Received: 29 March 2022 | Revised: 21 July 2022 | Accepted: 22 August 2022

DOI: 10.1002/ese3.1298

ORIGINAL ARTICLE

Comparative study of continuous hourly energy consumption forecasting strategies with small data sets to support demand management decisions in buildings

D. Mariano-Hernández[1,2] | L. Hernández-Callejo[2] | M. Solís[3] | A. Zorita-Lamadrid[4] | O. Duque-Pérez[4] | L. Gonzalez-Morales[5] | V. Alonso-Gómez[6] | A. Jaramillo-Duque[7] | F. Santos García[8]

[1]Área de Ingeniería, Instituto Tecnológico de Santo Domingo, Santo Domingo, Dominican Republic

[2]ADIRE-ITAP, Departamento Ingeniería Agrícola y Forestal, Universidad de Valladolid, Campus Universitario Duques de Soria, Soria, Spain

[3]Tecnológico de Costa Rica, Cartago, Costa Rica

[4]ADIRE-ITAP, Departamento de Ingeniería Eléctrica, Universidad de Valladolid, Valladolid, Spain

[5]Departamento de Ingeniería Eléctrica, Electrónica y Telecomunicaciones—DEET, Facultad de Ingeniería, Universidad de Cuenca, Ecuador

[6]Departamento de Física, Universidad de Valladolid, Campus Universitario Duques de Soria, Soria, Spain

[7]Grupo de Investigación en Manejo Eficiente de la Energía (GIMEL), Departamento de Ingeniería Eléctrica, Universidad de Antioquia, Medellín, Colombia

[8]Área de Ciencias Básicas, Instituto Tecnológico de Santo Domingo, Santo Domingo, Dominican Republic

Correspondence
Hernández-Callejo, Departamento Ingeniería Agrícola y Forestal, Universidad de Valladolid, Campus Universitario Duques de Soria, 42004 Soria, Spain.
Email: luis.hernandez.callejo@uva.es

Funding information
CYTED Ciencia y Tecnología para el Desarrollo, Grant/Award Number: 518RT0558

Abstract

Buildings are one of the largest consumers of electrical energy, making it important to develop different strategies to help to reduce electricity consumption. Building energy consumption forecasting strategies are widely used to support demand management decisions, but these strategies require large data sets to achieve an accurate electric consumption forecast, so they are not commonly used for buildings with a short history of record keeping. Based on this, the objective of this study is to determine, through continuous hourly electricity consumption forecasting strategies, the amount of data needed to achieve an accurate forecast. The proposed forecasting strategies were evaluated with Random Forest, eXtreme Gradient Boost, Convolutional Neural Network, and Temporal Convolutional Network algorithms using 4 years of electricity consumption data from two buildings located on the campus of the University of Valladolid. For performance evaluation, two scenarios were proposed for each of the proposed forecasting strategies. The results showed that for forecasting horizons of 1 week, it was possible to obtain a mean absolute percentage error (MAPE) below 7% for Building 1 and a MAPE below 10% for Building

2 with 6 months of data, while for a forecast horizon of 1 month, it was possible to obtain a MAPE below 10% for Building 1 and below 11% for Building 2 with 10 months of data. However, if the distribution of the data captured in the buildings does not undergo sudden changes, the decision tree algorithms obtain better results. However, if there are sudden changes, deep learning algorithms are a better choice.

KEYWORDS
building energy consumption, forecasting, learning algorithms, multistep forecasting, short-term forecasting

1 | INTRODUCTION

The search for energy efficiency in buildings is a mandatory advance to guarantee sustainable development as buildings consume a large amount of electric energy.[1] Therefore, approaches identified with the decrease in building electricity consumption and expenses have lately been presented.[2] Electricity consumption forecasting is important when using strategies for building energy management systems, such as model predictive control and demand-side management,[3] are attracting a great deal of attention in recent years.[4] Accurate forecasting models are useful in determining energy-efficient building plans and demand-side management programs.[5] With the improvement of the building automation system, a lot of operational information can be saved, which helps to forecast models to benefit from existing operational information.[6] To take full advantage of the amount of collected information, the most recent trend in building energy modeling is to move from conventional physical models to data-driven models.[7] In building electricity consumption, data-driven models rely on devices that produce large amounts of energy-related information.[8]

In data-driven forecasting models, the most important modules are the model inputs and the prediction algorithms.[9] The model inputs summarize the types of information found in the existing study and anticipate the electricity consumption pattern, emphasizing that these electricity consumption patterns are not only affected by the historical data but also by correlated variables.[10] Algorithms can be classified into shallow learning and deep learning.[11] Deep learning algorithms such as Long-Short Term Memory (LSTM), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Temporal Convolutional Network (TCN) regularly change the inputs multiple times, before conveying the outputs; while shallow algorithms, such as Gradient Boost Regressor (GBR), Support Vector Machines (SVM), K-nearest Neighbors, Random Forest (RF) and Extreme Gradient Boosting (XGBoost) mostly change the inputs only a couple of times. The main difference is the number of linear or nonlinear transformations of the input data.[9]

Building energy consumption forecasting models can be classified into a single forecast, ensemble forecast, and improved forecast.[12] The single forecast model uses only one algorithm to predict outcomes. The ensemble forecast model is characterized as an algorithm that consolidates the benefits of multiple single forecasting models to enhance overall performance.[13] An improved forecast model consists of increasing the quality of the data before being used by the forecasting model.[14] Table 1 shows the contributions and limitations of recent studies that have focused on the different types of approaches.

As can be seen in the literature review, there are recent studies that focus on electricity consumption forecasting in buildings using the data-driven method. However, a limitation observed in these investigations is that a large data set is required to obtain accurate forecasts, which implies that buildings with limited time-series data could not use this method. Due to this, the objective of this study is to analyze using continuous hourly electricity consumption forecasting strategies what would be the minimum data needed to obtain accurate results. The main contributions of this study are:

- A data-driven approach that can be used to forecast continuous hourly electricity consumption to support demand management decisions in buildings with limited time-series data.
- Comparison analysis between a method that forecasts the next 24 h for all hours of the day and a method that forecasts the next 24 h for a particular hour.

TABLE 1 Recent studies focus on single forecast, ensemble forecast, and improved forecast

Approach	Ref.	Contributions	Limitations
Single Forecast	[15]	Proposed an adaptive Long Short Term Memory (LSTM) driven by a Genetic Algorithm for building energy consumption forecasting.	The building occupancy variable used for the method was considered only based on working days.
	[16]	Proposed a Gradient Boost Regressor with a modified Particle Swarm Optimizer for energy consumption forecasting.	The proposed approach was not compared with other optimization methods.
	[17]	Presented an Artificial Neural Network along with a Genetic Algorithm for energy consumption forecasting in smart buildings.	The data set used was not large enough for training and validation, so the model did not obtain high accuracy.
	[18]	Presented a Convolutional Neural Network with two-dimensional input for short-term load forecasting.	For 24-hour ahead prediction Regression Trees produced better results.
Ensemble Forecast	[19]	Presented three ensemble learning methodologies for short-term energy consumption forecast in an office building.	The forecast model only used climatic variables and historical consumption as input variables.
	[20]	Proposed an ensemble model boosted by Particle Swarm Optimizer for energy consumption forecasting.	The problem of reduced diversity in the model prediction was not analyzed.
	[21]	Presented a deep learning framework for energy consumption forecasting based on Convolutional Neural Network and LTSM algorithms.	The performance of optimization algorithms for hyperparameter tuning was not tested.
	[22]	Proposed an ensemble model for short-term building energy consumption forecasting using five data-driven models.	The assembled model does not have algorithm diversity because it was built only with neural networks.
	[23]	Proposed a hybrid ensemble model to estimate short-term energy consumption by using sequential dependencies.	Neither climatic variables nor optimization methods were considered in the model.
	[24]	Proposed a hybrid forecasting method based on Orthogonal Maximum Correlation Coefficient feature selection and Convolutional Gated Recurrent Unit.	Interactivity between variables was not considered in choosing the best feature set.
Improved Forecast	[25]	Change a time-dependent database into a structure that machine learning algorithms can process and then apply various types of feature selection techniques.	The approach was not tested with deep learning algorithms to see if they improved performance.
	[26]	Proposed a methodology that coordinates a pre-handling step utilizing domain knowledge with an insight-based feature selection process.	The comparison analysis shows that domain knowledge is not the determining factor.
	[27]	Build a model suitable for short-term forecasting by utilizing small data sets, the performance of the model was increased by utilizing projected sample generation.	The method was implemented and analyzed only in neural networks.
	[28]	Proposed a fine-grained attention mechanism to enhance the performance of deep learning models for multistep forecasting.	The study did not consider analyzing the proposed approach with shallow learning algorithms.

