

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

fEECe, an energy use and occupant behavior dataset for net-zero energy affordable senior residential buildings

Frederick Paige^{1*}, Philip Agee² & Farrokh Jazizadeh^{1*}

The behaviors of building occupants have continued to perplex scholars for years in our attempts to develop models for energy efficient housing. Building simulations, project delivery approaches, policies, and more have fell short of their optimistic goals due to the complexity of human behavior. As a part of a multiphase longitudinal affordable housing study, this dataset represents energy and occupant behavior attributes for 6 affordable housing units over nine months in Virginia, USA which are not performing to the net-zero energy standard they were designed for. This dataset provides researchers the ability to analyze the following variables: energy performance, occupant behaviors, energy literacy, and ecological perceptions. Energy data is provided at a 1 Hz sampling rate for four circuits: main, hot water heater, dryer, and HVAC. Building specifications, occupancy, weather data, and neighboring building energy use data are provided to add depth to the dataset. This dataset can be used to update building energy use models, predictive maintenance, policy frameworks, construction risk models, economic models, and more.

Background & Summary

Building occupant education is an important research direction providing an opportunity to maximize the efficacy of energy efficiency investments in the housing industry¹. Despite increasing efforts by property managers to educate tenants on technology in housing units, many residents lack an understanding of energy efficient technologies. As technology becomes further integrated into housing units, user education will become more of a factor in the optimization of residential energy use. For example, McCoy et. al. reported that residents that received education on their apartments had a lower average energy usage monthly and annually (over 3 years) by almost 15% (14.8%) and a lower energy bill by \$10.56 per month².

Providing an education to residents is not only the appropriate ethical choice, but also it can be a powerful business strategy for builder-developers. In the context of this dataset, electricity pricing in Virginia is trending upwards with United States and global rates. Virginia electricity prices have increased by an average of 1.5% per year over 25 years and 3% annually over the last ten years³. Traditionally, utilities provide education in the form of energy usage feedback with a “lag” which becomes more detrimental as prices rise. Human factors researchers have reported that people are generally poor at managing systems with lags in information and delayed feedback loops^{4,5}. For energy efficient housing systems, energy monitoring systems exist which can report energy use in real-time using an Energy Feedback Display (EFD) for occupants (meetnexus.com).

Using the same hardware, building managers can utilize cloud-based software platforms to monitor and create performance reporting, instantly being aware of system failures or inefficiencies. Studies have shown that EFDs have improved behavior towards a more energy efficient lifestyle and resulted in 10–15% monthly energy use reductions given: (a) frequent feedback, (b) provided over long periods of time, (c) with appliances characterized individually, (d) presented in clear, appealing ways, and (e) utilizing computerized, interactive tools^{6–9}.

In this paper, we release our dataset fEECe¹⁰. This dataset is the outcome of a unique case study within a longitudinal multiphase study to explore the impacts of feedback and education on energy-use behaviors of seniors in

¹Charles E. Via Jr. Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg VA, 24061, United States. ²Myers-Lawson School of Construction, Virginia Tech, Blacksburg VA, 24061, United States. *email: freddyp@vt.edu; jazizade@vt.edu

affordable housing units. The units were designed to be net-zero energy and followed the EarthCraft Multifamily design guidelines¹¹, adhering to the Low-Income Housing Tax Credit requirements¹². The participants, in this study, participated in an educational workshop and were provided with a budget-based (giving feedback based on remaining budget) energy monitoring system to receive feedback on their electricity usage in real-time. The following two questions motivated the collection of the presented data.

- How does targeted energy education and energy consumption feedback impact energy literacy of seniors in net-zero energy apartment units?
- How does targeted energy education and energy consumption feedback impact energy consumption of seniors in net-zero energy apartment units?

The use of “how” in the questions reflects the qualitative portion of this study which is not presented in this dataset. The dataset described here is solely the quantitative portion of this mixed methods case study. The quantitative data collected can be analyzed in multiple ways to expand beyond the original objectives of our study.

How occupants shift their energy use behaviors is linked to their environmental perceptions, previous energy behaviors, and level of education^{13,14}. Targeted occupant education has shown potential to be an effective method for reducing energy consumption^{15,16} and in a previous phase of this study, education strongly correlated with reduced energy consumption². A unique aspect in our study is the focus on senior residents who are not financially incentivized to conserve energy – a topic which needs further investigation¹⁷ and is critical in the context of affordable housing. Furthermore, this dataset reflects a case study framed according to the concepts in the U.S. Department of Energy (DOE) Energy Literacy guide¹⁸ through coded survey responses.

