



Deep learning for prediction of energy consumption: an applied use case in an office building

Roberto Morcillo-Jimenez¹ · Jesús Mesa¹ · Juan Gómez-Romero¹ · M. Amparo Vila¹ · Maria J. Martin-Bautista¹

Accepted: 7 April 2024 / Published online: 30 April 2024
© The Author(s) 2024

Abstract

Non-residential buildings are responsible for more than a third of global energy consumption. Estimating building energy consumption is the first step towards identifying inefficiencies and optimizing energy management policies. This paper presents a study of Deep Learning techniques for time series analysis applied to building energy prediction with real environments. We collected multisource sensor data from an actual office building under normal operating conditions, pre-processed them, and performed a comprehensive evaluation of the accuracy of feed-forward and recurrent neural networks to predict energy consumption. The results show that memory-based architectures (LSTMs) perform better than stateless ones (MLPs) even without data aggregation (CNNs), although the lack of ample usable data in this type of problem avoids making the most of recent techniques such as sequence-to-sequence (Seq2Seq).

Keywords Buildings · Energy consumption forecasting · Time series · Deep learning · XGBoost · Multilayer perceptron · Recurrent neural networks · Convolutional neural networks · Sequence to sequence networks

1 Introduction

The current unstable situation in Europe has led to a decrease in the amount of energy supplied to the European continent, causing an exorbitant increase in prices across the continent. It is therefore necessary to reduce energy consumption in general, to prevent inflation from increasing further and the energy crisis from continuing to grow. Our work focuses on increasing energy efficiency in office buildings, as this is one of the most energy intensive areas due to the number of

hours they are in operation. Reducing energy consumption by controlling consumption through algorithmic artificial intelligence techniques is one of the goals of our society in the coming years and our case study.

According to [26], residential, office, and industrial buildings account for 20–40% of global energy consumption. In 2010 in the USA, buildings energy consumption accounted for 41% of their total energy consumption, with 75% of this energy coming from fossil fuels [17]. Meanwhile, in 2012 in the European Union, buildings consumed approximately 40% of the total energy used [17]. More than two-thirds of the energy consumed by buildings goes to heating systems (37%), water heating (12%), air conditioning (10%) and lighting (9%). Several case studies have shown that the operational phase is the most energy-consuming stage of the buildings life cycle, accounting for 90% in conventional buildings and 50% in low-energy buildings [29].

In the current context in Europe, it is important to reduce energy consumption due to the high prices being achieved in the various energy markets. It is not surprising that interest in improving the energy efficiency of buildings has increased. According to [6] and [4], this interest is driven by three factors: rising energy prices, increasingly restrictive environmental regulations and increased environmental awareness among citizens. All over the world, public policies are being

✉ Roberto Morcillo-Jimenez
robermorji@decsai.ugr.es

Jesús Mesa
sulimesa@gmail.com

Juan Gómez-Romero
jgomez@decsai.ugr.es

M. Amparo Vila
avila@decsai.es

Maria J. Martin-Bautista
mbautis@decsai.ugr.es

¹ Department of Computer Science and Artificial Intelligence, Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación, Universidad de Granada, Granada 18071, Spain

developed to increase this efficiency, as reflected in the Unsustainable Development Goals and the European Green Deal.

One of the first steps towards improving building energy efficiency is studying how consumption happens and knowing the factors that significantly impact it. In particular, predicting energy consumption from historical data allows building operators and managers to anticipate peaks in energy demand, modify building uses to shift this demand and plan equipment operation appropriately. Likewise, accurate consumption prediction allows assessing the improvement in building performance when making improvements or implementing new energy policies by comparing the actual consumption with the estimated consumption (or baseline). Data Science has emerged as an effective tool to address these objectives [21].

In the literature, there can be found numerous proposals for estimating energy consumption in buildings throughout time series prediction. Traditionally, these approaches were based on numerical regression or moving average models, which have limitations regarding multivariate series and series with changing trends. In contrast, modern machine learning methods based on neural networks have shown more effective in those scenarios [34]. However, the application of these techniques is often hampered by the noisy and incomplete nature of building energy data in real environments [9], which results in a gap between theoretical and practical works.

The rationale behind this work is to perform a comprehensive evaluation of neural network methods in a real-world scenario and to draw conclusions for practical application in similar contexts. More specifically, this paper studies the accuracy of several methods for heating consumption prediction, namely XGBoost, MLPs, RNNs, CNNs and Seq2Seq. The need and impact of preprocessing, which is applied to remove noisy and missing values and for data reduction, is also discussed.

The dataset was collected in the ICPE office building located in Bucharest, one of the pilot buildings considered in the Energy IN TIME¹. Our starting hypothesis is that modern neural network techniques improve the performance of other approaches, and within them, memory-based architectures (RNNs, Seq2Seq) are superior. The results show that the hypothesis holds, despite the risk of overfitting these techniques when applied to not very large datasets.

The remainder of this paper is structured as follows. We first provide a review of related works on prediction of building energy consumption (Section 2). Next, we describe the data used in the study (Section 3), the methodology (Section

4), and the experimentation (Section 5). At the end of the paper conclusions and directions for future research work (Section 6) will be exposed.

2 Related work

Energy consumption forecasting models can be differentiated into categories based on their respective energy end-uses, such as cooling, heating, space heating, primary, natural gas, electricity, and steam load consumption [35]. Regarding the application of the models, [33] identified two primary categories: (1) model-based control, demand response, and optimization of energy consumption in buildings; and (2) design and modernization of building parameters, including energy planning and assessing the impact of buildings on climate change.

Also in [33], numerous factors impacting energy consumption were also identified, mainly the number and type of buildings under consideration. The temporal horizon of the prediction and the resolution of the sensor data are also relevant. Remarkably, natural time-based groupings (i.e., hours, days, months, and years) were proved superior in [35] to the more generic short-, medium-, and long-term ranking schemes proposed in [1].

Many data-driven techniques have proved effective for estimating building energy consumption, ranging from classical statistical regression to modern deep learning architectures. Regarding the former, [15] evaluated four models incorporating exogenous inputs, specifically autoregressive moving average models with exogenous inputs (ARMAX). [10] proposed a system utilizing a seasonal autoregressive integrated moving average model (SARIMA) and a least squares support vector regression model based on firefly metaheuristic algorithms (MetaFA-LSSVR). The prediction system yielded highly accurate and reliable day-ahead predictions of building energy consumption, with an overall error rate of 1.18%.

Another interesting study is [19], which investigated two stochastic models for short-term time series prediction of energy consumption, namely the conditional constrained Boltzmann machine (CRBM) and the factored conditional constrained Boltzmann machine (FCRBM). In the comparison, the results showed that the FCRBM outperformed the artificial neural network, support vector machine, recurrent neural networks and CRBM. The work was extended to include a deep belief network with automated feature extraction for the short-term building energy modeling process [20]. Other relevant work applying classical learning techniques is [24], which used a decision tree method (C4.5). In addition to obtaining accurate results, this algorithm was able to identify the factors contributing to building energy use. A sophisticated regression tree algorithm (Chi-Square

¹ Energy IN TIME was an European project running in 2013-2017. The aim of the project was to implement a model-predictive control system to improve the energy efficiency of non-residential buildings [13].

