IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 September 2024, accepted 10 October 2024, date of publication 17 October 2024, date of current version 29 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3482572

RESEARCH ARTICLE

Machine Learning Pipeline for Energy and Environmental Prediction in Cold Storage Facilities

NASSER ALKHULAIFI[1,2], ALEXANDER L. BOWLER[3], DIRENC PEKASLAN[2], (Member, IEEE), GULCAN SERDAROGLU[4], STEVE CLOSS[4], NICHOLAS J. WATSON[3], AND ISAAC TRIGUERO[5,6,1], (Senior Member, IEEE)

[1]Computational Optimisation and Learning (COL) Laboratory, School of Computer Science, University of Nottingham, NG8 1BB Nottingham, U.K.
[2]Laboratory for Uncertainty in Data and Decision Making (LUCID), School of Computer Science, University of Nottingham, NG8 1BB Nottingham, U.K.
[3]School of Food Science and Nutrition, Faculty of Environment, University of Leeds, LS2 9JT Leeds, U.K.
[4]Faculty of Engineering, University of Nottingham, NG7 2RD Nottingham, U.K.
[5]Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain
[6]DaSCI Andalusian Institute in Data Science and Computational Intelligence, 18071 Granada, Spain

Corresponding author: Isaac Triguero (triguero@decsai.ugr.es)

This work was supported in part by U.K. Engineering and Physical Sciences Research Council (EPSRC) for the University of Nottingham under Grant EP/S023305/1; in part by the EPSRC Centre for Doctoral Training in Horizon: Creating Our Lives in Data; in part by the Knowledge Generation Project funded by the Ministry of Science, Innovation and Universities of Spain, under Grant PID2023-149128NB-I00; in part by the Research, Development and Innovation [Investigación+Desarrollo + Innovación (I+D+i)] Project co-funded by ''Consejería de Universidad, Investigación e Innovación'' under Grant C-ING-250-UGR23; and in part by European Union Related to the European Regional Development Fund (FEDER) Andalucía Program (2021–2027). The work of Isaac Triguero was supported by the María Zambrano Fellowship funded by Spanish Ministry of Universities and Next Generation Funds from European Union.

ABSTRACT As energy demands and costs rise, enhancing energy efficiency in Food and Drink Cold Storage (FDCS) rooms is important for reducing expenses and achieving environmental sustainability ambitions. Forecasting electricity use in FDCSs can help optimise operations and minimise energy consumption by enabling door opening frequency, maintenance, and restocking to be better scheduled. Although Machine Learning (ML) has been applied to forecast energy use in various domains such as commercial and residential buildings, its use in addressing the specific challenges of FDCS, which require stringent temperature and humidity control for food safety and quality, has been less explored. This work addresses this gap by proposing a tailored ML pipeline for FDCS settings capable of predicting one-week into the future and is suitable for small dataset sizes. It provides comparative analysis by employing two distinct real-world FDCS datasets for training, validation, and testing of the developed models. Moreover, in contrast to existing studies predominantly concerned with energy consumption prediction, this study includes the forecasting of indoor temperature and humidity, given their essential role in preserving the quality and longevity of stored food items. Ensemble-based methods, particularly Random Forest, excelled and achieved the lowest electricity MAEs of 150.65 and 384.88 for each dataset, respectively.

INDEX TERMS Energy forecasting, feature engineering, food and drink cold storage rooms, machine learning, sustainability.

I. INTRODUCTION

A. BACKGROUND

With escalating energy demands and costs, the UK's food and beverage industry is increasingly motivated to enhance energy efficiency, driven by the need to reduce operational costs and meet consumer expectations for sustainability. Food and Drink Cold Storage (FDCS) rooms, as a component of food systems, play a crucial role in the food supply chain by preserving a wide array of perishable goods, including dairy, meat, and fresh produce, and ensuring their safety by maintaining acceptable temperature and humidity levels.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhengmao Li.

For example, wine maintenance in cellar rooms demands a consistent temperature and abrupt changes can either hasten its ageing process in warmer conditions or inhibit its development in cooler environments [1]. Lower humidity can lead to dehydration, particularly of natural elements such as cork, resulting in air infiltration in bottles, while conversely, excessively high humidity can encourage mould development and damage bottle labels [2]. Accurate forecasts of electricity consumption, indoor temperature and humidity (driven by cooling and humidity control systems) in FDCSs can enhance operations and scheduling (e.g. managing door opening frequency to minimise energy loss, aligning maintenance with lower demand periods, determining the optimal times for restocking) leading to reduced energy consumption and ensure food product perseveration and quality [1], [3]. The energy consumption in FDCS is influenced by weather; temporal; operator activity; and less measurable factors such as the type, size, quantity, and packaging of food [4], [5].

The recent surge in data availability fuelled by lower sensor costs and improved data processing capabilities has ushered in numerous data-driven methods for modelling and predicting dynamic behaviours such as energy consumption forecasting [6], [7]. These approaches are particularly useful when the underlying physics of the system is not well understood or is difficult to model [7], [8]. In a data-driven model, data gathered from regular operations or specific tests is analysed using algorithms such as statistical regression to understand the relationship between input and output variables [9], [10]. While statistical methods such as Autoregressive Moving Average [11], Autoregressive Integrated Moving Average [12], [13], [14], [15], and Autoregressive Integrated Moving Average with eXogenous variables [16], [17], [18] have been used for forecasting energy consumption in buildings, they depend on restrictive assumptions such as linearity and stationary input data (constant statistical properties over time). These limitations have led to the exploration of machine learning (ML) techniques, algorithms that learn from data to make predictions without explicit programming, thereby overcoming such constraints and demonstrating growing interest in this field [19], [20], [21], [22].

