Data Article

High resolution dataset from a net-zero home that demonstrates zero-carbon living and transportation capacity

Borui Cui\textsuperscript{a}, Sangkeun Lee\textsuperscript{a}, Piljae Im\textsuperscript{a,}\textsuperscript{*}, Michael Koenig\textsuperscript{b}, Mahabir Bhandari\textsuperscript{a}

\textsuperscript{a} Oak Ridge National Laboratory, One Bethel Valley Rd. Oak Ridge, TN 37831, United States
\textsuperscript{b} American Honda Motor Co. Inc., 1919 Torrance Blvd. Torrance, CA 90501, United States

\textbf{A R T I C L E I N F O}

\textit{Article history:}
Received 9 August 2022
Revised 17 October 2022
Accepted 21 October 2022
Available online 26 October 2022

\textbf{Dataset link:} https://bbd.labworks.org/ (Original data)

\textbf{Keywords:}
Smart Home
Energy Efficiency
Sustainable Home
Electrification
Renewable Energy
Ground Source Heat Pump
Net-zero home

\textbf{A B S T R A C T}

This dataset includes high resolution, detailed end use data from a net-zero occupied home that demonstrates zero-carbon living and transportation capacity. The house is located in Davis, California, U.S., and the dataset includes full year data from 2020 with 1 minute time resolution. The data has been monitored with more than 230 sensors installed in the house, and there are total 332 channels available. The data includes detailed end use electricity data (e.g., HVAC system, lighting, plug load including major appliances), building’s interior thermal conditions (e.g., indoor air temperatures in multiple rooms and relative humidity), HVAC system operation data (e.g., soil temperatures around ground bores and supply water temperatures), on-site power generation system data (e.g., PV power supply and PV surface temperatures) and etc. The original dataset from the house has been curated, and the data has been carefully reviewed for quality check. The data quality check revealed there are 156 minutes of data were missing in the month of April, and around 1,404 minutes of data was missing in August. The data gap was filled with linear interpolation in case the gap is less than continuous 6 hours. Otherwise, the data is filled with -9999. The data curation has been processed using

\* Corresponding author.
\textit{E-mail address:} impi@ornl.gov (P. Im).
Social media: @GreenSim (P. Im), @MahabirBhnadari (M. Bhandari)

https://doi.org/10.1016/j.dib.2022.108703
2352-3409/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
the Tsdat framework (https://github.com/tsdat/tsdat). In addition, a semantic description for the dataset was generated by leveraging the Brick (https://brickschema.org/). The final curated and processed data as well as raw data are currently available through https://bbd.labworks.org/ds/bbd/hshus.

© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Specifications Table

| Subject | Energy |
|---------|--------|
| - Renewable Energy, Sustainability and the Environment |

Specific subject area

This dataset includes rich data from an occupied net-zero residential smart home. Corresponding specific subject areas could be

1. Building energy
2. Sustainable energy
3. Energy efficiency and conservation
4. Building optimal operation
5. Building demand response
6. Optimized Smart control and sub-metering

Type of data

The data types include temperature, relative humidity, solar radiation, flowrate, pressure, status, power, and energy consumption. The details are shown in section “Data description”.

How the data were acquired

The data were collected by multiple kinds of sensors and recorded in multiple data loggers. The details of the sensors and loggers are shown in section “Experimental design, materials, and methods”.

Data format

- Raw data and processed data in the CSV (comma-separated values file) format
- Semantic metadata description for the dataset in TTL (Terse RDF Triple Language) format

Description of data collection

This dataset was collected from a fully occupied home with normal operation condition. The dataset includes full-year data for 2020, and the data resolution is 1 minute. The original data was stored at 6-month interval. The target home had been occupied with one family (three occupants) from Jan 1 to Apr 6, 2020. Due to the impact of Covid-19 pandemic, they left the city (Davis, CA) on Apr 6. The home was then un-occupied from Apr 7 to Aug 21, 2020. It was finally occupied with another two residents for the remainder of the year. Efforts had been paid to find short-term occupants to fill the vacant time, but the house was kept unoccupied for several months due to the pandemic.

Data source location

- City/State: Davis, California
- Country: United States
- Latitude and longitude: 38.542540380258735, -121.77106512161058

Data accessibility

Repository name: Benchmark Datasets of Building Environmental Conditions and Occupancy Parameters
Data identification number: 10.17041/1856495
Direct URL to data: https://bbd.labworks.org/ds/bbd/hshus
Contact email: HondaSmartHomeUser@gmail.com
Password: HondaSmartHome1!