Automatic Interaction Automatic Detector) was used in [11] to predict short-term heating and cooling load.

Raza and Khosravi [28] provided a comprehensive review of artificial intelligence-based load demand forecasting techniques for smart grids and buildings. The authors explored various machine learning algorithms used in load forecasting, including artificial neural networks, fuzzy logic, genetic algorithms, and support vector regression. [3] also reviewed the applications of artificial neural networks (ANNs) and support vector machines (SVMs) for building electrical energy consumption forecasting. The authors compared the performance of ANNs and SVMs with traditional statistical methods used in energy forecasting. [30] explored the different ML algorithms used in the field, including artificial neural networks, support vector regression, decision trees, and clustering methods. They also discussed the challenges associated with accurate data acquisition and modelling and the limitations of different ML algorithms. More recently, [14] emphasized that applying machine learning and statistical analysis techniques can lead to significant energy savings and cost reductions. [37] examined the advantages and limitations of different ML algorithms, including artificial neural networks, decision trees, and support vector machines, and explore their application in different building load prediction scenarios. [27] proposed using deep recurrent neural networks (DRNNs) for predicting heating, ventilation, and air conditioning (HVAC) loads in commercial buildings. The authors describe the architecture and training process of the DRNN model, which included a combination of convolutional neural networks (CNNs) and long short-term memory (LSTM) layers.

Finally, other works showed the importance of preprocessing in building energy forecasting, e.g., data cleaning and feature selection. For instance, [2] reviewed the current development of machine learning (ML) techniques for predicting building energy consumption and discussed the challenges associated with data acquisition, feature selection, and model validation.

Table 1 summarises the related works mentioned in this section.

3 Data

Our data was collected from the ICPE building (Institute of Technologies for Sustainable Development) in Bucarest (Romania). This is a three-building, each one divided into areas. For the experiments, we defined a pilot zone through a transversely cut of the building covering three areas (D1, D2, D5/2) of the three floors (see Figs. 1, 2).

The building is equipped with sensors to measure room temperatures and meters measuring electricity, heating, and water consumption in distinct zones of the building. We

focused on predicting heating meter values, given the higher contribution of this subsystem to the total energy consumption. Due to the low temperatures in these countries, buildings are powered by a system called district heating, which is used to prevent pipes from freezing outside, and then energy is consumed inside the building to raise the temperature of the water before it is distributed to the radiators inside.

We considered as prediction targets the three heating meters that respectively cover the area D1, D2 and D5/2 from buildings. However, we found a problem with area D2 data was not available because of errors in sensors, we only selected zone D1 and zone D5/2 of our building. The variables that we select for the D1 zone of our building are H-F123-D1, where all the heating consumption of the D1 zone in the floor 1, 2, and 3 are grouped, and for the D5/2 zone are H-F123-D5/2W and H-F123-D5/2E, where the heating consumption are grouped as the previous one, both for the West zone and for the East zone. We have other external variables that are related to our building, these variables are the outdoor temperature and the building occupancy, which are crucial for our experiments.

We collected data in 2017 before the implementation of the new control system. We focused on heating, and only the cold season (January, February and March) was considered in the experimentation. We re-sampled all variables at 15-minute intervals and calculated the cumulative consumption values. As a result, we obtained a data set with 8640 samples. The dataset is not publicly available but an excerpt can be obtained on demand.

4 Methodology

The methodological approach of the experiments followed the workflow of the Data Science process applied to energy data presented in [21]. As shown in Fig. 3, we retrieved the data through an API REST and carried out data preprocessing, including outliers removal, missing value imputation, normalization and feature selection. Afterwards, we continued splitting these data by months to apply the learning algorithms and get the best prediction models for each period. Next, we describe how data is partitioned, as well as the parameters used to train the models and the evaluation metrics.

4.1 Definition of the prediction problem and data partition

The problem we aim to solve is to predict the energy usage of a building for the next 12 hours. We need to partition the original dataset into training and validation samples in chronological order to train our model. Given the different energy consumption patterns at the beginning and the end of

Table 1 Summary of related works

Authors	Approach	Methods	Dataset
Khalil et al. [14]	Machine Learning Algorithms	Artificial neural networks, decision trees and SVR	Multiple datasets
Pachauri and Ahn [24]	Machine Learning Algorithms	Decision tree	Japanese residential buildings
Molina-Solana et al. [21]	Building operation, Fraud detection Applications	MPC	Sanomatato Building
Wang and Srinivasan [35]	Artificial Intelligence, Artificial Neuronal Networks	MLR, ANNs, SVR, ensemble models	Multiple datasets
Zhang et al. [37]	Machine Learning	MLR, ANNs, SVR	Multiple datasets
Arpanahi and Javadi [3]	Machine Learning, forecasting methods	ANNs, SVR, GMDH, LSSVM	Multiple datasets
Raza and Khosravi [28]	Short term, load forecasting	ANNs	Multiple datasets
Lago et al. [15]	Forecasting	SVR with chaotic gravitational search	Historical electric data from the Northern China
Chou and Ngo [10]	Machine Learning	Time Series Analytics	Multiple datasets
Mocanu et al. [19]	Artificial Neuronal Networks	CRBM, FCRBM	Multiple datasets
Rahman et al. [27]	Artificial Neuronal Networks	Multiple algorithms	Multiple datasets
Seyedzadeh et al. [30]	Machine Learning	ANNs, SVM, Gaussian-based regressions, clustering	Energy benchmarking
Ahmad et al. [2]	Building energy modeling Building optimal control	Agent-Based Model, System Identification	Building Energy model
Ahmad and Chen [1]	Load forecasting, Data mining based approaches	TB, BoostedT, GPR, NN and BaggedT	Office building in Beijing, China
Ezan et al. [11]	Energy consumption: Pattern prediction, Time-series technique, Metaheuristic optimization, Machine learning	ARIMA, SARIMA, LSSVR, ML, SVR, ANNs, FA	Smart grid infrastructure
Mocanu et al. [20]	Reinforcement learning, Deep Belief Networks, Machine learning	DBN, SARSA, Q-learning algorithm	Multiple datasets
Tien et al. [33]	Machine learning techniques such as supervised, unsupervised, and reinforcement learning	Linear regression, decision trees, random forests, and neural networks, clustering and anomaly detection, Q-learning	Multiple datasets

the data collection period, we decided to split the full dataset into smaller chunks (namely, subproblems) and then make the training and validation partitions within each one. This implies that we have a different model trained specifically for each subproblem, which must be conveniently selected during the test phase to calculate the predictions depending on the date of the observations.