Current studies on energy consumption forecasting using ML methods have predominantly concentrated on various types of buildings, such as institutional and educational buildings [23], [24], [25], [26], commercial and residential buildings [27], [28], [29], [30], office and governmental buildings [31], [32], community buildings [33], [34], factory building [35], and industrial distribution complexes [36]. However, there remains a notable lack of research specifically targeting energy consumption prediction in FDCS. While there are parallels between predicting energy consumption in buildings and FDCS, FDCS presents unique challenges, particularly the stringent requirements for maintaining temperature and humidity levels to ensure food safety and quality. This work addresses this gap in the literature by trialling ML techniques, commonly used for building energy consumption prediction, to determine a tailored pipeline to the specific characteristics of FDCS. It focuses on the prediction of electricity usage, indoor humidity, and indoor temperature within FDCS environments, whereas most of previous studies on buildings have only focused on predicting energy consumption. The following related work section provides a comprehensive review of the input features, feature extraction, feature selection techniques, and prediction methods from building energy prediction studies, setting the foundation for their trialling in the FDCS context.

TABLE 1. Nomenclatures.

Nomenclature	Definition
CNN	Convolutional Neural Network
FDCS	Food and Drink Cold Storage
FFNN	Feed-Forward Neural Network
FS	Feature Selection
GRU	Gated Recurrent Units
KNN	K-Nearest Neighbours
LSTM	Long Short-Term Memory
ML	Machine Learning
MTL	Multi-task Learning
RF	Random Forest
RNN	Recurrent Neural Network
SVM	Support Vector Machine
SVR	Support Vector Regression
XGB	Extreme Gradient Boosting

B. RELATED WORK

Energy consumption prediction can be formulated as a supervised ML regression problem where the model learns the relationship between the input (dependent variable) and the output (independent variable) [37], [38]. Appropriate model inputs are essential for developing accurate ML energy usage prediction models [31]. As weather heavily influences energy use in buildings through heating needs in cold climates and cooling needs in warm ones, numerous studies have utilised weather data such as temperature, dew point, humidity, precipitation, wind speed, air pressure, and solar radiation as input features for energy prediction models [25], [26], [27], [30], [31], [32], [33], [39], [40]. Historical energy data, capturing the complexities of actual consumption patterns influenced by various factors such as abnormal events and human activities, is also commonly used as inputs because it numerically indicates both the pattern and trend of the load profile [23], [24], [29], [41]. In addition to weather and historical data, very few studies have incorporated indoor features such as the number of occupants, zone air temperature, zone relative humidity [42], indoor humidity, indoor temperature, and indoor carbon dioxide levels [39]. However, obtaining such data often requires specialised sensors that may be unavailable due to privacy concerns, logistics, and cost [41]. Table 2 and Figure 1 summarise the inputs, outputs, and prediction methods used in related work for energy consumption prediction in buildings. In this study, weather data and the working hours of two FDCSs are used as inputs, owing to their significant relevance and potential impact on energy consumption and indoor conditions.

TABLE 2. Features, building type, and methods used in related work to predict energy consumption.

Raw Data	Engineered Features	Building Type	Algorithm	Ref.
	Summarised energy classes feature	Institutional	ANN	[23]
	Number of occupants	Office	Ensemble Bagging Trees	[58]
	Internal environmental condition	MLP & SVR	MLP & SVR	[42]
	Internal environmental condition Number of occupants	Office	FFNN, Extreme Learning Machine, and ensemble models	[73]
		Office	SVR	[42]
Weather Data		Educational	Multiple Linear Regression, Elastic net, Gradient Boosting Tree, RF, SVR, EGB, Deep Neural Network	[24]
		Commercial and Residential	RNN	[27]
		Educational	ANN	[25]
		Residential	CNN-LSTM	[29]
		Industrial	Recurrent Inception Convolutional Neural Network	[36]
		Office, Factory & Educational	Stacking ensemble approach	[40]
Historical Data	Temporal Feature	Educational	KNN, RF, XGB, Gradient Boosted Decision Trees, SVM, Stacking	[41]
		Industrial Park	MTL-SVM	[67]
		Office	LSTM with an attention mechanism	[32]
		Offices, Gymnasium, Exhibition Hall	KNN	[33]
		36 different types (office, education, library, dormitory, and hospital)	Seq2seq with an attention mechanism	[34]
		Educational	CNN-GRU	[26]
		Educational	MTL	[60]
		Commercial	MTL	[30]

Having representative features that establish the relationship between inputs and outputs is key in ML. Feature engineering plays a crucial role in developing these representative features by creating informative attributes from raw data, more effectively capturing the underlying patterns and relationships, and thereby developing more robust models [37], [43]. These attributes often encompass time-based indicators, such as the hour of the day, whether it is a weekday or weekend, and the month of the year [24], [25], [29], [44]. Applying sine and cosine transformations to the cyclic features can create a more nuanced representation of temporal patterns, hence enhancing the model's performance [35]. Additionally, some studies included special occasions such as holidays, and statistical attributes such as mean, minimum, and maximum values for temperature and humidity [32], factored in seasons [30], and performance indicators such as on-peak, off-peak, and mid-peak statuses, which were represented as binary variables [36]. Time-lag features are also used as inputs with varying window sizes for past data, including the most recent five hours to capitalise on short-term trends [23] or up to the past 24 hours to integrate more comprehensive historical patterns [24]. Although some automated feature-extracting techniques such as Principal Component Analysis (PCA) [45], [46], Wavelet Decomposition [47], and Autoencoders [48] have been used in a few studies, they possess certain limitations. For example, while PCA is powerful for dimensionality reduction, it can obscure the interpretability of variables [37], The same applies to wavelet decomposition, as the wavelet coefficients are less intuitive than raw data and often require specialised knowledge for interpretation. With autoencoders, the encoded features are abstract and lack a clear intuitive relationship with the original data. This is an important consideration in this research context because the interpretability of features can help in understanding the factors that drive energy use and indoor conditions, which can inform policies and solutions. Despite these efforts, a comprehensive approach for feature engineering in FDCS is not yet established, indicating a gap in current research. Addressing this, the methodology section delves into this process within the FDCS context, concentrating on the relevance of features and their potential impact on model outcomes.