Scholars may use the fLEECe dataset to investigate:

- Energy use behaviors in the context of the US DOE Energy literacy concepts
- Long-term resident consumption patterns
- Evaluate the performance of energy efficient water-heaters, mini-splits, and dryers in affordable senior housing
- Understand the impact of climate on consumption
- Non-intrusive energy monitoring techniques for data with different resolution

Methods

Case study context. This dataset was collected in the context of a singular case study in continuation of a multiphase state-wide longitudinal study in Virginia¹. Building upon findings in the last phase of the longitudinal study, this case study utilizes interviews, field observations, surveys, and energy consumption data to explore the impact of targeted education on residents’ energy consumption. The dataset presented in this paper has detailed information on participants energy consumption, perceived behavior, beliefs, and energy literacy. This dataset is an addition to the increased amount of building monitoring data we need for improving residential energy use through advanced metering and monitoring^{19,20}.

The property under examination is a senior living Low Income Housing Tax Credit (LIHTC) project in Richmond, Virginia that has been being monitored since 2013. This property was selected for its unique occupant energy consumption patterns and design features. A unique feature of this study pertains to that fact that the occupants are not financially incentivized to conserve energy given that the energy bills in these buildings are paid by the property managers. While the property was performing better than average building nationally and in Virginia, the property was falling short of its net-zero energy goal². As reflected in Tables 1 and 2, the one-bedroom housing units were designed to be grid-tied net-zero housing units with all electric energy efficient appliances, tight building envelope, energy efficient glazing, and solar PV arrays on site. The windows and their shadings are manually operable and there is no known outside factors (such as noise or air pollution) to prevent occupants from opening the windows. Building characteristics can be found within the dataset in the tabular data excel file. A color coded illustration of the building and unit layout, material properties, window to wall ratio, and heating and cooling specifications can be found on the “Architectural Data” tab of the dataset¹⁰. The 6 units which are the focal point of this study come together to form a 7-unit building. Although all 7 units were monitored, due to unreliable data from one unit, the fLEECe dataset only includes circuit level data from 6 units. Also provided in the dataset is monthly data for the adjacent 32-unit building which is a part of the same housing property.

Human data statement. The study protocol used for this study was approved by the Virginia Tech Institutional Review Board. Participants of this study were not subject to any known risks and informed consent was provided from all subjects.

Energy monitoring protocol. This dataset includes energy use data, collected by advanced circuit level energy monitors and standard utility energy meters. The NEXI energy feedback device⁶ captured second level data (i.e., @1 Hz) for 6 units in this study on four circuits: main, hot-water heater, dryer, and HVAC (mini-split) system using 100A non-invasive split core current transformer clamps with $\pm 1\%$ accuracy. NEXI uses a proprietary data collection and aggregation process designed to work with their unique color-wheel display that provided users with simplified feedback. The detailed energy data from NEXIs including power, voltage, current, and phase angle for four circuits are included in the dataset as raw CSV files. The calibration process for the NEXI data is provided in the dataset as well. Each unit was sub-metered and energy use data was collected per month through a utility benchmark tracker, WegoWise (wegowise.com), which logged five years of monthly energy consumption

Characteristics	Building 1	Building 2
Climate zone	US 4A	US 4A
# of apartments	7 (6 were presented in fLEECe)	32
# of buildings	1	1
PV-system	16.9 m ² , south-facing, 2.43 kWp	166.1 m ² , south-facing, 2.43 kWp
Living area	668 ft ² (62.1 m ²)	708 ft ² (65.9 m ²)
Building volume	6,178 cf (175 m ³)	6,372 cf (180 m ³)
Heating system	Air-source heat pump 18 kBtuh, 9 HSPF	Air-source heat pump 18 kBtuh, 9 HSPF
Cooling system	Air-source heat pump, 18 SEER	Air-source heat pump, 18 SEER
Distribution	2, Ductless air systems	2, Ductless air systems
Water heating	Standard electric storage (0.92 EF)	Standard electric storage (0.92 EF)
Ventilation	Energy recovery ventilator	Energy recovery ventilator
Windows U-value SHGC (g-value)	0.30 BTUh/ft ² °F (0.95/W/m ² K) 0.28	0.30 BTUh/ft ² °F (0.95/W/m ² K) 0.28
Wall U-value	0.04 BTUh/ft ² °F (0.13/W/m ² K)	0.04 BTUh/ft ² °F (0.13/W/m ² K)
Roof/Attic U-value	0.03 BTUh/ft ² °F (0.11/W/m ² K)	0.03 BTUh/ft ² °F (0.11/W/m ² K)
Air tightness	4.1 ACH ₅₀	6.1 ACH ₅₀

Table 1. Building specifications. All electric 1-bedroom units.