- A comparative study between shallow and deep learning algorithms using a multistep ahead prediction strategy with limited time-series data.

The rest of this paper is organized as follows: Section 2 presents the data set, strategies, and algorithms used. Section 3 shows the results obtained with the different strategies and algorithms. The conclusions of this study are described in Section 4.

2 | METHODOLOGY

In summary, the methodology used for this study was:

- Data sets preparation. The electricity consumption data from two nonresidential buildings were used for this study. A data set with a calendar, past series values, weather, and historical data was created for each building.

- Proposed approaches. A data-driven forecasting strategy was used to predict continuous hourly energy demand, which was compared using two different methods.
- Learning algorithms. Four algorithms were selected to analyze the proposed forecasting approaches. The selected algorithms were RF, XGBoost, CNN, and TCN.
- Performance evaluation. Different scenarios were carried out to validate the performance of the different algorithms using mean absolute percentage error (MAPE) as the evaluation metric.

2.1 | Data sets preparation

The data sets used for training the forecasting models were composed of weather data, past series values, calendars, and historical data. The weather data used in the data sets were selected using correlation analysis between climatological parameters such as precipitation, relative humidity at 2 m, average temperature at 2 m, minimum temperature at 2 m, maximum temperature at 2 m, heating degree days below 18.3°C, cooling degree days above 0°C, cooling degree days above 10°C, and all-sky surface longwave downward irradiance. The selection of these variables was based on the importance they represent in sustainable buildings, such as all-sky surface longwave downward irradiance, which is used for building simulation and passive cooling design, since it helps to consider both sunny and cloudy conditions.[29] Based on Pearson's correlation coefficient, the variables of all-sky surface longwave downward irradiance, the maximum temperature at 2 m, and the average temperature at 2 m were selected due to their strong correlation with the electrical energy consumption.

Past series values correspond to the values of the hours before the hour to be predicted, which would be useful to forecast the values of the following hours. To determine these values, an autocorrelation function and partial autocorrelation function analysis were carried out, obtaining the result that the previous 25 h would be useful for the forecast. The calendar data were created from the variables year, months, days of the week, and holidays. Historical data correspond to the active energy consumed (kWh) of two complete buildings with a continental Mediterranean climate located at the University of Valladolid, Spain. Building 1 corresponds to the Faculty of Science, which is dedicated to administrative offices while Building 2 corresponds to the Faculty of Economics, which is dedicated to learning activities (see Table 2). The electrical consumption data were collected from smart meters installed in each building from 2016 to 2019 at intervals of 15 min.

TABLE 2 General description and characteristics of each of the buildings

Information	Building 1	Building 2
Name	Faculty of Science	Faculty of Economics
Coordinates	41.663411°, −4.705539°	41.658586°, −4.710667°
Climate	Mediterranean	Mediterranean
Area	16,006.66 m^2	15,456.51 m^2
Floors	5	3
Built Date	2005	1986
Current Use	Offices and laboratories	Offices and classroom
Estimated occupancy	1070	1340

Since the raw data had missing values, it was necessary to preprocess them. A linear interpolation approach was used because the missing information was less than 0.3% of the total information. The reason these buildings were selected is due to their different electricity consumption behavior during the years mentioned above (see Figure 1). Building 1 presents similar electrical consumption during the years evaluated while Building 2 presents variable behavior due to the replacement of low-efficiency equipment for high-efficiency equipment and the incorporation of renewable energy in the building, reducing annual electricity consumption.

2.2 | Proposed approaches

Based on multistep ahead predictions, which learn a simple parametric function from input time series and estimate a series of values.[30] Two methods were proposed for this investigation, which is visually represented in Figure 2. Method 1 forecasts the next 24 h for each hour of the day, allowing the method to be used to forecast at any hour of the day. Method 2 forecasts the next 24 h for a single hour of the day. For method 1 to be able to forecast energy consumption in this way, it was necessary to use the data for the hour to be forecast and the data for the next previous 24 h.

Method 1 was analyzed in previous research,[31] in which data from 2016 to 2018 were used for the training phase and data from 2019 for the testing phase. For the multistep ahead prediction strategy used in the aforementioned research, the method has to learn 26,304

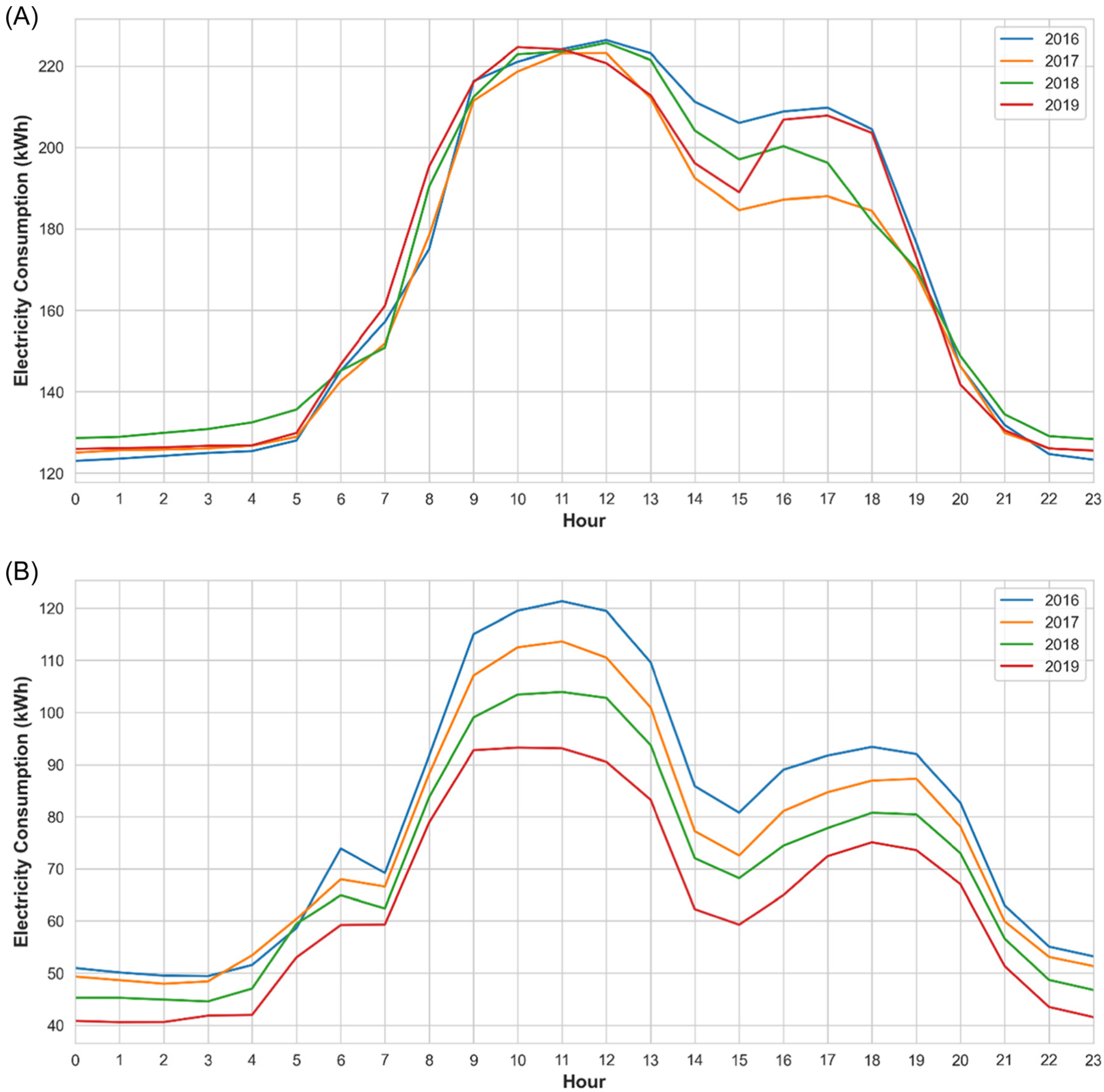

FIGURE 1 (A) Hourly average electricity consumption for Building 1 by year. (B) Hourly average electricity consumption for Building 2 by year.

