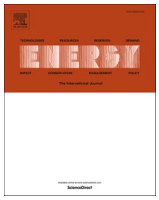
Energy 316 (2025) 134381

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy



Exploring energy flexibility from smart electrical water heaters to improve electrification benefits in residential buildings: Findings from real-world operation

Rui Amaral Lopes [a,*], Francisco Silva [a], Rafael Menezes-Barros [a], Nuno Amaro [a], Ana Gonçalves de Carvalho [b], João Martins [a]

[a] School of Science and Technology, UNINOVA-CTS and LASI, NOVA University Lisbon, Portugal
[b] CNET - Centre For New Energy Technologies, S.A, Portugal

ARTICLE INFO

Keywords:
Smart electric water heater
Internet-of-Things
Forecasting
Energy flexibility
Electrification

ABSTRACT

The integration of renewable energy sources, which often have limited dispatchability, along with the rising electrification of energy demand, is increasing the complexity of managing power systems. To address this challenge, smart solutions that leverage energy flexibility in buildings can support power system operations by adapting energy consumption according to existing needs. This study presents a forecasting and IoT-based solution for electric water heaters designed to improve energy flexibility, lower electricity costs, and enhance self-consumption of locally generated energy. Deployed in a real-world setting on Terceira Island, Portugal, the solution successfully replaced a gas boiler, achieving a 56 % decrease in electricity costs and a 34 % reduction in CO_2 emissions (compared to a scenario with a 75 % efficiency gas boiler). These benefits were realized while maintaining acceptable levels of comfort and user engagement.

1. Introduction

Power grids have traditionally relied on large, centralized fossil fuel-based plants, pumping energy through a one-direction system. This traditional approach is undergoing a significant transformation, as distributed renewable energy sources, like solar and wind, are rapidly gaining ground. The expansion of renewable energy presents a significant opportunity to combat climate change and create a more sustainable future. It is a transformation driven by both innovation and policy, paving the way for a cleaner and more diversified energy system [1].

While the benefits of integrating energy conversion systems based on renewable energy resources are evident, the time-varying and demand-independent nature of these power sources increases the complexity of power grids management. This calls for innovative solutions to explore energy flexibility and therefore increase matching between demand and supply [2]. Buildings, which account for 34 % of global energy demand and 37 % of energy and process-related CO_2 emissions [3], can help manage this complexity by leveraging energy flexibility through demand response measures [4].

As reported in the literature [5], exploring energy flexibility from buildings requires local or remote control of electrical and thermal loads whose operation can be modified with limited to no impact on users' comfort. Examples of real-world energy flexibility harvesting range from building-level control of space-heating loads to reduce energy consumption during fixed peak hours [6] to large-scale management of domestic electric vehicle charging processes to support grids operation [7]. Among different candidates, the study reported in this article focuses on the energy flexibility of Electric Water Heaters (EWH) in real-world conditions. This choice results from i) the thermal storage capabilities of these devices [8], allowing decoupling of water and electric energy consumption profiles; ii) the relatively simple and low-cost smartification of existing EWHs to allow remote control using Internet-of-Things (IoT) solutions [9]; and iii) the possibility to replace existing gas boilers in order to reduce CO_2 emissions and decarbonize water heating energy consumption [10]. Additionally, as water heating represents almost 25 % of household energy use in developed countries [11], this study contributes to future harvesting of a non-negligible energy flexibility market.

This article is part of a special issue entitled: SpliTech2024 published in Energy.

* Corresponding author.
E-mail address: rm.lopes@fct.unl.pt (R.A. Lopes).

https://doi.org/10.1016/j.energy.2025.134381
Received 14 August 2024; Received in revised form 2 January 2025; Accepted 3 January 2025
Available online 4 January 2025
0360-5442/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Several works have addressed the usage of energy flexibility provided by electric water heaters. At building-level, examples include the solution reported in Ref. [12], which resulted in cost savings by reducing heating operations during expensive periods while considering historical energy consumption patterns to avoid user discomfort, or the one described in Ref. [13], where improvement of PV self-consumption using energy storage capabilities of EWHs is also taken into consideration. Studies addressing the aggregated use of EWHs are mostly developed in simulation environment and the tackled challenges include peak shifting and frequency response services [14] or the characterization of available flexibility and rebound effects that might result from demand response activations [15]. However, a real-world implementation focusing on i) reducing electricity costs when Time-of-Use (ToU) tariffs are applied and ii) improving PV self-consumption in real-world operation, using energy consumption (electric and thermal) and generation time-series forecasting, while also addressing benefits of water heating electrification (costs and CO_2 emissions), is still missing in the literature. This study addresses this research gap by proposing a solution (hardware and software) that can be applied to EWHs to explore existing flexibility in real-world operation using forecasting methods. It includes a hybrid deep learning architecture composed of a Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and a fully connected layer, to predict not only water/thermal energy demand, as in Ref. [16], but also PV surplus to be stored in the respective storage tank and therefore reducing energy import from the distribution grid. The referred solution has been deployed on Terceira Island, Portugal, under the context of IANOS H2020 project [17]. Given the geographic context of this case study, using PV surplus to satisfy part of water heating energy demand is of special importance. This can be verified in scenarios where gas boilers are replaced by electric solutions, avoiding gas and naval transportation emissions, or when the operation of already installed EWHs is improved to reduce energy import from the grid, which commonly relies heavily on fossil fuels [18].

The main contributions of this study can be outlined as follows.

- The development of an IoT-based solution compatible with new and existing electric water heaters with mechanical thermostats, enabling the conversion of passive EWHs into smart energy systems with real-time and predictive control capabilities.
- The design of an EWH control methodology which relies on the forecasting of electricity consumption and generation based on a CNN-LSTM architecture, as well as a day-ahead prediction of hot water demand.
- The validation of the proposed methodology in a real-world environment, moving beyond simulations and limited scopes present in many studies. The proposed solution achieved significant reductions in both costs and CO_2 emissions.

The remaining of this article is structured as follows. Section 2 describes the methodology followed in this study, including the proposed solution. Section 3 will describe the IoT usage in a real household, presenting results that show a significant increase in PV self-consumption and cost savings due to the smart control of the EWH enabled by the IoT system. Lastly, Section 4 presents the conclusion and future works left open by this study.

2. Methodology

This section describes the IoT based solution proposed to explore the energy flexibility provided by EWHs to increase PV self-consumption and reduce electricity costs (Subsection 2.1), detailing in Subsection 2.2 the developed forecasting methods. Additionally, the power grid where the proposed solution is currently deployed is presented in Subsection 2.3 and the metrics used to assess the achieved performance are described in Subsection 2.4.

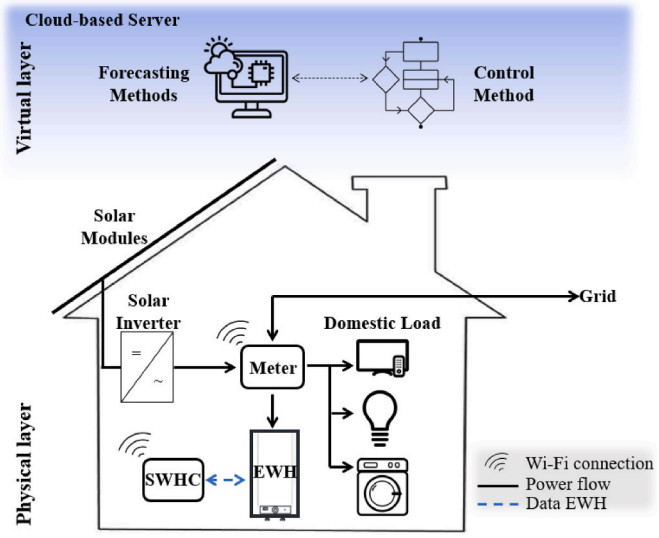


Fig. 1. Representation of the system's architecture for the control of a domestic EWH, highlighting the components belonging to the physical and virtual layers.

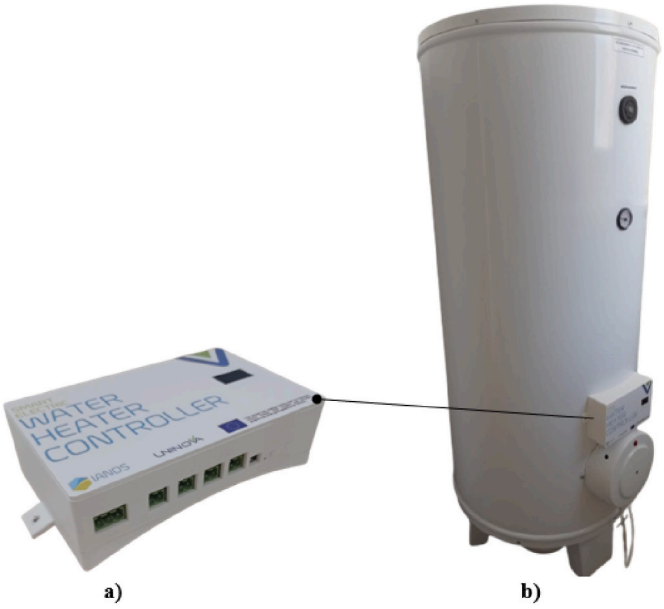


Fig. 2. Assembly diagram of the SWHC with the electric water heater. a) Developed SWHC. b) Integration of the SWHC with an existing electric water heater.

2.1. Proposed solution

The proposed solution combines a physical layer with a virtual one operating together. The physical layer interacts directly with the EWH, while also monitoring building consumption and generation profiles in real-time with 1-min. time resolution. Regarding the virtual layer, it consists of two main parts. The first part operates locally, with the devices of the physical layer, following a typical IoT based layout, where data series are encrypted and sent to the server, while also receiving the control setpoints imposed to the EWH. The second part operates on the cloud-based server, where the data series from the physical layer are received, processed, and the power control of the water heater is defined, based on the rules imposed by the algorithm operating in this layer (see Fig. 1).

Table 1
Data collected at the physical layer.

Measurand and unit	Observations
Total consumption power [W] EWH consumption power [W] Generation power [W]	Measured at the building's point of common coupling every minute using a set of Shelly EM solutions [23].
Inlet water temperature [°C] Outlet water temperature [°C] Top water temperature [°C] Inlet water flow [L]	Measured by the SWHC every minute. Inlet and outlet water temperature values are measured at the respective water pipes, that are external to the EWH. These transductors also allow the integration of water flow transductors. See Subsection 2.1.1 above for more details about the installed transductors.

2.1.1. Physical layer

The core component of the physical layer is the Smart Electric Water Heater Controller (SWHC), an IoT based device, which monitors and controls the EWH, as illustrated in Fig. 2. The SWHC not only collects all the EWH related information needed for virtual layer every minute (see Table 1) but also controls the supply of energy to the resistors responsible to heat-up the water using dedicated relays whose state is defined at each minute. The solution can be integrated in new or existing EWHs that allow the measurement of inside water temperature and rely on a mechanical thermostat. This feature allows the involvement of (previously) passive EWHs in smart energy solutions aiming, for instance, the deployment of Demand Response measures. In this study, the SWHC was integrated on an existing off-the-shelf EWH with no previous remote monitoring or control features.

The SWHC board is shown in Fig. 3, where the 230 V_{AC} power supply can be seen, along with a surge protection element (varistors), a overcurrent protection element (fuse), a transformer with an integrated rectifier bridge 230 V_{AC}/5 V_{DC}, a DC-DC converter 5 V_{DC}/12 V_{DC} for the relays (three relays for controlling the considered 150 L EWH presented in Fig. 2), an input for the EWH thermostat, an electric current meter, a microcontroller (control board), two analog-to-digital converters, seven analog inputs for temperature measurement (if several sensors are available), a digital input for measuring hot water consumption, and an output for a display. This study uses data from a single inside water temperature sensor despite the SWHC allowing the measurement of up to four sensors (plus the measurement of inlet and outlet water temperatures).

A microcontroller ESP32 [19] was used as control board, comprising two processing cores and several analog and digital input/output ports. The first core is dedicated to data acquisition and EWH power control. The second one manages connectivity with the local Wi-Fi network and the communication with the virtual layer.

In terms of data acquisition, temperature measurements are performed by Negative Temperature Coefficient (NTC) sensors, coated in stainless steel [20]. The signals sent by the NTCs are conditioned by two 16-bit resolution analog-to-digital converters [21]. A dig ital flow meter was used at the cold-water inlet to measure the water consumption [22].

The EWH power control is carried out by individually activating the three relays, each supporting a current of 10 A, considering that each activated relay represents an increase of 500 W in the consumption power of the EWH. The power supply for power control by the SWHC is provided by the EWH thermostat, serving as a second layer of protection to ensure operational safety. Additionally, a display shows relevant information, such as temperature, water consumption, and Wi-Fi connectivity status, offering a user interaction interface.