Meta Data	Total Units	Targeted Education (TEI)	TEI + In-home Display	Control
Number of Units	38	20	6	18
Energy Data	Main aggregate data per month	Main aggregate data per month	Main aggregate data, 3 sub-circuits (dryer, water heater, and air conditioning) – 1-sec, 1-min, and 15-min resolution	Main aggregate data per month
Survey Data		Pre and Post intervention energy perceptions, behaviors, and literacy	Pre and Post intervention energy perceptions, behaviors, and literacy	

Table 2. Behavioral data sample summary.

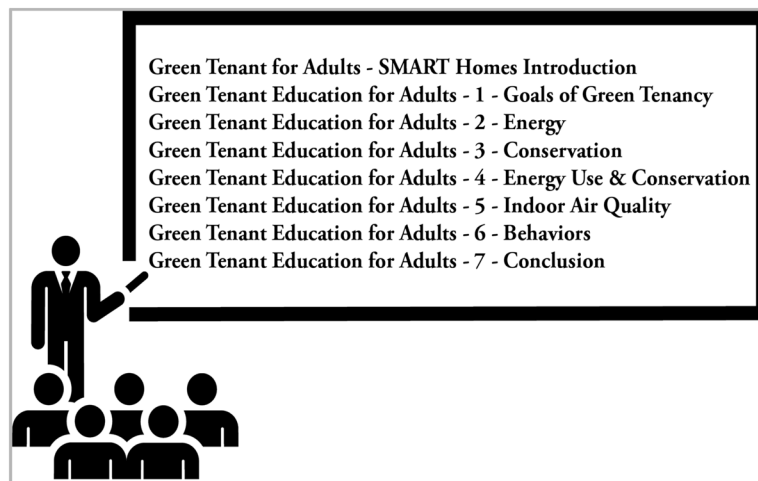
data for the property. Paper utility bills were used to spot check the data collected by both the NEXI devices and WegoWise platform.

Given the targeted community for this study, we opted for a hard-copy survey which was later digitized for data analysis. Participants were surveyed on site in a community room typically used for social events at the property. The survey combined instruments from literature to measure energy behaviors, perceptions, and literacy^{2,21,22} adapted for use in this case. Surveying in person requires a great deal of flexibility and effort. Surveys were designed to be senior-friendly utilizing large fonts, high contrast, and large formatting for recording responses. Even so, multiple participants needed assistance while taking the survey. A team of two researchers was present to administer the survey with one researcher reading the questions aloud to the group and the other researcher assisting participants one on one when necessary. Property managers helped with recruitment and provided logistics during the education and survey data collection. A \$25 gift card was provided by the research team during the pre and post survey, to participating residents to aid in recruitment. The survey questions were coded to the energy literacy guide providing descriptive data on the lessons learned by the study participants.

The recruitment efforts resulted in twenty residents attending a one-hour community meetings and receiving a Tenant Education Intervention (TEI). Of the twenty residents, twenty successful records were developed for analysis. Through stratified sampling, six residents who received the TEI, also were provided with an in-home display. It is important to note that the researchers were able to leverage the existing energy use data from previous work to compare monthly energy usage and the efficacy of resident education efforts. An overview of the behavioral data sample is provided in Table 2.

Educational intervention. There are two interventions in this study delivered in the following formats: 1) residents who received a Targeted Energy Intervention (TEI) (see Fig. 1a) on March 28th, 2017 and 2) residents who received a TEI + an energy feedback device (see Fig. 1b–d) with the EFD installation occurring on July 6th, 2017. The TEI consisted of the authors guiding residents through seven educational videos and a ten-minute PowerPoint presentation that featured technologies specific to their apartment unit.

Principles and concepts in the DOE Energy Literacy guide have a great deal of overlap and the most critical concepts were selected as a scope for this study. Concepts were prioritized in regard to their connection to the TEI and in-home experiences. After the concepts were prioritized, the scope of this study was set to fit the limitations of the data collection and analysis methods utilized. The research team selected five concepts to bound this study which link to daily energy decisions and the targeted energy education we provided: (1) human use of energy is subject to limits and constraints, (2) conservation is one way to manage energy resources, (3) electricity is



(a)



(b)



(c)



(d)

Fig. 1 Educational and feedback instruments leveraged in the case study. (a) six residents received Targeted Education Intervention (TEI) and an Energy Feedback Display (EFD). The EFD was programmed with a daily energy budget that reflected energy behavior with a dynamic color display; (b) EFD interface reset at midnight showing minimal instantaneous use on left with green color, and full daily budget remaining on right with green color; (c) high real-time consumption on left with fuchsia color, and 20% of daily budget consumed on right with yellow color; (d) low real-time consumption on left with green color, and 70% of daily budget consumed on right with red color.

generated in multiple ways, (4) social and technological innovations affect the amount of energy used by society, and (5) energy use can be calculated and monitored.