After defining and extracting features, additional preprocessing techniques such as feature selection (FS) may applied where appropriate, to identify the most relevant variable inputs for developing an ML model [49]. While this is not compulsory, this preprocessing technique can be useful for building energy consumption prediction [50]. Neglecting to filter inputs can result in larger datasets and slower training speeds and may either adversely or positively affect the model's performance and accuracy [37], [51]. For energy consumption predictions in buildings, FS is predominantly conducted manually, guided by domain knowledge as demonstrated in prior research [52], [53], [54]. Nevertheless, certain studies have incorporated the FS methods [55], [56], [57], [58], [59] as explained in Table 3. Recently, various FS methods for buildings' energy prediction, encompassing filters, wrappers, and embedded

TABLE 3. Feature selection methods.

Feature selection method	Description	Ref.
Filter-driven with SVR kernels	Best features are chosen based on acquisition feasibility and performance scores using filter methods. These features are tested on datasets using support vector regression with radial basis and polynomial kernels.	[55]
Embedded recursive feature elimination	This backward selection technique involves training a model with all variables, ranking them for importance, and iteratively removing the least important ones until no further reduction is possible.	[56]
Wrapper Genetic Search	This method explores feature subsets using a genetic search strategy, focusing on predictive ability, and minimising redundancy among features.	[57]
Data Permutation-Based	Optimal features are identified by assessing their role in enhancing predictive performance, with a focus on the impact of introducing irrelevant or noisy information.	[58]
Hybrid Filter-Wrapper: Two-Stage Approach	Starts with a filter method to remove irrelevant features, reducing dimensionality without compromising accuracy. Then, a wrapper method conducts an exhaustive search for the most effective feature set, balancing accuracy, and simplicity.	[59]

techniques, were reviewed [60]. In brief, the filter technique, rooted in statistical procedures, assigns a value to each feature, and ranks them, determining whether they should be retained or discarded; the wrap-per approach evaluates the predictive capabilities of models by assessing various subsets of potential features; and the embedded technique integrates feature selection directly into ML algorithms such as Random Forest (RF). Such studies indicate that the effectiveness of FS is somewhat context-dependent (e.g., data characteristics and model type), emphasising that no single method is universally superior. Recognising both the variability and the lack of investigation into FS within the FDCS context in the existing literature, this work seeks to bridge this knowledge gap. However, it is important to clarify that this work does not attempt to universally address the gap in FS approaches for energy consumption prediction across all settings. Instead, it focuses on examining and evaluating their impact specifically within the FDCS environments, all of which are explained in the methodology section.

Different ML algorithms have been used in predicting building energy use. Traditional ML algorithms such as K-Nearest Neighbours (KNN) [33], [41] have been used for buildings' energy consumption prediction. Given their capabilities in addressing complex forecasting problems, there is a notable increasing trend in using neural network-based algorithms for building energy use prediction, such as using Feed-forward neural networks (FFNN) for daily energy consumption forecasting of institutional buildings [23], Recurrent Neural Networks (RNN) for predicting 24-hour sequences of electric load [27], a Multilayer Perceptron (MLP) for 1-hour ahead prediction of office building [28], Long Short-Term Memory (LSTM) for office building energy consumption prediction [32]. Hybrid methods have also been studied, where models combine the feature extraction capabilities of Convolutional Neural Networks (CNNs) with the sequence modelling capabilities of LSTMs for predicting building energy consumption [29], CNN-Gated Recurrent Units (CNN-GRU) for predicting hourly energy usage in educational buildings [26], and Recurrent Inception Convolutional Neural Networks for predicting power consumption in large distribution complexes [29]. Additionally, ensemble-based methods such as Extreme Gradient Boosting (XGB) and RF have shown promising results when utilised for building energy consumption prediction [41], [44], [61], [62], [63]. In contrast to these single-task methods, Multi-task Learning (MTL) [64] leverages existing ML algorithms to simultaneously model multiple input-output relationships and task interdependencies, capitalising on these to improve prediction accuracy. For example, MTL was used to predict a building's electrical load and outdoor temperature simultaneously leveraging outdoor temperature forecasting as a secondary task and employing a hyperparameter c, to balance the auxiliary task's weight [30]. Utilising MTL in combination with a Temporal Convolutional Network (TCN) for short-term multi-energy load predictions has shown promising results [65], and reduced training times were observed when MTL was combined with a Support Vector Machine (SVM) for similar predictions [66]. The summary of previous studies, detailed in Table 2, shows algorithm selection is context-dependent, influenced by dataset challenges, computational efficiency, model interpretability and potentially by the researchers' preference for methods with which they are most acquainted. To address this challenge, the methodology section details the selection and justification of the prediction methods applied in this work for the FDCS context, encompassing traditional machine learning, ensemble learning, and deep learning algorithms.

ML model performance is heavily influenced by the training dataset size [67], [68]. The size of training datasets for energy consumption prediction of buildings varies widely, ranging from 5.5 months in some studies [69] to 6-12 months in others [24], [27], and even extending beyond two years in other works [29], [70]. A comprehensive study conducted by [71] indicates that out of 83 ML studies focused on predicting buildings' energy consumption, 43 used historical data ranging from one to two years, and 24 studies extended beyond two years. In contrast, the current study utilises comparatively smaller datasets, with the first dataset for training (FDCS 1) comprising only 53 days of data and the second (FDCS 2) containing 75 days. Given the very small size of the datasets used in this study compared to previous research, the proposed pipeline that is also suitable for other small dataset sizes, especially where acquiring extended