To perform the splits, we applied time series decomposition to analyze the evolution of the energy consumption variables and identify potential cut points, focusing on trend change. This study, which is explained below, showed that we could safely split the data in months since the time series have consistent behaviour in each of these periods. Furthermore, a preliminary analysis of the stability of the training, in terms of errors obtained with slightly different splits, showed that the differences were not significant. While this is a rough approximation of a more fine-grained splitting, it has the advantage of facilitating the selection of the model to be

applied for prediction. It remains for future work to apply a more sophisticated time series splitting algorithm [18] and a more comprehensive analysis of the impact of the splitting.

The analysis of the trend change to assess the monthly splitting was based on the decomposition of the time series with STL technique (Seasonal-Trend decomposition using LOESS), which has been previously used for energy demand time series [25]. Figure 4 depicts the decomposition of the heating consumption variables into three components: trend, i.e., long-term evolution of values; seasonality, i.e., repetitions in fixed periods; and irregularity, i.e., random patterns that remain after removing the other two components. It can be seen that there is a clear pattern of increasingly high consumption in January, followed by a steady decline in consumption in February and a plateau in March. The irregularity component throughout the whole series remains quite stable, meaning that the trend and season components capture quite well the changes in the series. Therefore, we proceeded to



Fig. 1 Aerial view

split the dataset in two parts, training and validation. Specifically, the selected training samples were 70% of each month’s observations (e.g., from 2016-01-05 00:15:00 until 2016-01-21 23:45:00 in January), while the remaining values were used for validation.

4.2 Methods

The methods used to build the predictions models are XGBoost [8] and a selection of neural networks including CNNs [7], RNNs [31], MLPs [32] and Seq2Seq [12]. The results with Seq2Seq suggested that other techniques, such as transformers, would not be very useful given the relatively small size of each partition, in line with the recent literature [36].

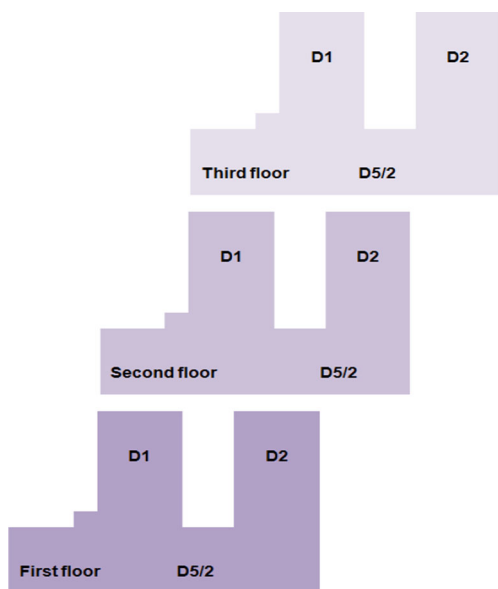


Fig. 2 Pilot zone

Table 2 summarises the configuration of the models and the hyperparameters probed with each technique. For XGBoost, we reflect the number of trees, maximum depth, and learning rate. For CNN, we describe the number of convolutional layers, max pooling 1D, and the number of filters. For MLP, RNN and Seq2Seq, the table lists the number of hidden layers, dropout layers, and layer size. The best configurations obtained with the training data are highlighted in section 5.2.

4.3 Metrics

To validate and compare the different results of our experiments, we must define the error metrics to be minimized. Since we have a regression problem, we use MAE (Mean Absolute Error), which aggregates at m the absolute difference between the predictions and the actual values of $delay$ data points:

$$MAE(m) = \frac{\sum_{i=0}^{delay-1} |y_i - \hat{y}_i|}{delay} \tag{1}$$

We also use the normalized MAE, namely NMAE, which averages the error for a batch of size N .

$$NMAE = \frac{\sum_{m=0}^{N-1} MAE(m)}{N} \tag{2}$$

5 Experiments and results

This section shows the results of the experiments after data preprocessing and model training and validation with the algorithms mentioned above: decision trees, XGBoost, and neural networks (MLPs, CNNs, RNNs and Seq2Seq). The implementation of the preprocessing and the prediction algorithms was developed with TSxtend [23], our open source library for batch analysis of sensor data.

5.1 Preprocessing and data preparation

In our experiment, we used processing data techniques such as aggregation, data modification and removal, data transformation, outliers detection, missing values detection, normalization and feature selection. We discarded heating consumption with null values to discard rows with no heating consumption. Then, we eliminated the variables with extreme values and finally joined the variables measuring the same consumption type on the same floor.

The variables in the collected dataset store aggregated values, e.g., summing up new values at each instant. The transformation of these variables was done with the differences between each instant of time, thus obtaining the actual

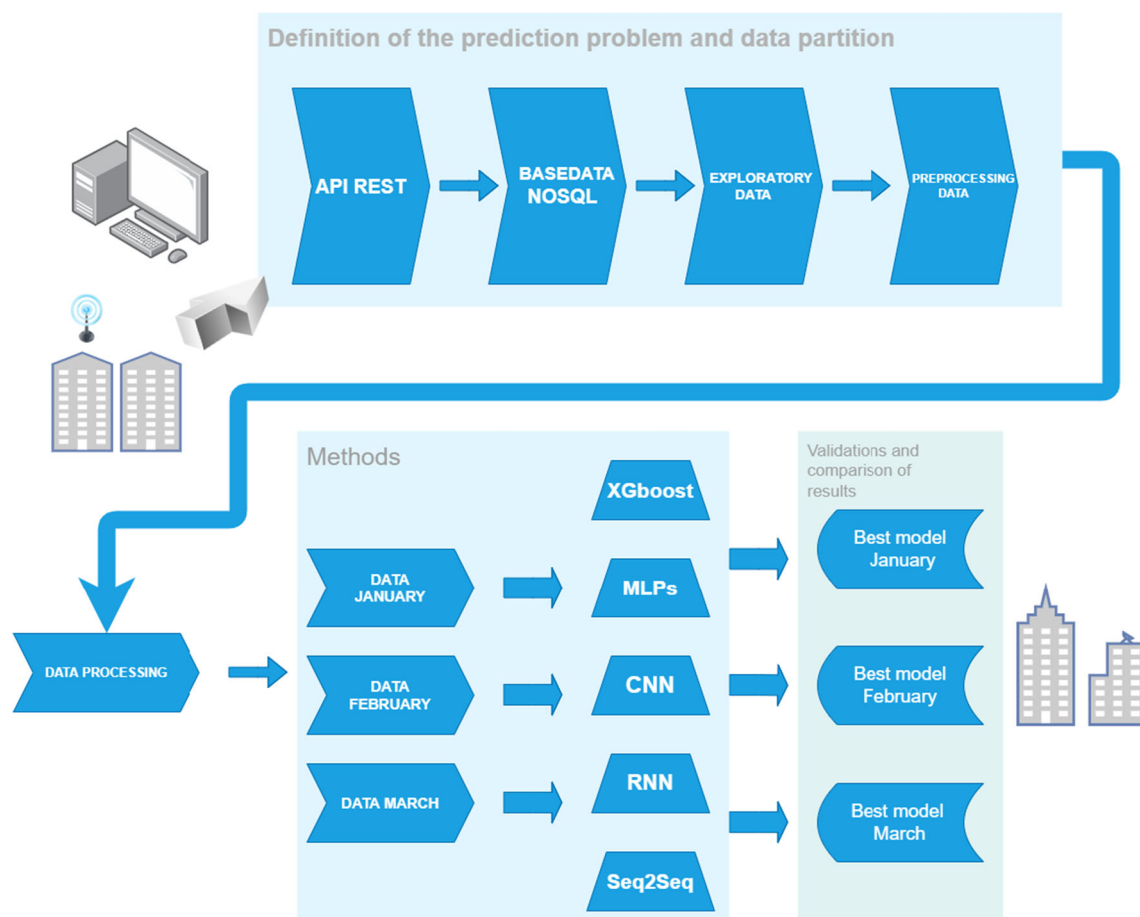


Fig. 3 Methodology

consumption of each variable at each instant of time. These variables were used as inputs for the execution of our prediction algorithms.