The research team considered several factors when developing the daily energy budget for the EFDs. First, the RESNET accredited home energy rating (HERS) energy simulations that were developed during design were reviewed. Second, five years of historical energy use data for the selected apartments were analyzed to develop a measured average kWh/day per apartment. Finally, the team established the energy reduction goals (budgeted kWh/day) to be programmed into the NEXI energy feedback display by multiplying the simulated consumption by the EFD manufacturer's suggested multiplier for each unit ranging from 1.7 to 1.8. Table 3 provides an overview of the estimated, measured, and budgeted energy goals.

Weather and climate data. Weather data for the project location is publicly available through the National Oceanic Atmospheric Administration (NOAA)²³ and Weather Underground²⁴. Historic weather data allows for a researcher to assess the impact of year to year weather variance and climate to weather variances on the site. A baseline of 65 °F/18 °C is used to calculate Heating Degree Day (HDD) and Cooling Degree Day (CDD) with outdoor temperatures measured at the NOAA Richmond International Airport Weather Station (RIAWS). The station is located at 37.51151°, -77.32344° and is approximately 5 miles (8 kilometers) from the case study site. We utilized Weather Underground to retrieve ten years of monthly HDD and CDD from RIAWS.

Next, the team referenced the Typical Meteorological Year (TMY3) data for Richmond, VA, USA to develop a climate benchmark. TMY3 data represents a 30-year average benchmark and was retrieved from the ASHRAE Handbook of Fundamentals²⁵. TMY3 data is commonly employed in energy simulation models. The authors leveraged the TMY3 data to unpack 1) simulated performance versus observed performance, 2) yearly weather variance compared to TMY3, and 3) HDD/CDD during the survey period versus the TMY3.

Description	Unit A	Unit B	Unit C	Unit D	Unit E	Unit F
Simulated kWh/day	3.94	3.94	3.94	3.78	3.62	3.62
Measured kWh/day	7.99	8.96	9.39	11.03	12.07	9.40
Budgeted kWh/day	6.78	6.78	6.78	6.78	6.20	6.20

Table 3. Energy Feedback Display budgeting.

Sample	Number of Samples	Temporal range	Description	Data
Tabular Data	38 units, 20 Building occupants	July 2013–March 2018	Occupant Demographics, Monthly-time series power data, Pre and Post educational intervention energy literacy, behaviors, and environmental beliefs	fEECe Dataset 5_15_19.xlsx
NEXI data	6 units	July 7, 2017–March-22-2018	Second data at the circuit level for the entire unit, dryer, hot water heater, and HVAC (mini-split) unit	A-F_EST.csv

Table 4. Dataset Description.

Data Records

The fEECe dataset can be found on the Open Science Framework¹⁰. Energy data from each unit was collected utilizing three methods: 1) utility bills in.pdf format, 2) WegoWise utility tracking service, and 3) NEXI energy monitor data loggers. The WegoWise and NEXI data are present in this data set. The utility bills were not stored or shared publicly to protect participant identities. Utility bills were used to spot check the other two records. The WegoWise data is provided in a Microsoft Excel table at a monthly interval (kWh/month/unit). The WegoWise data represents 38 apartments, with 59 contiguous months of data from July 2013 to March 2018. Tables 4 and 5 describe the dataset files and the monitored variables.

The NEXI data is provided as raw (uncalibrated) CSV files and calibrated CSV files. Each CSV has 11 columns. Column A is the timestamp. Columns B-F represent the measured electrical current. These numbers need to be processed through the calibration formula to get amperage values. Column G represents the voltage readings and must be processed through a conversion formula for voltage. A guide for the calibration process is included in the dataset for those who would like develop scripts for managing uncalibrated data from a NEXI device. This will be particularly useful in the future for cloud computing processes using the NEXI's WIFI capabilities which allow for network connected NEXI's to upload data directly to a server.