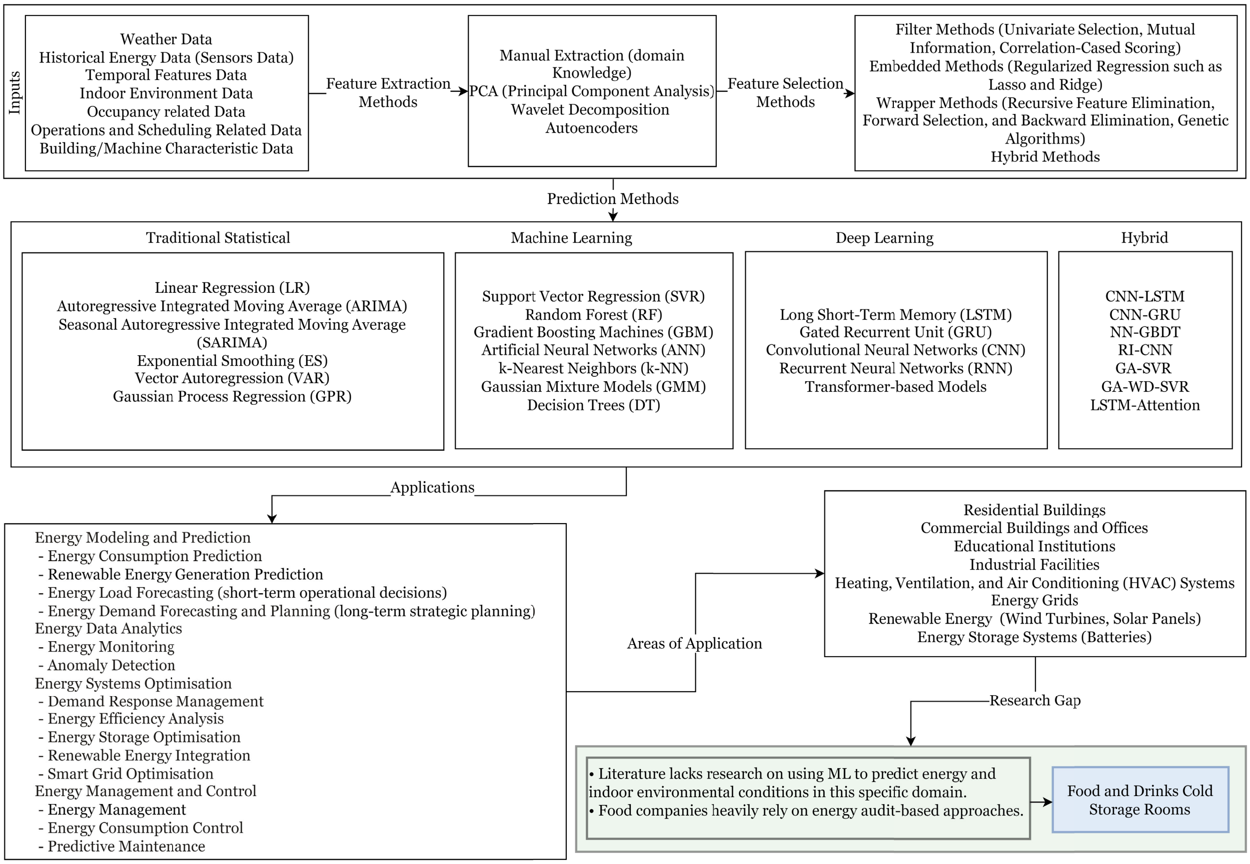


FIGURE 1. Overview of literature on predicting energy consumption, including input features, feature engineering, prediction methods, applications, areas of application, and identified research gaps for this study.

historical data may be infeasible due to time, cost, or data availability constraints.

C. RESEARCH GAPS AND CONTRIBUTIONS

The novelty and contribution of this study lies in:

- Using ML methods for forecasting essential parameters (electricity consumption, indoor temperature, and humidity) in the unique context of FDCS. This specificity is critical because, although previous studies have explored forecasting energy using ML methods in various domains such as commercial and residential buildings, Heating, Ventilation, and Air Conditioning (HVAC) systems, smart grids, wind turbines, and solar panels, FDCS presents its own set of unique challenges, particularly their stringent requirements for maintaining temperature and humidity levels to ensure food safety and quality.
- Propose a detailed pipeline for ML techniques in forecasting one week (hourly) into the future of electricity consumption, temperature, and humidity specific to FDCS environments and suitable for small dataset sizes. The proposed pipeline was validated using two newly collected datasets from different FDCS rooms located in Nottingham, UK. The inclusion of these datasets enables a comparative analysis, facilitating a more robust evaluation of ML methods in the specific context of FDCS.
- Emphasising multi-variable forecasting, in contrast to existing studies that often focus solely on energy consumption. This work underlines the importance of also forecasting indoor temperature and humidity, factors crucial for maintaining the quality and lifespan of stored items in FDCS.
- Investigating the often-overlooked aspect of feature selection methods. This involves examining the impact of eight different methods including filter-based, embedded, wrapper-based, and hybrid methods on different ML regression algorithms in FDCS settings, providing insights for future research in a similar context even though such investigation can be considered data-dependent.
- Trialling of different dataset sizes to demonstrate the impact of dataset volume on model accuracy in FDCS environments. Such environments often face limitations in data collection due to time, cost, or operational challenges. By evaluating model performance across varying dataset volumes, the research not only highlights the influence of dataset size on accuracy but also offers