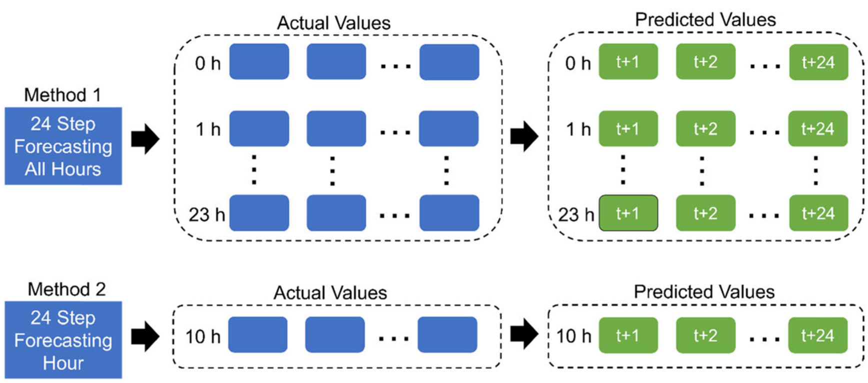

FIGURE 2 Schematic diagram of the two methods used for the comparison study

patterns, which correspond to the total hours comprised in the 3 years used in the training phase.

Due to the large number of patterns in the training phase used by method 1, method 2 was proposed. In the case of method 2, only 1095 patterns are used, corresponding to the number of times an hour is repeated in 3 years. However, since method 2 only forecasted the next 24 h for a specific hour, 24 models were trained simultaneously to learn the performance for each of the hours of the day. To run all 24 models simultaneously, this method was programmed in Python using the Joblib library.

2.3 | Learning algorithms

To analyze the proposed approaches, two shallow learning algorithms and two deep learning algorithms were used to forecast electricity consumption. For shallow learning algorithms, RF and XGBoost were used, and for deep learning algorithms, CNN and TCN were used. These algorithms were selected based on their good performance obtained in the aforementioned research.[31] A brief description of each of the algorithms is presented below:

- RF is a mixture of numerous decision trees created using a bootstrapping approach coming from the learning data set samples of the predictor and picking randomly at every node. RFs are completed regarding classification and regression trees model methodology.[32]
- XGBoost is an ensemble approach created based on Gradient Boosting, it learns a set of regression trees in parallel and acquires the outcome by summing the score of each one. Some of its improvements are the regularized objective to the loss function and instead of applying a stochastic gradient descent strategy to complement the corresponding optimization method, XGBoost adds the best tree model.[33]
- CNN includes a convolutional layer, a pooling layer, and a fully connected layer to imitate complex information. As a general rule, CNN has a few progressive systems of convolutional and pooling layers, wherein a few convolution runs are performed to extract the significant features from the input data. In the convolutional layer, neurons from various layers of the network are locally associated through a weight-sharing procedure.[21]
- TCN is architecture-dependent on two main ideas: the network convolutions are easy to perform as no past data goes into the future and the architecture likewise allows variable input sizes, mapping any output array. An important feature is that allows parallelism, due to its adaptability of input size and productive memory use.[34]

The four selected algorithms were programmed in Python using the XGBoost, Scikit-Learn, and TensorFlow libraries. The architectures and hyperparameters used for the comparison study are shown in Table 3. To obtain these values, randomized search and grid search techniques were used, obtaining the values presented as the best combinations of parameters.

2.4 | Performance evaluation

Several accuracy metrics are available to assess the performance of the algorithm based on the contrast between the actual and predicted value. In this study, MAPE has been used to evaluate the performance of the methods due to its easy interpretation. MAPE is a metric that indicates the accuracy of the expectation by contrasting the remaining and observed values. It is usually expressed in percentages and is feasible to evaluate the performance of the forecasting model by presenting the idea of absolute values. The MAPE is characterized by the following equation[35]:

TABLE 3 Learning algorithms architecture and hyperparameters

Algorithm	Architectures and hyperparameters
RF	• max_depth = 45 • n_estimators = 200 • min_samples_leaf = 1
XGBoost	• n_estimators = 50 • eta = 0.1 • max_depth = 5 • colsample_bytree = 0.8 • subsample = 0.8 • gamma = 1
CNN	• One convolutional hidden layer. • First layer: 1D convolution with 64 filters, kernel size = 2, linear activation function, MaxPooling1D of size 2. • Flatten. • Output layer with 24 units. • loss function = mean squared error • optimizer = adam • learning rate = 0.001 • batch size = 1 • The model with the best epoch in the loss function was selected.
TCN	• filters = 200 • kernel_size = 4 • dilations = [1, 2, 4, 8, 16, 32] • batch size = 1 • activation function = linear

Abbreviations: CNN, Convolutional Neural Network; RF, Random Forest; TCN, Temporal Convolutional Network.

$$MAPE = \frac{1}{n} \times \sum_{t=1}^{n} \left| \frac{x_t - y_t}{y_t} \right| \times 100\%, \tag{1}$$

Where, x_t is the actual value, y_t is the predicted value, and n is the total number of estimations.

To evaluate the two methods with the four selected learning algorithms, two different scenarios were considered. In the first scenario, the methods are trained with data from 1 month to 1 year, and the hours of the following week after training are forecast. In the second scenario, the methods are trained with data from 1 month to 1 year, and the hours of the next month after training are forecast. These scenarios were approached from the point of view of buildings with a short history of record keeping but requiring an electricity consumption forecast as soon as possible for decision-making.

Since method 2 focuses on a particular hour, several trials were carried out with different hours to find out which hour of the day had the best and which had the worst result, so only these hours were compared between the methods.

3 | RESULTS AND DISCUSSION

3.1 | Preliminary analysis

As method 2 forecasts from a particular hour, a previous analysis was carried out to know which hours of the day perform best with the selected algorithms. After analyzing the results of method 2 with all 24 models simultaneously, it was observed that the lowest percentage values were found between hours 10 and 12, while