Columns H through L are a representation of phase angle. This was an experimental feature, it may provide some information, but the feature remains in alpha and hasn't been tested for reliability to the same degree as the Voltage and Current readings. The NEXI calculates time stamps the number of seconds since July 7, 2017. The data provided starts at July 7, 2017 and ends at March 22, 2018.

In the data validation folder of the dataset, timeseries analysis are provided exemplifying one of our processes to ensure the data was valid before advanced analysis and validation processes were executed. It is important to note that there were no single missing seconds, and under a day's worth of time gaps in the data signifying the NEXI's were properly recording data consistently as long as they were powered. The identified time gaps most likely reflect power outages as the gaps were registered on multiple independent NEXI devices at the same time. Other data validation processes have been exemplified showing the ability for the data to be disaggregated to discover the use of the dryer at multiple settings.

Technical Validation

In this section, we have presented the results of analyzes on the dataset to ensure that they follow the expected patterns. As shown in Fig. 2, the energy data that was captured at the appliance (sub-circuit) level for the heat pump, dryer, and water heater, manifests rationale variability and reflects the expected patterns in the units and validates the quality of the data. Figure 2a,b shows the variation of energy consumption for the heat pump. Figure 2a shows the variations of daily energy consumption of heat pumps across the period of data collection. The heat pump represents both cooling and heating load variations. As expected, the elevated energy consumption around July and August (representing the cooling load) decreases due to seasonal transition between August and November, where the heating load causes a sharp increase in energy consumption. Furthermore, these graphs show the variability across different units reflecting the diversity in occupant behavior when it comes to operating the loads. Figure 2b shows the average daily variation of the heat pump energy averaged across all units for all days. This graph illustrates the expected variation of energy consumption which reflects the activities of occupants at home. As noted, the occupants of the units are senior citizens and therefore, during the day energy consumption remains relatively high.

On the other hand, the seasonality pattern is not observed in the energy consumption of dryers as shown in Fig. 2c. In other words, the observations show the distributed use of the dryers and the variations across different units. The interactional behavior impact could be better observed in Fig. 2d, in which the differences in dryer use including the time of use and the demand from each unit has been highlighted. The observed variations of the appliance use patterns support the validity of energy use with respect to occupant behavioral patterns across different units.

Data Category	Subcategory	Monitored Variables
Energy	Hot Water Heater	Electricity
Energy	Dryer	Electricity
Energy	Heating/cooling (HC)	Electricity
Inhabitants	Attitudes (AD)	Ecological beliefs
Inhabitants	Attributes (AT)	Consumption Behaviors

Table 5. Dataset Mapping to Mahdavi and Taheri's (2017) ontology¹⁹.

