition to any other relief which may be awarded. Subject to Article VIII hereof, all actions to enforce shall be brought in the state or Federal courts to the State of [REDACTED].

2. If any applicable federal, state, or local law, regulation, or rule requires the Data Providing Party to give written notice of a security breach to affected persons, the Receiving Party shall bear the cost of the notice.

ARTICLE VI OWNERSHIP, TERM AND TERMINATION

1. Except for Residuals, all Confidential Information shall be and remain the property of Disclosing Party, and no right or license is granted to Receiving Party with respect to any Confidential Information.
2. This Agreement shall be in full force and effect for the period of performance of [REDACTED]. Each Party may terminate its participation in this Agreement by giving thirty (30) days' prior written notice to the other Parties.
3. The treatment of information classified as Confidential Information pursuant to this Agreement will be in effect for five years after the Confidential Information is furnished by the Disclosing Party, unless the Parties agree to extend the effectiveness as it relates to specific Confidential Information prior to the termination of the five-year period.
4. Upon termination of this Agreement and upon a written demand by the Disclosing Party, the Receiving Party shall return to the Disclosing Party all written materials, including copies thereof, that contain or evidence Confidential Information. The Receiving Party shall also deliver to the Disclosing Party written statements signed by the Receiving Party certifying that all materials have been returned within 15 business days of receipt of the request. That portion of the Confidential Information that may be found in analyses, compilations, studies or other documents prepared by or for the Receiving Party and all Confidential Information that is oral will be kept by the Receiving Party subject to the terms of this Agreement or destroyed.

ARTICLE VII INDEMNIFICATION

1. The Receiving Party shall at times indemnify and hold the Disclosing Party and its subsidiaries and affiliates, and their respective officers, directors, agents, and employees (“**Indemnified**

Parties”) harmless from and against any and all damages (including consequential damages), liabilities, costs and expenses, including reasonable counsel fees and expenses (whether incurred in an action or proceeding between the Receiving Party or otherwise), arising out of or relating to any claim or cause of action asserted against the Disclosing Party or any of the Indemnified Parties due to material breach or alleged breach by the Receiving Party or Parties of its obligations hereunder; provided that the Disclosing Party shall give prompt written notice, cooperation, and assistance to the Receiving Party relative to any such claim or suit; provided further, that no settlement of any such claim or suit (other than between the Receiving Parties) shall be made without the prior written consent of the Receiving Party or Parties, which consent shall not be unreasonably withheld.

ARTICLE VIII GOVERNING LAW

1. This Agreement shall be governed by and construed in accordance with the laws of the State of (without giving effect to its choice of law principles).

ARTICLE IX MISCELLANEOUS

1. Except as otherwise provided, this Agreement shall inure to the benefit of and shall be binding upon the Parties hereto, their successors and permitted assigns, except that no assignments of any rights in Confidential Information shall be permitted without the prior written consent of the Disclosing Party and the Receiving Party. In the event that any one of the provisions contained in this Agreement should be found to be invalid, illegal or unenforceable in any respect by a court of competent jurisdiction, the validity, legality or enforceability of the remaining provisions contained in this Agreement shall not in any way be affected or impaired by such a finding. No waiver of any provisions of this Agreement shall be valid unless the same is in writing and signed by the Party against whom such waiver is sought to be enforced. A waiver or consent given by either Party on any one occasion is effective only in that instance and will not be construed as a bar to or waiver of any right on any other occasion.
2. This Agreement as well as all Appendixes, hereto contains the entire agreement of the Parties, supersedes any and all prior agreements, written or oral, between them relating to the subject matter hereof, and may not be amended unless agreed to in writing by each Party.
3. Neither Party is under legal obligation to continue negotiations. The Agreement does not indicate any formal relationship or partnership and either Party can terminate discussions at any point. Nothing in this Agreement is intended to or shall operate to create a partnership or joint venture of any kind between the Parties (or any of them), or to authorize any Party to act as agent for another, and no Party shall have authority to act in the name or on behalf of or

otherwise to bind any other in any way.

