energies

MDPI

Article

Machine-Learning-Based Prediction of HVAC-Driven Load Flexibility in Warehouses

Farzad Dadras Javan [1,†], Italo Aldo Campodonico Avendano [2,†], Behzad Najafi [1,*], Amin Moazami [2,3] and Fabio Rinaldi [1]

[1] Dipartimento di Energia, Politecnico di Milano, Via Lambruschini 4, 20156 Milano, Italy; farzad.dadras@polimi.it (F.D.J.); fabio.rinaldi@polimi.it (F.R.)
[2] Department of Ocean Operations and Civil Engineering, Faculty of Engineering, NTNU, 6009 Ålesund, Norway; italo.a.c.avendano@ntnu.no (I.A.C.A.); amin.moazami@sintef.no (A.M.)
[3] Department of Architectural Engineering, SINTEF Community, SINTEF AS, Børrestuveien 3, 0373 Oslo, Norway
* Correspondence: behzad.najafi@polimi.it
† These authors contributed equally to this work.

Abstract: This paper introduces a methodology for predicting a warehouse's reduced load while offering flexibility. Physics-based energy simulations are first performed to model flexibility events, which involve adjusting cooling setpoints with controlled temperature increases to reduce the cooling load. The warehouse building encompasses office and storage spaces, and three cooling scenarios are implemented, i.e., exclusive storage area cooling, exclusive office area cooling, and cooling in both spaces, to expand the study's potential applications. Next, the simulation data are utilized for training machine learning (ML)-based pipelines, predicting five subsequent hourly energy consumption values an hour before the setpoint adjustments, providing time to plan participation in demand response programs or prepare for charging electric vehicles. For each scenario, the performance of an Artificial Neural Network (ANN) and a tree-based ML algorithm are compared. Moreover, an expanding window scheme is utilized, gradually incorporating new data and emulating online learning. The results indicate the superior performance of the tree-based algorithm, with an average error of less than 3.5% across all cases and a maximum hourly error of 7%. The achieved accuracy confirms the method's reliability even in dynamic scenarios where the integrated load of storage space and offices needs to be predicted.

Keywords: load forecasting; warehouse buildings; machine learning; flexibility in buildings; demand response; multi-layer perceptron

Citation: Dadras Javan, F.; Campodonico Avendano, I.A.; Najaf, B.; Moazami, A.; Rinaldi, F. Machine-Learning-Based Prediction of HVAC-Driven Load Flexibility in Warehouses. *Energies* **2023**, *16*, 5407. https://doi.org/10.3390/en16145407

Academic Editor: Holger Hesse

Received: 16 June 2023
Revised: 4 July 2023
Accepted: 13 July 2023
Published: 16 July 2023

1. Introduction

The unpredictable nature of renewable energy sources leads to intermittent generation that gives rise to balancing issues between supply and demand in the electrical grid [1]. Therefore, smart grids have been introduced for dynamically balancing the load. Demand Side Management (DSM) is a possible approach to adjust the consumption based on the rate of renewable energy production [2,3]. Efficiency improvements, controlling the consumption load, and the control of distributed energy resources are among DSM's possible actions [4]. Demand-side strategies that incentivize/penalize customers for shifting or reducing their consumption during peak periods are called Demand Response (DR) [5]. DR programs contribute to improving grid reliability, reducing aggregate electricity prices, and providing opportunities for participants to achieve cost savings [6].

Buildings are excellent candidates for contributing to grid stability by offering flexibility due to their considerable energy consumption [7]. Within this setting, a building's capacity to reduce, shed, shift, modulate, or generate electricity provided by on-site distributed energy resources (DERs) is often called demand flexibility or energy flexibility.

The building sector is responsible for a notable share of 36% of global energy consumption, causing about 40% of the total emissions worldwide [8]. Among this load, heating, ventilation, and air conditioning (HVAC) consumption with a substantial share of 38% [9] holds a promising potential to contribute to providing flexibility in buildings. Among different categories of buildings, warehouses, which are the buildings that serve as facilities to receive, store, and dispatch goods [10], are of substantial significance. Warehouses hold a share of 11% of the emissions in the logistics sector [11], where a significant portion of their emissions is due to HVAC consumption [12]. HVAC loads are among the thermostatically controlled loads (TCLs) [13–15] that offer promising potential for delivering demand flexibility to the grid. Despite the continued increase in warehouse demand, the importance of their energy efficiency and consumption is often neglected [16]. Warehouses have a high thermal inertia due to their specific architecture with high ceilings and vast indoor spaces, providing them with a high amount of flexibility [17]. Moreover, enormous uninterrupted rooftop spaces in warehouses provide them with a high capacity for photovoltaic (PV) panel installation [18], which can be manipulated to deliver flexibility to the grid.

Cooling consumption is extremely important among HVAC loads in buildings, given that its demand has doubled since 2000, as shown in the International Energy Agency (IEA) assessments [19]. It has emerged as the fastest-growing demand in buildings, and projections indicate that this demand is expected to triple by 2050 [20–22]. Moreover, advancements in technologies and higher levels of welfare require more refrigeration loads as well as HVAC loads in warehouses. A lack of refrigeration accounts for approximately 20% of total food losses [23], which indicates an upward demand for refrigeration load, especially in developing countries. Moreover, due to the significant share of the cooling load in the warehouse facility's energy consumption, in contrast to the majority of commercial/industrial customers, the corresponding energy demand varies notably with a change in weather conditions. Accordingly, the cooling load holds extreme significance when it comes to warehouses.

Deploying DR strategies has been facilitated through Internet of Things (IoT) technologies and Artificial Intelligence (AI) [24] in recent years. By recognizing patterns and solving non-linear functions, AI and machine learning (ML) can handle extensive data and contribute to complex computations required to enable DR [25,26]. In this context, grid-interactive efficient buildings (GEBs) leverage smart technologies and on-site distributed energy sources to provide DR while optimizing energy costs and meeting occupants' comfort and productivity requirements [27]. Consumption and generation in GEBs can be altered based on signals from the grid [28]. Before receiving a signal, the building follows the typical consumption, also known as the baseline load. In general, the baseline load represents the regular energy consumption of the building in the absence of any DR measures [29]. Upon receiving the grid signal, the load of the building is reduced for a certain period, followed by a rebound effect. Prediction of the load of the building levels allows planning for participation in DR and responding to these signals. Accurate prediction of building energy consumption ensures grid safety and mitigates financial risks in electricity market management [30]. Load predictions can be performed on the baseline load during the flexibility event to quantify the energy flexibility offered by the building to compensate the participants [29]. Numerous works in the literature exist in this field, from averaging methods [31–33], control groups [34], and regression methods [29,35–37]. Alternatively, the load of the building during the demand response can be predicted in advance. Campodonico et al. [38] employed deep learning neural networks to predict the demand flexibility of the HVAC load in a modeled office building. When the penalty-aware signal was received, predictions for up to 3 h in advance were made. Results were reported considering the Mean Absolute Percentage Error (MAPE), with a maximum value of 3.55% reported during testing.