In the subsequent step, we employed techniques to detect and remove outliers for each consumption variable. We calculated percentiles for each variable and searched for values outside this range. Such values were substituted with more conventional values (mean, median, etc...). This approach, however, leads to a loss of information. Fortunately, the number of outliers in our case is limited.

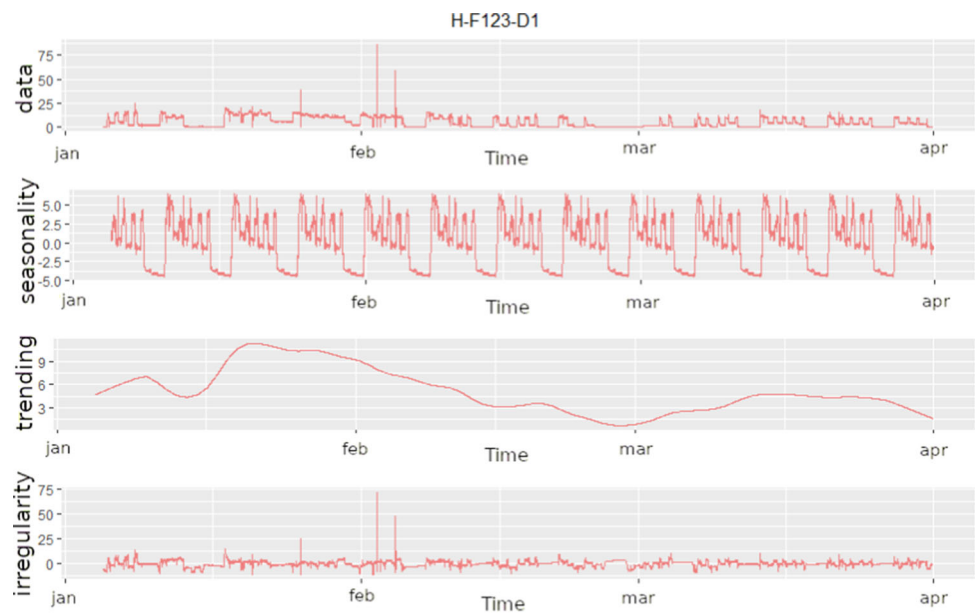
We had in our dataset missing values in February when the system did not gather data, for which we applied missing values imputation algorithms. Particularly, we used Interp to fill gaps by interpolation of known values. Interpolation techniques are utilized to fill gaps by interpolating known values, as mentioned in work by [22]. Specifically, this approach involves the utilization of Interp, a commonly employed software for performing interpolation. This methodology aims to accurately estimate values within a given range based on the available data. The interpolation process involves estimating each data point's value by considering its neighbouring data points. Applying this technique makes the resulting

dataset complete and can be analyzed more effectively for the intended purpose. Figure 5 depicts data after imputation for three variables.

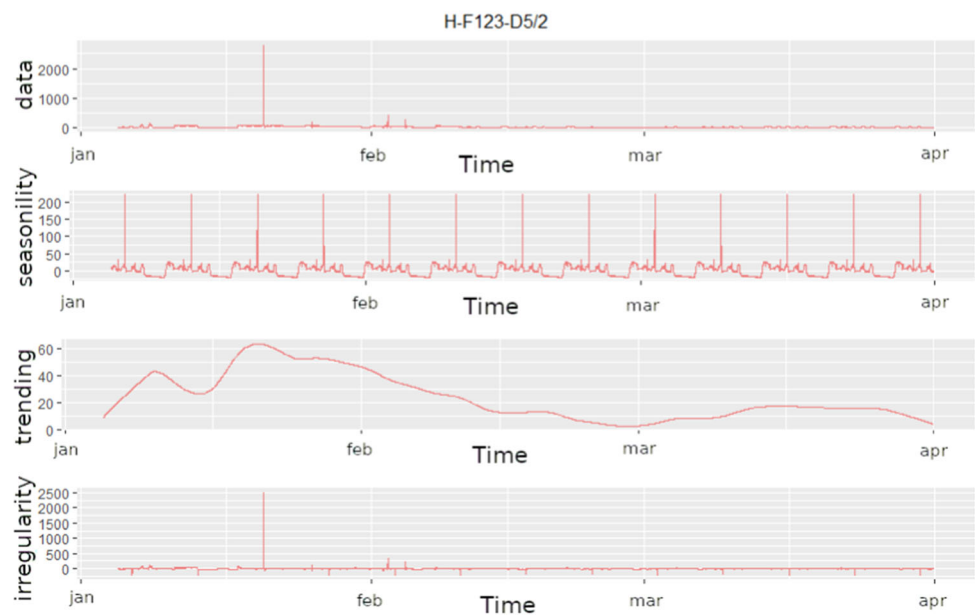
Then, we selected the most relevant variables for predicting energy consumption as Energy Zone D1 and Energy Zone D5/2. We calculated the cross-correlation between pairs of variables and grouped them by levels according to their correlation to the prediction target. Figure 6 indicates few dependencies between the electrical sensors. There is a high correlation between the heating consumption variables, while water consumption does not correlate. In the case of the occupancy variables, there are more correlations, but it is not high because the values are estimated. We can also observe energy consumption when the building is empty.

Then we used XGBoost algorithm for the assessment of variable importance. In Fig. 7, we can see that the most relevant variables for energy consumption are occupancy, outside temperature and heating consumption (Energy Zone D1 and Energy Zone D5/2). With this information, we performed the selection of variables, being the most important ones (not surprisingly) Outside Temperature, Occupation, Energy Zone D5/2 and Energy Zone D1.

Fig. 4 Decomposition of the heating consumption variables of plants 1, 2, and 3 in the pilot zone of the ICPE building



(a) H-F123-D1.



(b) H-F123-D5/2.

The last step in order to preprocess data is to normalize the dataset. To do this, we calculated the mean and standard deviation of each variable, and then transformed the values to the range $[0, 1]$.

5.2 Results and discussion

We experimented with the January, February and March data using the selected algorithms. We applied grid search to

obtain the best configurations and hyperparameter, yielding the values depicted in Table 3. We can observe that the best configurations are similar for each subproblem.