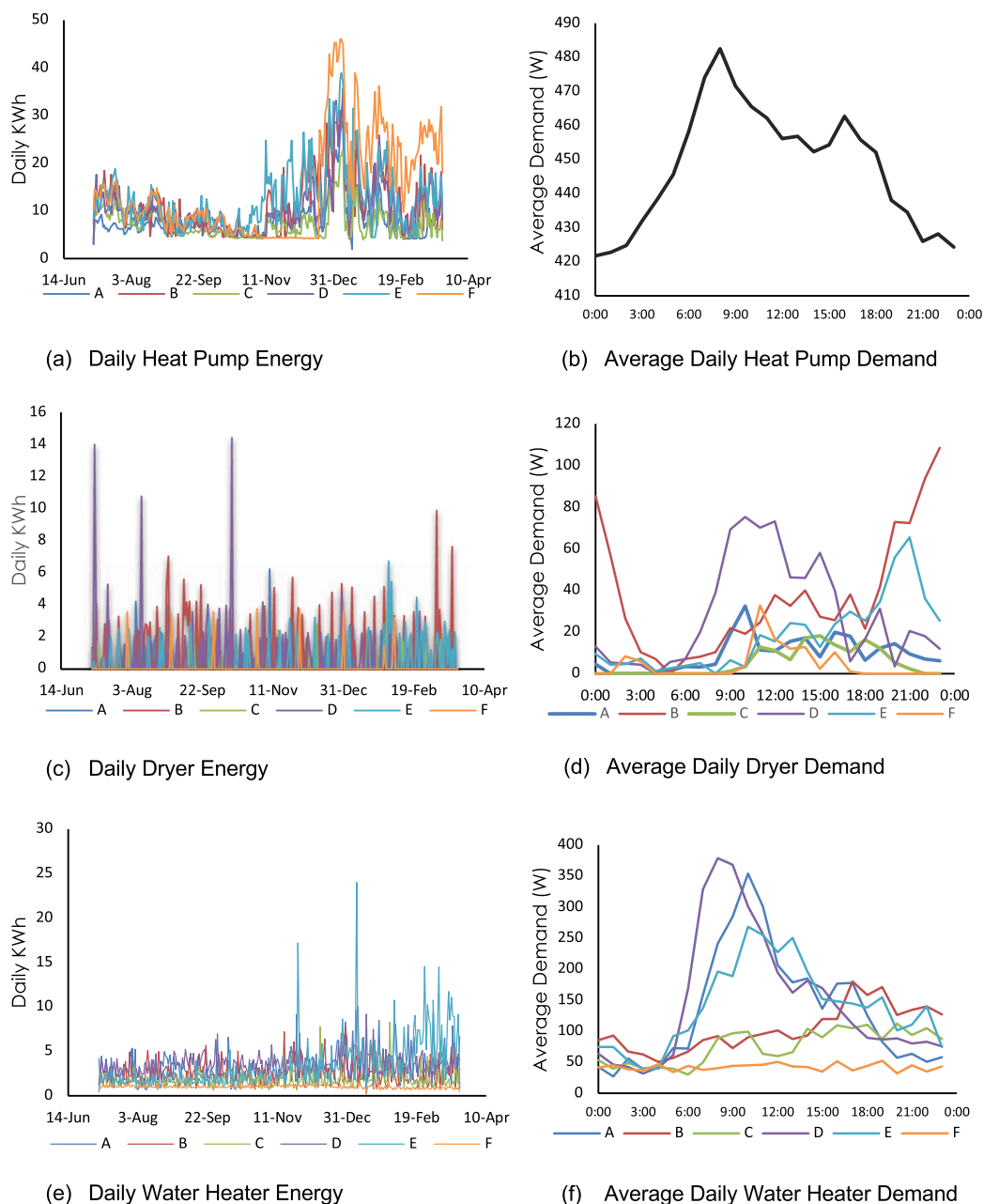


Fig. 2 Energy consumption patterns at appliance level. The variations reflect the changes in occupant behavior in these units that were monitored.

Furthermore, Fig. 2e,f represent the daily energy consumption and demand of water heaters at different time scales. The former presents the daily variations for the water heater energy consumption across the entire period of data collection. Although most of the units show similar load behavior with relatively constant energy

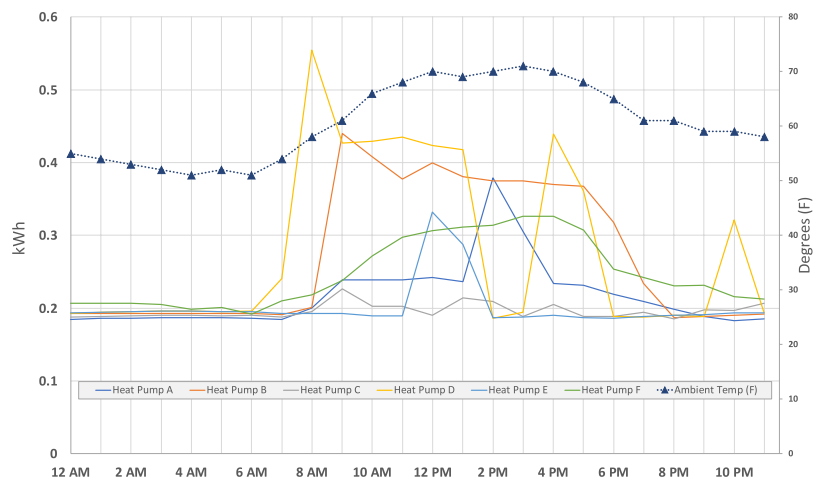


Fig. 3 Relationship between heat pump energy consumption and temperature. Sample NEXI data for the 6 units in the fEECe dataset for a day.