Identified Research Gap and Contributions of the Current Work

Buildings have a high potential to provide flexibility to the grid through their HVAC consumption and thermal mass. Various studies have tried to address this issue in residential and non-residential buildings. However, the flexibility of warehouses needs to be further investigated. No previous work has investigated demand flexibility prediction in warehouses. Furthermore, with the increasing demand for HVAC cooling loads in warehouses, there are opportunities for interventions that can provide flexibility to the grid that need to be further explored. Prediction of hourly demand flexibility in warehouses could provide warehouse managers with opportunities to respond to DR programs, save costs, enhance the comfort of the occupants, and incorporate the charging of electric vehicles (EVs). Moreover, most of the literature on load prediction has utilised a batch learning approach without investigating the effect of gradually incorporating new data.

Motivated by the identified research gap, this work introduces a method to predict the demand flexibility of the HVAC load (reduced electric load under a flexibility event) enabled through adjusting the cooling system's setpoints in a modeled warehouse. Modeling was conducted from one hour prior to the initiation of a flexibility event, extending for up to four hours after its onset. Therefore, the contributions of the present study can be listed as follows:

- Performing physics-based modeling on warehouse buildings, including office spaces, for three different cooling scenarios aiming to cover various case studies.
- Performing adjustments of cooling setpoints to reduce consumption and provide flexibility.
- Implementing ML-based pipelines to perform multi-step predictions on five hourly consumption values of the building one hour ahead of the flexibility event. This allows for the simultaneous prediction of the baseline load (the preceding hour), the load reduction resulting from the setpoint adjustment, and the load behavior for up to four hours after the flexibility event.
- Comparing the performance of a tree-based machine learning algorithm (random forest (RF) regressor) and an ANN algorithm (MLP) in predicting the electrical load under DSM scenarios.
- Employing multi-step forecasting with an expanding window training scheme to emulate the real-case scenarios where flexibility event data are generated progressively over time.

HVAC-based flexibility can effectively be utilized in warehouses for balancing (increasingly utilized) EV vehicle's charging load, along with participation in DR programs and flexibility markets (available for all types of buildings). The growing utilization of electric trucks in the transport/warehousing sector (suggested in decarbonization programs) raises concerns about the impact of the resulting charging load on the grid, especially in real-world case studies where EVs arrive in the evening and require partial charging for nighttime deliveries [39]. The proposed methodology can effectively assess the possibility of charging trucks without causing a surge in consumption by reducing the HVAC load. The results of the implemented methodology can thus proactively recommend the optimal number of trucks and how much charging can be provided to them (in terms of kWh) in advance, owing to the fact that predictions of the expected load reduction are provided an hour before the setpoint adjustments are imposed.

Moreover, DR (and participation in flexibility markets) is becoming increasingly popular in the industrial sector owing to the corresponding notable incentives. However, without a tool that determines the extent to which the setpoint adjustment strategy can reduce the load and the resulting load reduction duration, these facilities cannot effectively participate in these programs. This research work was initiated with the intention of addressing the challenges faced by conditioned warehouse facilities to balance the growing EV charging load while participating in demand response programs.

2. Methodology

This section outlines the details of the methodology for predicting short-term electricity consumption under flexibility schemes in a warehouse in Los Angeles, CA. This study integrates a physics-based energy simulation for the building (Figure 1) with state-of-the-art regressive ML models for multiple forecasting of short-term time series. Accordingly, the details of the case study buildings, the physics-based co-simulation, the adopted cooling scenarios, and the procedure for emulating the flexibility events are presented. Finally, the methodology for predicting five hourly load values an hour ahead of the setpoint adjustment (the baseline consumption an hour before the initiation of the event and the penalty aware load up to 4 h after the flexibility event) using ML pipelines is explained. Additionally, a schematic representation of the adopted methodology is presented in Figure 2.

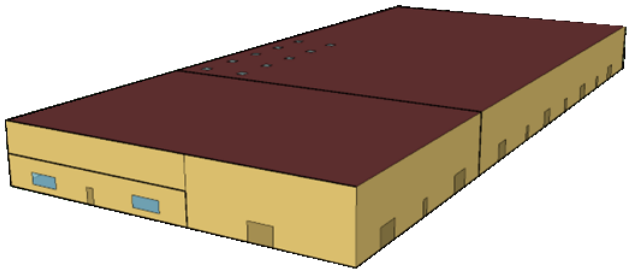

Figure 1. Sample of the warehouse building used in physics-based simulations.

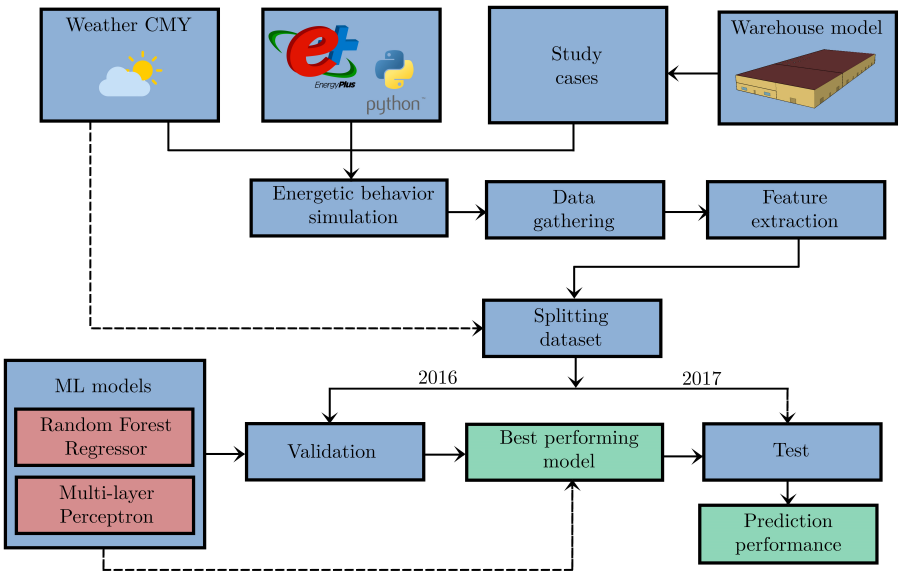

Figure 2. Schematic representation of the proposed methodology.