4. The Receiving Party attests that it does not and will not require its employees or Subawardees to sign internal nondisclosure or confidentiality agreements or statements prohibiting or otherwise restricting its employees or Subawardees from lawfully reporting waste, fraud, or abuse to a designated investigative or law enforcement representative of a Federal department or agency authorized to receive such information.
5. These provisions are consistent with and do not supersede, conflict with, or otherwise alter the employee obligations, rights, or liabilities created by existing statute or executive order relating to (1) classified information, (2) communication to Congress (3) the reporting to an inspector general of a violation of any Law, rule, or regulation, or mismanagement, a gross waste of funds, an abuse of authority, or a substantial and specific danger to public health or safety, or (4) any other whistleblower protection. The definitions, requirements, obligations, rights, sanctions, and liabilities created by controlling executive orders and statutory provisions are incorporated into this Agreement and are controlling.
6. The limitation in Article IX Section 4 above shall not contravene requirements applicable to Standard Form 312, Form 4414, or any other form issues by a Federal department or agency governing the nondisclosure of classified information.
7. Notwithstanding provision listed in Art. IX Section 4, a nondisclosure or confidentiality policy form or agreement that is to be executed by a person connected with the conduct of an intelligence or intelligence-related activity, other than an employee or officer of the United States Government, may contain provisions appropriate to the particular activity for which such document is to be used. Such form or agreement shall, at a minimum, require that the person will not disclose any classified information received in the course of such activity unless specifically authorized to do so by the United States Government. Such nondisclosure or confidentiality forms shall also make it clear that they do not bar disclosures to Congress, or to an authorized official of an executive agency or the Department of Justice, that are essential to reporting a substantial violation of law.
8. Where appropriate to give effect to the Purpose, references to the Disclosing Party includes affiliates, subsidiaries, agents, directors, officers, employees and consultants, to the extent that each is bound by the obligations set forth herein..
9. Communications under this Agreement shall be valid only if in writing and delivered personally, by mail, by facsimile or by electronic mail, and sent by the sender to each other Party at its address or number listed for a Party's representative. A validly given notice, consent or approval will be effective when received if delivered.
10. This Agreement shall be executed in number of counterparts, each of which when executed and delivered shall constitute an original of this Agreement, but all the counterparts shall together constitute the same agreement. No counterpart shall be effective until each Party

has executed at least one counterpart.

11. This Agreement contains the entire understanding between the Parties with respect to the confidentiality and non-disclosure obligations relating to the Confidential Information.

IN WITNESS, WHEREOF and intending to be legally bound hereby, the Parties by and through their authorized representative execute this document as of the Effective Date set forth above:

[DISCLOSING PARTY]

Signature:

Printed Name:

Title:

Company:

Address:

Email:

Phone:

Fax:

[RECEIVING PARTY]

Signature:

Printed Name:

Title:

Company:

Address:

Email:

Phone:

Fax:

APPENDIX A

Pursuant to Art.III.1 of the Agreement, the Disclosing Party hereby certifies that it will provide the Receiving Party with the following Confidential Information, in compliance with the relevant terms and clauses of the Agreement:

Acknowledgements and Disclaimers

This material is based upon work supported by a Department of Energy Office of Energy Efficiency and Renewable Energy Cities Leading through Energy Analysis and Planning (Cities-LEAP) grant, No. DE-EE0007738.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

The work of the Vermont Law School Institute for Energy and the Environment is intended to assist and expedite professional assessment. The IEE is an academic organization and its work does not purport to be, and is not the equivalent of, the work of a licensed professional with expertise in this area. The IEE does not provide legal services, advice, or consultation. Any legal assistance must be the subject of a separate agreement with a licensed attorney. Before making significant decisions based on this work, it would be appropriate to consider consultation with a licensed professional with expertise in this field.