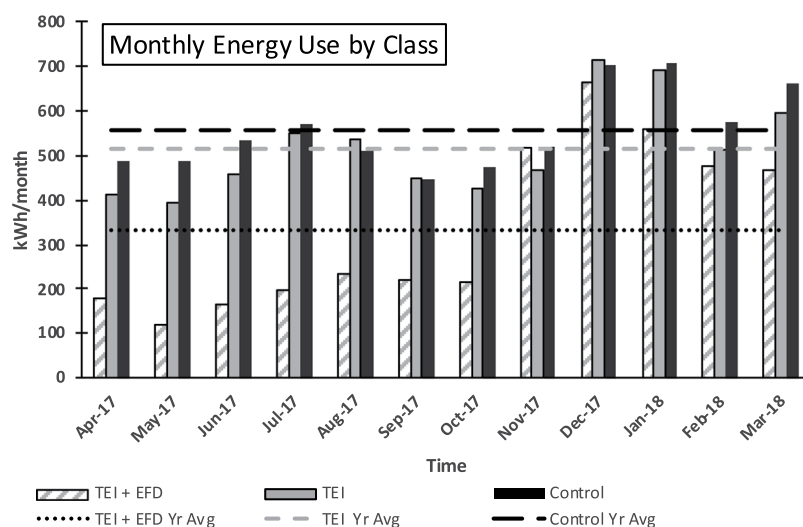


Fig. 4 Monthly average energy use (kWh). Representing complete 38 unit sample with interviewed units, TEI units, and control units. Lines represent a yearly average for each group.

consumption, the occupant behavior impact could be also observed as it has been reflected on the variations of consumption across different days. For unit E the change in consumption behavior could be observed across different months. Although these units are single occupancy units, some units have been occupied by up to two occupants at different times of the data collection period. Figure 2f shows the average daily variation of energy demand for different units. Again, this graph shows the differences in behavior of the occupants in the targeted units and shows reflects the impact of behavioral differences in interacting with loads.

Figure 3 illustrates the relationship between energy use for the heat pump circuit during a cooling day. As expected, the heat pumps begin to draw energy as the temperature begins to rise in the morning and loads drop with temperature decline in the evening. The difference in behaviour can also be observed in the energy use.

Figure 4 illustrates the difference in monthly energy use for the 38 unit sample (from which the final 20 unit survey sample, and 6 unit NEXI sample were selected) over a year. The TEI + EFD units, TEI units, and control units, were compared to analyze the impact of the TEI on energy consumption over a year. The TEI + EFD units average monthly energy consumption was much lower than the TEI units and control group which was expected due to the TEI + EFD units having solar generation offsetting their use.

Usage Notes

The fEECe dataset is focused around a set of time-series datasets which can be analyzed with a variety of software packages. We encourage the use of Python, which was used in our example files via the Jupyter Notebook Platform to create very accessible HTML files. The use of Plotly is also recommended to create interactive visualizations of the time-series datasets. The fleece dataset also provided access to detailed information about the occupants which to our knowledge has not been combined with timeseries data for seniors in affordable energy

efficient housing. The contextual data (perceptions, energy literacy, environmental beliefs), are being prepared for another publication and we encourage users of this dataset to contact the corresponding authors for assistance in leveraging this data. The fEECe dataset will continue to be updated as the longitudinal multiphase study which allowed for the collection of the fEECe dataset is continuing. The authors encourage dialogue from the community to allow for the most useful data to be contributed to the greater scientific community.

Code availability

An example of the Python code used to analyze the raw energy monitor data is publicly available on the Open Science Framework data repository¹⁰. The data can also be analyzed using software which handles tabular timeseries data such as R or MATLAB. The code used to aggregate and plot the data for this paper are publicly available on the data repository on OSF.

Received: 30 May 2019; Accepted: 19 September 2019;