2.1. Case Study

The reference warehouse building model [40] utilized in the EnergyPlus physics-based co-simulation was developed under the ANSI/ASHRAE/IES Standard 90.1 [41]. A detailed description of the building is provided in Table 1. The simulations are performed for Los Angeles, California, which is classified as Climatic Zone 3B. The warehouse building integrates office space and storage space, and three different cooling scenarios are considered: (1) exclusively providing cooling for the storage space, (2) only cooling the office spaces with storage spaces unconditioned, and (3) conditioning both storage and office spaces. The details of the methodology applied in modeling and emulating the flexibility events are provided in Figure 2.

Table 1. Description of the buildings used in the physics-based simulations in *EnergyPlus*.

Building Type	Warehouse
Location	Los Angeles, CA, USA
Simulation period	May–September
Frequency	15 (min)
Status	New building
Total floor area	4836 (m^2)
Cooling type	Packaged air conditioning unit
Cooling setpoint—offices	24 °C
Cooling setpoint—storage	26 °C
Flexibility temperature threshold	2 °C (±0.2)
Cooling setback—all zones	30 °C

2.2. Physics-Based Energy Simulation

A physics-based co-simulation utilizing EnergyPlus V9.4 [42] and its Python API [43] was conducted to simulate the penalty-aware electricity consumption of the warehouse during a flexibility event. Simulations were performed with a sub-hourly resolution of four timestamps per hour (15 min intervals) during the summer cooling season (from mid-May to the middle of September) of two consecutive simulated years. Simulations were executed using the weather conditions of Climatic Zone 3B, California (LA). The weather file for the exact location was converted from the typical meteorological year (TMY) weather file into the current meteorological year (CMY) for 2016 and 2017 using the *diyepw* tool [44]. This is to ensure the reliability and feasibility of the validation and testing approach and to deploy the expanding window method similar to real-world scenarios. The choice of California as the location of the simulated buildings was firstly motivated by the fact that the flexibility simulation is performed for the cooling systems, which can be implemented in moderately warm areas (the implementation in cold regions is not feasible due to the non-existing/low cooling load, while the implementation in extremely warm/hot areas (e.g., middle east) is not possible as it will result in low durations of flexibility measures). Secondly, the emergency load reduction program of California State [45] is an established demand response program, and the specifications of the performed predictions (up to 5 h) have been partially inspired based on this existing program.

Flexibility events were introduced by temporarily relaxing the cooling setpoints on all working days in the three aforementioned scenarios, resulting in a decreased electricity consumption associated with cooling. In scenario 1, where the cooling load is provided in storage, the flexibility event terminates before the 4 h duration limit in case the temperature in storage exceeds 2 °C above the setpoint. In scenario 2 (conditioned offices only), the maximum duration of the event is up to one hour (between 5 p.m. and 6 p.m.) based on the schedules of the buildings or if the temperature increase exceeds the 2 °C threshold. In scenario 3 (cooling load in both offices and storage), should the temperature increase exceed 2 °C in any of the conditioned zones, cooling is resumed only in that zone, and the flexibility event continues for up to 4 h (between 5 p.m. and 9 p.m.). In the case of the offices, a setpoint flexibility range of 2 °C was suggested in a study conducted by Nicol and Humphreys [46]. In their research, aimed at determining the upper thresholds of maximum temperatures attainable within European (France, Greece, Portugal, Sweden, and the UK) office buildings while maintaining thermal comfort, the researchers noted that with temperature deviations within the range of ±2 K from the established comfort level, a significant majority (exceeding 80% of the subjects) reported experiencing a state of comfort. A similar indication (setpoint flexibility of 2 °C for the cooling season) was also suggested in a study conducted in office buildings in Japan (conducted by Tan et al. [47]). In the case of the warehouse, a conditioned warehouse that is utilized for storing medical devices has been considered in the present study; the default setpoint and the acceptable setpoint flexibility have been both suggested by the energy management team of a warehouse facility for medical devices (based on the required temperature range for the storage of the corresponding devices and field experience). It is also worth mentioning that the setpoint

adjustment procedures simulated in the present study result in the temperature increase reaching the maximum of 2 °C only in a few days, and these modifications are of limited duration, lasting for a maximum of four hours.

2.3. Demand Response Prediction

The following part of this work incorporates collecting the obtained data from the simulation and developing predictive pipelines for flexible load, considering a non-autoregressive approach with multiple hourly predictions (up to 5 h). An RF regressor algorithm was selected owing to the superior performance of tree-based algorithms in load prediction [29], utilizing Python and the Scikit-learn package [48]. Moreover, due to its performance in predicting load under demand response and to provide a comparison with the tree-based algorithm, an Artificial Neuron Network (ANN) was also chosen to perform prediction algorithms [38] employing TensorFlow [49]. Next, the predictive models were trained using the sequential data of electrical consumption, outdoor temperature, and solar radiation from 13:00 to 16:00 (the moment predictions are made). The obtained simulated data have a sampling frequency of 15 min, which is then summed up (for consumption data) or averaged (in the case of outdoor temperature and solar radiation) to achieve a sampling frequency of one hour. Additionally, to capture the cyclic nature of cooling consumption, the time of the year (encoded with a cosine function) is incorporated, resulting in a total of 10 training features. The utilized features in the study are listed in Table 2, where the number of lagged values was determined through a binary approach. Testing various lagged values ranging from 2 to 24 h revealed that increasing the number of features does not substantially improve the prediction performance. Hence, the optimal number of lagged values, providing a balance between prediction accuracy and computational costs, has been included in the analysis.

Table 2. The list of the features used for prediction.

Features	Values
Day of Year	-
External Temperature	Last 3 hourly values before the prediction moment
Radiation	Last 3 hourly values before the prediction moment
Electrical Consumption	Last 3 hourly values before the prediction moment

Training is performed using an expanding window training scheme as explained in Section 3.3, similar to the actual deployment scenario, where new data are incorporated each time the building undergoes a flexibility event (online learning). These new data are included in retraining the model. The results obtained from the first year of simulation (2016) are employed for validation, while the following year's data are used for testing. The validation process is initiated when ten flexibility events are gathered. New data are included in subsequent predictions until the end of the first year, with the total length of the training data set (first year) incorporating 89 flexibility events. Moreover, the testing commenced with the data available from the first year, incorporating 89 testing data points (second year) with a similar approach. Then, the model was retrained, incorporating the new data points after making each test prediction, and the average prediction accuracy in the second year is reported. Moreover, the number of estimators in the RFR was set to 700 and the number of epochs in the MLP was equal to 300.

3. Complementary Concepts

This section will present an overview of the concepts utilized in the development of this study, including the machine learning algorithms, training methodology, and the metrics employed to assess the reliability of the models.