In the communication core, the SWHC communicates with the server using the HyperText Transfer Protocol Secure (HTTPS). At this level, all information collected in the current minute is encrypted and sent in JSON format, where the control method decides how many heating elements should be turned on or off in the next minute. In case of communication failure with the server (e.g., due to lack of Wi-Fi connection), a local control based on maximum and minimum temperature rules takes over operations, maintaining the user's thermal comfort. The water temperature range selected for this study was defined according to the values observed when the original thermostat-based control was still in operation (i.e., 50–74 °C).

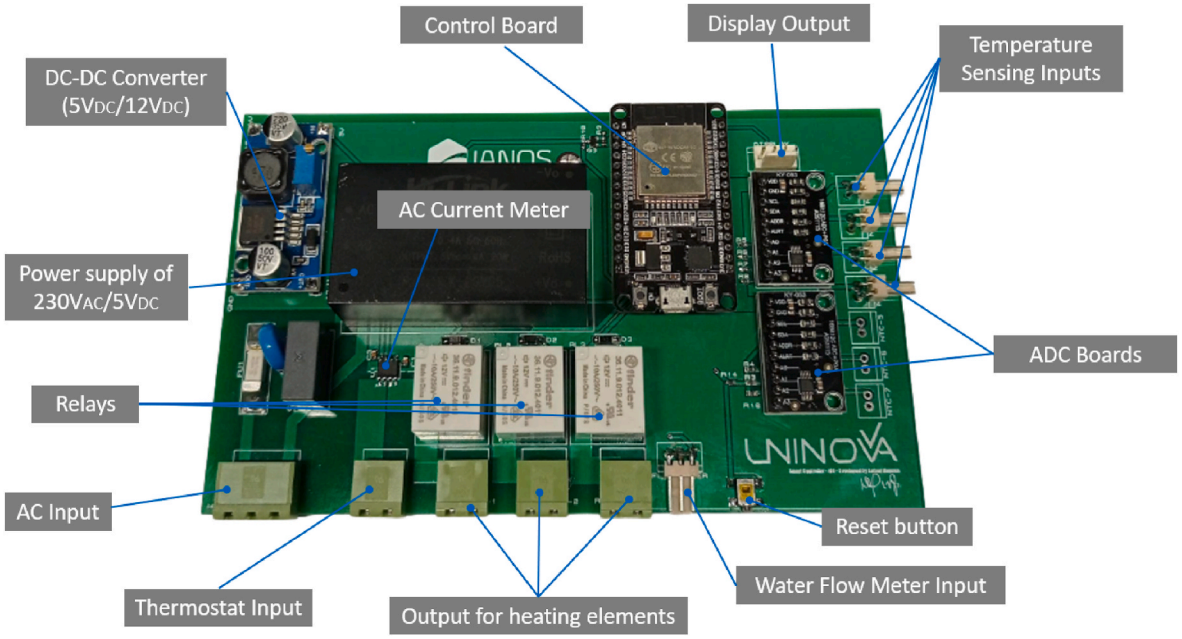


Fig. 3. Physical assembly of the Smart Electric Water Heater Controller board, highlighting the main components.

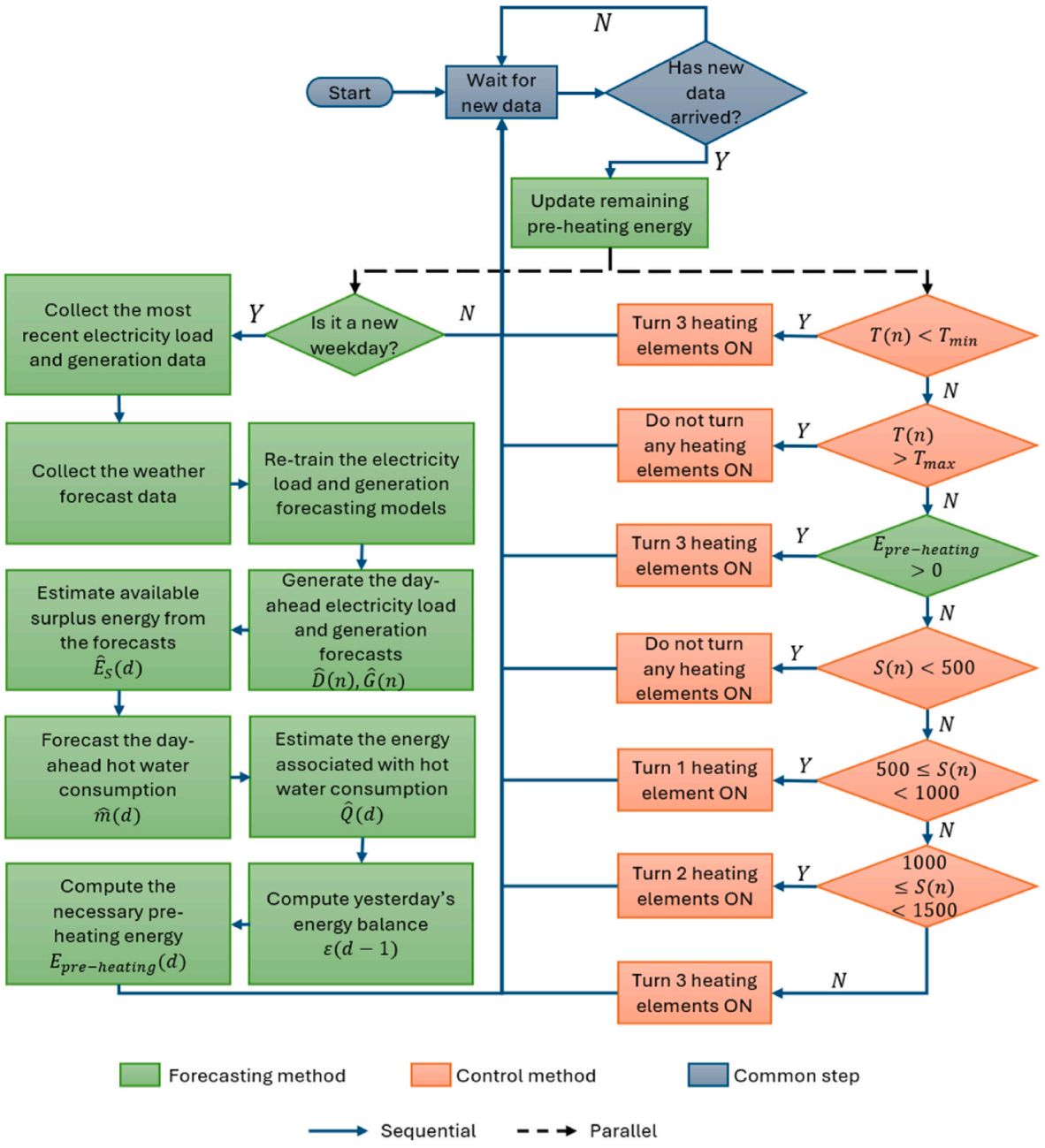


Fig. 4. Flowchart describing the interaction between the control and forecasting methods taking place at the virtual layer.

2.1.2. *Virtual layer*

The virtual layer processes data to determine the appropriate commands for the EWH. Every data point is processed to establish the appropriate command for the following time-step in real time. As mentioned in the previous section, the SWHC can activate three relays, each corresponding to a different heating element capable of providing 500 W of power. Thus, the decision carried out at the virtual layer consists of specifying the number of heating elements, from 0 to 3, that should be active in the next minute.

A rule-based control strategy, relying on the measured net power, is applied to absorb solar energy surplus and effectively increase self-consumption. This is achieved by activating the number of heating elements whose cumulative energy consumption can be met by the measured available energy surplus from the previous timestep. Due to the granularity in power control, commands for 1, 2, or 3 heating elements can only be issued when there is at least 500 W, 1000 W, or 1500 W of available surplus, respectively. Additionally, to ensure that the temperature of the water inside the tank is kept within a comfort range, predefined thresholds T_{min} and T_{max} are used (50 °C and 74 °C in this study). When the water temperature is lower than T_{min}, the system activates 3 heating elements (except during peak periods, as per the users' preferences). Conversely, when the temperature exceeds T_{max}, the heater blocks any attempt to further heat the water, even if surplus energy is available.

Equations Eq. (1) and Eq. (2) describe how the net power, P_{net}, and surplus power, S, for a given timestep n, are computed, while Eq. (3) describes the control scheme for the SWHC (i.e., the number of heating elements, $N_{heating}$, to be activated during time-step n).

$$\mathrm{P_{net}}\,(\mathrm{n}) = \mathrm{D(n)} - \mathrm{G(n)} \tag{1}$$

$$\mathrm{S(n)} = |\min(\mathrm{P_{net}(n)}, 0)| \tag{2}$$

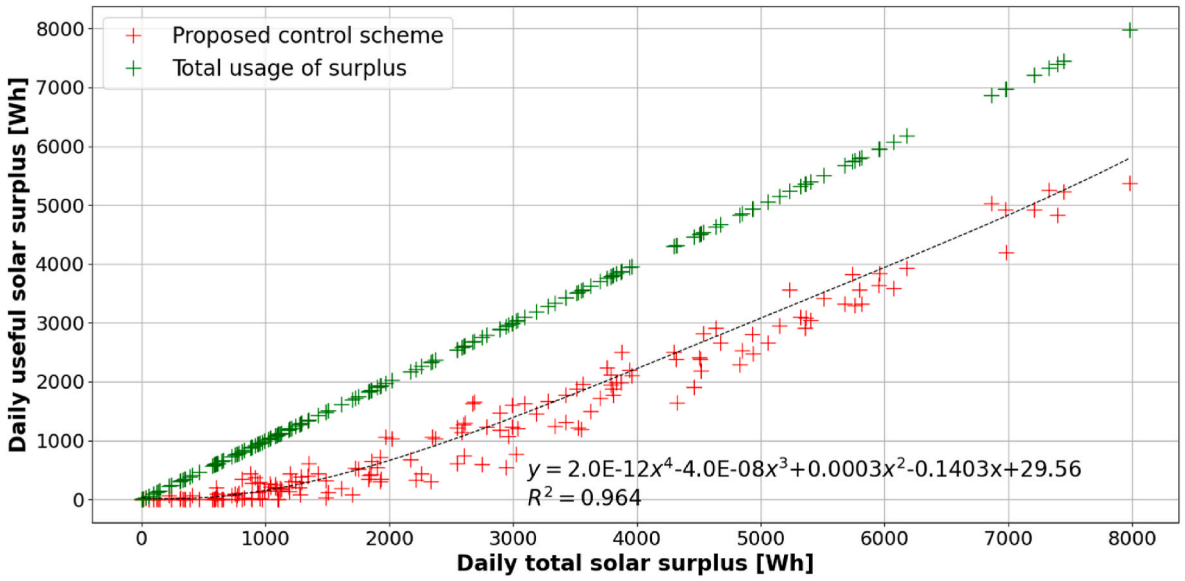


Fig. 5. Relation between the total PV surplus energy and the PV surplus energy that can effectively be consumed by the EWH.

$$N_{heating}(n)\begin{cases} 0, \text{if } S(n-1) < 500 \quad \text{or} \quad T(n-1) > T_{max} \\ 1, \text{if } 500 \le S(n-1) < 1000 \quad \text{and } T_{min} < T(n-1) < T_{max} \\ 2, \text{if } 1000 \le S(n-1) < 1500 \quad \text{and } T_{min} < T(n-1) < T_{max} \\ 3, \text{if } S(n-1) > 1500 \quad \text{or} \quad T(n-1) < T_{min} \end{cases} \tag{3}$$

In addition to the control strategy described above, the proposed control method also makes use of load, generation and hot water demand forecasting to further reduce electricity costs taking advantage of the applied ToU tariffs. By predicting the energy surplus that will be available for heating the water in the following day, as well as the expected hot water demand of the household, it is possible to estimate the remaining energy that should be used to fully meet the water heating energy demand. It is then possible to shift this heating load to a period with lower costs (see Table 6).

Concretely, every day at 00:00 (Azores time) and if it is a new weekday, the solution starts generating the necessary forecasts to compute the pre-heating energy for the day. When the pre-heating energy is yet not totally consumed, and the temperature of the water is within the defined thresholds, the virtual layer issues the command to activate 3 heating elements. In the following timestep (next minute), the resulting heating energy from the previous command is subtracted from the pre-heating energy. This is repeated until the pre-heating energy is fully depleted or T_{max} is reached. Fig. 4 shows a flowchart representing the operation of the proposed solution, where control and forecasting related actions are identified.