Published online: 26 November 2019

References

- McCoy, A. P. *et al.* *The Impact of Energy Efficient Design and Construction on LIHTC Housing in Virginia*. 102 (A Report by the Virginia Center for Housing Research (VCHR) at Virginia Tech for Housing Virginia 2015).
- McCoy, A., Zhao, D., Paige, F. & Tech, V. *Sustaining Energy Efficiency: Longitudinal Evidence of Virginia's Low-Income Housing Tax Credit Properties*. 33 (A Report by the Virginia Center for Housing Research (VCHR) at Virginia Tech for Housing Virginia 2017).
- U.S. Energy Information Administration. *Electric Power Monthly with data for February 2019*. 266 (EIA, 2019).
- Brehmer, B. Dynamic decision making: Human control of complex systems. *Acta Psychol. (Amst.)* **81**, 211–241 (1992).
- Sterman, J. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. (Irwin/McGraw-Hill, 2000).
- Ehrhardt-Martinez, K., Donnelly, K. A. & Laitner, J. A. Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities. 2010. In Report Number E105 for the *American Council for an Energy-Efficient Economy* 140 (2010).
- Fischer, C. Feedback on household electricity consumption: a tool for saving energy? *Energy Effic* **1**, 79–104 (2008).
- Ouyang, J. & Hokao, K. Energy-saving potential by improving occupants' behavior in urban residential sector in Hangzhou City, China. *Energy Build* **41**, 711–720 (2009).
- Stinson, J., Willis, A., Williamson, J. B., Currie, J. & Smith, R. S. Visualising Energy Use for Smart Homes and Informed Users. *Energy Procedia* **78**, 579–584 (2015).
- Paige, F. & Agee, P. fEECe, an Energy Use and Occupant Behavior Dataset for Net Zero Energy Affordable Senior Residential Buildings. *Open Science Framework*, <https://doi.org/10.17605/OSF.IO/2AX9D> (2019).
- ECMF Team. *EarthCraft Multifamily Technical Guidelines* (2016).
- VHDA. *VHDA Low Income Housing Tax Credit Manual* (2019).
- Nair, G., Gustavsson, L. & Mahapatra, K. Factors influencing energy efficiency investments in existing Swedish residential buildings. *Energy Policy* **38**, 2956–2963 (2010).
- Poortinga, W., Steg, L. & Vlek, C. Values, Environmental Concern, and Environmental Behavior: A Study into Household Energy Use. *Environ. Behav.* **36**, 70–93 (2004).
- Delmas, M. A., Fischlein, M. & Asensio, O. I. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy* **61**, 729–739 (2013).
- Zografakis, N., Menegaki, A. N. & Tsagarakis, K. P. Effective education for energy efficiency. *Energy Policy* **36**, 3226–3232 (2008).
- McMakin, A. H., Malone, E. L. & Lundgren, R. E. Motivating Residents to Conserve Energy without Financial Incentives. *Environ. Behav.* **34**, 848–863 (2002).
- DOE. *Energy Literacy: Essential Principles and Fundamental Concepts for Energy Education* (2012).
- Mahdavi, A. & Taheri, M. An ontology for building monitoring. *J. Build. Perform. Simul* **10**, 499–508 (2017).
- Jazizadeh, F., Afzalan, M., Becerik-Gerber, B. & Soibelman, L. EMBED: A Dataset for Energy Monitoring through Building Electricity Disaggregation. In *Proceedings of the Ninth International Conference on Future Energy Systems - e-Energy 2018*, <https://doi.org/10.1145/3208903.3208939>, 230–235 (ACM Press, 2018).
- Jan DeWaters; *Energy Literacy Survey: A Broad Assessment of Energy-related Knowledge Attitudes and Behaviors* (2009).
- Dunlap, R. E. The New Environmental Paradigm Scale: From Marginality to Worldwide Use. *J. Environ. Educ* **40**, 3–18 (2008).
- Local Climatological Data Station Details: RICHMOND INTERNATIONAL AIRPORT, VA US, WBAN:13740 | Climate Data Online (CDO) | National Climatic Data Center (NCDC). Available at, <https://www.ncdc.noaa.gov/cdo-web/datasets/LCD/stations/WBAN:13740/detail> (Accessed: 15th May 2019).
- Richmond, V. A. History. *Weather Underground* Available at: /history/daily/us/va/richmond/KRIC/date/2019-5-15. (Accessed: 15th May 2019).
- ASHRAE, A. S. of H., Refrigerating and Air-Conditioning Engineers. *2013 ASHRAE Handbook: Fundamentals*. (American Society of Heating, Refrigerating and Air-Conditioning Engineers Inc, 2013).

Acknowledgements

The authors would like to thank the Virginia Center for Housing Research (VCHR), which initiated this study in partnership with Housing Virginia, Virginia Housing Development Authority, and Dominion Energy. Andrew McCoy, Director of VCHR, played an instrumental role in developing the idea for the original study which allowed for the creation of this dataset. The authors would also like to thank and acknowledge Santhosh Thirukkumaran, Emma Coleman, Greg Pfothenauer, Dong Zhao, Byung-Jun Kim and Lata Kodali who assisted in installation of the energy monitoring devices and preliminary analysis of this dataset. Lastly, we would like to acknowledge all of the study participants in this study who cannot be named to protect their personal privacy. This data would not exist without their willingness to share their experience and knowledge with the research team and the world.

Author contributions

F. Paige, as the project PI, Paige coordinated with all involved parties for the study, provided guidance on preparing the dataset, provided supplementary occupant data via the survey instrument such as demographics, environmental perceptions, energy literacy, prepared sections of the manuscript and reviewed all elements of the manuscript. P. Agee, collected and organized all of the energy data for this study, prepared sections of this

manuscript, provided review for co-authored elements of this manuscript. F. Jazizadeh, reviewed the complete manuscript, provided guidance on data representation for an open access audience, developed the technical validation section of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.P. or F.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2019