3.1. Random Forest Regressor

As a machine learning algorithm, *random forest* utilizes an ensemble of decision trees to generate predictions. Training each tree on a different random sample of the data makes the final prediction more accurate and less prone to overfitting. This approach creates a robust predictor less sensitive to specific training data and provides more generalized insights [50,51].

3.2. Multi-Layer Perceptron

MLP is a type of ANN that uses multiple interconnected layers with many neurons to uncover hidden relationships between the outputs and features used for training the model [52,53]. During the training of the MLP model, data are passed through hidden layers until they reach the output layer. The weights of each neuron are then calculated through forward propagation, and the model is optimized using Stochastic Gradient Descent (SGD) with backward propagation. The loss gradient is recalculated multiple times to improve the accuracy and minimize the loss.

3.3. Multi-Step Forecasting with Expanding Window Training Scheme

The sequential nature of the flexibility events and load behavior calls for a time series approach in training the predictive models and assessing their performance. Additionally, in real-case scenarios, new data are introduced progressively as the building undergoes more flexibility events, thus returning more training opportunities to the model. This behavior can be represented in the validation and testing process by the expanding window scheme [54]. Moreover, incorporating new data and retraining allows the model to include recent changes in the behavior of the load or environmental features in its predictions. This study's prediction output consists of five consecutive time stamps representing the local time series of electrical consumption. Therefore, multi-step forecasting [55,56] is utilized to achieve this in addition to the expanding window training scheme.

3.4. Evaluation Metrics

The current work takes advantage of three different accuracy metrics (Mean Absolute Percentage Error (MAPE), Normalized Mean Bias Error (NMBE), and the Coefficient of Variation of the Root Mean Squared Error (CV(RMSE))) to report and benchmark the forecasting results.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \tag{1}$$

$$NMBE = \frac{100\%}{\bar{y}} \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \tag{2}$$

$$CV(RMSE) = \frac{100\%}{\bar{y}} \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-1}} \tag{3}$$

The MAPE measures the average prediction error and considers significant variations without being influenced by the output scale. This allows for comparing different case studies with varying amounts of electrical load. ASHRAE Guideline 14 [57,58] recommends using NMBE and CV(RMSE) as benchmarking metrics. The NMBE measures the average error of a sample, with negative values suggesting over-predictions and positive values indicating under-predictions. However, if there are errors that can cancel each other out, the NMBE may not accurately reflect the overall error. To address this, the CV(RMSE) is used to measure the model's ability to predict the overall data pattern regardless of error cancellation [57–59]. Moreover, y_i in these equations denotes the actual values in the dataset, $\hat{y}$ represents the predicted value, and $\bar{y}$ represents the mean of the actual values in the dataset. Finally, i returns the count of the element in the dataset.

As per ASHRAE Guideline 14, hourly predictions of a model are considered calibrated if the NMBE value falls within the range of ±10%, while the maximum acceptable error for CV(RMSE) is below 30%. Additionally, the International Performance Measurement and Verification Protocol (IPMVP) [60] establishes more strict performance evaluation standards, with the maximum acceptable threshold of ±5% for NMBE and indicating that CV(RMSE) values must not exceed 20%.

4. Results and Discussion

This section initially details the results of physics-based simulations for the warehouse building regarding three studied cooling scenarios: (1) only in the storage area, (2) only in the offices, and (3) in both storage and offices. The effect of the proposed setpoint adjustment strategy and the resulting reduction in the cooling load is illustrated and discussed. Next, the results of predictive models for multi-step load forecasting under a demand flexibility event are provided and discussed. The considered scenarios assume a flexibility measure will be imposed an hour after the moment that the predictions occur. The five hourly forecasted loads include an hour ahead baseline consumption and four consecutive hours of reduced load under the flexibility event. The performance of RFR, a tree-based machine learning algorithm, and an ANN algorithm (MLP) are then compared in terms of the MAPE during the validation process. Finally, the best-performing algorithm is selected for each study case, and the results are presented in terms of the CV(RMSE) and NMBE.

4.1. Flexibility Events Simulations

As explained in Section 2.2, flexibility events are simulated in a warehouse building through cooling setpoint adjustments for three different cooling scenarios. Figure 3 presents the energetic behavior of the investigated cases under moderate environmental conditions. As previously stated, the offices' schedule terminates the cooling load at 18:00, which is why the cooling setpoint increases to 30 °C. A comparison of the first and third cases clearly demonstrates that the cooling load of the storage spaces significantly exceeds that of the offices. The flexibility provided by the office lasts for a short time, as the temperature in thte offices quickly exceeds the threshold after 15 min. However, also providing cooling in the storage spaces, as in case 3, extends this duration, and the temperature increase in the offices remains within the accepted threshold for 45 min, followed by the conclusion of the working hours in the office.

On the other hand, the temperature in the storage space during this specific day does not exceed the established limit in both cases 1 and 3. Yet, the reduction in the load compared to the baseline occurs mainly in the first hour after the flexibility event starts, as the baseload consumption (even without flexibility measures) would have also become zero after a particular time owing to the environmental conditions.

A visual representation of the trend in energy consumption under the flexibility event for a sample day characterized by more severe weather conditions is provided in Figure 4. Higher temperatures and radiation during this day extend the demand for cooling in the building into the later evening hours, potentially benefiting more from the proposed setpoint adjustment strategy. This is evident in cases 1 and 3, where there is a notable disparity between the baseline and flexible consumption. The thermal mass of the storage spaces provides flexibility, resulting in longer hours of deactivated cooling, while the demand for cooling is evident in the baseload. Nevertheless, the offices follow a similar trend in energy consumption, while a more significant consumption is generally observed due to extreme environmental conditions.

