



OPEN ACCESS

EDITED BY

Thomas Alan Adams,
Norwegian University of Science and
Technology, Norway

REVIEWED BY

Zhihong Pang,
Louisiana State University, United States
Valentina Zaccaria,
Mälardalen University, Sweden
Elisa Marrasso,
University of Sannio, Italy

*CORRESPONDENCE

Gustavo Felipe Martin Nascimento,
✉ gustavofmn@hotmail.com

RECEIVED 13 September 2022

ACCEPTED 21 July 2023

PUBLISHED 09 August 2023

CITATION

Martin Nascimento GF, Wurtz F,
Kuo-Peng P, Delinchant B,
Jhoe Batistela N and Laranjeira T
(2023), GreEn-ER—Electricity consumption
data of a tertiary building.
Front. Sustain. Cities 5:1043657.
doi: 10.3389/frsc.2023.1043657

COPYRIGHT

© 2023 Martin Nascimento, Wurtz, Kuo-Peng, Delinchant, Jhoe Batistela and Laranjeira. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GreEn-ER—Electricity consumption data of a tertiary building

Gustavo Felipe Martin Nascimento^{1,2*}, Frédéric Wurtz¹, Patrick Kuo-Peng², Benoit Delinchant¹, Nelson Jhoe Batistela² and Tiansi Laranjeira¹

¹Univ. Grenoble Alpes, CNRS, Grenoble INP, G2Elab, Grenoble, France,

²GRUCAD, Universidade Federal de Santa Catarina, Florianópolis, Brazil

The increased use of intermittent renewable energy sources makes the use of machine learning methods combined with demand-side management more and more frequent. Machine learning algorithms rely on data to identify patterns and learn insights. Hence, data availability is of utmost importance, and the more, the merrier. Therefore, this data report aims to present a dataset concerning the electricity consumption of a tertiary building located in the French Alps region (Grenoble) in 2017 and 2018. It is a massively monitored and controlled building with about 330 electricity meters, whose measurement data constitute the dataset. The data were collected directly from the building management system and correspond to raw data, without any pre-treatment. The dataset also includes Python notebooks that allow for understanding the system design, navigating the data, and performing some simple analyses. This is a publicly available dataset that tries to fill the gap of the availability of electricity consumption data, especially regarding tertiary buildings.

KEYWORDS

training data, electricity consumption dataset, tertiary buildings, buildings' electricity consumption, data report

1 Introduction

The energy consumed in buildings accounts for a significant share of global energy consumption, indicating that they play a central role in the energy transition. In France, for instance, approximately 67.8% of electricity is consumed in buildings, both residential and tertiary ([Réseau de Transport d'électricité, 2019](https://www.reseau-energie.fr/)).

The increased use of intermittent renewable energy sources, such as solar and wind, makes the use of machine learning methods combined with demand-side management more

Abbreviations: AMPds, the Almanac of Minutely Power Dataset; AREE, automatic reporting for energy efficiency; BMS, building management system; COM, common areas; CSV, comma-separated values; ECL, éclairage—French word for lightning; ECS, eau chaude sanitaire—French for domestic hot water; EE3, Ense³; FM, forces motrices—French for driving forces used for dedicated outlets; G2Elab, Grenoble Génie Electrique laboratory; GreEn-ER, Grenoble Energie—Enseignement et Recherche; iAWE, Indian Dataset for Ambient Water and Energy; NILM, non-intrusive load monitoring; Nzeb, nearly-zero energy building; PC, prises de courant—French for current electrical outlets; PLC, programmable logic controller; REDD, Reference Energy Disaggregation Dataset; SQL, Structured Query Language; TD, tableau de distribution—French for distribution switchboards; TGBT, tableau général de basse tension—French for general low voltage switchboard; UK-DALE, United Kingdom Domestic Appliance-Level Electricity.