This approach uses a daily energy perspective to model the system's energy balance, as described in Eq. (4), where $E_{heater}(d)$ corresponds to the energy used heating the water the water at day d, $Q(d)$ is the energy demand associated with hot water consumption, and $L(d)$ represents the losses in the water tank. Therefore, according to the above description, $E_{heater}(d)$ should be satisfied by the PV surplus, $E_S(d)$, and by the pre-heating energy, $E_{pre-heating}(d)$, that needs to be provided in order to fully compensate for the losses and energy demand (hot water consumption), aiming to achieve a neutral energy balance on a daily basis, as described by Eq. (5). This pre-heating energy is then obtained using the forecasts of Q, L and $E_{surplus}$ for day d, represented in Eq. (6) by $\widehat{Q}$, $\widehat{L}$ and $\widehat{E}_{surplus}$, respectively, and a compensation factor ε_B that is used to correct the energy balance considering the forecasting errors from day d-1, providing a feedback mechanism to the proposed approach. The remaining of this subsection provides information about $\widehat{E}_S(d)$, $\widehat{Q}(d)$ and $\widehat{L}(d)$, while the specific forecasting methods developed during this study are presented in Subsection 2.2.

$$E_{heater}(d) = Q(d) + L(d) \tag{4}$$

$$E_{heater}(d) = E_S(d) + E_{pre-heating}(d) \tag{5}$$

$$E_{pre-heating}(d) = \widehat{Q}(d) + \widehat{L}(d) - \widehat{E}_S(d) + \varepsilon_B(d-1) \tag{6}$$

Due to the granularity in the heating power and the various shapes that the power surplus profile can assume, the total surplus energy that is available, $\widehat{E}_{S_T}(d)$, does not correspond to the energy that can be effectively used by the EWH, $\widehat{E}_S(d)$. For instance, if PV surplus is lower than 500 W during several time-steps, no heating element is activated, resulting in a discrepancy between $\widehat{E}_{S_T}(d)$ and $\widehat{E}_S(d)$. Thereby, historical data ranging from 15th of December of 2023 to 1st of June of 2024 were used to construct a polynomial linear regression that models the relation between the useful and total energy surplus at the daily level, as shows Fig. 5, acting as a correction function, Eq. (7), between $\widehat{E}_S(d)$ and $\widehat{E}_{S_T}(d)$.

Considering a finite foresting horizon of length N (24 h in this study), $\widehat{E}_{S_T}(d)$ is obtained by Eq. (8), where $\widehat{S}(n)$ corresponds to the forecasted PV surplus at time-step n computed using Eq. (9), which relies on the forecasted net power $\widehat{P}_{net}$ that is obtained by the difference between the building's forecasted demand, $\widehat{D}(n)$, and generation, $\widehat{G}(n)$, using Eq. (10). The specific methods developed to obtain $\widehat{D}(n)$ and $\widehat{G}(n)$ are described in Subsection 2.2.

$$\widehat{E}_S(d) = CF(\widehat{E}_{S_T}(d)) \tag{7}$$

$$\widehat{E}_{S_T}(d) = \int_0^{N-1} \widehat{S}(n)dn \tag{8}$$

$$\widehat{S}(n) = \min(\widehat{P}_{net}(n), 0) \tag{9}$$

$$\widehat{P}_{net}(n) = \widehat{D}(n) - \widehat{G}(n) \tag{10}$$

As per the users' preferences, the next day's consumption of hot water is set to happen at an average T_{outlet} of 60°C, where the same volume of water is replaced at the expected inlet temperatures, which can be estimated by taking the average inlet temperature, $\overline{T}_{inlet}$, of the previous day. The change in the system's thermal energy due to consumption of hot water at a certain day d can therefore be estimated by

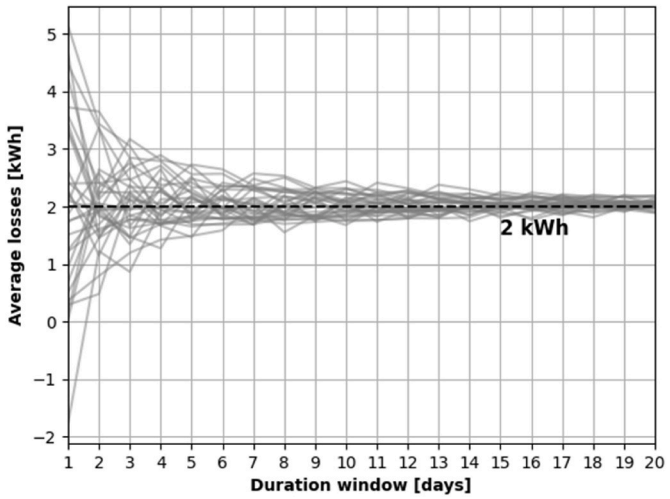


Fig. 6. Average losses for different duration windows.

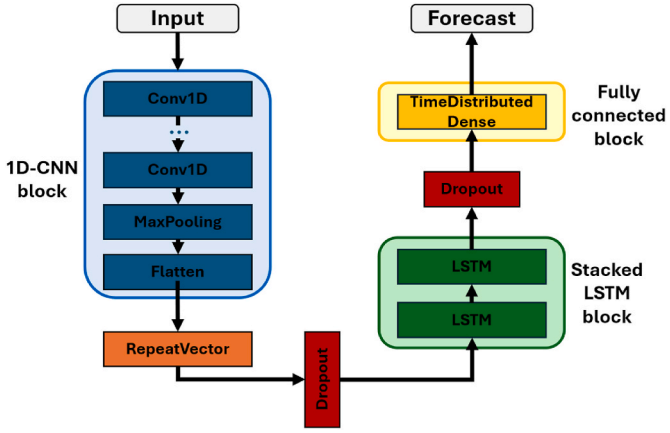


Fig. 7. Main blocks of the method developed for the short-term electricity consumption and generation forecasting.

equation Eq. (11), in kWh, where $\widehat{m}(d)$ is the forecasted hot water consumption volume for day d, in m^3, described in Subsection 2.2, C_s is water's heat specific capacity (1.16 Wh/kg·°C) and ρ is the mass density of water in (1000 kg/m^3).

$$\widehat{Q}(d) = \widehat{m}(d) \times C_s \times \Delta\widehat{T}(d) = \widehat{m}(d) \times C_s \times \rho \times (60 - \overline{T}_{inlet}(d-1)) \quad (11)$$

Following a data-driven approach, the natural heat losses due to exchanges with the environment, $\widehat{L}$, are estimated by the average of the difference between the daily energy used heating the water and the daily energy tied to the consumption of hot water over a certain number of past days. Fig. 6 shows that, by computing this difference for every day of the month of May 2024, using Eq. (12), the average losses tend to approximately 2 kWh. A negative value in Fig. (6) refers to a period where more energy was extracted from the EWH than supplied to, taking advantage of the thermal storage capacities of this device. As the duration of the average computing window increases, the average losses tend to the referred 2 kWh equilibrium.

$$\widehat{L}(d,k) = \frac{1}{k}\sum_{i=1}^{k}(E_{heater}(d-i) - Q(d-i)) \quad (12)$$

2.2. Forecasting

As outlined in the previous subsection, accurate day-ahead forecasts of domestic electricity consumption, generation, and domestic hot water (DHW) demand are essential for estimating pre-heating energy requirements. This section details the forecasting methods used to achieve these day-ahead forecasts.

Table 2
Proposed forecasting method and grid search intervals.

		Proposed models		Grid search intervals
		Consumption	Generation	
1D-Convolution and MaxPooling	Filters	32	32	–
	Kernel size	3	3	–
	Activation function	ReLU	ReLU	–
	Number of layers	3	2	[1, 2, 3]
	Pool size	2	2	–
Dropout	Rate	0.1	0.3	[0, 0.1, 0.3]
LSTM	Units	128	64	[32, 64, 128]
	Activation function	tanh	tanh	–
	Number of layers	2	2	[1, 2, 3]
Dropout	Rate	0.1	0.3	[0, 0.1, 0.3]
Time Distributed Dense	Hidden neurons	128	64	[32, 64, 128]
	Activation function	ReLU	ReLU	–
Output	Hidden neurons	1	1	–
	Activation function	ReLU	ReLU	–
-	–	–	–	–
Training parameters	Epochs	200	200	[200, 400, 600]
	Batch size	64	64	[32, 64, 128]
	Optimizer	Adam	Adam	–
	Learning rate	0.001	0.0005	[0.0005, 0.001]
	Time period	60 days	60 days	[30, 60, 90]

2.2.1. Electric energy consumption and generation forecasting

Electricity generation and consumption forecasts enable the system to anticipate periods of PV energy surplus, resulting in better scheduling of the EWH operation. This ensures that the EWH primarily uses excess solar energy for water heating while also taking advantage of lower electricity costs during the night. By prioritizing the use of locally generated energy, this approach leads to higher rates of PV self-consumption and self-sufficiency. In this study, we propose the use of a hybrid deep learning architecture composed of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and a fully connected layer for short-term electricity consumption and generation forecasting at the single household level. By combining 1D-CNN and LSTM models, it is expected that the resulting hybrid model benefits from their individual qualities, such as the capacity for feature extraction of the CNN model and the ability to capture temporal dependencies provided by the LSTM model. There exist instances of comparison between the CNN-LSTM hybrid and its constituent models regarding forecasting short-term electricity consumption at both the individual household [24] and aggregated levels [25], as well as PV generation [26], where CNN-LSTM commonly outperforms the individual models. The proposed model architecture is represented in Fig. 7.

The hybrid model's input is a Convolutional Neural Network (CNN), composed of a variable number of one-dimensional (1D) convolutional layers, with the number of layers treated as a hyperparameter to be tuned. Each convolutional layer consists of 32 filters with a kernel size of 3. This is followed by a max pooling layer with a pool size of 2, ensuring that only half of the values from the generated feature maps are propagated to the subsequent layers, in specific, the maximum values. The resultant feature maps are then flattened and passed to a repeat vector,

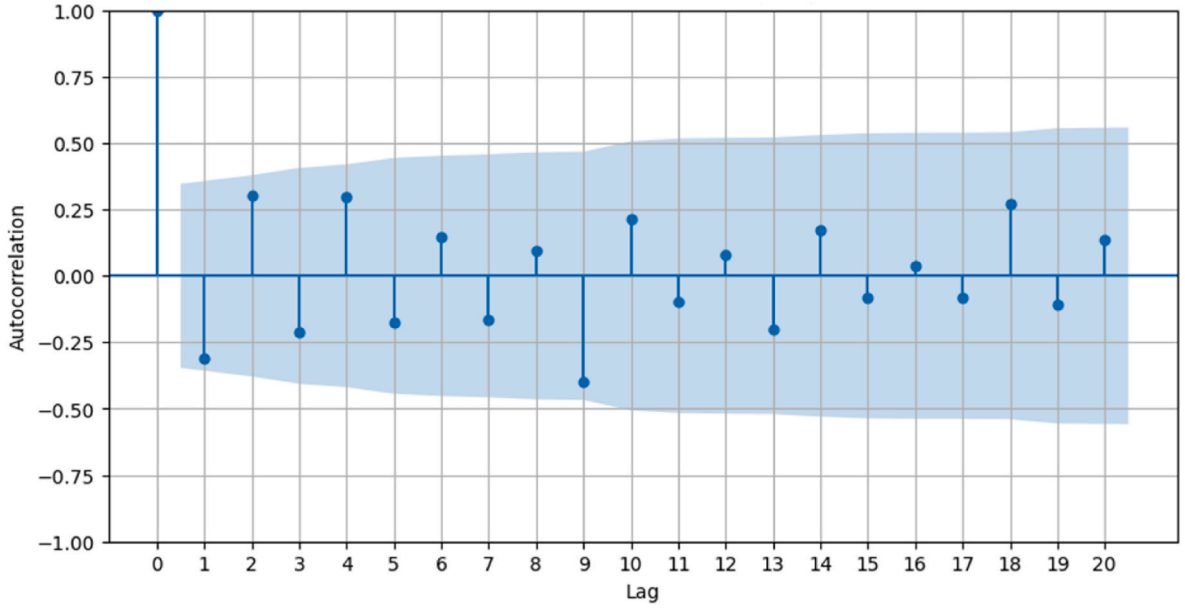


Fig. 8. Autocorrelation function for the hot water consumption in the month of May during the validation period.

which replicates them according to the number of expected outputs. This step ensures compatibility between the CNN and the subsequent Long Short-Term Memory (LSTM) layers. This architecture is often referred to as an encoder-decoder. In this specific instance, the CNN layer constructs an internal representation of the inputs through the feature maps, while the subsequent LSTM and fully connected layers decode this representation to generate a prediction. The LSTM layers have a *tanh* activation function and a variable number of units, also regarded as one of the tuneable hyperparameters. The fully connected layers are also time distributed, meaning that the same dense layer is applied to each time step of the LSTM output. Dropout layers are added between the three blocks to avoid overfitting. The generated output has 24 time-steps relating to each of the hours for the forecasted day.