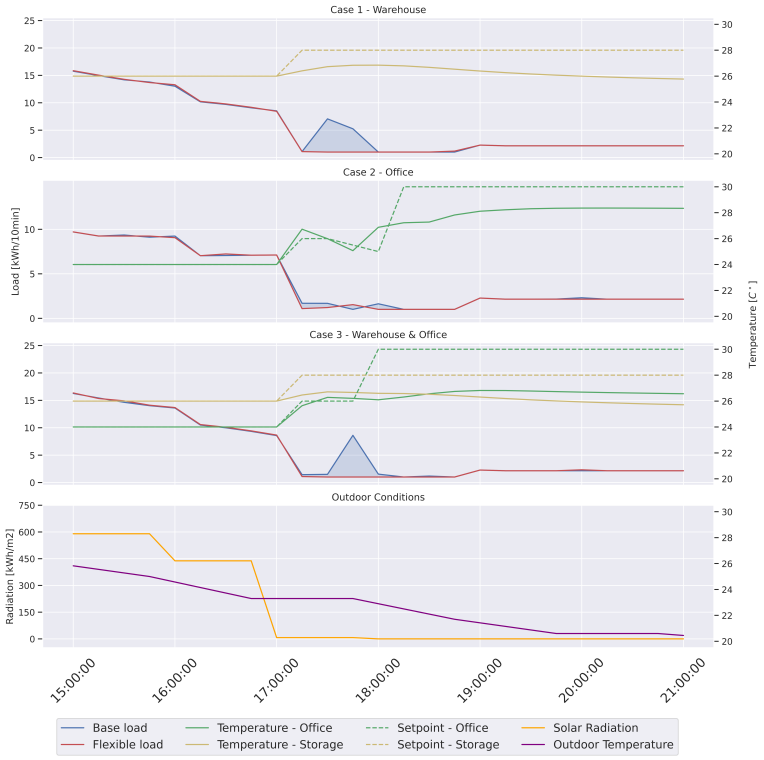

Figure 3. Flexibility event on 07–20 that represents the behavior of all cases under normal environmental conditions.

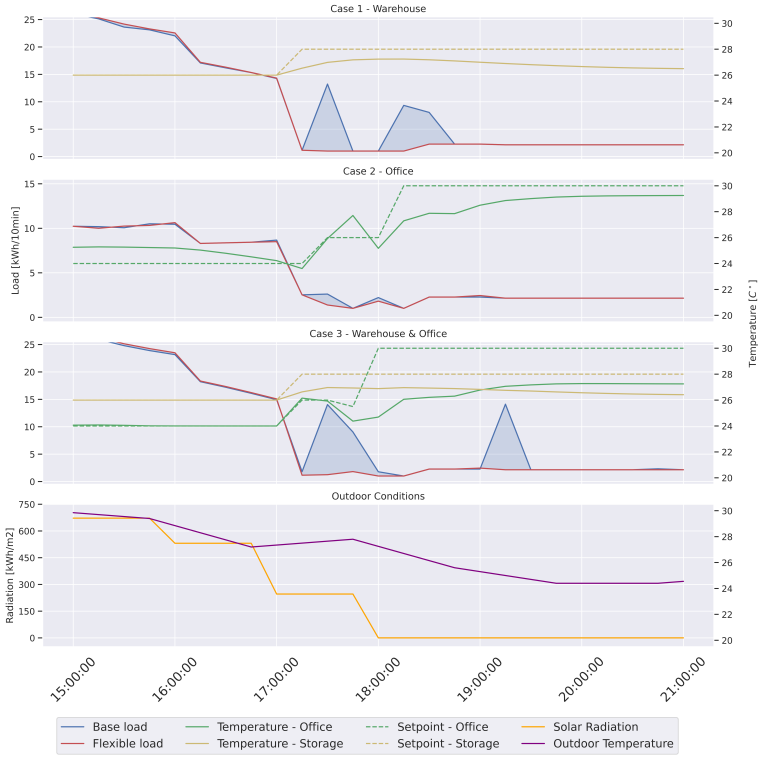

Figure 4. Flexibility event on 09–01 that represents the behavior of all cases under extreme environmental conditions.

4.2. Predictive Modeling

The following subsections present a comprehensive assessment of the performance of the predictive models (RFR and MLP) in forecasting the electrical load during the simulated flexibility events.

4.2.1. Case 1—Cooling in the Storage Area Only

The accuracy obtained by the two predictive algorithms during the validation process (explained in Section 2.3) is provided in Table 3. The hourly load is forecasted up to 5 h in advance, and the resulting hourly and average accuracy is reported. "Hour 0" signifies the one-hour ahead baseline consumption before the flexibility event starts. "Hour 1" coincides with the reduced consumption due to the setpoint adjustment 2 h ahead of the prediction timestamp and 1 h into the flexibility event, and so on. Accordingly, temperature could exceed the defined threshold anytime after the initiation of the event, and the transition back to typical consumption can occur at any of the subsequent timestamps. The superior performance of RFR is evident at all the timestamps (with an average MAPE of 3.95%) compared to the MLP. Additionally, it can be observed that the prediction of the second hour of the flexibility event, "Hour 2", is specifically more challenging for both prediction models. This can be attributed to the transitions into normal behavior being mainly experienced during this hour. Overall, it can be concluded that RFR performs reliably in predicting the hour-ahead baseline load, the consumption reduction due to setpoint adjustments, the transition to normal consumption, and the rebound effect when the flexibility event terminates. Overall, it performs well in predicting the load behavior during the flexibility event from an hour before it starts.

Table 3. MAPE accuracy for validation process in Case 1—only storage.

MAPE (%)	Hour 0	Hour 1	Hour 2	Hour 3	Hour 4	Total
RF Regressor	4.49	4.32	7.55	2.89	0.47	3.95
MLP	6.32	7.42	14.43	11.82	8.60	9.72

Testing, which involves the second year data, is achieved by employing the best-performing algorithm from the validation phase. An improvement is evident in most prediction horizons compared to the previous phase, as shown in Table 4. This could be credited to the expanding window training method adopted in this work, which provides an increased training opportunity for the algorithms, including a broader scope of external temperatures and solar radiation. Additionally, considering the CV(RMSE) and NMBE, it can be observed that the obtained results fall within the thresholds defined by the ASHRAE guideline 14 and IPMVP explained in Section 3.4. Therefore, the models are considered calibrated.

Table 4. Test accuracy metrics (%) for Case 1—only storage.

Metric	Hour 0	Hour 1	Hour 2	Hour 3	Hour 4	Total
MAPE	3.24	3.22	6.92	2.34	0.59	3.26
CV(RMSE)	12.83	6.94	15.81	3.91	0.88	15.46
NMBE	−1.18	0.36	−4.70	−0.61	−0.05	−0.97

4.2.2. Case 2—Only Office

In the second scenario (Case 2), where the cooling load is provided in the office spaces only, the predictions are limited to two consecutive timestamps as the flexibility event terminates after an hour due to the schedule (internal gain and lighting calendars), which ends at 18:00. Therefore, "Hour 0" consists of the prediction of the normal consumption of the building in the next hour, at the end of which the setpoint adjustments lead to

load reductions. Regarding the performance of models on the validation set, a notable improvement is observed in general compared to the previous case (Table 5). The behavior observed in Section 4.1, where the office's low thermal inertia leads to a more predictable load behavior, can be attributed as the underlying cause. Moreover, RFR outperforms the MLP model in all prediction horizons, similar to the previous case study.

Table 5. MAPE accuracy for validation process in Case 2—only office.