and more frequent. For example, anomaly detection using machine learning techniques can help identify the unusual electricity consumption of assets and detect equipment faults. For instance, Zhou et al. (2021) focused on identifying anomalies in the energy consumption of central air-conditioning systems. Gaur et al. (2019) and Himeur et al. (2021) sought to detect unusual energy consumption in buildings, while Lee used artificial intelligence to reduce the maintenance cost of chillers (Lee et al., 2021). In addition, studies show that it is possible to reduce electricity consumption more effectively when information of individual appliance consumption is available, if compared to simple monthly bills (Wood and Newborough, 2003). Therefore, disaggregated consumption data measurement and processing are of great importance. Information about individual appliance consumption can be obtained through comprehensive monitoring, when most loads are measured individually, or through machine learning methods that use energy disaggregation techniques. The latter type of monitoring is known as non-intrusive load monitoring (NILM) (Hart, 1992). Most of the research related to NILM methods is based on supervised learning techniques, whose algorithms require well-identified data (Zoha et al., 2012). It is therefore important to have real data from electricity consumption measurements so the research in this field can move forward.

Several buildings' datasets of electricity consumption are available. However, most of them relate to the residential environment. It can be cited as examples the Reference Energy Disaggregation Dataset (REDD) (Kolter and Johnson, 2011), the Almanac of Minutely Power Dataset (AMPDs) (Makonin et al., 2013), the Indian Dataset for Ambient Water and Energy (iAWE) (Batra et al., 2013), the United Kingdom Domestic Appliance-Level Electricity (UK-DALE) dataset (Kelly and Knottenbelt, 2015), and the DataPort dataset (Parson et al., 2015). These datasets are largely used as training and testing data for several machine learning algorithms related to the electricity consumption in buildings. The existence of so many datasets regarding the residential sector allowed the research to go further in this type of building.

Because of the nature of the loads and even the difference between the behavior of the buildings' occupants when residential and tertiary buildings are compared, the need for the availability of datasets regarding tertiary buildings emerges. The scarcity of datasets on electricity consumption in tertiary buildings decreases the reproducibility of research in the area. Therefore, this paper aims to provide a dataset with electricity consumption data of a tertiary building located in the region of the French Alps. The papers associated with this dataset are licensed under a Creative Commons Attribution 4.0 International License.

The first section presents the building, its general information, and its main loads. It also presents the meters and the building management system (BMS) from where the data were collected. This section also details the structure of the dataset, its metadata, etc. Section 3 presents some insights over the data, as the building main loads, annual consumption and reports some data quality problems.

2 Data collection and structure

The process of collecting the data and assembling the dataset starts with the knowledge of the building structure. It is vital to know

how and which loads are measured and how the meter hierarchy is organized. The accessibility to the data is also important. Whether it is from a measurement campaign or extracted from meters already installed, the data should be easily accessible to facilitate analyses. Finally, the assembly of a dataset is what makes the data usable, even to those who do not know the facility. The data structure, the metadata, and even examples help people in the data analysis. This section details the building structure, how the data were collected, and the dataset structure.

2.1 GreEn-ER building

The GreEn-ER building is in the Polygone Scientifique, located at the Presqu'île of Grenoble, France. It comprises the Grenoble-INP engineering school Ense³, the G2Elab laboratory, and training and research platforms. The building has more than 22,000 m² of floor space, which is divided over six floors and the roof. There are about 1,500 students and hundreds of professors, researchers, and staff using it. There are more than 1,500 meters, including more than 300 electricity consumption ones. The other meters concern internal and external conditions and thermic energy data among other variables. The measured data are used to control the internal conditions regarding the comfort of the occupants and to monitor the consumption (Delinchant et al., 2016).

The building houses a diverse mix of loads from common office ones such as personal computers, monitors, and printers, to typical industrial loads such as air compressors, fans, and pumps. There are also air handling units, a data center, and a university restaurant within the facility.

Since the building houses a teaching facility and does not have student entry control, it is not trivial to have a precise occupancy schedule. However, it is possible to state that the occupancy is concentrated on weekday daytime periods, which is reflected on the electricity consumption.

2.1.1 Electric scheme

The grid delivers the electricity to the building at three-phase 20 kV. Two 2-MVA transformers (TR1 and TR2) step down the voltage at 400 V. Each transformer leads the electricity to a main switchboard, called TGBT, French acronym to general low-voltage switchboards. Each one of these boards has its own meter to measure its consumption. A switch that is normally open interconnects these boards. In that way, the two TGBTs are normally independent. Thus, all the building's loads are connected to these two main switchboards, either directly or by some sub boards. Each one of the branches of the scheme has also its own meter. Figure 1 illustrates the electric scheme of the building.