The validation errors of the best models are shown in Tables 4 and 5, while Figs. 8, 9, 10, 11, 12, and 13 depict the predicted vs the real values. In January, the algorithm that offered the best results was Seq2Seq, which achieved a validation NMAE of 0.21 for the prediction of consumption in zone D1 and 0.20 for zone D5/2, significantly improving the

Table 2 Hyperparameter grids used to tune the learning algorithms

Algorithms	Parameters	Range
XGBoost	Number of trees	{50,100,50}
	Maximum Depth	{51,100,50}
	Learning Rate	{0.05,0.001,0.1}
CNN	Number of convolutional layers	{1,2}
	Max pooling 1D	{3,5,7}
	Number of filters	{16,32,64,128}
	Number of hidden layers	{1,2}
MLP, RNN, Seq2Seq	Number of dropout layers	{0.15,0.3,0.5}
	Number of neurons	{16,32,64,128}

Fig. 5 Imputation of missing values ((end February) with Interpolate for variables Energy Zone D1, E-F1-D5/2-C03 and W-ALL

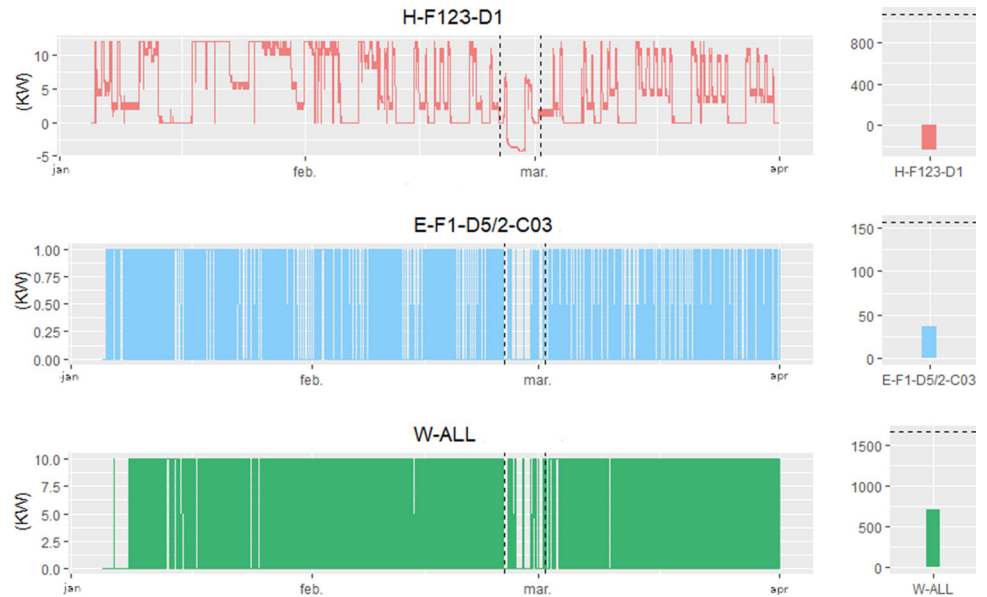


Fig. 6 Heatmap showing correlations between variables

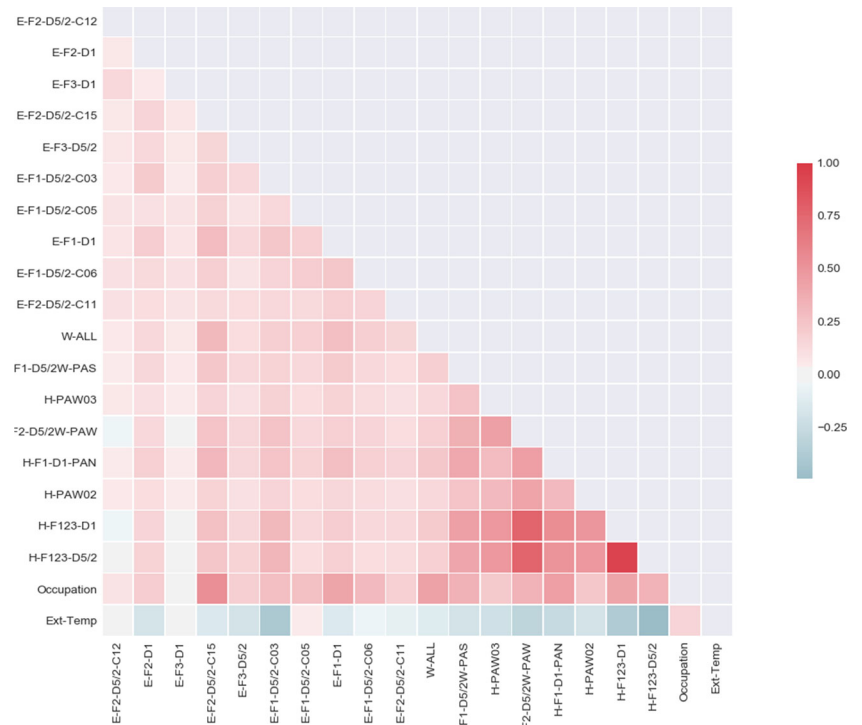


Fig. 7 Ranking of variables by importance with XGBoost

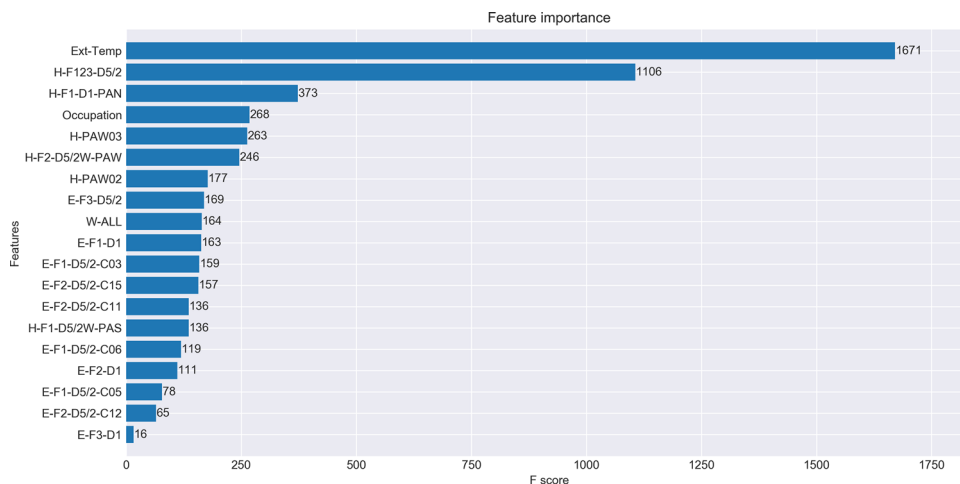


Table 3 Best configurations and hyperparameters for each algorithm and subproblem