MAPE [%]	Hour 0	Hour 1	Total
RF Regressor	1.38	1.89	1.63
MLP	1.99	3.69	2.84

Table 6 presents the performance of RFR on the test dataset (second year), where a further improvement in the MAPE is observed for both prediction horizons. Similar to the previous case, the reason for this is incorporating additional training points using the expanding window training scheme. Moreover, models are considered calibrated based on ASHRAE 14 and IPMVP.

Table 6. Test accuracy metrics [%] for Case 2—only office.

Metric	Hour 0	Hour 1	Total
MAPE	0.97	1.72	1.35
CV(RMSE)	1.53	3.62	2.08
NMBE	−0.30	−0.34	−0.31

4.2.3. Case 3—Both Storage and Office

The results obtained in the validation process for the third case that provides cooling both in the storage and office area are presented in Table 7. Similarly, the office working schedule concludes after the first hour of the flexibility event (corresponding to the second hour of the prediction horizon), and the cooling system ceases operation in those zones. However, the storage area has uninterrupted working hours, indicating that flexibility can continue until 21:00, the end of investigated hours. The obtained results indicate a notably superior performance of RFR (MAPE score of 4.18%) compared to the MLP model in all prediction horizons. Moreover, a lower prediction accuracy in the second hour is observed (especially in the case of MLP with 18.61%), which can be explained by the transient behaviors caused by the termination of office working hours and rebound effects in the storage spaces.

Table 7. MAPE accuracy for validation process in Case 3—both storage and office.

MAPE [%]	Hour 0	Hour 1	Hour 2	Hour 3	Hour 4	Total
RF Regressor	4.88	4.63	7.94	2.94	0.51	4.18
MLP	9.77	9.01	18.61	12.45	11.06	12.18

Finally, the results of predictions of RFR on the test set for case 3 are delivered in Table 8. Expectedly, improvements in the forecasts are observed as the testing is performed in the second year with an expanding window. Consequently, including new data points enhances the training process, providing more training opportunities for the models. Hence, an impressive MAPE score of 3.26% achieved during the testing of the third case demonstrates the model's reliability in forecasting loads during demand response scenarios, even in dynamic situations that involve zones with different thermal behaviors and schedules. Moreover, the discussion regarding ASHRAE 14 and IPMVP calibration criteria is valid here, as the results align with the predefined thresholds.

Table 8. Test accuracy metrics (%) for Case 3—both storage and office.

Metric	Hour 0	Hour 1	Hour 2	Hour 3	Hour 4	Total
MAPE	3.38	3.45	6.82	2.38	0.55	3.32
CV(RMSE)	11.95	7.61	15.33	3.98	0.85	14.61
NMBE	−1.17	0.26	−4.23	−0.79	−0.03	−0.97

5. Conclusions

The performance of ML-based pipelines in performing short-term predictions on reduced load, under flexibility events, in a modeled warehouse building was assessed in the current work. Physics-based energy simulations were performed using EnergyPlus on a modeled warehouse in Los Angeles, California, under a summer cooling scenario. Flexibility events were imposed using setpoint adjustments in the cooling system at 17:00, leading to a reduced electrical consumption while containing the temperature increase within a limited threshold of 2.1 °C. Moreover, the maximum duration of the flexibility event was set to 4 h. Three different cases were introduced, broadening the applicability of the proposed approach to various cooling scenarios. The first case considered the cooling demand only in the storage areas, while the second case included non-conditioned storage with cooling provided for the offices. Finally, the last case considered the cooling load in both storage and office zones. The results of the simulation were employed to train machine-learning-based pipelines to predict (at 16:00 with a total of five outputs) the five hourly values of the building's consumption, including one hour before the flexibility event and up to four hours after the setpoint adjustment. Therefore, the first hour prediction contained the baseline load consumption of the building at the end of which flexibility measures are imposed (through setpoint adjustment) and the subsequently reduced consumption and rebound effect during the next four hours after the initiation of the flexibility event. Two different forecasting pipelines of RFR and MLP were implemented to compare their performance. Moreover, simulations were performed for two consecutive years and an expanding window scheme was selected as the training method. The first year simulations were employed for validation and second year data were used for testing.

The obtained results indicated that in all the simulated case studies, RFR outperformed the MLP algorithm. Therefore, it can be inferred that the tree-based machine learning algorithm is better suited than the MLP model in this application, which involves expanding windows and limited observations. Moreover, it was shown that the forecasting pipelines with RFR could achieve an average error of less than 3.5% in all three cases with a maximum error of 7%. The accuracy achieved, particularly in the third scenario that incorporated both offices and storage spaces with distinct thermal behavior and demands, demonstrates the reliability of the proposed approach and the implemented machine-learning-based pipelines. It was also observed that flexibility events in case 2 (only offices) would terminate quickly owing to the low thermal inertia observed in the simulation stage, leading to less transient behavior and a higher prediction accuracy for that specific case. Next, the results were reported considering error metrics suggested by ASHRAE standard 14 and IPMVP for calibrated models and were shown to fall within the defined threshold. Furthermore, it was shown that expanding window training that gradually incorporates new data points to emulate online learning enhances the model's performance consistently as new training opportunities are introduced and also allows retraining the models on the latest shifts in load behavior or environmental characteristics.

Including the extra hour of forecasting ahead of the flexibility event, involving setpoint adjustment, establishes the baseline load consumption in the next hour, after which the flexibility measures are imposed. This will provide an accurate assessment of the load reduction that can be obtained through the proposed setpoint adjustment. Additionally, predicting load values an hour in advance, with the accuracy offered in the current work, provides a reliable tool and sufficient planning time for integrating EV charging or planning to engage in demand response programs. Moreover, the load values for up to four hours

after initiating the flexibility measure represent the rebound effect and the resulting load penalty, which can effectively be used in the decision-making mechanism.

Author Contributions: F.D.J.: software, formal analysis, methodology, validation, data curation, writing—original draft; I.A.C.A.: software, formal analysis, methodology, validation, data curation, writing—original draft; B.N.: conceptualization, methodology, supervision, writing—review and editing; A.M.: conceptualization, methodology, supervision, writing—review and editing; F.R.: supervision, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: The present study was financially supported by National Centre for Sustainable Mobility, 10th Spoke (Centro Nazionale per la Mobilità Sostenibile, spoke 10) financed by the Italian National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza).