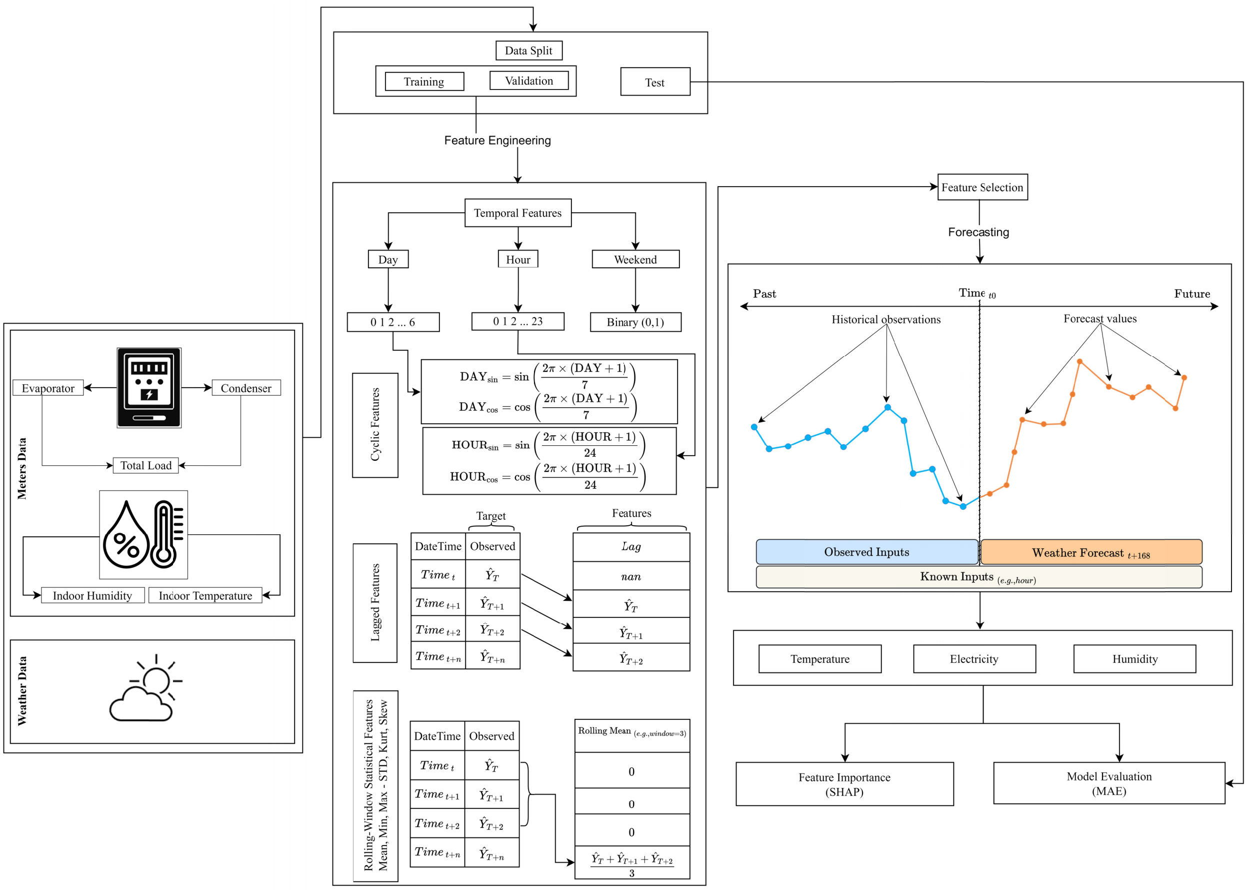


FIGURE 2. **Proposed ML pipeline for predicting electricity consumption, indoor temperature, and indoor humidity one week into the future in FDCS settings using weather forecast.**

an estimation of the minimum dataset size needed for forecasting in FDCSs.

II. METHODOLOGY

This section explains the methodology used in this work and the proposed pipeline to predict electricity consumption, temperature, and humidity one week (hourly) into the future, as shown in Figure 2. The methodology includes collecting electricity consumption data and indoor environment conditions through metering systems and weather. Different feature engineering methods were examined, owing to their effectiveness in previous related works and to assess their influence within the specific context of FDCS. Various ML methods were explored, focusing on those that have shown superior performance in comparable settings as explained in the preceding related work section.

A. DATA COLLECTION

In this study, electricity consumption data for the condenser and evaporator units, as well as internal ambient conditions (temperature and humidity) were collected from two FDCSs based in Nottingham, United Kingdom, as shown in Figure 3 and Table 4. Electricity consumption data were collected at 4-second intervals for FDCS 1 from 12 November 2021 up to 31 January 2022, and for FDCS 2 from 21 October 2021 up to 31 January 2022. The internal temperature and internal humidity were recorded at 10-minute intervals within the storage room and close to the evaporator unit. Figure 4 and Figure 5 present the collected datasets of the three variables from the FDCS datasets over time, resampled to a 1-hour resolution. The hourly weather data utilised in this study was sourced from the NASA Langley Research Centre's POWER Project, a repository of solar and weather data sets produced by NASA to support renewable energy and building energy efficiency research [73]. In this study, the exact locations of each FDCS location were identified for retrieving weather data using geographical coordinates (latitude and longitude). These observations are historical instead of forecasts to reduce uncertainty from forecast errors and better understand the impact of various input features. The data were partitioned into three subsets: 70% for training the ML models, 15% for validation (to fine-tune hyperparameters and monitor performance), and 15% for testing to evaluate the models' performance on previously unseen data. A 5-fold cross-validation approach, based on trial and error, was also employed, and for the final model training, the training and validation sets were combined to maximise the data utilised. This approach supports the development of a robust ML model and aligns with the methodologies adopted in related studies [30], [65], [74], [75].

B. FEATURE EXTRACTION

Given that no previous work has identified the optimal features to extract in the context of FDCSs, and considering their relevance and potential impact, this work uses a comprehensive set of extracted features. These include indicator variables accounting for categorical events such as the hour of the day and weekends, cyclic features that capture temporal patterns, time-lag features, and rolling-window statistical features. The following subsections describe the feature extraction process used in this work, and Table 5 presents a summary of all the input variables employed.

1) INDICATOR VARIABLES AND OPERATION HOURS

Indicator variables and temporal features are crucial in time series analysis to account for the impact of categorical events, such as weekdays and weekends [51], [71]. For example, in predicting the electricity consumption of FDCS, these encoded indicator variables enhance the model's ability to capture the effects of these events on energy use. A 'weekend' variable, for instance, can indicate a potential increase or decrease in demand for FDCS systems on weekends, which may lead to lower or raised electricity consumption. Additionally, variables that represent the daily operation hours (working hours) of the two FDCSs have been employed. Such integration could be important as different operation hours can significantly influence the patterns of electricity consumption and other metrics being predicted.