Algorithms	Parameters	Month January	Month February	Month March
XGBoost	Number of trees	{50}	{100}	{50}
	Maximum Depth	{51}	{100}	{51}
	Learning rate	{0.05}	{0.001}	{0.005}
CNN	Number of convolutional layers	{1}	{1}	{1}
	Max pooling 1D	{3}	{3}	{3}
	Number of filters	{32}	{16}	{16}
MLP	Number of hidden layers	{1}	{1}	{1}
	Number of dropout layers	{0.3}	{0.15}	{0.15}
	Number of neurons	{128}	{128}	{128}
RNN	Number of hidden layers	{1}	{1}	{1}
	Number of dropout layers	{0.15}	{0.3}	{0.5}
	Number of neurons	{128}	{128}	{128}
Seq2Seq	Number of hidden layers	{2}	{2}	{2}
	Number of dropout layers	{0.3}	{0.3}	{0.3}
	Number of neurons	{64}	{64}	{64}

Table 4 Performance of the best models for Energy Zone D1 (best result for each month and metric in bold font)

Models	January		February		March	
	NMAE	MAE (kw)	NMAE	MAE (kw)	NMAE	MAE (kw)
XGBoost	0.48	2.5	1.17	5.88	0.56	2.24
MLPs	0.38	1.97	0.82	4.12	0.53	2.12
RNNs	0.27	1.40	0.43	2.16	0.29	1.16
CNNs	0.34	1.77	0.43	2.16	0.34	1.36
Seq2Seq	0.21	1.09	0.43	2.16	0.39	1.56

Table 5 Performance of the best models for Energy Zone D5/2 (best result for each month and metric in bold font)

Models	January		February		March	
	NMAE	MAE (kw)	NMAE	MAE (kw)	NMAE	MAE (kw)
XGBoost	0.56	16.05	1.10	26.125	0.60	8.25
MLPs	0.31	8.88	0.61	14.48	0.64	8.8
RNNs	0.30	8.6	0.31	7.36	0.29	3.98
CNNs	0.36	10.32	0.30	7.12	0.33	4.53
Seq2Seq	0.20	5.73	0.36	8.55	0.37	5.08

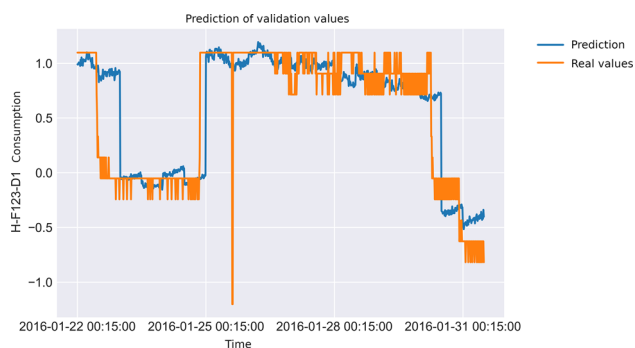


Fig. 8 Prediction of the Validation Values in January with Seq2Seq Energy Zone D1

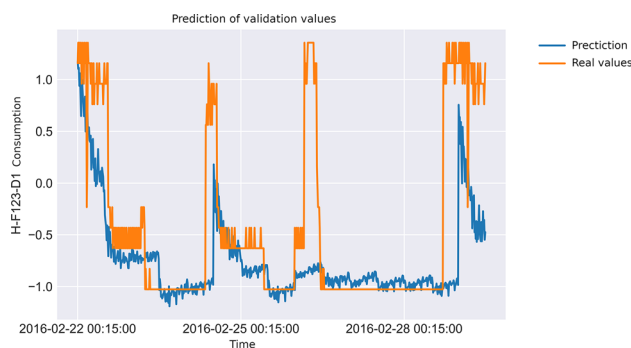


Fig. 10 Prediction of the Validation Values in February with CNNs Energy Zone D1

performance of the other models. The best model for February in both zones was CNNs. Overfitting was lower than that with other algorithms, at the cost of having worse validation metrics —the model tends to ignore or lessen the consumption peaks, as it can be seen in the chart. A second reason to explain these results is that we had many more missing values to impute in this period. The best method for March was RNNs, with an NMAE of 0.29 in both zones, almost twice better than MLP and XGBoost. Overall, we confirmed the ability of memory-based models (Seq2Seq and RNN) to extract characteristics from the input time series and learn the predictions in more diverse scenarios, while CNNs were slightly better when there was a uniform trend and fewer data. In all cases, the best models' MAE was around 5 kW, which is small enough for this type of applications.

Regarding the drawback of having limited data for training and validation, it resulted in that the more complex models did not learn the prediction of heating consumption as well as it could be expected. Hence, we suggest that more sophisticated techniques (Seq2Seq, but also transformer-based architectures) might not be necessary in this kind of problems or under similar circumstances. Instead, RNNs or CNNs with proper data pre-processing could be precise enough and less prone to overfitting.

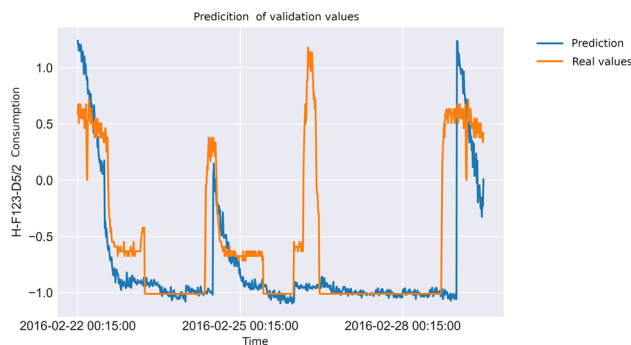


Fig. 11 Prediction of the Validation Values in February with CNNs Energy Zone D5/2

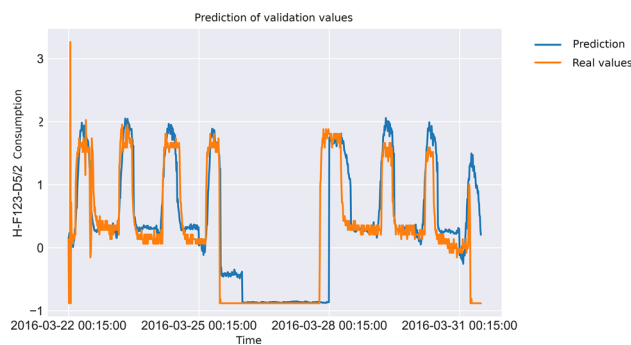


Fig. 12 Prediction of the Validation Values in March with RNNs Energy Zone D1

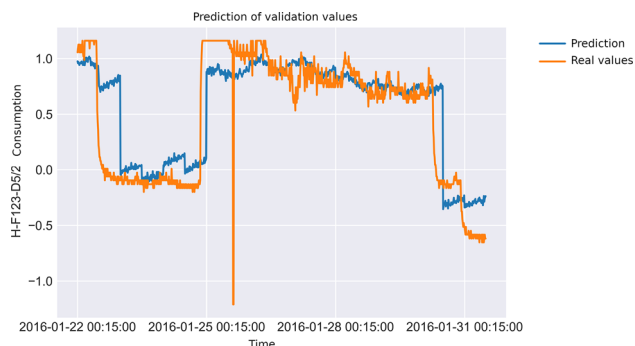


Fig. 9 Prediction of the Validation Values in January with Seq2Seq Energy Zone D5/2