Related research article

There is no research article available at this moment.
Value of the Data

This paper contains data from a zero-carbon residential home, called Honda Smart Home, in California, USA. It is a 2-story single family residence with detached garage. Honda Smart Home includes multiple advanced systems, such as adaptive circadian lighting, passive solar design, radiant geothermal heating & cooling, and pre-cooling system.

The building has a living area of 180 m² (1,944 ft²). This house serves a living laboratory for implementing, demonstrating and measuring the impact of various energy efficient technologies and practices for creating net zero homes. The home may deploy a particular technology for a limited period of time for measuring the performance and impact of that technology or combination of technologies.

The high-resolution data (i.e., 1-minute) allows researchers and data users to investigate the performance of individual energy efficient technologies and overall highly efficient and/or net zero energy homes.

The usefulness of such datasets to scientific community includes:

- The data set can establish the ground truth of a building’s operation.
- Correlations in the data can suggest least-cost pathways to accomplish tasks like non-intrusive load monitoring, virtual sensing, building energy model calibration, forecasting, benchmarking, control optimization, fault detection, and many others.
- Processing lower-resolution versions of this data can help us understand the achievable accuracy and precision of derived conclusions when sensor data is not nearly as ubiquitous.
- Researchers, facility managers, sensors and controls platform developers can benefit from this data.

1. Data Description

The Honda Smart Home has diverse measurements, including HVAC system, individual lighting and plug loads, water, renewable power, batteries, plug-in EV etc. Data were recorded on a minutely basis from a range of sensors that monitored electricity usage, equipment performance, indoor conditions, and outdoor conditions.

There are more than 300 measured data channels. Based on the physical features, these data can be categorized into different types, like temperature, flowrate, and power, as listed in Table 1. The examples of data variables are also listed.

For exhibition purpose, the following figure (Fig. 1) shows several processed time-series data of typical summer and winter month (i.e., Aug and Feb) in the year of 2020. \( T_{OA} \) means the measured outdoor dry bulb temperature, \( T_{MB} \) refers to the measured air temperature in the master bedroom, \( T_{Kit} \) means the measured air temperature in the kitchen. It is worth noticing that the data logger did erroneous re-zeroing on Feb 17,2020, so that some variables, like TMB and TKit, were not recorded correctly on that day, as shown in Fig. 1.
Table 1
Data types and example of data variables.

| Data type (unit) | Example of data variables |
|------------------|---------------------------|
| Temperature (degF) | Outdoor air temperature  
Indoor air temperatures  
Ceiling panel supply air temperature  
Cold water supply temperature  
Desuperheater supply/return temperature  
Surface temperature of walls/ceilings/slabs/insulations  
Temperature inside walls  
Heat pump chilled/cooling water temperature  
Dry bore temperature  
Wet bore temperature  
Soil temperature |
| Relative humidity (%) | Outdoor air relative humidity  
Indoor air relative humidity  
Ceiling panel supply air relative humidity  
Air relative humidity in wall cavity |
| Solar radiation (W/m²) | Incident Solar Radiation |
| Flowrate (Gal) | Cold water flowrate  
Heat pump refrigerant flowrates  
Heat pump chilled/cooling water flowrate  
Ceiling panel flowrate  
Domestic hot water flowrate  
Sink flowrate  
Shower flowrate  
Clothes washer water flowrate |
| Pressure (hPa) | Outdoor barometric pressure |
| Status | Fan status  
Status of window actuator  
Lighting status  
Heat pump heating/cooling status  
Ceiling loop valve opening |
| Power (KW or Btu/h) | Plug power  
PV supply power  
Appliance power  
Fan power  
HVAC power  
Water heater power  
Battery supply/sink power  
Grid supply/sink power  
Lighting power  
Delivered heating/cooling energy from heat pump  
Heat pump energy consumption  
Delivered energy from floor  
Delivered energy from ceiling panel |
| Energy consumption (kWh) | HVAC  
Inverter  
Lighting  
Appliance |
Fig. 1. Example data – Outdoor temperature, room temperatures for master bedroom and kitchen.
### Table 2
Sensor list.