In Figure 1, the branches that have TD in their names are, in fact, other boards ("Tableau de distribution" in French) that distribute the electricity to different zones. The third character in these branches' names, 1 or 2, stands for the TGBT to which the board is connected. The G2E stands for G2Elab and represents the boards that distribute electricity to that area. At the same time, EE3 stands for Ense³, and those boards distribute electricity to the classrooms and other facilities of the Engineering School Ense³. COM stands for the common areas, and PRE represents the boards that distribute the electricity to PREDIS charges, a training and research platform for

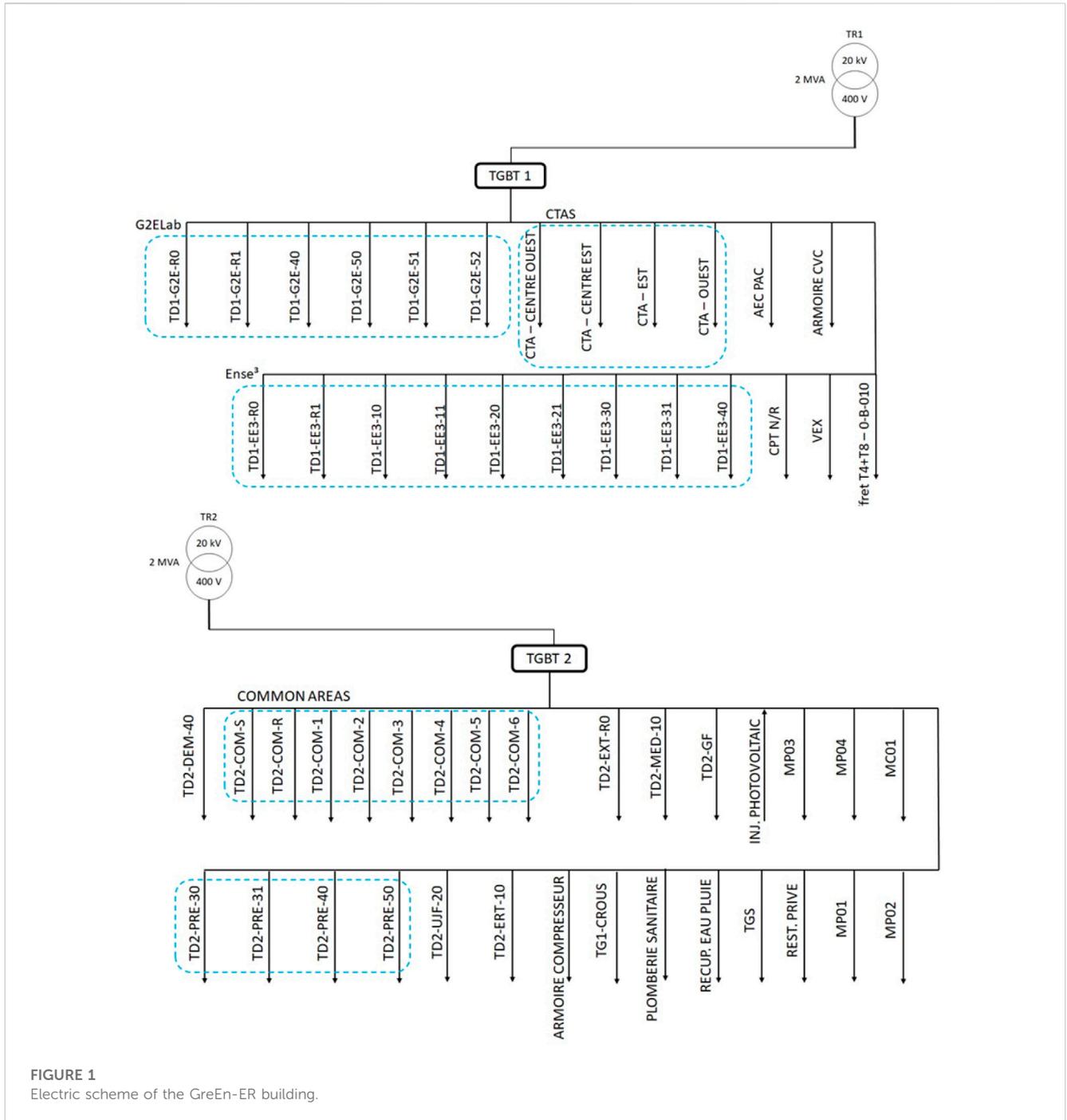


FIGURE 1 Electric scheme of the GreEn-ER building.

smart grids. The name’s ninth character (considering the hyphens “-”) of each board is linked to the floor where it is located [R stands for ground floor (Rez-de-Chaussée in French)]. The loads in the TD switchboards are generally divided into three or four loads, which represent lightning (ECL), electrical outlets (PC), water heater (ECS), or dedicated outlets (FM). These loads also have their own meters.

Within the building, there is a platform, called “PREDIS-MHI,” conceived to be a nearly zero-energy building (Nzeb). It is a 600-m² platform energetically independent from the rest of the building. This platform, represented in the early drawing by the branch with

the acronym “TD2-DEM-40,” is even more monitored than the rest of the building. In this sector, the lightning and the outlets of each room is measured independently (Delinchant et al., 2016).

2.1.2 Meters and building management system

The electricity consumption of the building is measured using Socomec meters models such as E13, E23, E33, E43, E63, I30, I35, and I60, according to their specifications (SOCOMECE–Innovative Power Solution, 2022; SOCOMECE–Innovative Power Solution). Each meter has Modbus communication via RS485 with a programmable logic controller (PLC) installed in the switchboard

to which the measured load is connected. The PLCs, in turn, send the measurements to the storage and to the BMS.

The BMS is based on the StruxureWare environment, developed by Schneider Electric Company (Schneider Electric). It gathers the data coming from the PLCs and stores the measured data into a Structured Query Language (SQL) server, where the data are logged. It also enables the control of some parameters, such as the internal temperature of some rooms and the air pressure and flow of the air handling units. Energy consumption