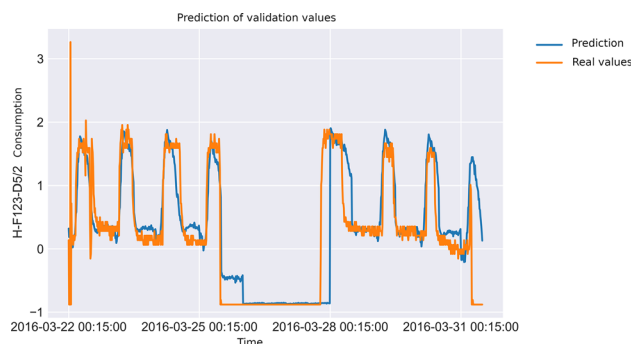


Fig. 13 Prediction of the Validation Values in March with RNNs Energy Zone D5/2

6 Conclusions and future work

Concluding our discussion, we can see how the XGBoost model shows worse results, as XGBoost does not extract the dependencies between the variables it receives as input to perform the energy consumption prediction. With MLP models we have shown that they do not provide better results than other models because they do not allow to follow a chronological order in our dataset. With respect to the use of RNNs on our dataset, we have been able to reduce the NMAE error function by half, as these algorithms remember the information by processing the data in chronological order. Furthermore, we have applied Seq2Seq, which has allowed us to observe that, like RNNs, they obtain good results in general but do not perform very well on datasets with too many missing values (i.e., smaller in size). Finally, convolution networks (CNNs) have been found to perform better than RNNs and Seq2Seq algorithms on sections of the dataset with a large number of missing values and steady trends.

One of the problems we have in this study has been the limited amount and quality of data, since we have only worked with the ICPE building sensors for three months including many missing values. In the future we can take the ICPE building sensors from other years to have a more extensive training and validation. Another option may be to generate artificial values with building simulation models. We can also use recurrent neural network models using attention mechanisms to improve synthetic data generation and missing values imputation [5]. Additionally, peak changes could be addressed with noise reduction techniques to smooth abrupt oscillations, e.g., as the filter proposed in [16].

Acknowledgements This work was funded by the European Union NextGenerationEU/PRTR through the IA4TES project (MIA.2021.M04.0008); by MICIU/AEI/10.13039/501100011033 through the SINCERGY project (PID2021.125537NA.I00); by ERDF/Junta de Andalucía through the D3S project (P21.00247); by the FEDER programme 2014-2020 (B-TIC-145-UGR18 and P18-RT-1765); and by the European Union (Energy IN TIME EeB.NMP.2013-4, No. 608981).

Funding Funding for open access publishing: Universidad de Granada/CBUA.

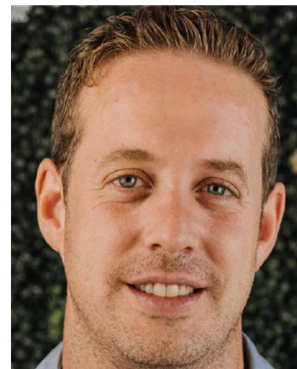
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad T, Chen H (2018) Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy Build* 166:460–476. <https://doi.org/10.1016/J.ENBUILD.2018.01.066>
- Ahmad T, Chen H, Guo Y et al (2018) A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy Build* 165:301–320. <https://doi.org/10.1016/j.enbuild.2018.01.017>
- Arpanahi G, Javadi M (2018) A review on applications of artificial neural networks and support vector machines for building electrical energy consumption forecasting. *Renew Sustain Energy Rev* 82:1814–1832. <https://doi.org/10.1016/j.rser.2014.01.069>
- Benedetti M (2015) A proposal for energy services classification including a product service systems perspective. *Procedia CIRP* 30:251–256. <https://doi.org/10.1016/j.procir.2015.02.121>
- Bülte C, Kleinebrahm M, Yilmaz HÜ et al (2023) Multivariate time series imputation for energy data using neural networks. *Energy and AI* 13:100,239. <https://doi.org/10.1016/j.egyai.2023.100239>, <https://www.sciencedirect.com/science/article/pii/S2666546823000113>
- Chau C, Leung T, Ng W (2015) A review on life cycle assessment, life cycle energy assessment and life cycle carbon emissions assessment on buildings. *Appl Energy* 143:395–413. <https://doi.org/10.1016/j.apenergy.2015.01.023>
- Chauhan R, Ghanshala KK, Joshi RC (2018) Convolutional Neural Network (CNN) for Image Detection and Recognition. *ICSCCC 2018 - 1st International Conference on Secure Cyber Computing and Communications* pp 278–282. <https://doi.org/10.1109/ICSCCC.2018.8703316>
- Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system pp 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen W, Zhou K, Yang S et al (2017) Data quality of electricity consumption data in a smart grid environment. *Renew Sustain Energy Rev* 75:98–105. <https://doi.org/10.1016/j.rser.2016.10.054>
- Chou JS, Ngo NT (2016) Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns. *Appl Energy* 177:751–770. <https://doi.org/10.1016/J.APENERGY.2016.05.074>
- Ezan MA, Uçan ON, Kalfa M (2017) Predicting short-term building heating and cooling load using regression tree algorithm. *J Build Perform Simul* 10(5):487–502. <https://doi.org/10.1080/19401493.2016.1202888>
- Gong G (2019) Research on short-term load prediction based on Seq2Seq model. *Energies* 12:3199. <https://doi.org/10.3390/en12163199>
- Gómez-Romero J, Fernández-Basso CJ, Cambronero MV et al (2019) A probabilistic algorithm for predictive control with full-complexity models in non-residential buildings. *IEEE Access* 7:38,748–38,765. <https://doi.org/10.1109/ACCESS.2019.2906311>
- Khalil M, McGough AS, Pourmirza Z et al (2022) Machine learning, deep learning and statistical analysis for forecasting building energy consumption - a systematic review. *Eng Appl Artif Intell* 115(105):287. <https://doi.org/10.1016/j.engappai.2022.105287>
- Lago J, Marcjasz G, De Schutter B et al (2021) Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Appl Energy* 293(116):983. <https://doi.org/10.1016/j.apenergy.2021.116983>