| Sensor type                        | Sensor manufactory | Sensor model                             | Number |
|------------------------------------|--------------------|------------------------------------------|--------|
| Thermocouple (Immersion probe)     | Omega              | TQSS-18U-6                                | 13     |
| Thermocouple (Twisted/Shielded)    | Omega              | EEPP-T-20-TWSH-SLE                        | 75     |
| Thermistor                         | Omega              |                                          | 33     |
| Power meter                        | Schneider Electric |                                          | 44     |
| Flow meter                          |                    |                                          |        |
| Omega                              | FTB431             |                                          | 20     |
| Onicon                             | F1300-1” brass     |                                          | 5      |
| Dwyer                              | WMT2-A-C-03        |                                          | 2      |
| RH                                 | Setra              | SRH1-2P-W-2C-T5-N/ SRH1-2P-D-2C-T0-N     | 6      |
| RTD                                | Setra              | SRH1-2P-W-2C-T5-N/ 2651R25WD2BT1C        | 5      |
| Pyranometer                         | Licor              | Li 200Sz                                 | 1      |
| Anemometer                          | Vaisala            |                                          | 2      |

### Table 3
Logger list.

| Logger manufactory | Logger model            | Number |
|--------------------|-------------------------|--------|
| dataTaker          | DT85 Series             | 74     |
|                    | CEM20 #1                | 40     |
|                    | CEM20 #2                | 36     |
|                    | CEM20 #3                | 40     |
| Vaisala            | WXT 520                 | 7      |
| Schneider Electric | Powerlink MVP BCPMA    | 43     |
| Dent               | PoweScout 3             | 9      |
| Cantara            | AMX                     | 93     |
| Advantech          | ADAM 4150               | 21     |

### 2. Experimental Design, Materials and Methods

#### 2.1. Data collection

All the measured data in this target home were collected by multiple kinds of sensors, as listed in Table 2. The sensors were connected with multiple data loggers and all the measured data were then transferred to the data loggers, as listed in Table 3.

One whole year measured data (from Jan 1, 2020 to Dec 31, 2020) recorded in 12 monthly data files were downloaded from its official website and formed as raw data. All the systems were with normal operation conditions and the data resolution is 1-minute.

#### 2.2. Data processing

The raw data were then checked, curated, and integrated into one single whole-year data file (called “honda_latest_curated_20220322.csv”) and presented in this paper.
A complete list of variables (channels) in the dataset is listed in the raw data files. This variable list introduces the description for each variable (channel), including the subsystem, measurement location, and measured parameters. The missing data in the raw data files were filled by using a programmatical approach. The approach that handles the raw data is introduced as following:

1. We used the Tsd (https://tsdat.readthedocs.io/en/latest/), an open-source Python framework for processing and standardizing time-series data [1]. We created a data pipeline configuration for Tsd and executed the pipeline with the raw data. The execution of the data pipeline generates a curated data.
2. For all the columns, linear interpolation of missing data gaps was done for the data gap with the gap of less than continuous 6 hours. All the rows that contain linearly interpolated values were also marked so that the data user can track if the value is either actually coming from the sensor or linearly interpolated by data curator.
3. For the data gaps longer than continuous 6 hours, the value of -9999 was filled.

We also developed a Brick model, which is a semantic metadata description for the dataset [2]. Complex relationships across locations, sensors, equipment, and data labels are represented in a graph using standard terms and concepts provided by the Brick schema (https://brickschema.org/). The representation is saved as an RDF TTL (Terse Triple Language, pronounced Turtle) formatted file. Fig. 2 shows a part of the developed Brick model and a visualization of the developed model.

**Ethics Statements**

As part of the Occupancy Agreement, all occupants expressly agree and understand that the house is monitored and that the data is recorded, and that the house uses prototype systems for HVAC, Lighting, and Energy Management.
CRediT Author Statement

Borui Cui: Data curation, Writing – original draft preparation; Sangkeun Lee: Data processing, Tsdat pipeline configuration, Brick ontology development; Michael Koenig: Original data monitoring; Piljae Im: Funding acquire, supervision, Writing – original draft preparation; Mahabir Bhandari: Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

https://bbd.labworks.org/ (Original data).

References

[1] Carina Lansing, et al., Tsdat: An Open-Source Data Standardization Framework for Marine Energy and Beyond, OCEANS 2021: San Diego–Porto, IEEE, 2021.
[2] Bharathan Balaji, et al., Brick: Towards a unified metadata schema for buildings, in: Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments, 2016.