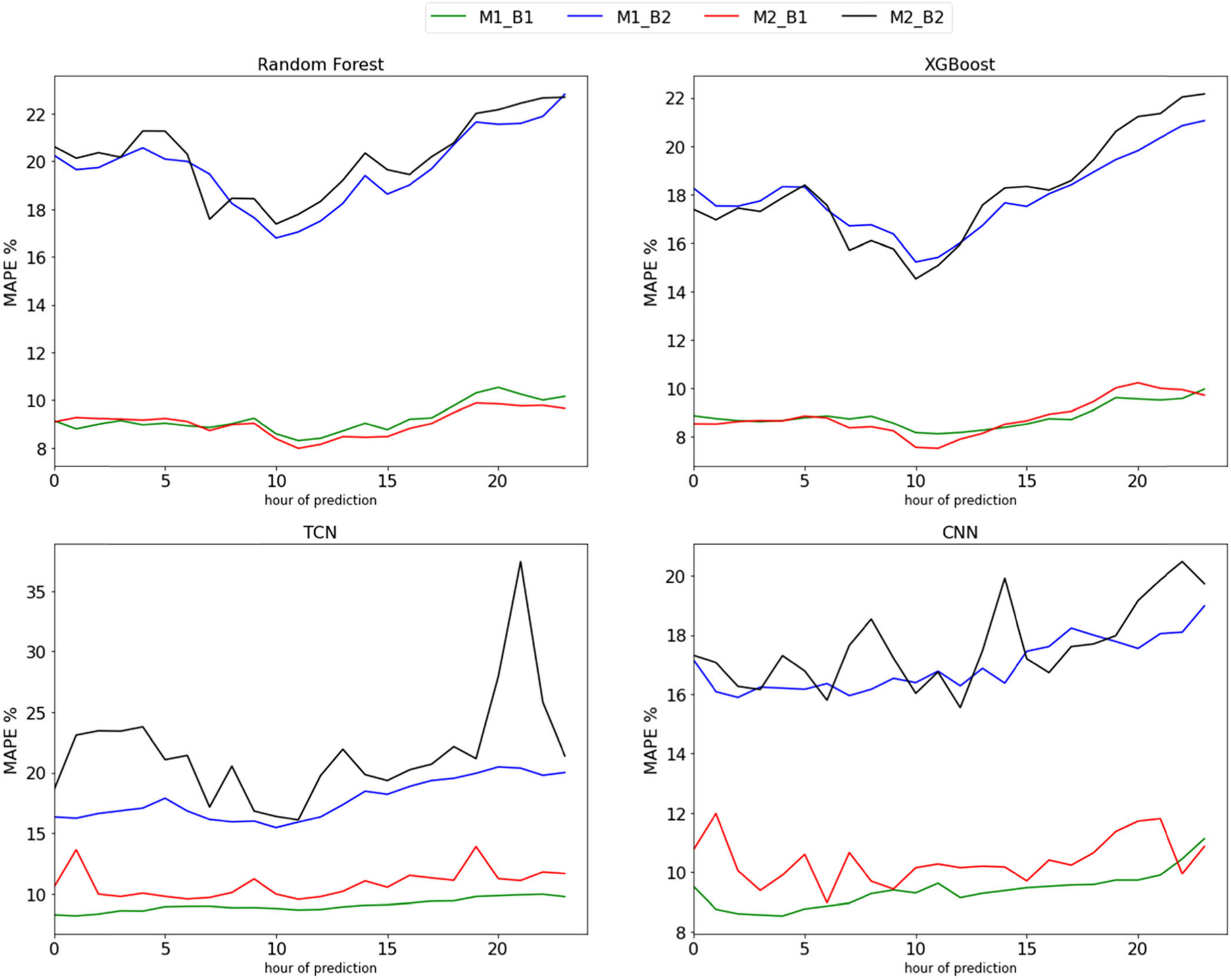

FIGURE 3 Methods performance of each of the algorithms for the 2019 hourly energy consumption forecast. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; TCN, Temporal Convolutional Network.

the highest percentage values were found between hours 20 and 23. Due to these results, hours 10 and 22 were selected to compare the methods.

When working with all the data (3 years), method 1 presents a lower error percentage for the deep learning algorithms in both buildings. However, for the shallow learning algorithms, the results of both methods tend to be more similar, and even when it comes to Building 1, method 2 presents a lower error percentage. This finding tells us that deep learning algorithms assimilate better the variability that may exist in the data, which is why they perform better with method 1. In addition, they learn better the more data they have, which is the case with method 1.

The outcomes for method 1 in Building 1 and Building 2 (M1_B1 and M1_B2) and method 2 in Building 1 and Building 2 (M2_B1 and M2_B2) can be seen in Figure 3.

TABLE 4 Average MAPE for each of the buildings using both strategies

	Building 1 MAPE (%)		Building 2 MAPE (%)	
Algorithm	**Method 1**	**Method 2**	**Method 1**	**Method 2**
RF	9.22	9.05	19.68	20.15
XGBoost	8.83	8.81	17.92	18.06
CCN	9.02	10.77	17.74	21.64
TCN	9.38	10.38	16.96	17.59

Abbreviations: CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; RF, Random Forest; TCN, Temporal Convolutional Network.

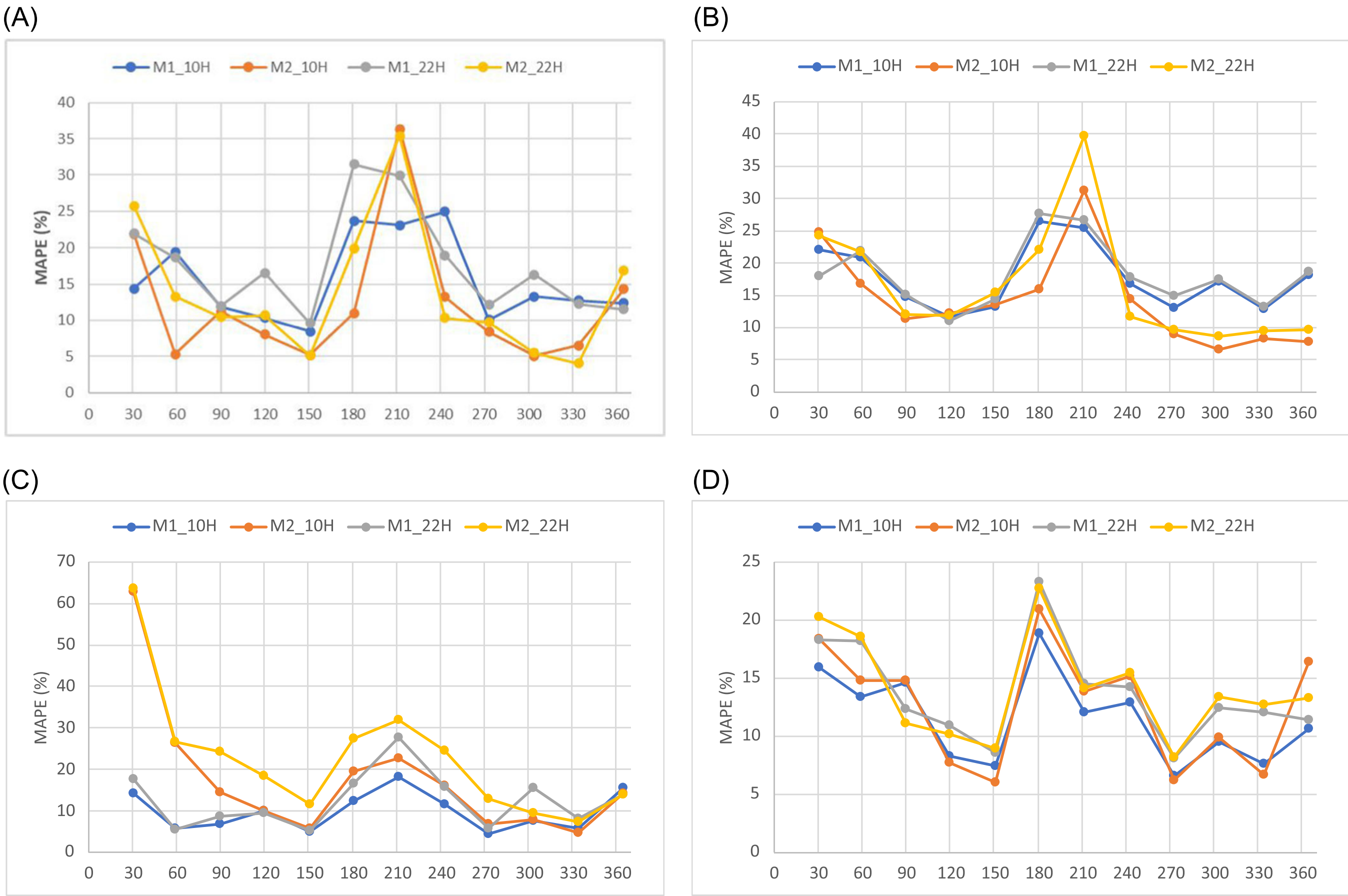

FIGURE 4 (A) MAPE results of Scenario 1 in Building 1 for RF. (B) MAPE results of Scenario 1 in Building 1 for XGBoost. (C) MAPE results of Scenario 1 in Building 1 for CNN. (D) MAPE results of Scenario 1 in Building 1 for TCN. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; RF, Random Forest; TCN, Temporal Convolutional Network.

3.2 | Scenario 1 evaluation

Based on the comparison between methods, we observed that for Building 1 the shallow learning algorithm obtained a better average MAPE than the deep learning algorithms. In the case of Building 2, the deep learning algorithms obtained a better average MAPE than the shallow learning algorithm. This indicates that when the energy behavior of the building is stable over the years, as is the case of Building 1, shallow learning algorithms present a lower error percentage, while if the energy behavior varies over the years due to improvements in energy efficiency such as the replacement of less efficient equipment with high-efficiency equipment and the integration of solar panels in the building as is the case of Building 2, deep learning algorithms present lower error percentage. Table 4 shows the average MAPE of method 1 and method 2 for each of the buildings.