management software, called Automatic Reporting for Energy Efficiency (AREE) building, developed by Inneasoftware (Inneasoftware), organizes the meter hierarchy and the trends and can show several performance indicators. It also enables the access to the logged data and can easily export the data into text files. The data available in this dataset are extracted from the SQL server with the help of AREE building software.

2.2 Dataset structure

The dataset was separated into four main contents: global consumption, TGBT1, TGBT2, and PREDIS-MHI. In the dataset main folder, there is a folder named “Data.” Inside this folder, two sub-folders represent each year of data available, which are 2017 and 2018. Each sub-folder contains three other sub-folders, and each one corresponds to content cited earlier.

Inside the sub-folders of each content, there are files that contain the electricity consumption data. The data are stored in comma-separated value (CSV) files, with semicolon as a separator. Each file contains the timestamp, with 10-min sampling and the cumulative electricity consumption, in kWh. As the meters are cumulative and the resolution is 1 kWh, the consumption sample will only increase after 1 kWh of consumption of the respective load. There is one CSV file for each meter, and they are all named according to their respective meter numbers. These numbers can be retrieved in tables and drawings available in Jupyter Notebook.

Four Jupyter Notebook files, a format that allows combining text, graphics, and code in Python, are also available in the main folder. These files allow exploring all the data within the dataset. These Jupyter Notebook files also contain the metadata necessary for understanding the system, such as drawings of the system design and of the building. A couple of CSV files with the system design are also available. They are named “TGBT1_n.csv,” “TGBT2_n.csv,” and “PREDIS-MHI_n.csv.” In these files, each column stands for a switchboard. The head contains the names of the boards, and the value in the first row represents the respective meter number. The values in the following rows represent the number of the sub-meters that are located downstream of the meter described in the first row. So, for example, in the file “TGBT1_n.csv,” there is a column whose heading is “TD1-G2E-51.” The value in the first row is “776,” which represents the number of the meter of this switchboard. The values located in the following rows, “517,” “518,” and “519,” represent the meters of the loads located downstream of the “TD1-G2E-51” switchboard.