The adequate hyperparameters were obtained through a grid search and are presented in Table 2 for both electricity consumption and generation forecasting models. It should be noted that the considered training data period was also optimized in the grid search. Additionally, comparison with common models, such as SARIMAX and LSTM, was performed on a validation period, as addressed in Subsection 3.4.1, where the CNN-LSTM model achieved better closed-loop performance for both load and generation forecasts (i.e., more accurate estimation of the day-ahead energy surplus than its counterparts). It should be noted that the choice of this forecasting model stems from the referred model comparison results and its relevant presence in the short-term forecasting literature, as previously described. The trade-off between performance and computational cost is not explored in this paper thereby representing an opportunity for future related works. The proposed models were trained and validated using the TensorFlow deep learning framework on Python version 3.10 in a NVIDIA GeForce RTX 3070 Laptop GPU and 16 GB of RAM hardware environment.

2.2.2. *Domestic hot water consumption forecasting*

Methodologies for forecasting hot water demand at the residential level have been previously used in the context of optimizing the

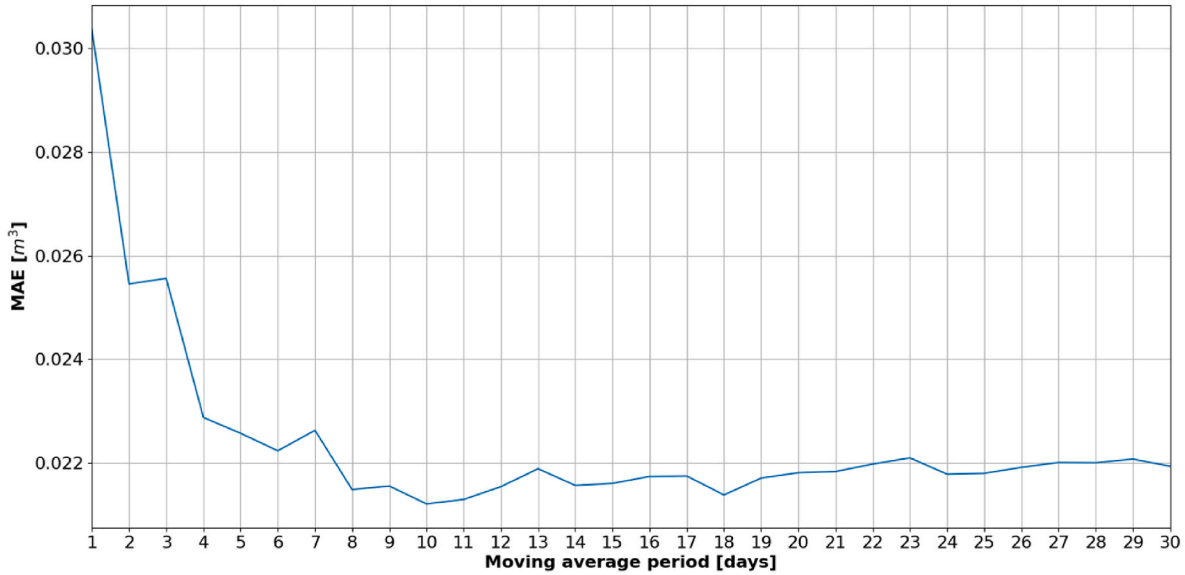


Fig. 9. Variation of the mean absolute error of daily hot water consumption.

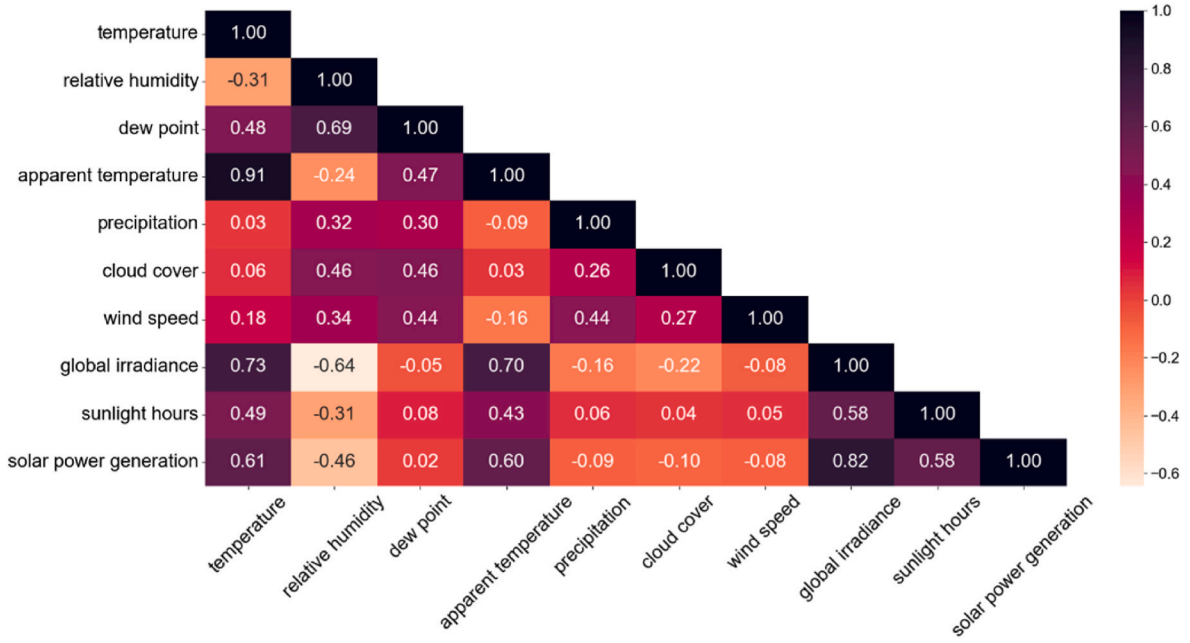


Fig. 10. Correlation matrix of weather variables and solar power generation.

operation of EWHs. In Ref. [16], the authors use XGBoost and LSTM models, trained on historical data, to predict a household's timing of usage and respective volume of hot water during a given day. Similarly [27], forecasts hot water demand, but at an aggregated level, using a PROPHET model to use in a simulated Model Predictive Control (MPC) for scheduling the operation of EWHs. Artificial Neural Networks have also been used in the forecasting of domestic hot water demand to compare MPC schemes with perfect and imperfect forecasts [28]. In all the above-mentioned cases, the forecasting targets the demand profile at different data granularities. Since the solution proposed in this study follows a daily energy perspective, the forecasted hot water consumption will also refer to a daily total, warranting a different approach.

Analysis of the autocorrelation function plot for daily hot water consumption in May, shown in Fig. 8, indicates that all lagged correlations fall within the Bartlett confidence interval. This implies that there are no statistically significant autocorrelations, making the behavior of this time series akin to white noise. As so, a conservative strategy to predict hot water consumption is to use the moving average of the water consumption for the last N days. A generalization of the moving average is represented in Eq. (13), where $N = 1$ corresponds to a random walk model.

$$\widehat{Y}_t = \frac{\sum_{i=1}^{N} Y_{t-i}}{N} \tag{13}$$

Fig. 9 shows that, for the period from the 1st of January of 2024 until the 1st of June of 2024, the moving average of the last 10 days minimizes the mean absolute error in the prediction of the daily hot water consumption. The 4-day moving average ($N = 4$) was selected for hot water consumption forecast since it is more responsive to changes in certain variables, such as the number of people in the household, than the 10-day moving average.

2.2.3. Exogenous features

To aid in the forecasting tasks, some exogenous features were selected to be part of the developed models. Real weather forecast variables are obtained from the open-data global numerical weather prediction model ARPEGE. This model's weather forecast data was chosen due to its coverage of the case-study location, with a reasonable resolution of 0.1°.

Table 3
Selected exogenous features for the electricity consumption and generation forecasting methods.

	Exogenous features
Consumption	"Day of the week", "Weekday or Weekend"
Generation	"temperature", "relative humidity", "precipitation", "cloud cover", "global tilted irradiance"

An analysis of Pearson's correlation coefficient between the target and relevant weather variables was carried out for the month of May 2024, which is represented in Fig. 10. It is possible to verify some strong correlations (above 0.4) between the solar power generation and some of the considered features, such as temperature, global tilted irradiance, relative humidity and, as expected, sunlight hours. Certain features that showcase lower correlation coefficients, such as cloud cover and precipitation may have a varying level of impact on the prediction of solar power generation depending on the time of year.

Artificial features like day of the week and a binary indicator of whether it is a weekday or weekend were included as exogenous features of the electricity consumption model. Table 3 lists the exogenous features that are used in the electricity consumption and generation forecasting methods.

2.2.4. Forecasting validation

The developed methods were validated in a walk forward scheme like the one illustrated in Fig. 11. It consists in training the models with time-series data of a fixed length and forecasting the subsequent 24 h out of sample. After storing the forecast, the training data is shifted by one day, keeping the same length but now including the real data from the same period as the last forecast, and thus repeating the model training and successive forecasting. This process is repeated for a selected number of days in a validation window.

For this study, a walk forward validation window of 15 days, starting in 1st of May of 2024 and ending in 15th of May of 2024, was considered. This allows for a realistic evaluation of the electricity consumption and generation forecasting methods, since it is identical to the real time operation of the proposed system.

To assess the quality of the developed methods, and to effectively

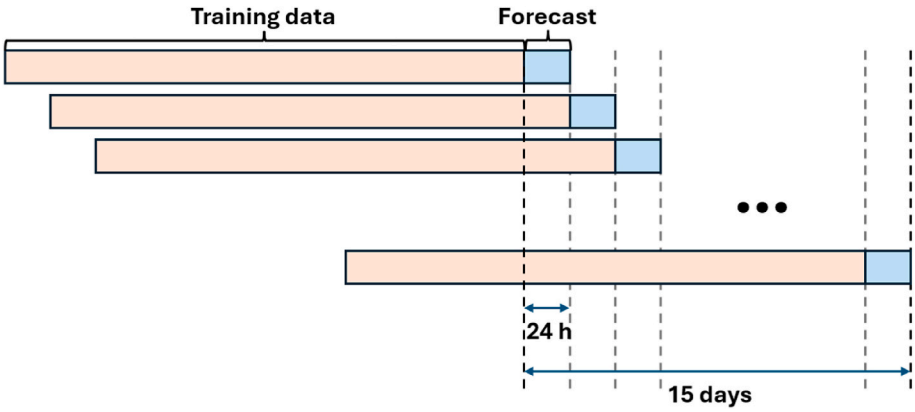


Fig. 11. Illustration of the walk forward validation approach for the forecasting methods.

Table 4
Forecasting validation metrics.

Metric	Expression
The root mean squared error (RMSE)	$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}$
Mean absolute error (MAE)	$MAE = \frac{1}{k}\sum_{i=1}^{k}\lvert y_i - \hat{y}_i\rvert$
Mean absolute percentage error (MAPE)	$MAPE = \frac{1}{k}\sum_{i=1}^{k}\frac{\lvert y_i - \hat{y}_i\rvert}{y_i} \times 100$

select the best fits to integrate the pre-heating control, three common prediction evaluation metrics were adopted, namely: the root mean squared error (RMSE), the absolute error (MAE) and the mean absolute percentage error (MAPE). These metrics are described in Table 4, where y_i and $\hat{y}_i$ are the real and forecasted time-series, respectively.

Given that the electricity load and generation forecasts are used to predict the day-ahead energy surplus, the error associated with the electricity consumption forecasting method is calculated while assuming a perfect electricity generation forecast, and vice versa. This approach allows for a detailed assessment of each method's performance by isolating the errors in load and generation forecasts, thereby providing a more accurate evaluation of their impact on day-ahead energy surplus forecasts.

The CNN-LSTM model, used for both consumption and generation forecasts, is compared to a Naive model baseline – weekly for consumption and daily for generation, where each performed best in the considered period. Additionally, it is compared to a less complex LSTM model with a fully connected output layer, whose hyperparameters were optimized through a grid search. Other models considered in this comparison include SARIMAX, with its order and seasonal order determined by minimizing the Akaike and Bayesian information criteria, and the PROPHET model, known for its simplicity and reasonable performance.