According to the results of scenario 1 for Building 1, the shallow learning algorithms MAPE range was between 4% and 39%. For RF, the best average results were 7.08% in 151 days and 8.88% in 334 days, while XGBoost obtained 11.66% in 273 days and 12.47% in 304 days. In the case of deep learning algorithms, the MAPE range was between 6% and 63%. For CNN, the best average results were obtained with 6.85% in 151 days and 6.47% in 334 days while TCN was 7.75% in 151 days and 7.27% in 273 days. Although the deep learning algorithms obtained a wider range than the shallow learning algorithms in terms of average, they obtained a lower error percentage (see Figure 4).

In Building 2, the shallow learning algorithms MAPE range was between 9% and 81%. For RF, the best average results were obtained for training with 20.15% in 151 days and 22.63% in 304 days, while the XGBoost with 18.93% in 151% and 22.58% in 304 days. In the case of deep learning algorithms, the MAPE range was between 9% and 55%. For CNN, the best average results were obtained with 14.92% in 151 days and 13.62% in 273 days, while TCN with 20.11% in 151 days and 15.94% in 365 days. For this case, deep learning algorithms obtained a better range and better average results than shallow learning algorithms (see Figure 5).

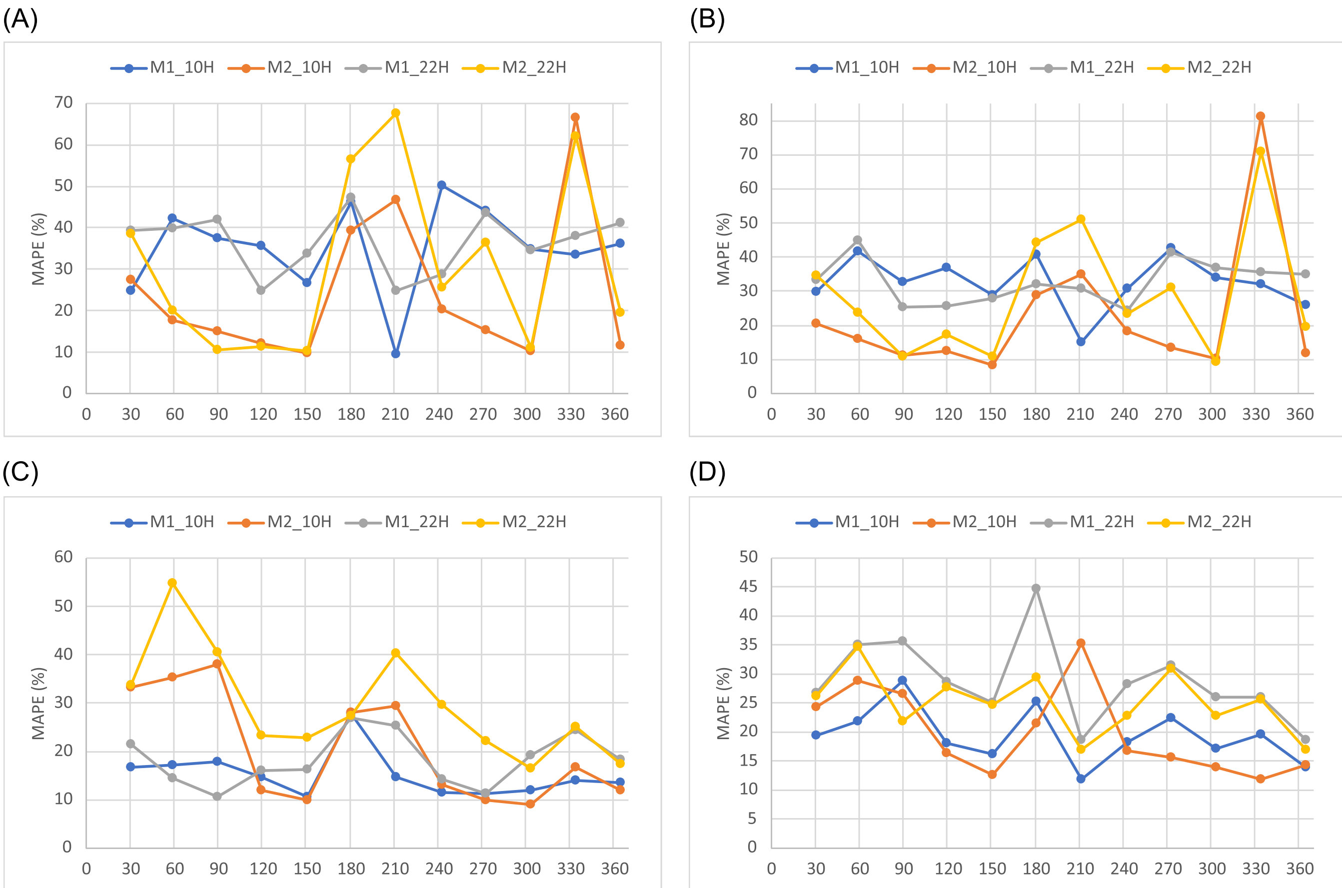

FIGURE 5 (A) MAPE results of Scenario 1 in Building 2 for RF. (B) MAPE results of Scenario 1 in Building 2 for XGBoost. (C) MAPE results of Scenario 1 in Building 2 for CNN. (D) MAPE results of Scenario 1 in Building 2 for TCN. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; RF, Random Forest; TCN, Temporal Convolutional Network.

Based on the average results for each method, for the first scenario where we forecast one week after training, the shallow algorithms tend to perform better with method 2. For deep learning algorithms in general it is better to use method 1, the only exception was when we trained the TCN for Building 2. Overall, CNN obtained the best performance with 5.04% for Building 1% and 11.2% for Building 2.

3.3 | Scenario 2 evaluation

According to the results of scenario 2 for Building 1, the shallow learning algorithms MAPE range was between 6% and 42%. For RF, the best average results were obtained with 11.7% in 273 days and 11.13% in 304 days while XGBoost with 11.66% in 273 days and 12.47% in 304 days. In the case of deep learning algorithms, the MAPE range was between 7% and 80%. For CNN the best average results were obtained with 11.43% in 304% and 16.18% in 334 days while TCN with 8.5% in 120 days and 9.94% in 304 days. For this case, the shallow learning algorithms obtained a better range than deep learning algorithms, however, TCN obtained better average results than the shallow learning algorithms (see Figure 6).

For Building 2, the shallow learning algorithm MAPE ranged from10% to 91%. For RF, the best average results were 24.69% in 243 days and 22.83% in 304 days, while XGBoost was 19.71% in 243 days and 23.43% in 304 days. In the case of deep learning algorithms, the MAPE range was between 11% and 81%. For CNN, the best average results were 14.7% in 243 days and 15.23% in 304 days, while TCN was 21.84% in 151 days and 18.39% in 304 days. For this case, the deep learning algorithms obtained a better range and better average results than shallow learning algorithms (see Figure 7).

Based on the average results for each method, for the second scenario where we forecast one month after training, the shallow algorithms tend to perform better with method 2. For deep learning algorithms in general it is better to use method 2, the only exception was when we trained the CCN on Building 2. For Building 1, RF performed the best with 7.2%. For Building 2, XGBoost obtained the best results with 11.2%.