The folder “Data” also contains the CSV files with the electricity consumption data of the whole building and a file named “Temp.csv” with the temperature data. The temperature data are in Celsius degrees °C, with 1 hour sampling, a semicolon as a separator, and comma as a decimal marker.

TABLE 1 Time zone changes within the data.

Timestamp interval	Time zone
01/01/2017 00:00–26/03/2017 01:50	UTC+1
26/03/2017 03:00–29/10/2017 01:50	UTC+2
29/10/2017 02:00–25/03/2018 01:50	UTC+1
25/03/2018 03:00–28/10/2018 01:50	UTC+2
28/10/2018 02:00–31/12/2018 23:50	UTC+1

The timestamp follows the format “dd/MM/yyyy HH:mm,” as shown in “29/10/2017 03:40:00,” for instance. The timestamp data observe the local time in France, i.e., UTC+1 or UTC+2, considering daylight saving time periods. The raw data are not time zone aware; thus, when there is a step forward, from UTC+1 to UTC+2, in March, there is a gap in the data. However, when there is a step back, in October, the data regarding the duplicated timestamps are overwritten. Table 1 presents the timestamps when changes in the time zone occur.

The GreEn-ER dataset is available in open-access in the Mendeley Data repository under the DOI “10.17632/h8mmnthn5w.1,” and under a Creative Commons Attribution 4.0 International License (Martin Nascimento et al., 2020).

3 Annual consumption and data quality

The measurements available in this dataset correspond to raw data without any preprocessing. This may present some data quality problems. The measuring and storing data system were designed to evaluate the electricity consumption rather than the load curve. Therefore, during a substantial period, the final electricity consumption tends to be accurate. However, sometimes, a loss of communication between the meters and the data storage system may occur. In those periods, the meter accumulates the measured electricity consumption, and when the communication resumes, it sends the accumulated value. Occasionally, it represents a high peak of consumption in a short time step. The loss of communication addresses the lack of completeness of the data and the peaks in the consumption and the poor accuracy of the measurements in some samples. Figure 2A shows the annual consumption of the two transformers of the building, while Figure 2B presents some examples of data quality problems in this dataset.

According to the results presented in Figure 2A, the annual consumption of the building in 2017 was 1791.67 MWh, which corresponds to an average power of 204.53 kW.

Another way to evaluate the electricity consumption is by analyzing the load curve, which can be reconstructed from the curves presented in Figure 2A and shown in Figure 3A. In these figures, it is possible to visualize an increase in the power consumption during the summer, probably due to cooling charges. Periods of scholar vacations, especially in August, are also distinguishable. Additionally, it is possible to see that there is a weekly pattern, with consumption dropping during weekends and likewise during nighttime.

Furthermore, it is also possible to visualize data quality problems, as shown in Figure 3A, by the presence of outliers at the beginning of the

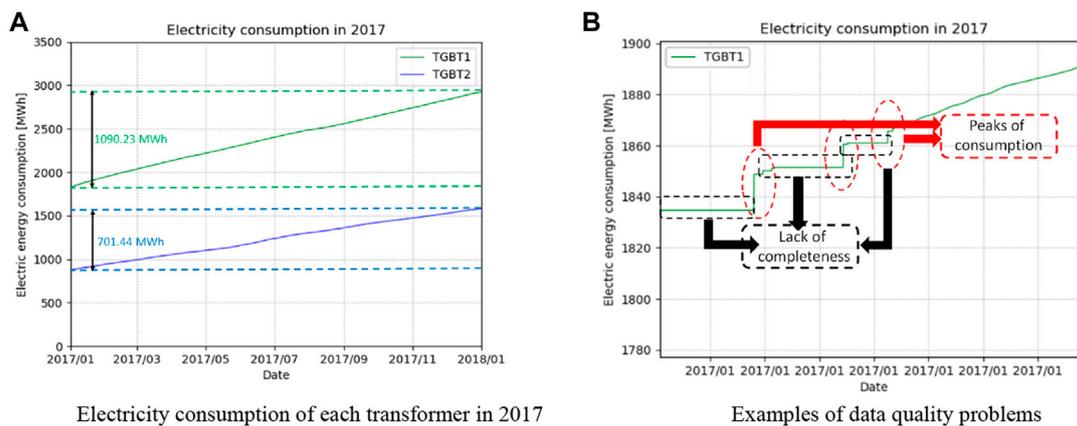


FIGURE 2 Raw electricity consumption data from the two main switchboard of the building.