2.3. Power system in Terceira Island

The Azores archipelago is composed of nine disperse islands. Due to their location and distance among them, they have insulated electric systems. These comprise different technologies, which try to explore as much as possible the endogenous resources of each island. Although there is an ongoing effort on increasing the penetration of renewable energy sources, thermal power plants, resorting to fossil fuels as diesel, are still the main technology. Each island has a thermal power plant, necessary to ensure the adequate levels of safety of supply and providing as well ancillary services. The most important renewable resource in Azores has been geothermal energy. There are a total of three geothermal power plants, one in Terceira Island. Wind, mini-hydro and solar generation follow behind in terms of renewable energy sources in the archipelago. In 2023, thermal energy was responsible for 63 % of total generation. From the 37 % of renewable generation, 22 % was from geothermal, 9 % from wind and 6 % from other renewable sources [29]. The system is vertically integrated, and EDA - *Electricidade dos Açores*, a public owned company, is the responsible utility.

The Terceira Island represents the second largest system in Azores, following behind São Miguel. In terms of consumption power, the Terceira Island is responsible for 23 % of the total consumption of the archipelago (São Miguel has a consumption of 56 % of the total). It has a total installed capacity of 100 MW, from which 78 MW are from a fuel thermal power plant (11 groups). Wind power plants have an installed capacity of 13 MW while geothermal (5 MW), urban residues (3 MW) and mini-hydro (1 MW) account for the remaining 9 MW [29].

As of December 2023, the Terceira power system was composed of a total of 8 generation power plants with the mentioned installed capacities, five substations and a battery storage system (lithium ion – 24 MW/16 MWh) which provides additional flexibility to the system. The network is operated with a transmission system operating at 30 kV, a medium-voltage distribution network operating at 15 kV and low-voltage networks.

In 2023, the total generation was 196.59 GWh with a peak demand of 33.2 MW on the Summer. On a typical summer day, the load factor is around 75 %, while the peak valley ration is 49 % [29]. These indicate a reasonable variation in the load profile, resulting in the need for system flexibility sources, as the one proposed in this article.

2.4. Assessment metrics

The performance of the proposed solution is assessed in terms of PV self-consumption (SC), water heater's self-sufficiency (SS), electricity costs for the households and CO_2 emissions due to electric energy import from the distribution grid. For this assessment, the total demand $D(n)$, local generation $G(n)$, water heater's consumption power $H(n)$, and the building's net transactions with the electrical grid $P_{net}(n)$ are expressed mathematically as time-series with 1440 values, representing the 24 h of the day with 1-min resolution.

Equation (14) is used to compute the household SC, considering local generation $G(n)$, and the total demand $D(n)$, both in W and n in minutes [30]. This SC indicator denotes the percentage of locally generated electrical energy that is used to directly meet the local demand within the considered time interval.

$$SC = \frac{\sum_{n=n1}^{n2} \min\{D(n), G(n)\}}{\sum_{n=n1}^{n2} G(n)} \cdot 100 \quad (14)$$

Regarding the SS of the water heater, as shown in Equation (15), it the available PV surplus, $S(n)$, and quantifies how much of the local

Table 5
Emission factors and energy source relative contribution.

Energy source	*Emission factor [kgCO_2/kWh]*	*Relative contribution [%]*
Diesel	0.615	64.4
Hydropower	0.00	3.5
Geothermal	0.199	21.5
Wind	0.0	8.5
Urban Solid Waste	0.722	1.6
Others	0.0	0.5

Table 6
Electricity tariffs applied to the considered household.

Day	Period	Cost [€/kWh]
Business days	Peak (10:30–15:29)	0.2415
	Half-peak (7:00–10:29, 15:30–23:59)	0.1736
	Off-peak (00:00–6:59)	0.1072
Saturday	Half-peak (11:00–14:29, 19:30–22:59)	0.1736
	Off-peak (00:00–10:59, 14:30–19:29, 23:00–23:59)	0.1072
Sunday	Off-peak (00:00–23:59)	0.1072

generation surplus meets the energy needs of the water heater within the considered time interval.

$$SS = \frac{\sum_{n=n1}^{n2} \min\{H(n), S(n)\}}{\sum_{n=n1}^{n2} H(n)} \cdot 100 \quad (15)$$

The equivalent CO_2 emissions (EI_CO_2), in $kgCO_2$, associated with energy import to supply the EWH, measured in kWh, is computed using a weighting of the emission equivalents of the available primary energy sources, as shown in Equation (16). In this equation, X represents the percentage of energy produced by each source during the considered period, and E_{eq} the equivalent emission associated with the respective energy source (i.e., hydropower H, geothermal G, diesel D, urban solid waste USW, and others O, see Table 5), all in $kgCO_2kWh^{-1}$. The equivalent emission factor of each energy source used in study was obtained from the local utility EDA - *Electricidade dos Açores* [31].

$$EI\_CO_2 = E_I\left(X_H E_{eqH} + X_G E_{eqG} + X_D E_{eqD} + X_{USW} E_{eqUSW} + X_O E_{eqO}\right) \quad (16)$$

Similarly, the equivalent emissions for the gas water boiler are computed using Equation (17) with the corresponding equivalent for butane gas (E_{eqBG}).

$$G\_CO_2 = E_{flow}\, E_{eqBG} \quad (17)$$

The energy provided by the gas water heater (E_{flow}), measured in kWh, during hot water consumption is shown in Equation (18), where ρ is the mass density of water in kgm^{-3}, $\dot{m}$ is the mass flow of water in m^3, cp is the specific heat of water in kWh/kg°C, $T_{in}(n)$ and $T_{out}(n)$ are the inlet and outlet water temperatures in the tank in °C, respectively, and η is the efficiency of the gas boiler.

$$E_{flow} = \sum_{n=n_1}^{n_2} \frac{\rho\, \dot{m}\,(n) c_p}{\eta} \left(T_{out}(n) - T_{in}(n)\right) \quad (18)$$

To calculate the costs associated with using butane gas, the low calorific value was used, with 1 kg of butane corresponding to 12.70 kWh [32]. The applied tariff was calculated based on the price of a 55 kg gas cylinder, costing 82.93 € [33], resulting in an energy cost of 0.1180 €/kWh.

3. results and analysis

The proposed solution has been deployed at a single-family household to explore the energy flexibility provided by a 150 L EWH, as described in Subsection 3.1. The collected results are presented and analyzed in terms of PV self-consumption improvement (Subsection 3.2) and costs and CO_2 emissions reductions (Subsection 3.3). The performance of the developed forecasting method is assessed in Subsection 3.4.

3.1. Case study

The proposed solution is currently in operation at a single-family household consisting of two adults and one child, located on the northern coast of Terceira Island, Azores, Portugal. This case study is conducted under the context of IANOS H2020 project, which aims to decarbonize the energy systems of two Lighthouse Islands (in Azores and in the Netherlands) while exploring the replicability potential of the tested technologies in three Fellow Islands (Bora-Bora, Lampedusa and Nisyros). IANOS aggregates the flexibility potential of storage technologies, like the one presented here, to increase renewable penetration, reduce CO_2 emissions, and boost the decarbonization of these islands.

Terceira Island is one of the nine islands in the Azores archipelago, Portugal. It is the second most inhabited island in the archipelago, with 53,311 inhabitants and the economy is mainly based on agriculture (dairy farming) and industries associated with dairy processing. In this island the climate is mostly rainy and humid throughout the year, with average monthly temperatures ranging from 21.5 °C in August to 13.2 °C in February [34].

The house is connected to the normal low-voltage distribution grid, operating at 230 V, 50 Hz, with a contracted power of 6.9 kVA, under a Time-of-Use (ToU) tariff scheme with three pricing periods [35], as described in Table 6 (plus VAT taxes). The building comprises a photovoltaic generation system with four 330 Wp modules, totaling 1320 Wp installed, with no electrochemical storage. It is worth mentioning that an electrical vehicle is present, being typical charged once a day during the off-peak period. User feedback was collected via periodic direct communication. At an initial stage, user preferences (see Section 2.2) were assessed and integrated in the control methodology. This case study considers data collected during the last two weeks of July 2024.

3.2. Improvement of PV self-consumption

The integration of the proposed solution conducted to a PV self-consumption improvement of 36 % during the period of analysis (from 31 % to 67 %). By analyzing the collected power consumption and generation profiles (see examples in Fig. 12), it is evident that the EWH operation was shifted to periods with generation surplus. For the sake of readability, the net power profile depicted in Fig. 12 does not comprise the power consumption of the EWH. Therefore, the power profile used by the supplier to compute the electricity bill would be represented by adding the brown and blue curves.

Regarding the EWH self-sufficiency, a rate of 75 % was obtained for the period under analysis. This metric did not reach a 100 % value due to i) pre-heating during the evening, ii) forced heating when the water temperature was below 50 °C and iii) mismatch between PV surplus during consecutive time-steps, given that H(n) is defined according to n-1 data. This self-sufficiency rate compares to a scenario prior to the deployment of the proposed solution where all energy (from gas) was imported, resulting in considerable costs and CO_2 reductions (see Section 3.3).

Fig. 12 also illustrate the pre-heating performed in some days based on the developed forecasting methods. This can be observed in Fig. 12 a), where 3.70 kWh were consumed during the off-peak period (green area in the figure) to avoid energy consumption during peak periods (red area). It is important to note that the referred pre-heating is not applied every weekday (e.g., no need for pre-heating on July 18, 2024, given the forecasted generation surplus) nor during weekends, as observed Fig. 12 b), given that not peak pricing periods are imposed to the household and

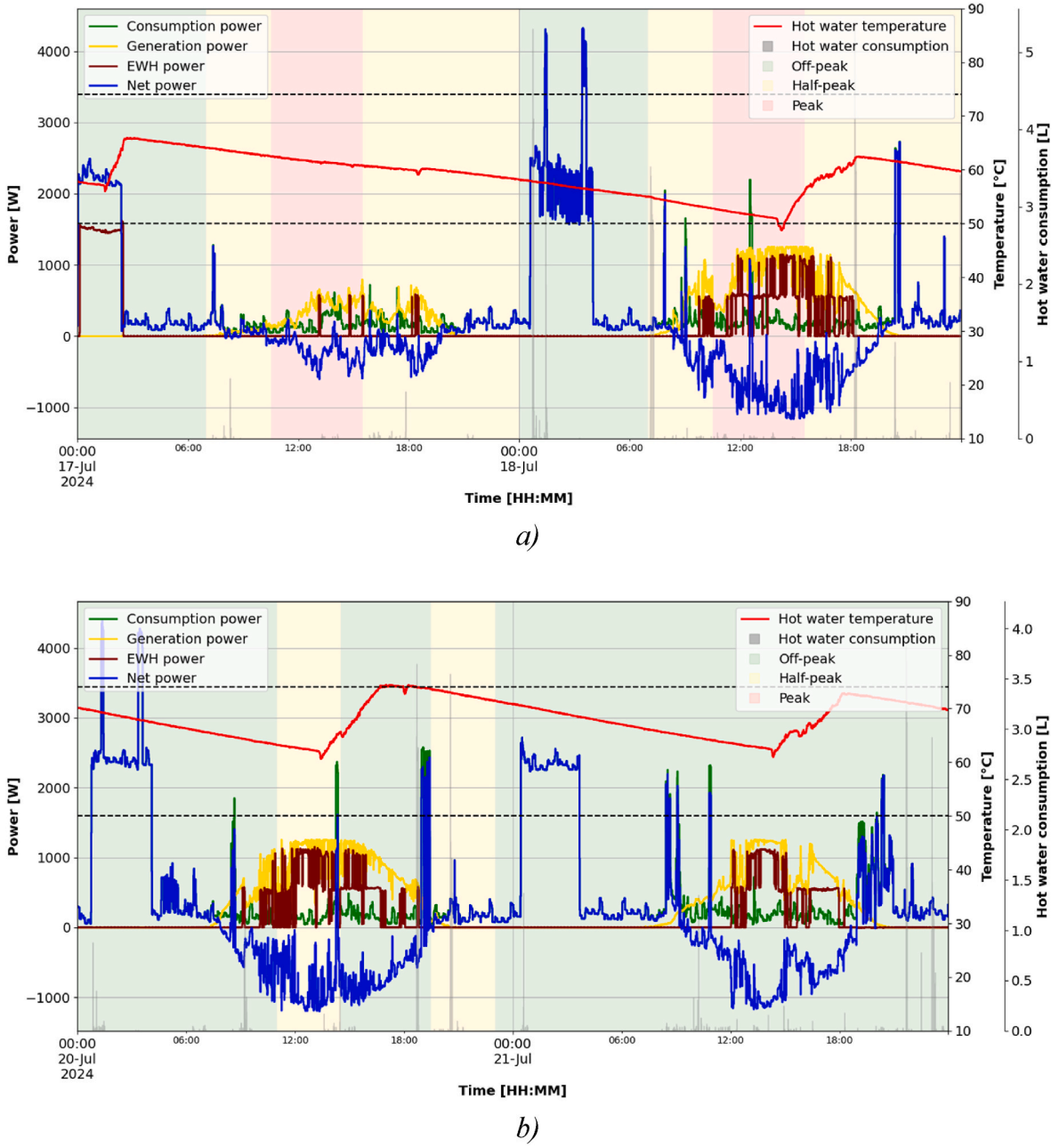


Fig. 12. Energy consumption and generation profiles, together with hot water consumption and temperature, collected during the case study. a) Example for weekdays. b) Example for a weekend.

therefore no strong incentive exist to shift the EWH load besides the PV improvement strategy conducted every day.