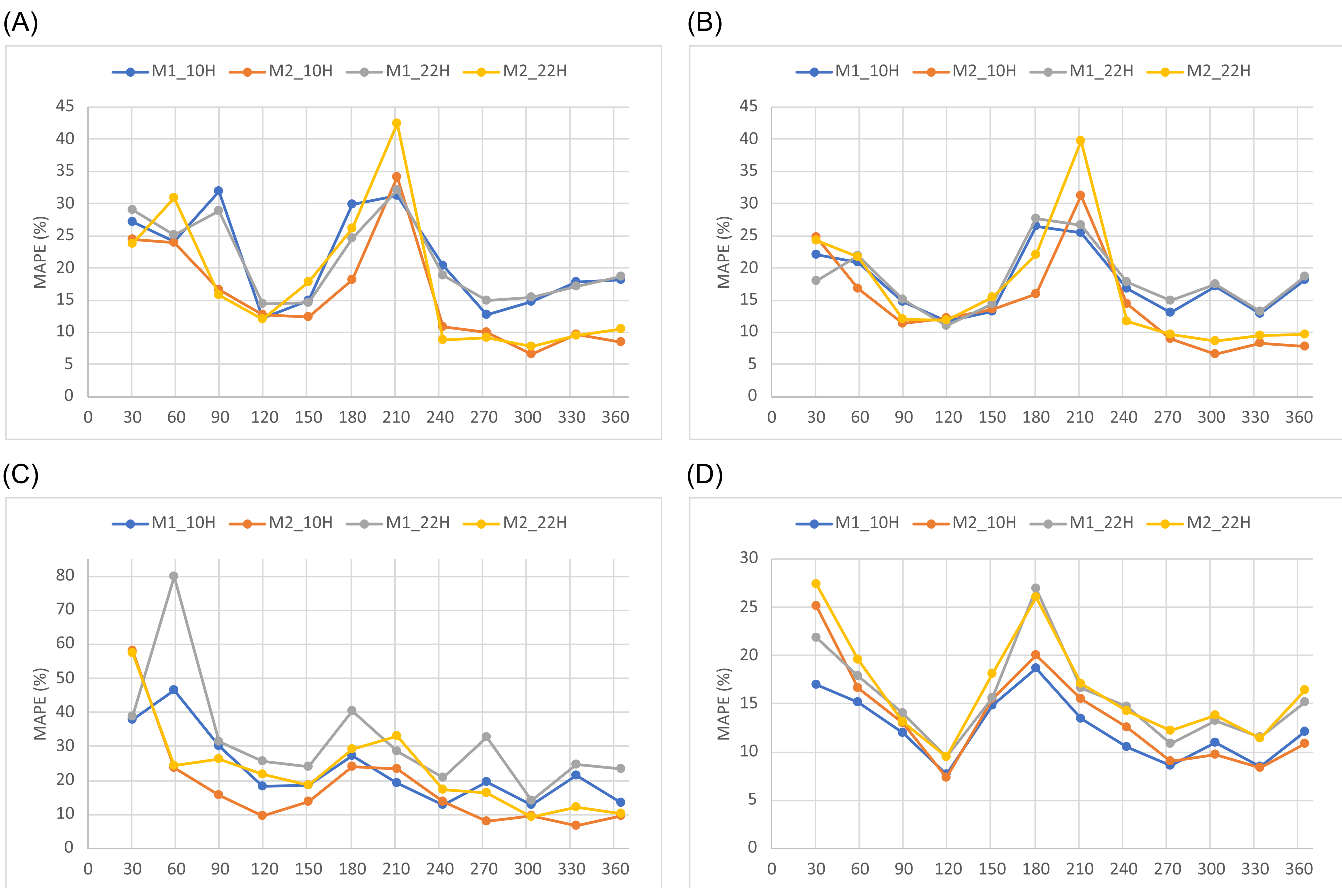

FIGURE 6 (A) MAPE results of Scenario 2 in Building 1 for RF. (B) MAPE results of Scenario 2 in Building 1 for XGBoost. (C) MAPE results of Scenario 2 in Building 1 for CNN. (D) MAPE results of Scenario 2 in Building 1 for TCN. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; RF, Random Forest; TCN, Temporal Convolutional Network.

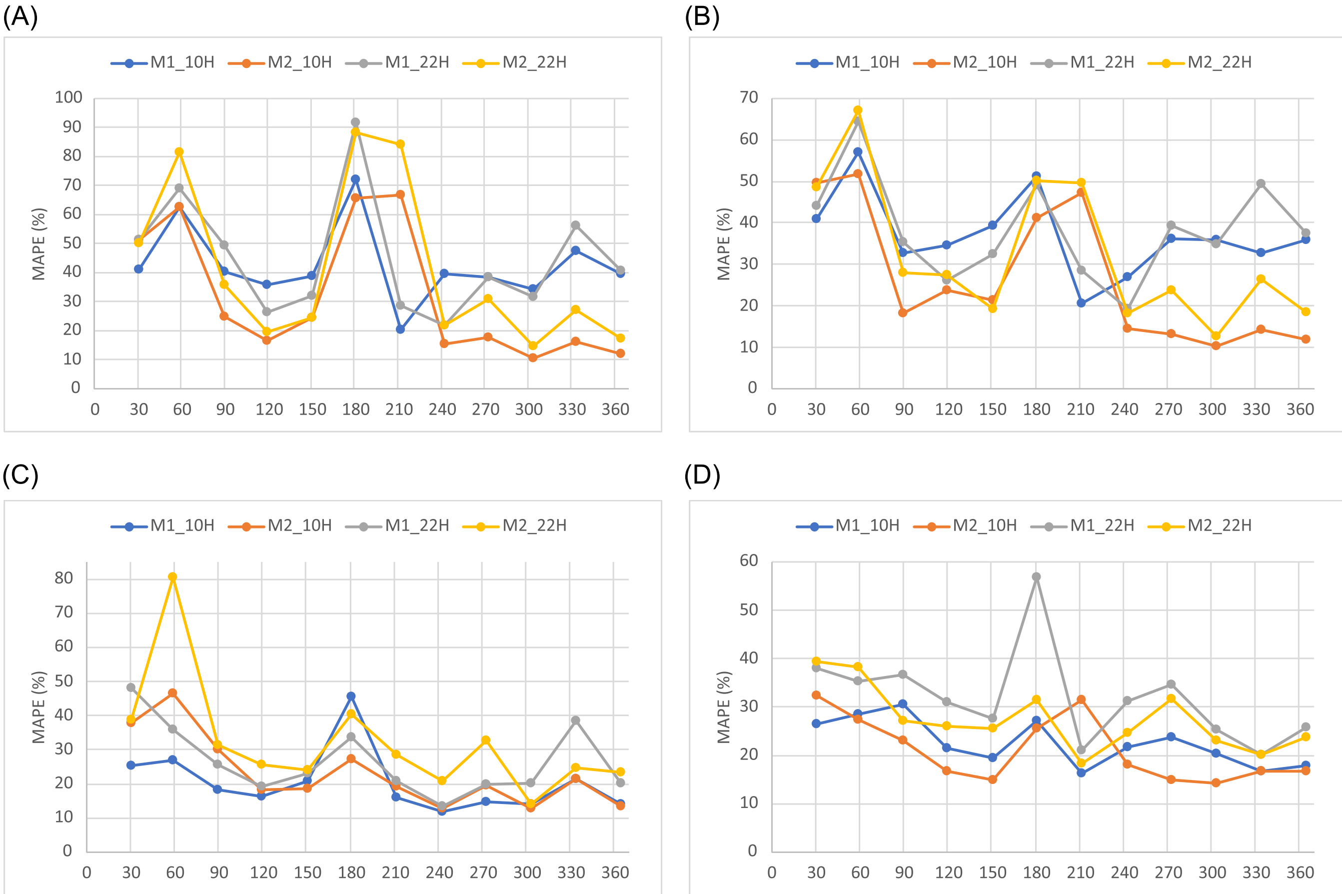

FIGURE 7 (A) MAPE results of Scenario 2 in Building 2 for RF. (B) MAPE results of Scenario 2 in Building 2 for XGBoost. (C) MAPE results of Scenario 2 in Building 2 for CNN. (D) MAPE results of Scenario 2 in Building 2 for TCN. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; RF, Random Forest; TCN, Temporal Convolutional Network.

In general, according to the results, decision tree algorithms trained with small data sets perform better with the method that focuses on forecasting from a particular hour, because the data used in the training stage are more specific. For deep learning algorithms trained with small data sets perform better with the method that focuses on forecasting from any hour of the day, because the data used in the training stage have a greater variety of data. It should be noted that the deep learning algorithms, regardless of the method used, adapted better to sudden changes in electricity consumption.