2) CYCLIC FEATURES

Although time-based features such as the hour of day and day of the week provide temporal information, these features may not always be effective in representing time-based patterns. To overcome this, sine and cosine transformations can be applied to these features to capture the temporal patterns in the data [35]. This enhances the model's ability to capture the cyclical and periodic patterns in the data, leading to improved prediction accuracy. Equations (1) and (2) calculate the sinusoidal and cosinusoidal transformation of the day of the week respectively, allowing for cyclical pattern representation. Similarly, Equations (3) and (4) perform the sinusoidal and cosinusoidal transformation of the hour of the day respectively. By adjusting the input by $+1$ and normalising by the period (7 for days and 24 for hours), these transformations capture cyclical temporal patterns in data.

$$DAYsin = \sin\left(\frac{2\pi \times (DAY + 1)}{7}\right) \tag{1}$$

$$DAYcos = \cos\left(\frac{2\pi \times (DAY + 1)}{7}\right) \tag{2}$$

$$HOURsin = \sin\left(\frac{2\pi \times (HOUR + 1)}{24}\right) \tag{3}$$

$$HOURcos = \cos\left(\frac{2\pi \times (HOUR + 1)}{24}\right) \tag{4}$$

3) TIME-LAG FEATURES

Time-lag features in time series data are values from previous time points. They are created by shifting the target variable back in time t by a certain number of steps k. The primary purpose of these features is to capture the relationship between the current value of the target variable and its past values. For example, in an FDCS system, the energy consumption at time t might be influenced by consumption levels at $t - k$ reflecting the inertia of cooling systems. Similarly, the indoor temperature or humidity at a given moment could be a result of conditions from previous hours. By incorporating such features, models can better account for historical influences, capture temporal dependencies and trends, and ultimately improve their predictive accuracy. In this study, as the forecasting horizon is 168 hours into the future (one week), the top 10 lags showing the highest correlation with the respective target variable were employed, selected from the last 168-336 lags, Figure 6, to ensure the use of only available data at the time of forecasting.

4) ROLLING-WINDOW STATISTICAL FEATURES

The extraction of statistical features from the target variables in this study employed a rolling-window technique applied to historical data. This approach involved segmenting the time series data into smaller windows, allowing for the capture of evolving data patterns and trends over time. To ensure that only past information was used for feature computation, and to avoid look-ahead bias, rolling statistics were assigned to a timestamp after the window and from the last 168-336 hours as the forecast is 168 hours into the future. These features included the rolling window of the mean, variance, skewness (a measure of asymmetry), and kurtosis (a measure of the distribution's tail heaviness). The mean can be computed using equation (5), while equation (6) can be utilised to calculate the variance, equation (7) enables the computation of the skewness, and equation (8) is employed to derive the kurtosis, where X_i represents the target variable (e.g., electricity consumption) during the i^{th} hour of the day, with i ranging from 0 to 23. The total number of hours is denoted by N. The symbols $M(\mu)$, V, S, and K represent the mean, variance, skewness, and kurtosis, correspondingly.

$$M = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{5}$$

$$V = \frac{1}{N}\sum_{i=1}^{N} (X_i - \mu)^2 \tag{6}$$

$$S = \frac{1}{N}\sum_{i=1}^{N} (X_i - \mu)^3 \tag{7}$$

$$K = \frac{1}{N}\sum_{i=1}^{N} (X_i - \mu)^4 \tag{8}$$

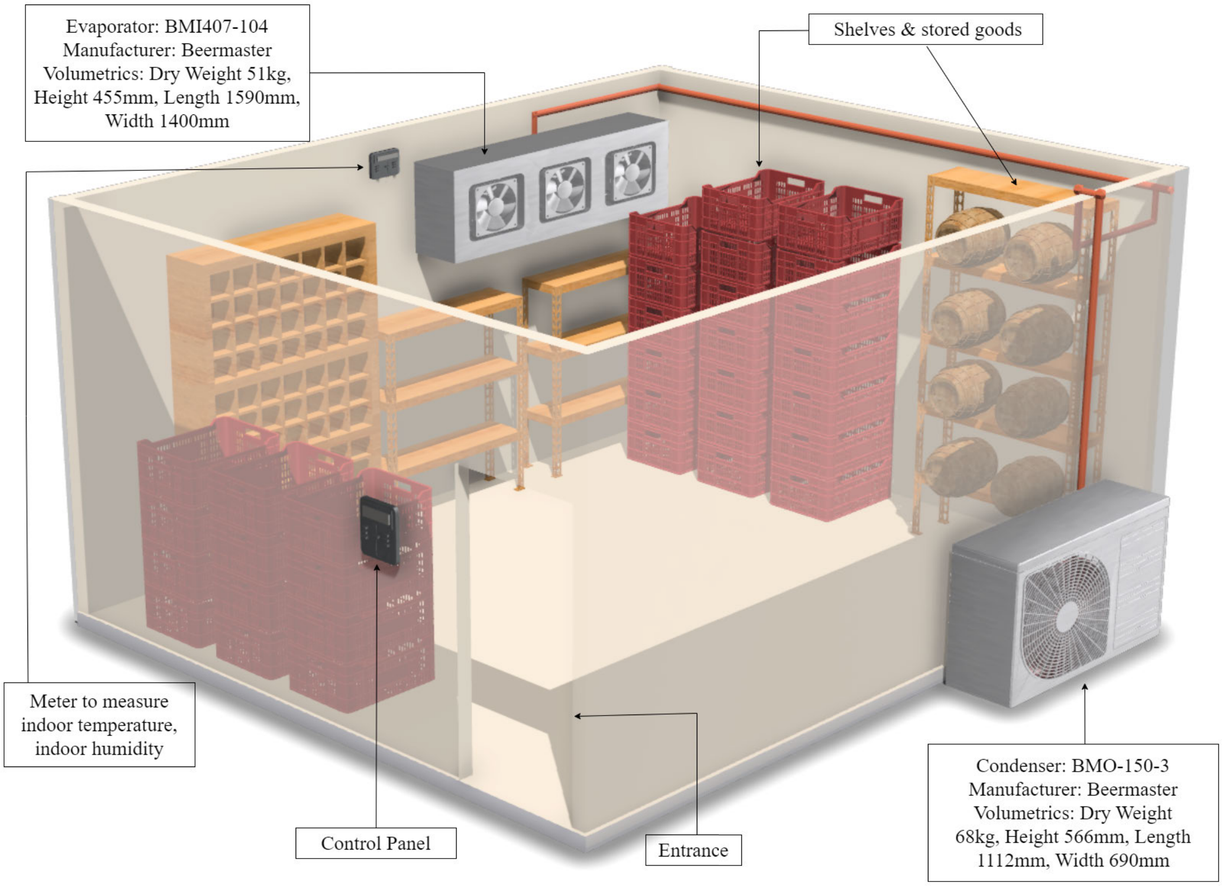


FIGURE 3. **3D schematic representation of the investigated FDCSs. Key components indicated include (i) evaporator unit, responsible for absorbing heat from the storage space and maintaining low temperatures; (ii) condenser unit, essential for releasing the absorbed heat outside the storage room and condensing the refrigerant back into a liquid; (iii) an integrated meter collecting indoor temperature and humidity values; and (iv) a control panel that facilitates monitoring and adjusting the storage condition.**