3.3. Reduction of energy costs and CO_2 emissions

The reductions in energy cost and CO_2 emissions were evaluated for the 2-week duration of the presented case study. Information about the efficiency of the gas boiler that was replaced by the EWH was not available. Therefore, the hot water consumption data, as well as the inlet and outlet temperatures of the electric water heater, were used to quantify the energy exchanges and using simulations and collected data it was decided to investigate the cost and CO_2 emissions, varying the gas boiler efficiency from 50 % (pessimistic scenario) to 100 % (ideal scenario). As shown in Fig. 13, at the extremes of the gas boiler efficiency, the total energy costs of the household are €6.35 with 12.72 $kgCO_2$ emissions for 50 % efficiency and €3.17 with 6.36 $kgCO_2$ emissions for 100 % efficiency. The cost (Fig. 13, left axis) and emissions (Fig. 13, right axis) curves are computed considering the fixed electricity consumption of the residence (excluding the consumption of the electric water heater), added to the energy equivalent to the use of butane gas during hot water use.

By introducing the EWH with the control described in Subsection 2.1.2, significant cost savings were achieved, ranging from €4.47 (−70.4 %) to €1.29 (−40.7 %), compared to the original scenario (i.e., water heating with gas boiler), with efficiencies at the previously mentioned extremes, respectively (50 %–100 %, see Fig. 14, left axis). Electric energy import during pre-loading and in periods with temperatures below 50 °C, represented an increase in electricity consumption of the household, corresponding to a cost of €1.88. Regarding the self-sufficiency of the storage water heater, a value of 75 % was obtained, as described in Subsection 3.2, which represents 14 kWh of energy import from the power grid (see Fig. 15). When stratifying the 14 kWh

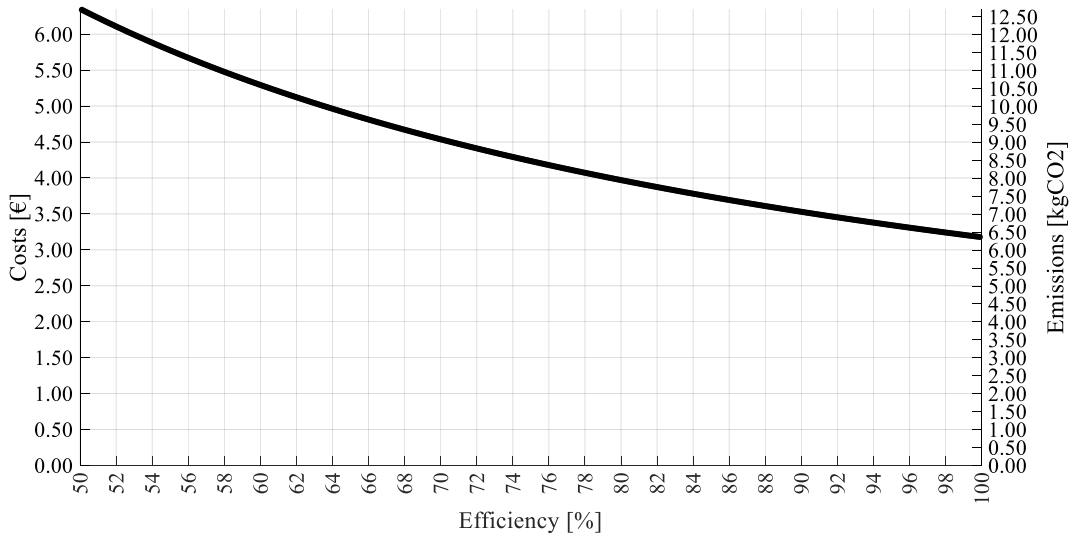


Fig. 13. Total energy costs of the household before the electrification of water heating energy demand for different efficiency values of the gas boiler.

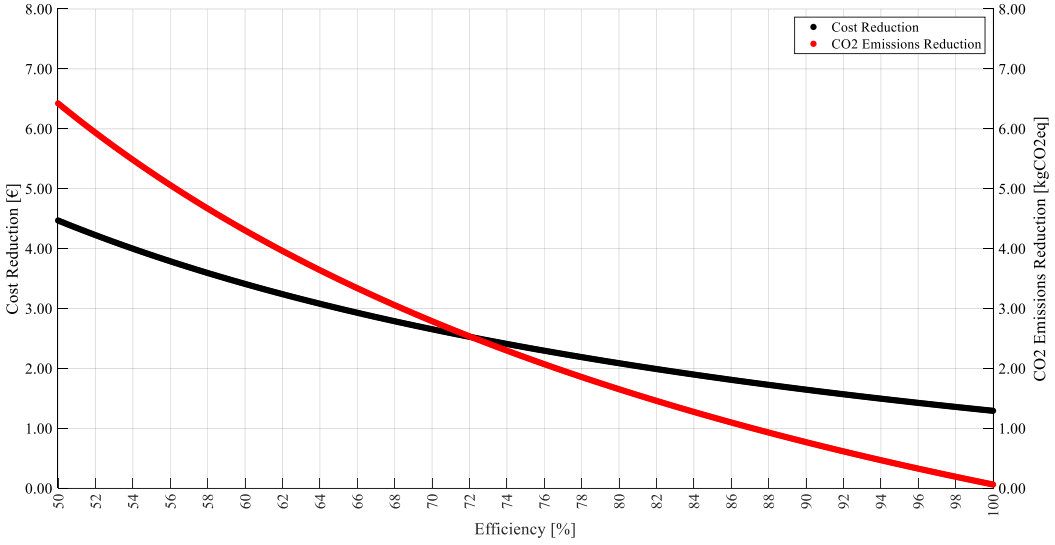


Fig. 14. Reduction of costs and CO_2 emissions resulting from the electrification of water heating energy demand resulting from the proposed solution.

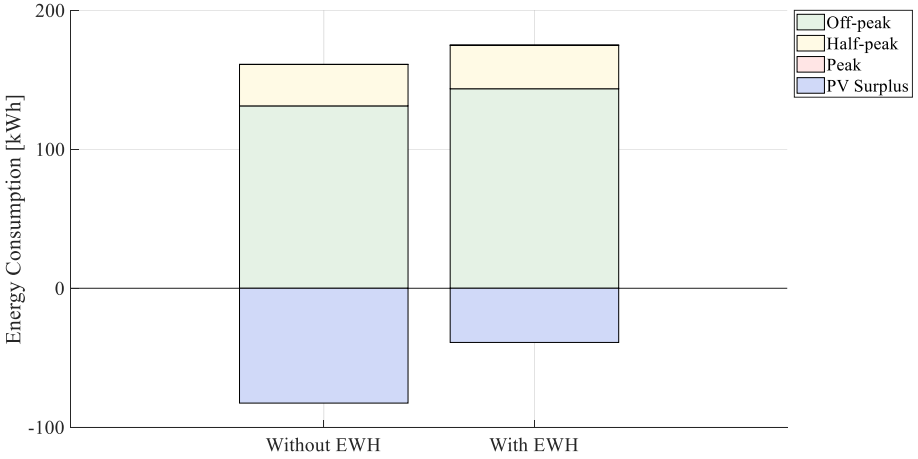


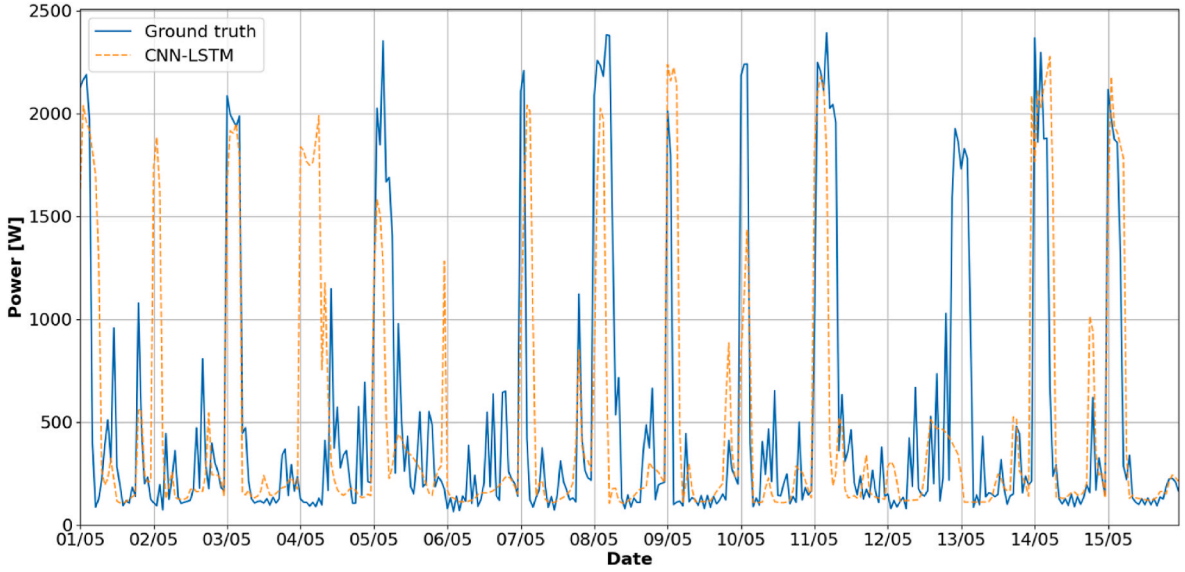
Fig. 15. Energy consumption per tariff period, with and without the EWH, during the case study.

Table 7
Comparison of the electricity consumption and generation forecasting performance of the chosen models for the period between May 01, 2024 and May 15, 2024.

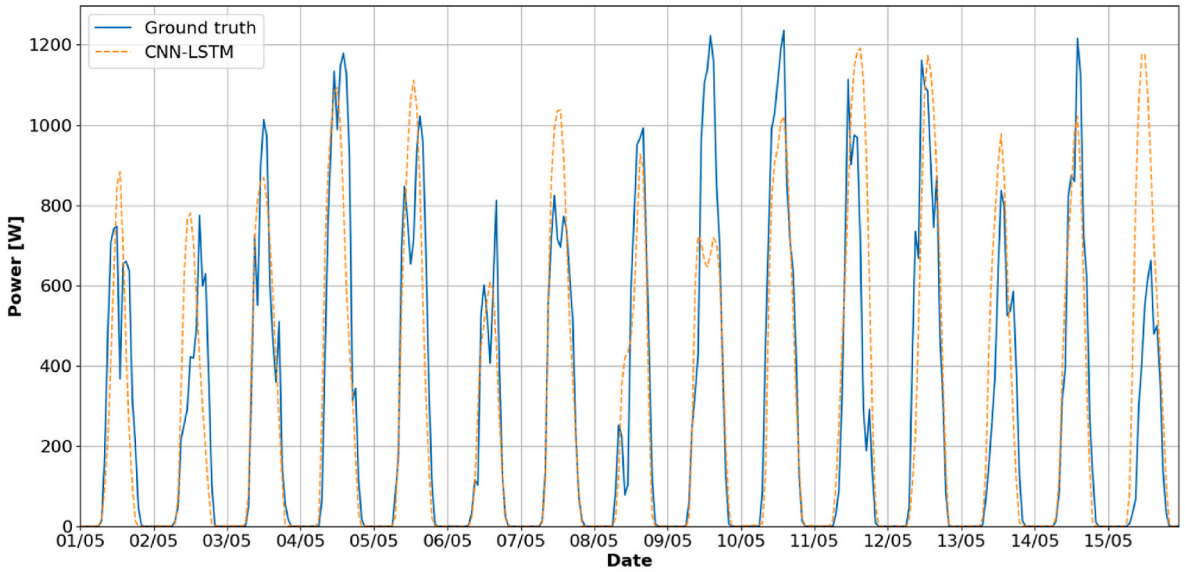
Model	Ground truth				Surplus energy	
	MAE (W)		RMSE (W)		MAPE (%)	
	D(n)	G(n)	D(n)	G(n)	D(n)	G(n)
Naive Weekly	422.7	118.5	747.3	214.3	19.8	50.3
LSTM	**312.2**	105.5	**533.6**	179.7	14.1	29.9
CNN-LSTM	315.8	**102.5**	587	184.7	**13.9**	**27.4**
SARIMAX	384.1	113	542	**175.7**	29.2	37.4
Prophet	387.1	122	553.8	177.5	26	31

consumed by the EWH from the grid, it was found that 12.33 kWh (88 %) were used during the off-peak period (preheating), with an associated cost of €1.53. The remaining 1.37 kWh (12 %), with a cost of €0.35, were consumed during peak and half-peak periods due to: i) internet connection failures with the SWHC, resulting in local control by the EWH thermostat; ii) incompatibility between the photovoltaic surplus during consecutive time intervals; and iii) delays in sending and receiving commands from the server.