Similarly, the results indicate that in buildings that are starting to record electricity consumption and the distribution of data remains without sudden changes, decision tree algorithms can be used to forecast electricity consumption to implement demand-side management strategies. However, if the data distribution could present abrupt changes due to energy improvements, the deep learning algorithm would be a better choice since they adapt better to sudden changes.

It should be noted that sudden changes can be caused by seasonal effects, as is the case in these buildings during the holiday months, which produce high levels of inaccuracy in the models. However, the abrupt changes between weekdays and weekends, because they occur constantly in a very short time, only initially affect the models, but later the models learn these patterns.

Figure 8 shows the mean and standard deviation of the MAPE for each model trained with each of the 12 short data set versions. According to the results for scenario 1, the performance is better with CNN using method 1, however, with method 2, Random Forest would provide the best results for Building 1, while the TCN would provide the best results for Building 2. In scenario 2, TCN performs better for building 1 with both methods, while CNN performs better with method 1 in Building 2.

The standard deviations suggest that the performance of the 12 models has more dispersion with method 2. Therefore, with this method, it is more relevant to pay attention to the amount of data available for training.

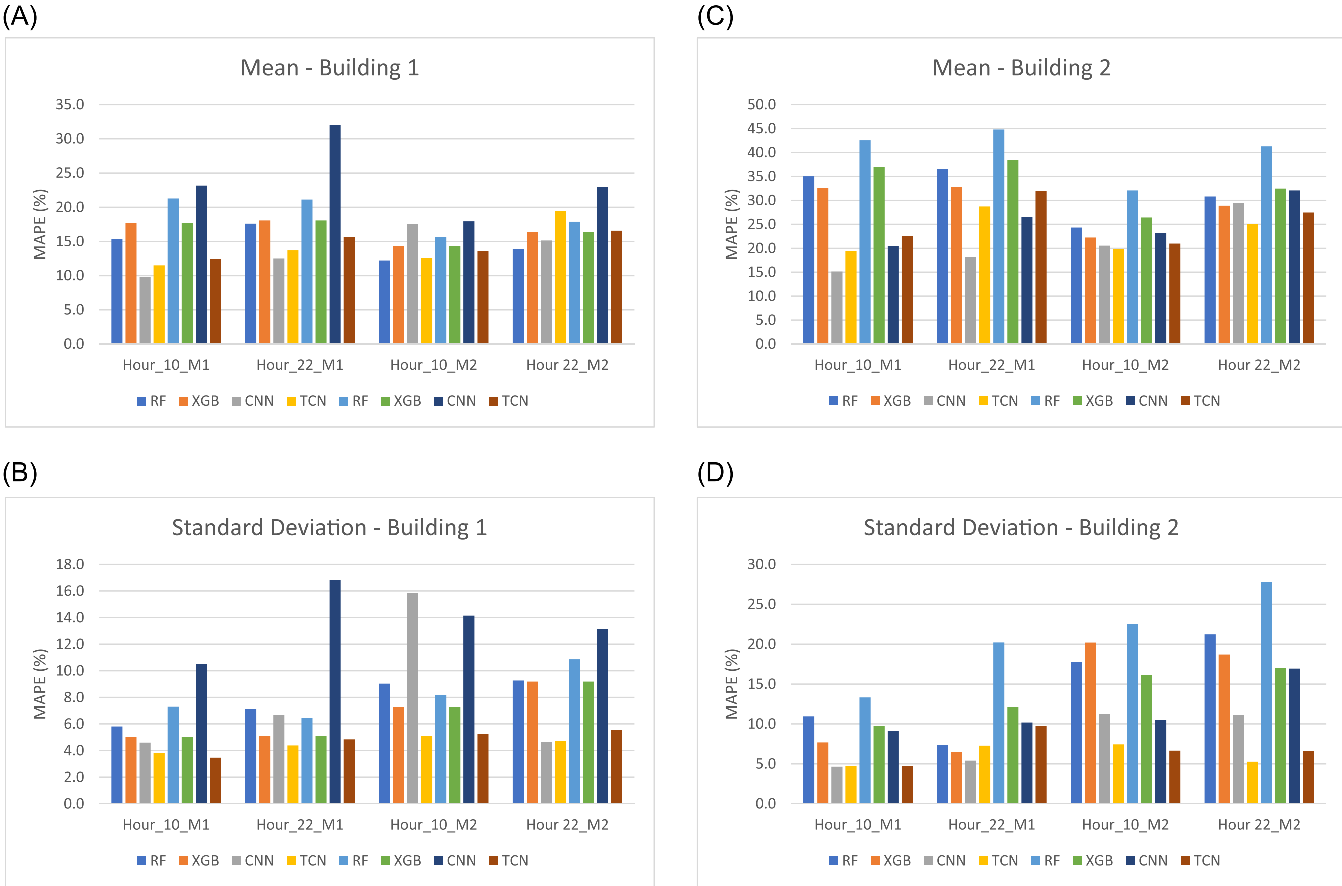

FIGURE 8 (A) Mean of MAPE for Building 1. (B) Standard Deviation of MAPE for Building 1. (C) Mean of MAPE for Building 2. (D) Standard Deviation of MAPE for Building 2. CNN, Convolutional Neural Network; MAPE, mean absolute percentage error; TCN, Temporal Convolutional Network.

This finding is even more visible with RF and XGB models because the dispersion with method 2 tends to have larger increases.

4 | CONCLUSION

This paper presents a comparison between two multistep ahead prediction strategies with different algorithms to determine which approach is better to use in buildings with limited time-series data to forecast electric consumption. For the analysis of the methods, the data from two buildings located at the University of Valladolid, Spain were used. In addition to historical data from the two buildings, the data set was composed of weather, calendar, and past values data. To analyze the methods, RF, XGBoost CNN, and TCN algorithms were used.

According to the results, for buildings where data distribution does not undergo sudden changes during data capture, the decision tree algorithm would be the best choice. However, if the data distribution undergoes sudden changes, such as improvements due to energy efficiency measures, deep algorithms would be a better choice since they adapt better to sudden changes compared to decision tree algorithms.

For the case where the forecasting horizon was 1 week, it was possible to obtain a MAPE below 7% for Building 1 and a MAPE below 10% for Building 2 with 6 months of data, using the forecasting method that considered only a particular hour. However, in the case where the forecasting horizon was 1 month, it was possible to obtain a MAPE below 10% for Building 1 and below 11% for Building 2 with 10 months of data using the method that considered all hours of the day. This indicates that to forecast a longer forecasting horizon, these algorithms would need more learning patterns and the seasonality of the data must be considered.

For future lines of research, the determination of the limitations to adapt to sudden changes that deep learning algorithms have in buildings that are starting to record data considering that they are the best option, as well as the analysis of the inclusion of energy efficiency improvements implemented in the building as input variables to help the models to improve their performance.

AUTHOR CONTRIBUTIONS

D. Mariano-Hernández: Conceptualization, methodology, data curation, software, data analysis, investigation, data curation, writing-original draft preparation, visualization. **L. Hernández-Callejo**: Conceptualization, methodology, validation, data analysis, writing-review and editing, supervision. **M. Solís**: Conceptualization, data curation, software, data analysis, investigation, visualization. **A. Zorita-Lamadrid**: Validation, writing-review, and editing. **O. Duque-Pérez**: Validation, writing-review and editing. **L. Gonzalez-Morales**: Validation. **V. Alonso-Gómez**: Software, data analysis. **A. Jaramillo-Duque**: Writing-review and editing. **F. Santos García**: Validation. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the University of Valladolid and the Instituto Tecnológico de Santo Domingo for their support in this study, which is the result of a co-supervised doctoral thesis. The authors also acknowledge the CITIES thematic network, a member of the CYTED program. CYTED, grant number: 518RT0558.