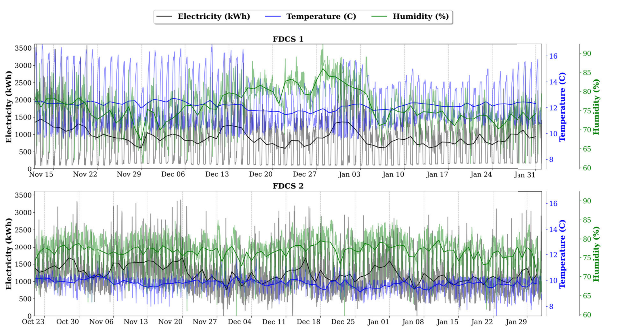


FIGURE 4. **Collected data from the two FDCS rooms plotted as a time series. Raw data points are represented by faint lines, while mean values are shown as solid lines. The mean is calculated by resampling the data into 24-hour blocks and taking the average value within each block, which smooths out short-term fluctuations and helps identify trends and anomalies.**

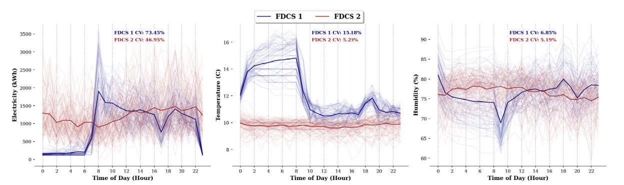


FIGURE 5. **Hourly electricity, temperature, and humidity patterns for the two FDCSs. Each hourly reading is shown with reduced opacity, while mean hourly values are depicted in darker lines for clarity. The Coefficient of Variation (CV) quantifies relative variability. FDCS 1 shows higher electricity, temperature, and humidity variability (higher CV) than FDCS 2, indicating greater sensitivity to internal and/or external factors in FDCS 1 versus a more regulated environment in FDCS 2.**

C. FEATURE SELECTION METHODS.

To examine the impact of FS methods on predicting electricity use, temperature, and humidity in FDCS environments, this study employed eight distinct FS methods, as shown in Table 6, including filter-based, embedded, wrapper-based and hybrid. These meth-ods were selected for their effectiveness in related work and potential suitability for FDCS challenges. Utilising diverse FS techniques allowed for a thorough examination, leveraging each method's strengths to comprehensively assess feature relevance. An analysis and comparison of these methods are presented in the Results and Discussion Section.

D. MACHINE LEARNING ALGORITHMS

In this work, electricity consumption, temperature, and humidity within two FDCSs were predicted using weather

TABLE 4. Summary statistics of the collected data for electricity consumption, indoor humidity, temperature and weather in both FDCSs.

Variable	Statistics						
	Mean	Std Dev	Min	25th Perc	50th Perc	75th Perc	Max
FDCS 1							
Electricity Consumption (kWh)	922.22	677.51	106.01	178.35	958.15	1385.56	3613.5
Humidity Closer to Evaporator (%)	76.21	5.22	61.17	72.67	76.25	79.75	92.42
Temperature Closer to Evaporator (°C)	12.16	1.84	9.5	10.67	11.42	13.5	16.92
Outdoor Temperature (°C)	4.62	3.44	−3.30	2.04	4.44	6.80	13.76
Outdoor Dew Point (°C)	3.96	3.34	−4.15	1.55	3.48	6.39	12.33
Outdoor Wet Bulb Temperature (°C)	4.29	3.35	−3.69	1.87	3.94	6.48	12.87
Specific Humidity (g/kg)	5.17	1.25	2.81	4.21	4.88	5.98	8.97
Outdoor Relative Humidity (%)	95.03	5.69	60.12	93.38	96.75	98.69	100.0
Precipitation (mm/hour)	0.06	0.18	0.0	0.0	0.01	0.04	2.09
Surface Pressure (kPa)	100.32	1.50	96.13	99.14	100.44	101.58	102.81
Wind Speed (m/s)	5.14	2.61	0.28	3.13	4.76	6.70	14.88
Wind Direction (Degrees)	238.70	74.11	0.25	212.86	249.14	292.39	359.66
FDCS 2							
Electricity Consumption (kWh)	1218.32	572.11	0	781.78	1134.07	1537.62	3361.45
Humidity Closer to Evaporator (%)	76.71	3.98	59.67	74.65	77.33	79.5	85.75
Temperature Closer to Evaporator (°C)	9.77	0.52	7.25	9.5	9.83	10.08	11.67
Outdoor Temperature (°C)	5.53	3.78	−3.23	2.56	5.44	8.26	16.47
Outdoor Dew Point (°C)	4.71	4.71	−4.15	2.02	4.81	7.28	12.99
Outdoor Wet Bulb Temperature (°C)	5.12	3.60	−3.61	2.28	5.12	7.83	14.68
Specific Humidity (g/kg)	5.47	1.37	2.81	4.39	5.34	6.39	9.36
Outdoor Relative Humidity (%)	94.16	6.68	47.07	92.08	96.36	98.48	100.0
Precipitation (mm/hour)	0.07	0.22	0.0	0.0	0.01	0.04	4.30
Surface Pressure (kPa)	100.19	1.45	96.14	99.11	100.30	101.25	102.81
Wind Speed (m/s)	5.35	2.64	0.38	3.29	4.99	7.02	14.55
Wind Direction (Degrees)	238.78	68.62	0.32	209.72	245.02	286.44	357.56

TABLE 5. Summary of input features used in this study.