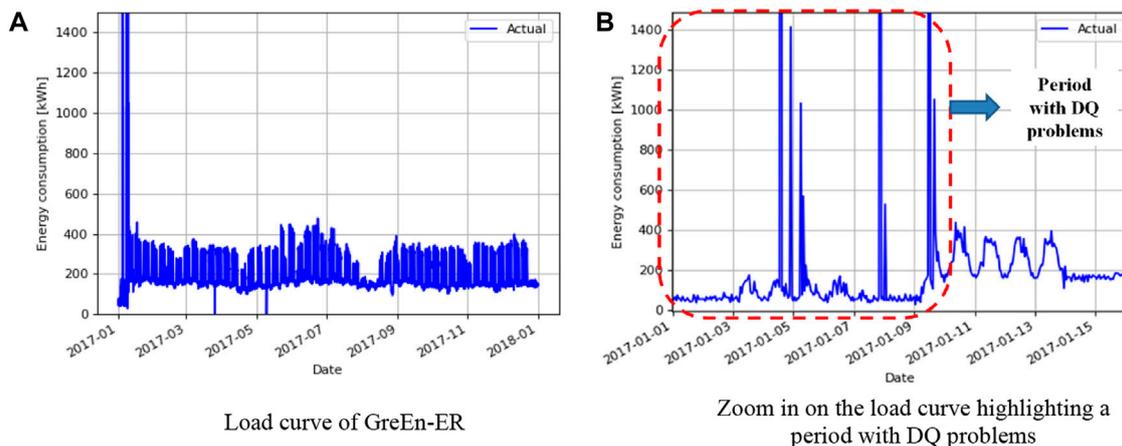


FIGURE 3 Load curve of the GreEn-ER building.

series, in January. That is another way to visualize the issues presented in Figure 2B and can be better seen in Figure 3B. These data quality problems due to the lack of communication between the meters and the BMS are not critical when dealing with consumption only, since when communication is restored, the meter sends the integrated value to the BMS. However, when there is an interest in reconstructing the load curve, those problems generate outliers and gaps in the time series, making it difficult to analyze these periods.

Therefore, these data quality issues need to be treated carefully to allow for more precise analysis. They can be present in every file that contains consumption data within the dataset. Hence, data quality problems lead to another research direction, such as the detection and quantification of these issues as well as their impact on different analyses.

Although the dataset presented in this paper represents a step forward in the availability of tertiary buildings datasets, it is also important to highlight limitations other than the data quality issues previously mentioned. The low-frequency sampling may make several analyses difficult; however, high frequency sampling data is not the

standard for most facilities. Additionally, the 1 kWh resolution for the energy meters impairs the timeliness, especially for low-consumption loads. Therefore, it is important to evaluate whether these limitations restrict the use of the data for the intended purpose.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://data.mendeley.com/datasets/h8mmnthn5w/1>.

Author contributions

Conceptualization: GM, PK-P, and FW; methodology: GM, PK-P, and FW; validation: GM; data curation: GM and TL; writing—original

draft preparation: GM; writing—review and editing: PK-P and FW; supervision: FW, PK-P, NJ, and BD; project administration: FW and PK-P; and funding acquisition: FW.

Université Grenoble Alpes (UGA), and the Federal University of Santa Catarina (UFSC) for supporting this research.

Funding

This work has been mainly supported by the Carnot Énergies du Futur Institute, by the project ORCEE. It has also been partially supported by the ANR (Agence Nationale de la Recherche) project eco-SESA (<https://ecosesa.univ-grenoble-alpes.fr/>) and Observatory of Transition of Energy–OTE (<https://ote.univ-grenoble-alpes.fr/>) ANR-15-IDEX-02.