ORCID

D. Mariano-Hernández http://orcid.org/0000-0002-4255-3450
L. Hernández-Callejo http://orcid.org/0000-0002-8822-2948

REFERENCES

1. Morán A, Fuertes JJ, Prada MA, et al. Analysis of electricity consumption profiles in public buildings with dimensionality reduction techniques. *Eng Appl Artif Intell*. 2013;26:1872-1880. doi:10.1016/j.engappai.2013.04.005
2. Kim J, Kim K-I. Data-driven hybrid model and operating algorithm to shave peak demand costs of building electricity. *Energy Build*. 2020;229:110493. doi:10.1016/j.enbuild.2020.110493
3. Mariano-Hernández D, Hernández-Callejo L, Zorita-Lamadrid A, Duque-Pérez O, Santos García F. A review of strategies for building energy management system: model predictive control, demand side management, optimization, and fault detect & diagnosis. *J Build Eng*. 2021;33:101692. doi:10.1016/j.jobe.2020.101692
4. Dong Z, Liu J, Liu B, Li K, Li X. Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification. *Energy Build*. 2021;241:110929. doi:10.1016/j.enbuild.2021.110929
5. Bassi A, Shenoy A, Sharma A, Sigurdson H, Glossop C, Chan JH. Building energy consumption forecasting: a comparison of gradient boosting models. *ACM Int Conf Proc Ser*. June, 2021. doi:10.1145/3468784.3470656
6. Li G, Zhao X, Fan C, Fang X, Li F, Wu Y. Assessment of long short-term memory and its modifications for enhanced short-term building energy predictions. *J Build Eng*. 2021;43:103182. doi:10.1016/j.jobe.2021.103182
7. Bourdeau M, Basset P, Beauchêne S, et al. Classification of daily electric load profiles of non-residential buildings. *Energy Build*. 2021;233:110670. doi:10.1016/j.enbuild.2020.110670
8. Bourdeau M, Zhai Xqiang, Nefzaoui E, Guo X, Chatellier P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain Cities Soc*. 2019;48:101533. doi:10.1016/j.scs.2019.101533
9. Wang Z, Hong T, Piette MA. Building thermal load prediction through shallow machine learning and deep learning. *Appl Energy*. 2020;263:114683. doi:10.1016/j.apenergy.2020.114683
10. Song W, Fujimura S. Capturing combination patterns of long- and short-term dependencies in multivariate time series forecasting. *Neurocomputing*. 2021;464:72-82. doi:10.1016/j.neucom.2021.08.100
11. Xu Y, Zhou Y, Sekula P, Ding L. Machine learning in construction: from shallow to deep learning. *Dev Built Environ*. 2021;6:100045. doi:10.1016/j.dibe.2021.100045
12. Wang R, Lu S, Feng W. A novel improved model for building energy consumption prediction based on model integration. *Appl Energy*. 2020;262:114561. doi:10.1016/j.apenergy.2020.114561
13. Opitz D, Maclin R. Popular ensemble methods: an empirical study. *J Artif Intell Res*. 1999;11:169-198. doi:10.1613/jair.614
14. Ding Y, Zhang Q, Yuan T, Yang F. Effect of input variables on cooling load prediction accuracy of an office building. *Appl Therm Eng*. 2018;128:225-234. doi:10.1016/j.applthermaleng.2017.09.007
15. Luo XJ, Oyedele LO. Forecasting building energy consumption: adaptive long-short term memory neural networks driven by genetic algorithm. *Adv Eng Inform*. 2021;50:101357. doi:10.1016/j.aei.2021.101357
16. Sepulveda LFM, Diniz PS, Diniz JOB, et al. Forecasting of individual electricity consumption using optimized gradient boosting regression with modified particle swarm optimization. *Eng Appl Artif Intell*. 2021;105:104440. doi:10.1016/j.engappai.2021.104440
17. Bourhnane S, Abid MR, Lghoul R, Zine-Dine K, Elkamoun N, Benhaddou D. Machine learning for energy consumption prediction and scheduling in smart buildings. *SN Appl Sci*. 2020;2:297. doi:10.1007/s42452-020-2024-9
18. Bendaoud NMM, Farah N. Using deep learning for short-term load forecasting. *Neural Comput Appl*. 2020;32:15029-15041. doi:10.1007/s00521-020-04856-0
19. Pinto T, Praça I, Vale Z, Silva J. Ensemble learning for electricity consumption forecasting in office buildings. *Neurocomputing*. 2021;423:747-755. doi:10.1016/j.neucom.2020.02.124
20. Larrea M, Porto A, Irigoyen E, Barragán AJ, Andújar JM. Extreme learning machine ensemble model for time series forecasting boosted by PSO: application to an electric consumption problem. *Neurocomputing*. 2021;452:465-472. doi:10.1016/j.neucom.2019.12.140
21. Somu N, Raman MRG, Ramamritham K. A deep learning framework for building energy consumption forecast. *Renew Sustain Energy Rev*. 2021;137:110591. doi:10.1016/j.rser.2020.110591
22. Li K, Tian J, Xue W, Tan G. Short-term electricity consumption prediction for buildings using data-driven swarm

intelligence based ensemble model. *Energy Build*. 2021;231:110558. doi:10.1016/j.enbuild.2020.110558
23. MustaqeemIshaq, M, Kwon S. Short-Term energy forecasting framework using an ensemble deep learning approach. *IEEE Access*. 2021;9:94262-94271. doi:10.1109/access.2021.3093053
24. Liu R, Chen T, Sun G, Muyeen SM, Lin S, Mi Y. Short-term probabilistic building load forecasting based on feature integrated artificial intelligent approach. *Electr Power Syst Res*. 2022;206:107802. doi:10.1016/j.epsr.2022.107802
25. González-Vidal A, Jiménez F, Gómez-Skarmeta AF. A methodology for energy multivariate time series forecasting in smart buildings based on feature selection. *Energy Build*. 2019;196:71-82. doi:10.1016/j.enbuild.2019.05.021
26. Zhang L, Wen J. A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy Build*. 2019;183:428-442. doi:10.1016/j.enbuild.2018.11.010
27. Chang CJ, Lin JY, Chang MJ. Extended modeling procedure based on the projected sample for forecasting short-term electricity consumption. *Adv Eng Inform*. 2016;30:211-217. doi:10.1016/j.aei.2016.03.003
28. Fang X, Yuan Z. Performance enhancing techniques for deep learning models in time series forecasting. *Eng Appl Artif Intell*. 2019;85:533-542. doi:10.1016/j.engappai.2019.07.011
29. Chang K, Zhang Q. Modeling of downward longwave radiation and radiative cooling potential in China. *J Renew Sustain Energy*. 2019;11:066501. doi:10.1063/1.5117319
30. Chalapathy R, Khoa NLD, Sethuvenkatraman S. Comparing multi-step ahead building cooling load prediction using shallow machine learning and deep learning models. *Sustain Energy Grids Netw*. 2021;28:100543. doi:10.1016/j.segan.2021.100543
31. Mariano-Hernández D, Hernández-Callejo L, Solís M, et al. A data-driven forecasting strategy to predict continuous hourly energy demand in smart buildings. *Appl Sci*. 2021;11:7886. doi:10.3390/app11177886
32. Ibrahim IA, Khatib T. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. *Energy Convers Manag*. 2017;138:413-425. doi:10.1016/j.enconman.2017.02.006
33. Ma J, Ding Y, Cheng JCP, Tan Y, Gan VJL, Zhang J. Analyzing the leading causes of traffic fatalities using XGBoost and grid-based analysis: a city management perspective. *IEEE Access*. 2019;7:148059-148072. doi:10.1109/ACCESS.2019.2946401
34. Lemos VHB, Almeida JDS, Paiva AC, et al. Temporal convolutional network applied for forecasting individual monthly electric energy consumption. 2020 IEEE Int. Conf. Syst. Man, Cybern. IEEE; 2020:2002–2007. doi:10.1109/SMC42975.2020.9282960
35. Wang Z, Wang Y, Zeng R, Srinivasan RS, Ahrentzen S. Random forest based hourly building energy prediction. *Energy Build*. 2018;171:11-25. doi:10.1016/j.enbuild.2018.04.008

How to cite this article: Mariano-Hernández D, Hernández-Callejo L, Solís M, et al. Comparative study of continuous hourly energy consumption forecasting strategies with small data sets to support demand management decisions in buildings. *Energy Sci Eng*. 2022;10:4694-4707. doi:10.1002/ese3.1298