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

References

1. Li, H.; Wang, Z.; Hong, T.; Piette, M.A. Energy flexibility of residential buildings: A systematic review of characterization and quantification methods and applications. *Adv. Appl. Energy* **2021**, *3*, 100054. [CrossRef]
2. Ramchurn, S.D.; Vytelingum, P.; Rogers, A.; Jennings, N.R. Putting the'smarts' into the smart grid: A grand challenge for artificial intelligence. *Commun. ACM* **2012**, *55*, 86–97. [CrossRef]
3. Wijaya, T.K.; Larson, K.; Aberer, K. Matching demand with supply in the smart grid using agent-based multiunit auction. In Proceedings of the 2013 Fifth International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 7–10 January 2013; pp. 1–6.
4. Palensky, P.; Dietrich, D. Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE Trans. Ind. Inform.* **2011**, *7*, 381–388. [CrossRef]
5. Liu, Z.; Zeng, X.; Meng, F. An integration mechanism between demand and supply side management of electricity markets. *Energies* **2018**, *11*, 3314. [CrossRef]
6. Siano, P. Demand response and smart grids—A survey. *Renew. Sustain. Energy Rev.* **2014**, *30*, 461–478. [CrossRef]
7. Aduda, K.; Labeodan, T.; Zeiler, W.; Boxem, G.; Zhao, Y. Demand side flexibility: Potentials and building performance implications. *Sustain. Cities Soc.* **2016**, *22*, 146–163. [CrossRef]
8. Khalil, M.; McGough, A.S.; Pourmirza, Z.; Pazhoohesh, M.; Walker, S. Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption—A systematic review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105287. [CrossRef]
9. González-Torres, M.; Pérez-Lombard, L.; Coronel, J.F.; Maestre, I.R.; Yan, D. A review on buildings energy information: Trends, end-uses, fuels and drivers. *Energy Rep.* **2022**, *8*, 626–637. [CrossRef]
10. Sarkis, J. A strategic decision framework for green supply chain management. *J. Clean. Prod.* **2003**, *11*, 397–409. [CrossRef]
11. Doherty, S.; Hoyle, S. *Supply Chain Decarbonization: The Role of Logistics and Transport in Reducing Supply Chain Carbon Emissions*; World Economic Forum: Geneva, Switzerland, 2009.
12. Ries, J.M.; Grosse, E.H.; Fichtinger, J. Environmental impact of warehousing: A scenario analysis for the United States. *Int. J. Prod. Res.* **2017**, *55*, 6485–6499. [CrossRef]
13. Perfumo, C.; Kofman, E.; Braslavsky, J.H.; Ward, J.K. Load management: Model-based control of aggregate power for populations of thermostatically controlled loads. *Energy Convers. Manag.* **2012**, *55*, 36–48. [CrossRef]
14. Najafi, B.; Di Narzo, L.; Rinaldi, F.; Arghandeh, R. Machine learning-based disaggregation of airconditioning loads using smart meter data. *IET Gener. Transm. Distrib.* **2020**, *14*, 4755–4762. [CrossRef]
15. Manivannan, M.; Najafi, B.; Rinaldi, F. Machine learning-based short-term prediction of air-conditioning load through smart meter analytics. *Energies* **2017**, *10*, 1905. [CrossRef]
16. Park, S.Y.; Cho, S.; Ahn, J. Improving the quality of building spaces that are planned mainly on loads rather than residents: Human comfort and energy savings for warehouses. *Energy Build.* **2018**, *178*, 38–48. [CrossRef]
17. Akerma, M.; Hoang, H.M.; Leducq, D.; Flinois, C.; Clain, P.; Delahaye, A. Demand response in refrigerated warehouse. In Proceedings of the 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA, 16–19 September 2018; pp. 1–5.
18. FacilitiesNet. Solar PV Rooftop Systems Can Help Defray Warehouse Electrical Costs. n.d. Available online: https://www.facilitiesnet.com/energyefficiency/article/Solar-PV-Rooftop-Systems-Can-Help-Defray-Warehouse-Electrical-Costs--13318 (accessed on 4 May 2023).
19. Agency, I.E. The Future of Cooling: Opportunities for Energy-Efficient Air Conditioning. 2018. Available online: https://www.iea.org/reports/the-future-of-cooling (accessed on 4 May 2023).
20. Hu, S.; Yan, D.; Qian, M. Using bottom-up model to analyze cooling energy consumption in China's urban residential building. *Energy Build.* **2019**, *202*, 109352. [CrossRef]
21. Najafi, B.; Depalo, M.; Rinaldi, F.; Arghandeh, R. Building characterization through smart meter data analytics: Determination of the most influential temporal and importance-in-prediction based features. *Energy Build.* **2021**, *234*, 110671. [CrossRef]