The import can eventually penalize the CO_2 emissions indicator if the energy matrix relies heavily on fossil fuels (i.e., diesel, gas, coal, etc.). During the analyzed period, the emissions recorded by the local utility represented a value of 0.45 $kgCO_2/kWh$, corresponding to 6.3 kgCO2 due to the 14 kWh imported by the electric water heater. When compared to the gas water heater solution (see Fig. 14, right axis), at its efficiency extremes, this reduction represented 6.42 $kgCO_2$ (−50.47 %)



a)



b)

Fig. 16. Comparation between forecasting methods' output and collected real data during the validation period (from May 01, 2024 to May 15, 2024). a) Electric energy consumption. b) PV generation.

Table 8
Comparison of the daily hot water consumption forecasting performance of the chosen models for the period between May 01, 2024 and May 15, 2024.

Model	Ground truth		DHW energy, $\widehat{Q}(d)$
	MAE (m^3)	RMSE (m^3)	MAPE (%)
Naive daily	0.042	0.049	48.8
SMA	**0.022**	**0.028**	**30.8**

to 0.063 $kgCO_2$ (−1%).

3.4. Assessment of forecasting performance

In this subsection, the performance of the developed forecasting methods is assessed over the previously described 15-day validation period during May 2024, and during the chosen timeframe for the case study (July 2024). It is important to note that the validation had to be conducted in May 2024 due to the need of choosing the forecasting models (CNN-LSTM, in this case) before the real-world deployment of the solution under context of IANOS H2020 project.

3.4.1. Prior to case study validation

Table 7 summarizes the results from the walk forward validation for both electricity consumption and generation forecasting models. In the case of the consumption forecast, $\widehat{D}(n)$, results show that the LSTM model outperforms the remaining models based on the error in relation to the ground truth, achieving a MAE of 312.2 W, not far from the CNN-LSTM model which has a slightly larger MAE of 315.8 W. From an energy perspective, the model with the better performance changes to the proposed CNN-LSTM model, with a MAPE of 13.9 %. Both SARIMAX and PROPHET models had their energy related metrics larger than the Naive weekly baseline, meaning that their current implementations may not be suitable for forecasting electricity consumption.

As for the generation forecast, $\widehat{G}(n)$, the CNN-LSTM outperformed the remaining models in the MAE, achieving an error of 102.5 W, while the SARIMAX model had the lowest RMSE of 175.7 W. From the surplus energy perspective, both deep learning models achieved a better performance than the remaining. CNN-LSTM achieved a MAPE of 27.4 %, while the second better model, LSTM, had a MAPE of 29.9 %. All models, regarding the surplus energy forecast, improved over the Naive daily baseline. For both generation and consumption forecasts, the CNN-LSTM models minimized the MAPE of the daily surplus energy. Fig. 16 shows the consumption and generation forecasts of the CNN-LSTM models, for the validation period, over their respective ground truths.

Regarding the hot water demand forecasting, Table 8 also features a comparison between the simple moving average (SMA) model and a naive daily baseline, which in this context is the equivalent to a random walk model. The moving average improves over the chosen baseline in all considered metrics, particularly with a MAE of 0.022 m^3 or 22 L. From the perspective of the energy associated with the hot water demand, SMA achieves a MAPE of 30.8 %. Fig. 17 presents the total daily hot water consumption and the respective forecast for each day of the considered validation timeframe.

3.4.2. Forecasting performance during the case study

The selected CNN-LSTM models and SMA were re-evaluated for their performance in the context of real-time operation within the case study. Fig. 18 illustrates the day-ahead electricity consumption and generation forecasts, respectively. The results, shown in Tables 9 and 10, indicate a decline in the performance of the electricity consumption forecasting model, with the MAPE increasing from 13.9 % during the validation period to 23.2 % in real-time operation. Conversely, the electricity generation forecasting model exhibited improved performance, with the MAPE decreasing from 27.4 % to 25.7 %.

Hot water consumption forecasting also showed an improvement in performance. However, it is noted that the demand is significantly below the verified in the validation period. Fig. 19 shows the real hot water demand and respective forecast for the considered case-study timeframe. An average 39 L daily water consumption was registered during the case study, while the application of the SMA method resulted on 42 L for the same metric. This figure also presents the daily average inlet water temperature, which varied from 23.7 to 27.6 °C.

The collected forecasting results, despite being acceptable for the proposed solution, which conducted to considerable savings and CO_2 reductions, show that users stochastic behavior can have a significant impact on the performance of the developed forecasting methods. This leaves room for further improvement and reveals the need for longer training data-collection, as the performance of deep learning models, like the one used for the consumption and generation forecasting, tend to improve with larger quantities of data [36]. Additionally, it was observed that the quality of the exogenous forecasting data related to the PV generation highly impacts the performance of the developed generation forecasting method.

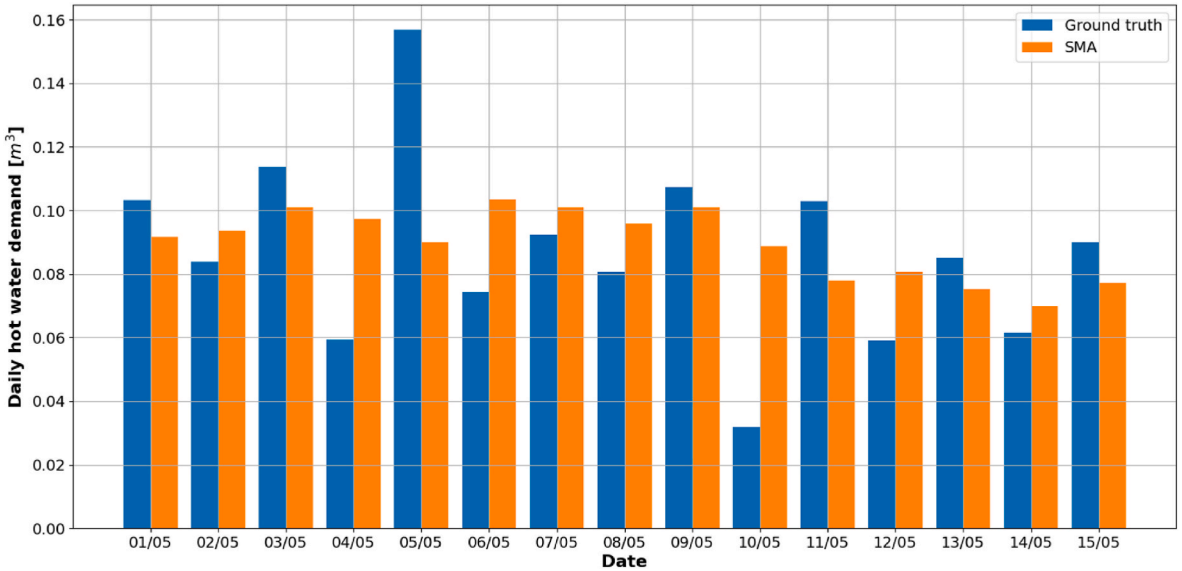


Fig. 17. Daily hot water consumption forecast and real data during the validation period (from May 01, 2024 to May 15, 2024).

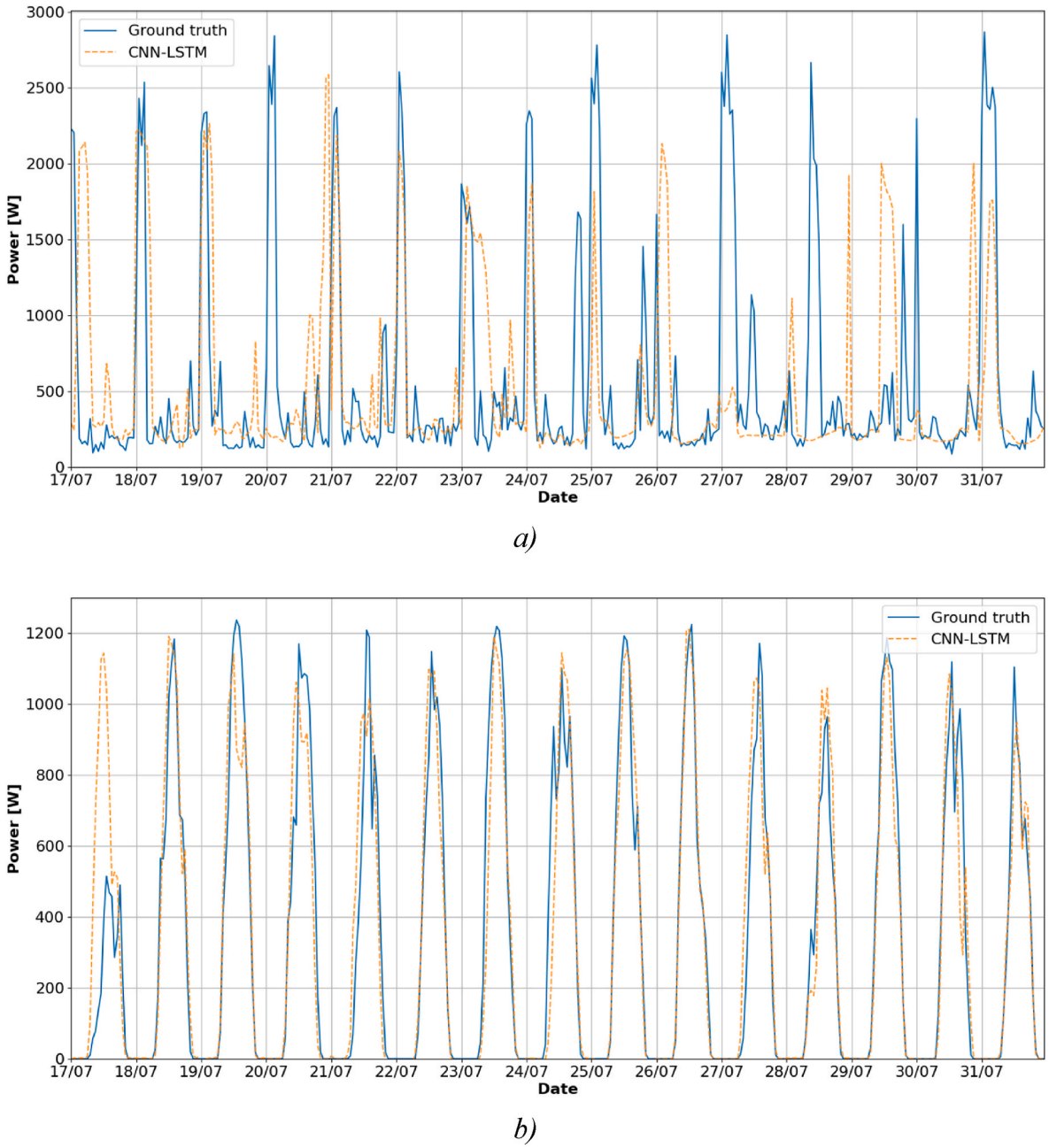


Fig. 18. Comparation between forecasting methods' output and collected real data during the case study (from July 17, 2024 to July 31, 2024). a) Electric energy consumption. b) PV generation.

Table 9
Forecasting performance of the CNN-LSTM models for the period between July 17, 2024 and July 31, 2024.

Model	Ground truth				Surplus energy, $E_s(d)$		
	MAE (W)		RMSE (W)		MAPE (%)		
	$\widehat{D}(n)$	$\widehat{G}(n)$	$\widehat{D}(n)$	$\widehat{G}(n)$	$\widehat{D}(n)$	$\widehat{G}(n)$	$\widehat{S}(n)$
CNN-LSTM	425.5	83.3	752.8	156.6	23.2	25.7	33.8

Table 10
Forecasting performance of the SMA model for the period between July 17, 2024 and July 31, 2024.