16. Ma W, Wang W, Wu X et al (2019) Control strategy of a hybrid energy storage system to smooth photovoltaic power fluctuations considering photovoltaic output power curtailment. *Sustain* 11(5). <https://doi.org/10.3390/su11051324>, <https://www.mdpi.com/2071-1050/11/5/1324>
17. Marino DL, Amarasinghe K, Manic M (2016) Building energy load forecasting using deep neural networks pp 7046–7051. <https://doi.org/10.1109/IECON.2016.7793413>
18. Micheletti A, Aletti G, Ferrandi G et al (2020) A weighted χ^2 test to detect the presence of a major change point in non-stationary markov chains. *Stat Methods Appl* 29(4):899–912. <https://doi.org/10.1007/s10260-020-00510-0>, <https://doi.org/10.1007/s10260-020-00510-0>
19. Mocanu E, Nguyen PH, Gibescu M et al (2016) Deep learning for estimating building energy consumption. *Sustain Energy, Grids Netw* 6:91–99. <https://doi.org/10.1016/j.segan.2016.02.005>
20. Mocanu E, Nguyen PH, Kling WL et al (2016) Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning. *Energy and Buildings* 116:646–655. <https://doi.org/10.1016/J.ENBUILD.2016.01.030>
21. Molina-Solana M, Ros M, Ruiz MD et al (2017) Data science for building energy management: A review. *Renew Sustain Energy Rev* 70:598–609. <https://doi.org/10.1016/j.rser.2016.11.132>
22. Montero-Manso P, Hyndman RJ (2021) Principles and algorithms for forecasting groups of time series: Locality and globality. *Int J Forecast* 37(4):1632–1653. <https://doi.org/10.1016/j.ijforecast.2021.03.004>
23. Morcillo-Jimenez R, Gutiérrez-Batista K, Gómez-Romero J (2023) Tsxtend: A tool for batch analysis of temporal sensor data. *Energies* 16(4). <https://doi.org/10.3390/en16041581>, <https://www.mdpi.com/1996-1073/16/4/1581>
24. Pachauri N, Ahn CW (2022) Regression tree ensemble learning-based prediction of the heating and cooling loads of residential buildings. *Build Simul* 15(11):2003–2017. <https://doi.org/10.1007/s12273-022-0908-x>
25. Phinikarides A, Makrides G, Zinsser B et al (2015) Analysis of photovoltaic system performance time series: Seasonality and performance loss. *Renew Energy* 77:51–63. <https://doi.org/10.1016/j.renene.2014.11.091>, <https://www.sciencedirect.com/science/article/pii/S0960148114008222>
26. Pérez-Lombard L, Ortiz J, Pout C (2008) A review on buildings energy consumption information. *Energy Build* 40(3):394–398. <https://doi.org/10.1016/j.enbuild.2007.03.007>
27. Rahman A, Srikumar V, Smith AD (2018) Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl Energy* 212:372–385. <https://doi.org/10.1016/j.apenergy.2017.12.051>, <https://www.sciencedirect.com/science/article/pii/S0306261917317658>
28. Raza MQ, Khosravi A (2015) A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renew Sustain Energy Rev* 50:1352–1372. <https://doi.org/10.1016/j.rser.2015.04.065>
29. Schmidt M, Åhlund C (2018) Smart buildings as cyber-physical systems: Data-driven predictive control strategies for energy efficiency. *Renew Sustain Energy Rev* 90:742–756. <https://doi.org/10.1016/j.rser.2018.04.013>
30. Seyedzadeh S, Rahimian FP, Glesk I et al (2018) Machine learning for estimation of building energy consumption and performance: a review. *Vis Eng* 6:1–20. <https://doi.org/10.1186/s40327-018-0064-7>
31. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenom* 404(132):306. <https://doi.org/10.1016/j.physd.2019.132306>
32. Taud H, Mas J (2018) Multilayer perceptron (MLP) pp 451–455. https://doi.org/10.1007/978-3-319-60801-3_27
33. Tien PW, Wei S, Darkwa J et al (2022) Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality - a review. *Energy AI* 10:100,198. <https://doi.org/10.1016/j.egyai.2022.100198>, <https://www.sciencedirect.com/science/article/pii/S2666546822000441>
34. Torres JF (2021) Deep learning for time series forecasting: A survey. *Big Data* 9:3–21. <https://doi.org/10.1089/big.2020.0159>
35. Wang Z, Srinivasan RS (2016) A review of artificial intelligence based building energy prediction with a focus on ensemble prediction models. pp 3438–3448. <https://doi.org/10.1109/WSC.2015.7408504>
36. Zeng A, Chen M, Zhang L et al (2023) Are transformers effective for time series forecasting?
37. Zhang L, Wen J, Li Y et al (2021) A review of machine learning in building load prediction. *Appl Energy*. <https://doi.org/10.1016/j.apenergy.2021.116452>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Dr. Roberto Morcillo-Jimenez

received a Degree in Computer Engineering and a Ph.D. in Computer Science from the University of Granada in 2014 and 2023, respectively. He has contributed to various national and international projects. His expertise includes computer engineering, time series analysis, architecture software, data science, and artificial intelligence (AI). His proficiency spans many domains, from designing efficient computer architectures to extracting insights from complex temporal data and creating immersive architectural designs using advanced software until the ability to seamlessly integrate theoretical knowledge with practical applications positions him as a dynamic force in shaping the future of technology.



M.Sc. Jesús Mesa received his degree in Computer Science from the University of Granada in 2018 and an MSc in Computer Engineering from the same university in 2020. He is currently a Senior Software Developer at NTT DATA. His interests include Data Science, time series analysis and blockchain technologies.



Dr. Juan Gómez-Romero has been a Full Professor at the Computer Science Department of Universidad de Granada (DECSAI-UGR) since 2023. Previously, Dr Gómez-Romero was an Associate Professor at UGR and a Lecturer at the Computer Science Department of Universidad Carlos III de Madrid (Spain). He received his PhD from UGR in 2008. His research interests include the application of Machine Learning to modelling, simulation and optimisation of energy systems.



Dr. M. Amparo Vila has been a Professor of Computer Science and Artificial Intelligence University of Granada since 1992. She was previously an Associate Professor and analyst programmer at the Computer Centre at the same university. She has developed her research and teaching primarily in databases and intelligent information systems. Her main lines of research are the treatment of imprecision in information systems using fuzzy logic, knowledge discovery in databases using

soft computing techniques, and ubiquitous computing and knowledge mobilization. She has been the advisor of 27 doctoral theses, has been responsible for more than 10 research projects and has published more than 100 papers in SCI journals. She has been responsible for the research group ICT-113 (Approximate Reasoning and Artificial Intelligence) from 1994 to 1997 and head of the group TIC174 (Databases and Intelligent Information Systems) since its inception in 2000 until today. At the University of Granada, she was Vice-Director of Organization at the School of Computer Science (1994-1997) and Director of the Department of Computer Science and Artificial Intelligence (1997-2004). She has also collaborated in quality assessment processes as an advisor for Information Technology and Communications in the Andalusian Research Plans, as an evaluator for the Spanish Evaluation Agency (ANECA) in the VERIFICA programme and as the president of various committees of research assessment in Andalusia.



Dr. Maria J. Martin-Bautista is a Full Professor at the Dept. of Computer Science and Artificial Intelligence at the University of Granada, Spain, since 2018. She is the head of the UGRITAI Lab (Information Technologies and Artificial Intelligence Laboratory at the University of Granada). Her current research interests include Federated Learning, Data Science, Big Data Analytics in Data, Text and Web Mining, Knowledge Representation and Uncertainty. She has supervised several Ph. D.

Thesis and published more than 150 papers in high impact international journals and conferences. She has supervised and participated in more than 40 R+D+i projects and international consortia, and research technology transfers with companies. She has been visiting researcher at the University of Houston (EEUU), and the University of Aarhus and University of Roskilde (Denmark).