Acknowledgments

The authors would also like to thank the French Centre National de la Recherche Scientifique (CNRS), the Grenoble-INP, the

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Batra, N., Gulati, M., Singh, A., and Srivastava, M. B. (2013). "It's Different: Insights into home energy consumption in India," in Proceedings of the Fifth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (ACM BuildSys), Roma, Italy, November 2013.
- Delinchant, B., Wurtz, F., Ploix, S., Schanen, J., and Marechal, Y. (2016). "GreEn-ER living lab: A green building with energy aware occupants," in 2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS), Rome, Italy, April 2016, 1–8.
- Gaur, M., Makonin, S., Bajic, I. V., and Majumdar, A. (2019). Performance evaluation of techniques for identifying abnormal energy consumption in buildings. *IEEE Access* 7, 62721–62733. doi:10.1109/access.2019.2915641
- Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proc. IEEE* 80 (12), 1870–1891. doi:10.1109/5.192069
- Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., and Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Appl. Energy* 287, 116601. ISSN 0306-2619. doi:10.1016/j.apenergy.2021.116601
- Inneasoftware building. Available at: <https://inneasoftware.com/en/aree-building-energy-efficiency/>.
- Kelly, J., and Knottenbelt, W. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* 2, 150007. doi:10.1038/sdata.2015.7
- Kolter, J. Z., and Johnson, M. J. (2011). "Redd: A public data set for energy disaggregation research," in Proceedings of the SustKDD Workshop on Data Mining Applications in Sustainability, San Diego, United States, August 2011, 1–6.
- Lee, D., Lai, C., Liao, K., and Chang, J. (2021). Artificial intelligence assisted false alarm detection and diagnosis system development for reducing maintenance cost of chillers at the data centre. *J. Build. Eng.* 36, 102110. ISSN 2352-7102. doi:10.1016/j.job.2020.102110
- Makonin, S., Popowich, F., Bartram, L., Gill, B., and Bajic, I. V. (2013). "AMPds: A public dataset for load disaggregation and eco-feedback research," in Electrical Power and Energy Conference (EPEC), Nova Scotia, Canada, August 2013, 1–6.
- Martin Nascimento, G. F., Delinchant, B., Wurtz, F., Kuo-Peng, P., Jhoé Batistela, N., and Laranjeira, T. (2020). GreEn-ER - electricity consumption data of a tertiary building. *Mendeley Data* 6. doi:10.17632/h8mmthn5w.1
- Parson, O., Fisher, G., Hersey, A., Batra, N., Kelly, J., Singh, A., et al. (2015). "Dataport and nilmtk: A building data set designed for non-intrusive load monitoring," in 2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP, Orlando, United States, December 2015, 210–214.
- Réseau de Transport d'électricité (2019). Bilan électrique 2018. [Online] Available: https://www.rte-france.com/sites/default/files/be_pdf_2018v3.pdf.
- Schneider Electric Power monitoring & control software. Available at: <https://www.se.com/us/en/product-subcategory/4170-power-monitoring-control-software/?filter=business-4-low-voltage-products-and-systems>.
- Socomec – Innovative Power Solution AC current measurement module. Multi-circuit Power Metering and Monitoring DIRIS Digiware I - measure and monitor at the closest point to the loads. Available at: <https://www.socomec.us/range/diris-digiware-i/>.
- Socomec – Innovative Power Solution (2022). Single-circuit metering, measurement & analysis. Available at: https://www.socomec.com/single-circuit-energy-meter_en.html.
- Wood, G., and Newborough, M. (2003). Dynamic energy-consumption indicators for domestic appliances: Environment, behaviour and design. *Energy Build.* 35 (8), 821–841. doi:10.1016/s0378-7788(02)00241-4
- Zhou, X., Yang, T., Liang, L., Zi, X., Yan, J., and Pan, D. (2021). Anomaly detection method of daily energy consumption patterns for central air conditioning systems. *J. Build. Eng.* 38, 102179. ISSN 2352-7102. doi:10.1016/j.job.2021.102179
- Zoha, A., Gluhak, A., Imran, M., and Rajasegarar, S. (2012). Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors* 12, 16838–16866. doi:10.3390/s121216838