Model	Ground truth		DHW energy, $Q(d)$
	MAE (m^3)	RMSE (m^3)	MAPE (%)
SMA	16.2	22.7	29.4

4. Conclusion and future work

This study demonstrates the potential of IoT-based solutions to enhance energy flexibility in residential buildings through smart control of Electric Water Heaters. The deployment on Terceira Island resulted in significant reductions in electricity costs and CO2 emissions. The study presents a new forecasting, and Internet-of-Things based solution to explore energy flexibility from electric water heaters in order to reduce electricity costs and CO_2 emissions associated to energy import from distribution grids. A case study under real-world operation was conducted in Terceira Island, Portugal, where an off-the-shelf electric water

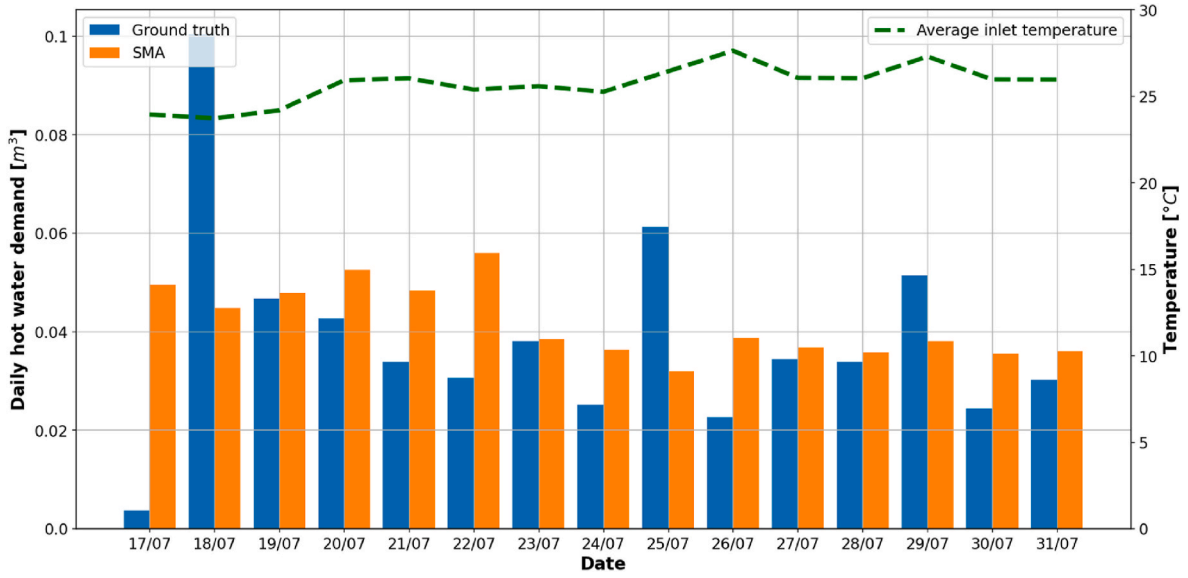


Fig. 19. Daily hot water consumption forecast and real data during the case study (from July 17, 2024 to July 31, 2024).

heater was equipped with the proposed solution to replace an existing gas boiler, increasing the electrification of energy demand in this household.

Relying on a CNN-LSTM based method for energy consumption and generation forecasting, the developed solution conducted to an increase of PV self-consumption (from 31 % to 67 %), while reaching a 75 % self-sufficiency in terms of water heating needs. Savings of €4.47 (−70.4 %) and €1.29 (−40.7 %) were obtained when considering 50 % and 100 % efficiency values for the replaced gas boiler, respectively, during the two weeks of the case study. For the same levels of efficiency, CO_2 reductions of 6.42 kgCO2 (−50.47 %) and 0.063 kgCO2 (−1%) were computed.

The extension of the developed solution from building to community-level, accounting for uncertainty in the development of the EWH scheduling methodology, as well as considering different time periods, represent limitations to be addressed in future works. The integration of the proposed solution in explicit Demand Response programs, while respecting the user needs, is also being considered by the authors for further development. Lastly, the impact of longer training data-collection periods and higher quality exogenous data on the performance of the used deep learning forecasting models should be assessed in the future, as well as the trade-off between performance and computational cost of selected models.

CRediT authorship contribution statement

Rui Amaral Lopes: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Francisco Silva:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Rafael Menezes-Barros:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Nuno Amaro:** Writing – original draft, Methodology, Investigation. **Ana Gonçalves de Carvalho:** Writing – review & editing, Writing – original draft, Project administration. **João Martins:** Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Conceptualization.

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.

Acknowledgements

This research was partly funded by European Union's H2020 programme as part of the IANOS (Integrated Solutions for Decarbonization and Intelligence of Islands) project, Grant Agreement ID 957810), and by the Portuguese "Fundação para a Ciência e a Tecnologia" (FCT) in the context of the Center of Technology and Systems CTS/UNINOVA/FCT/NOVA, reference UIDB/00066/2020. Authors also would like to acknowledge IEA EBC Annex 82 (Energy Flexible Buildings Towards Resilient Low Carbon Energy Systems).

Data availability

The authors do not have permission to share data.

References

[1] International Energy Agency. World energy Outlook 2023 [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2023. [Accessed 30 July 2024].
[2] Lopes RA, Magalhães P, Gouveia JP, Aelenei D, Lima C, Martins J. A case study on the impact of nearly Zero-Energy Buildings on distribution transformer aging. Energy Aug. 2018;157:669–78. https://doi.org/10.1016/j.energy.2018.05.148.
[3] United Nations Environment Programme. Global Status Report for Buildings and Construction: beyond foundations - mainstreaming sustainable solutions to cut emissions from the buildings sector. United Nations Environment Programme Mar. 2024. https://doi.org/10.59117/20.500.11822/45095.
[4] Junker RG, et al. Characterizing the energy flexibility of buildings and districts. Appl Energy Sep. 2018;225:175–82. https://doi.org/10.1016/j.apenergy.2018.05.037.
[5] Le Dréau J, et al. Developing energy flexibility in clusters of buildings: a critical analysis of barriers from planning to operation. Elsevier Ltd.; Dec. 01, 2023. https://doi.org/10.1016/j.enbuild.2023.113608.
[6] Clauß J, Brozovsky J, Georges L. Demonstrating the load-shifting potential of a schedule-based control in a real-life educational building. Energy Build 2024;316 (Aug). https://doi.org/10.1016/j.enbuild.2024.114321.
[7] Dudek E. The flexibility of domestic electric vehicle charging: the electric nation project. IEEE Power Energy Mag Jul. 2021;19(4):16–27. https://doi.org/10.1109/MPE.2021.3072714.
[8] Yildiz B, et al. Analysis of electricity consumption and thermal storage of domestic electric water heating systems to utilize excess PV generation. Energy 2021;235 (Nov). https://doi.org/10.1016/j.energy.2021.121325.
[9] Pereira TC, Lopes RA, Martins J. Exploring the energy flexibility of electric water heaters. Energies Dec. 2019;13(1). https://doi.org/10.3390/en13010046.
[10] Menezes-Barros R, Lopes RA, Martins J. Promoting decarbonization of islands: a case study on the replacement of gas water heaters in Terceira island, Azores, Portugal. 2024. p. 254–66. https://doi.org/10.1007/978-3-031-63851-0_18.
[11] Clift DH, Stanley C, Hasan KN, Rosengarten G. Assessment of advanced demand response value streams for water heaters in renewable-rich electricity markets. Energy 2023;267(Mar). https://doi.org/10.1016/j.energy.2022.126577.

[12] Tejero-Gómez JA, Bayod-Rújula AA. Energy management system design oriented for energy cost optimization in electric water heaters. Energy Build Jul. 2021;243. https://doi.org/10.1016/j.enbuild.2021.111012.
[13] Laguili O, Eynard J, Grieu S. Nonlinear model-based predictive control of electric water heaters in individual dwellings. In: 2024 IEEE congress on evolutionary computation (CEC). IEEE; Jun. 2024. p. 1–8. https://doi.org/10.1109/CEC60901.2024.10611874.
[14] Clarke T, Slay T, Eustis C, Bass RB. Aggregation of residential water heaters for peak shifting and frequency response services. IEEE Open Access Journal of Power and Energy 2020;7:22–30. https://doi.org/10.1109/OAJPE.2019.2952804.
[15] Lakshmanan V, Sæle H, Degefa MZ. Electric water heater flexibility potential and activation impact in system operator perspective – Norwegian scenario case study. Energy 2021;236(Dec). https://doi.org/10.1016/j.energy.2021.121490.
[16] Gough M, Rakhsia K, Bandeira T, Amaro H, Castro R, Catalão JPS. Design and implementation of a data-driven intelligent water heating system for an island community: a case study. Energy Convers Manag 2023;285(Jun). https://doi.org/10.1016/j.enconman.2023.117007.
[17] IANOS H2020 Project. Integrated solutions for decarbonisation and smartification of islands [Online]. Available: https://ianos.eu. [Accessed 6 August 2024].
[18] Groppi D, Feijoo F, Pfeifer A, Garcia DA, Duic N. Analyzing the impact of demand response and reserves in islands energy planning. Energy Sep. 2023;278. https://doi.org/10.1016/j.energy.2023.127716.
[19] Espressif systems, "ESP32 datasheet," [Online]. Available: https://espressif.com/documentation/esp32-wroom-32e_esp32-wroom-32ue_datasheet_en.pdf; 2023.
[20] Ametherm. NTC thermistor probe [Online]. Available: https://www.ametherm.com/thermistor/ntc-thermistors-panw. [Accessed 6 August 2024].
[21] Texas Instruments, "ADS1115 datasheet." Accessed: August. 6, 2024. [Online]. Available: https://www.ti.com/lit/ds/symlink/ads1115.pdf?ts=1713180304842&ref_url= https%253A%252F%252Fwww.google.com%252F.
[22] Robert mauser lda, "YF-S201." [Online]. Available: https://mauser.pt/catalog/product_info.php?products_id=095-0376. [Accessed 6 August 2024].
[23] Shelly Europe Ltd. Energy metering [Online]. Available: https://www.shelly.com/en/products/energy-metering. [Accessed 6 August 2024].
[24] Alhussein M, Aurangzeb K, Haider SI. Hybrid CNN-LSTM model for short-term individual household load forecasting. IEEE Access 2020;8:180544–57. https://doi.org/10.1109/ACCESS.2020.3028281.
[25] Rafi SH, Al-Masood N, Deeba SR, Hossain E. A short-term load forecasting method using integrated CNN and LSTM network. IEEE Access 2021;9:32436–48. https://doi.org/10.1109/ACCESS.2021.3060654.
[26] Agga A, Abbou A, Labbadi M, El Houm Y, Ou Ali IH. CNN-LSTM: an efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Elec Power Syst Res Jul. 2022;208:107908. https://doi.org/10.1016/j.epsr.2022.107908.
[27] Shen G, Lee ZE, Amadeh A, Zhang KM. A data-driven electric water heater scheduling and control system. Energy Build Jul. 2021;242:110924. https://doi.org/10.1016/j.enbuild.2021.110924.
[28] Shaad M, Momeni A, Diduch CP, Kaye ME, Chang L. Forecasting the power consumption of a single domestic electric water heater for a direct load control program. In: 2015 IEEE 28th Canadian conference on electrical and computer engineering (CCECE). IEEE; May 2015. p. 1550–5. https://doi.org/10.1109/CCECE.2015.7129511.
[29] EDA - Electricidade dos Açores. Caracterização das redes de transporte e distribuição de energia elétrica em 2023 (in Portuguese) [Online]. Available: https://www.eda.pt/EDA/DocsDistribuicao/CARE%202023.pdf. [Accessed 6 August 2024].
[30] Luthander R, Widén J, Nilsson D, Palm J. Photovoltaic self-consumption in buildings: a review. Elsevier Ltd.; Mar. 05, 2015. https://doi.org/10.1016/j.apenergy.2014.12.028.
[31] EDA - Electricidade dos Açores. A Origem da Nossa Energia em 2023 (in Portuguese) [Online]. Available: https://www.eda.pt/Regulacao/Rotulagem/Paginas/Folhetos.aspx. [Accessed 8 August 2024].
[32] Portgás. Equivalências energéticas (in Portuguese) [Online]. Available: https://www.portgas.pt/profissionais/apoio/equivalencias-energeticas/. [Accessed 6 August 2024].
[33] Açores SodiGás. As nossas garrafas de gás (in Portuguese) [Online]. Available: https://www.sodigasacores.pt/gas-bottles. [Accessed 8 August 2024].
[34] Azores Government. Caracterização e Diagnóstico - Terceira (in Portuguese) [Online]. Available: https://servicos-sraa.azores.gov.pt/grastore/DRA/PGRHA_2022-2027/PGRH-A_2022-2027_RT_Parte2_VOL3-TER.pdf. [Accessed 6 August 2024].
[35] Azores Government. Electricity tariffs (in Portuguese) [Online]. Available: https://portaldaenergia.azores.gov.pt/portal/Portals/0/Documentos/folhetos/tarifas.pdf. [Accessed 6 August 2024].
[36] Rajagukguk RA, Ramadhan RAA, Lee H-J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. Energies Dec. 2020;13(24):6623. https://doi.org/10.3390/en13246623.