22. Raymand, F.; Najafi, B.; Mamaghani, A.H.; Moazami, A.; Rinaldi, F. Machine Learning-based Estimation of Buildings' Characteristics Employing Electrical and Chilled Water Consumption Data: Pipeline Optimization. *Energy Build.* **2023**, *295*, 113327. [CrossRef]
23. Zilio, C. Moving toward sustainability in refrigeration applications for refrigerated warehouses. *HVAC&R Res.* **2014**, *20*, 1–2.
24. Dadras Javan, F.; Khatam Bolouri Sangjoeei, H.; Najafi, B.; Mamaghani, A.H.; Rinaldi, F. Application of Machine Learning in Occupant and Indoor Environment Behavior Modeling: Sensors, Methods, and Algorithms. In *Handbook of Smart Energy Systems*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–25.
25. Sharda, S.; Singh, M.; Sharma, K. Demand side management through load shifting in IoT based HEMS: Overview, challenges and opportunities. *Sustain. Cities Soc.* **2021**, *65*, 102517. [CrossRef]
26. Antonopoulos, I.; Robu, V.; Couraud, B.; Kirli, D.; Norbu, S.; Kiprakis, A.; Flynn, D.; Elizondo-Gonzalez, S.; Wattam, S. Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renew. Sustain. Energy Rev.* **2020**, *130*, 109899. [CrossRef]
27. Neukomm, M.; Nubbe, V.; Fares, R. *Grid-Interactive Efficient Buildings*; US Dept. of Energy (USDOE): Washington, DC, USA; Navigant Consulting. Inc.: Chicago, IL, USA, 2019.
28. Reynders, G.; Lopes, R.A.; Marszal-Pomianowska, A.; Aelenei, D.; Martins, J.; Saelens, D. Energy flexible buildings: An evaluation of definitions and quantification methodologies applied to thermal storage. *Energy Build.* **2018**, *166*, 372–390. [CrossRef]
29. Campodonico Avendano, I.A.; Dadras Javan, F.; Najafi, B.; Moazami, A.; Rinaldi, F. Assessing the impact of employing machine learning-based baseline load prediction pipelines with sliding-window training scheme on offered flexibility estimation for different building categories. *Energy Build.* **2023**, *294*, 113217. [CrossRef]
30. Chen, Y.; Guo, M.; Chen, Z.; Chen, Z.; Ji, Y. Physical energy and data-driven models in building energy prediction: A review. *Energy Rep.* **2022**, *8*, 2656–2671. [CrossRef]
31. Lake, C. PJM Empirical Analysis of Demand Response Baseline Methods. 2011.
32. Wijaya, T.K.; Vasirani, M.; Aberer, K. When bias matters: An economic assessment of demand response baselines for residential customers. *IEEE Trans. Smart Grid* **2014**, *5*, 1755–1763. [CrossRef]
33. Mohajeryami, S.; Doostan, M.; Asadinejad, A.; Schwarz, P. Error analysis of customer baseline load (CBL) calculation methods for residential customers. *IEEE Trans. Ind. Appl.* **2016**, *53*, 5–14. [CrossRef]
34. Wang, F.; Li, K.; Liu, C.; Mi, Z.; Shafie-Khah, M.; Catalão, J.P. Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description. *IEEE Trans. Smart Grid* **2018**, *9*, 6972–6985. [CrossRef]
35. Li, K.; Yan, J.; Hu, L.; Wang, F.; Zhang, N. Two-stage decoupled estimation approach of aggregated baseline load under high penetration of behind-the-meter PV system. *IEEE Trans. Smart Grid* **2021**, *12*, 4876–4885. [CrossRef]
36. Sha, H.; Xu, P.; Lin, M.; Peng, C.; Dou, Q. Development of a multi-granularity energy forecasting toolkit for demand response baseline calculation. *Appl. Energy* **2021**, *289*, 116652. [CrossRef]
37. Cerquitelli, T.; Malnati, G.; Apiletti, D. Exploiting scalable machine-learning distributed frameworks to forecast power consumption of buildings. *Energies* **2019**, *12*, 2933. [CrossRef]
38. Campodonico Avendano, I.A.; Dadras Javan, F.; Najafi, B.; Moazami, A. Predicting HVAC-Based Demand Flexibility In Grid-Interactive Efficient Buildings Utilizing Deep Neural Networks. In Proceedings of the ECMS 2023: 37th International ECMS Conference on Modelling and Simulation, Florence, Italy, 20–23 June 2023.
39. Dadras Javan, F.; Campodonico Avendano, I.A.; Najafi, B.; Fabio Rinaldi, A.M. Modulating the HVAC Demand of a Warehouse to Provide Load Flexibility for Charging Electric Trucks. In Proceedings of the ECMS 2023: 37th International ECMS Conference on Modelling and Simulation, Florence, Italy, 20–23 June 2023.
40. Deru, M.; Field, K.; Studer, D.; Benne, K.; Griffith, B.; Torcellini, P.; Liu, B.; Halverson, M.; Winiarski, D.; Rosenberg, M.; et al. US Department of Energy Commercial Reference Building Models of the National Building Stock. 2011. Available online: https://digitalscholarship.unlv.edu/renew_pubs/44/ (accessed on 4 May 2023).
41. ANSE/ASHRAE/IES Standard 90.1-2010. Energy Standard for Buildings Except Low Rise Residential Buildings. 2010. Available online: http://d-pma.org/wp-content/uploads/2018/07/1.-ASHRAE_90_1_2010_IP.pdf (accessed on 4 May 2023).
42. Crawley, D.B.; Lawrie, L.K.; Winkelmann, F.C.; Buhl, W.F.; Huang, Y.J.; Pedersen, C.O.; Strand, R.K.; Liesen, R.J.; Fisher, D.E.; Witte, M.J.; et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy Build.* **2001**, *33*, 319–331. [CrossRef]
43. Input Output Reference. EnergyPlus Version 9.6.0 Documentation. 2021. Available online: https://energyplus.net/assets/nrel_custom/pdfs/pdfs_v9.6.0/InputOutputReference.pdf (accessed on 4 May 2023).
44. Smith, A.D.; Stürmer, B.; Thurber, T.; Vernon, C.R. diyepw: A Python package for Do-It-Yourself EnergyPlus weather file generation. *J. Open Source Softw.* **2021**, *6*, 3313. [CrossRef]
45. Emery, A.; Erne, D.; Lyon, E. Demand Side Grid Support Program: Proposed Draft Guidelines First Edition. CEC-300-2022-008. 2022. Available online: https://www.energy.ca.gov/filebrowser/download/4373 (accessed on 4 May 2023).
46. Nicol, F.; Humphreys, M. Maximum temperatures in European office buildings to avoid heat discomfort. *Sol. Energy* **2007**, *81*, 295–304. [CrossRef]
47. Tan, C.K.; Ogawa, A.; Matsumura, T. *Innovative Climate Change Communication: Team Minus 6%*; Global Environment Information Centre (GEIC), United Nations University (UNU): Shibuya, Tokyo, 2008.

48. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
49. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX symposium on operating systems design and implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016; Volume 16, pp. 265–283.
50. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 4–16 August 1995; Volume 1, pp. 278–282.
51. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
52. Murtagh, F. Multilayer perceptrons for classification and regression. *Neurocomputing* **1991**, *2*, 183–197. [CrossRef]
53. Jain, A.K.; Mao, J.; Mohiuddin, K.M. Artificial neural networks: A tutorial. *Computer* **1996**, *29*, 31–44. [CrossRef]
54. Bergmeir, C.; Benítez, J.M. On the use of cross-validation for time series predictor evaluation. *Inf. Sci.* **2012**, *191*, 192–213. [CrossRef]
55. Masum, S.; Liu, Y.; Chiverton, J. Multi-step time series forecasting of electric load using machine learning models. In Proceedings of the Artificial Intelligence and Soft Computing: 17th International Conference, ICAISC 2018, Zakopane, Poland, 3–7 June 2018; Proceedings, Part I 17, pp. 148–159.
56. Abedi, S.; Kwon, S. Rolling-horizon optimization integrated with recurrent neural network-driven forecasting for residential battery energy storage operations. *Int. J. Electr. Power Energy Syst.* **2023**, *145*, 108589. [CrossRef]
57. ASHRAE Guideline 14-2002. *Measurement of Energy and Demand Savings*; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, Georgia, 2002; Volume 35, pp. 41–63.
58. ASHRAE Guideline. *Measurement of Energy, Demand, and Water Savings*; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, Georgia, 2014; Volume 4, pp. 1–150.
59. Ramos Ruiz, G.; Fernandez Bandera, C. Validation of calibrated energy models: Common errors. *Energies* **2017**, *10*, 1587. [CrossRef]
60. Cowan, J. *International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings*; Office of Energy Efficiency and Renewable Energy, U.S. Department Of Energy: Washington, DC, USA, 2002; Volume 1 .