Features	Type	Description
Weather data	Continuous	Outdoor Temperature (°C), Outdoor Dew Point (°C), Outdoor Wet Bulb Temperature (°C), Specific Humidity (g/kg), Outdoor Relative Humidity (%), Precipitation (mm/hour), Surface Pressure (kPa), Wind Speed (m/s), Wind Direction (Degrees).
Temporal features	Integer Value	Hour of the day (0-23), day of the week (0-6), day of the month (1-31), the month of the year (1-12), weekday vs weekend (0-1)
Operation hours	Binary	The 'Is_Open' feature serves as a binary indicator: 1 indicates that the storage is operating within working hours, while 0 indicates otherwise.
Cyclic features	Continuous	Sine and cosine transformation of temporal features (hour of day and day of week)
Time-lag features	Continuous	Previous values of the given target (e.g., electricity consumption from preceding hours/days)
Rolling-window statistical features	Continuous	Summary statistics (maximum, minimum, mean, kurtosis, skewness, and standard deviation) of a given target computed over a fixed-size window

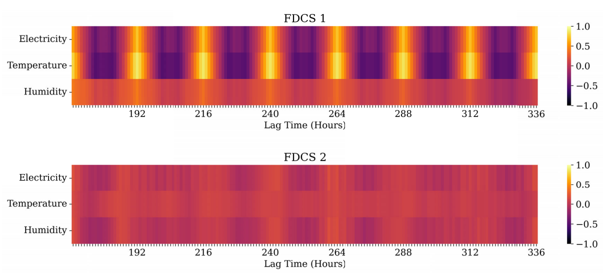


FIGURE 6. Correlation heatmaps of target variables against their time lags over one week for FDCS 1 and 2. The colour scale represents the Pearson correlation coefficients.

data and extracted features as input variables. The ML algorithms used in this work, as shown in Table 7, were chosen not only based on their established efficacy in predicting the energy consumption of buildings, as supported by the existing literature but also to maximise the strengths of each method, mitigate method-specific limitations, and enhance prediction accuracy and reliability. More precisely; KNN was employed for its simplicity and fast training speed [33], [41]; RF was chosen due to its robustness against overfitting, along with its embedded ability to provide insights into feature importance [58]; XGB was included for its rapid performance and high efficiency [62], [63]; MLP was selected for its high ability to capture complex non-linear relationships, necessary for modelling interactions within FDCS environments [42]; LSTM was incorporated due to its proficiency in handling sequential data and capturing time dependencies [32], [70], [75], [76]; and MTL was utilised for its potential to learn multiple related tasks simultaneously, aiming to enhance generalisation despite its inherent

TABLE 6. Overview of feature selection methods used in the study (k = number of features to select).

Method	Type	Description	Hyperparameters
Correlation (F-Test)	Filter	Select top-k features using univariate linear regression tests based on F-statistics.	k=20
Mutual Information	Filter	Chooses top-k features based on mutual information with the target, capturing non-linear relationships.	k=20
Lasso Regularisation	Embedded	Uses L1 regularisation in Lasso regression to eliminate less important features by driving their coefficients to zero.	alpha=0.01
Tree Importance (Extra Trees)	Embedded	Employs an ensemble of decision trees to rank features by importance, selecting those with higher importance.	n_estimators=300
ElasticNet Regularisation	Embedded	Combines L1 and L2 regularisation, eliminating features with coefficients that shrink to zero.	alpha=0.01,
Recursive Feature Elimination (RFE)	Wrapper	Uses RFE with Linear Regression to recursively eliminate the least important features.	k=20
Sequential Forward Selection	Wrapper	Starts with no features, and iteratively adds important features based on negative mean squared error evaluated by Linear Regression.	k=20
Filter + Embedded	Hybrid	Uses Correlation F-Test as a filter to select top-k features, followed by an embedded method with a Random Forest Regressor for final selection.	Filter: k=20, n_estimators=300

complexity and the potential for task interference [30], [65], [66]. Hyperparameter tuning, which involves setting model configurations before training, is crucial for optimising model performance. To fine-tune each model, a grid search of hyperparameter combinations was conducted using training and validation sets to identify the optimal values. The tuned models were then evaluated on the test set. Table 8 shows the results of the grid search for hyperparameter tuning across all models.

E. MODEL PERFORMANCE ASSESSMENT METRICS

The performance of the ML regression models was evaluated using the Mean Absolute Error (MAE) [82]. MAE, which measures average absolute differences between predicted and actual values in the original units, was selected for its interpretability. The equation of MAE (9) is shown below where M is the number of samples in the studied dataset, $predicted_i$ is the predicted i^{th} value, $observed_i$ is the true i^{th} value, and $m\_observed$ is the mean of the true values.

$$MAE = \frac{1}{M}\sum_{i=1}^{m}|predicted_i - observed_i| \qquad (9)$$

III. RESULTS AND DISCUSSION

This section presents an analysis and discussion of the experimental results, including the performance of ML models in prediction tasks within the FDCS settings on the test set, the influence of eight different FS methods, feature importance, and finally, the implications of dataset size.

A. PERFORMANCE EVALUATION

The results of the experimental evaluation, as summarised in Figure 7, Figure 8, and Figure 9, demonstrate that ensemble methods (XGBR and RFR), can achieve more accurate predictions for forecasting electricity consumption and internal environmental conditions one week into the future in FDCS settings when compared to other algorithms. For electricity predictions of FDCS 1, these two methods outperformed others, even without applying FS methods (i.e., using all features), achieving the lowest errors in the test set with MAEs of 150.75 and 157.79, respectively. The promising performance of the XGBR algorithm observed in this study aligns with findings from other studies, such as [62], [63]. However, while comparing these findings with existing literature is important, such a comparison may not be entirely appropriate due to the unique context of FDCSs, which differs significantly from other domains. Notably, the hybrid (filter + embedded) was the FS method that most improved the performance of LSTM, MLP, KNNR, and MTL. Similar patterns were observed as the ensemble-based methods produced the lowest errors in predicting both indoor temperature and humidity in FDCS 1.

In FDCS 2, the MLP model, combined with the Hybrid (filter + embedded) FS method, produced prediction errors almost matching those of ensemble-based methods for predicting electricity consumption, yet the prediction errors were noticeably high. For predicting indoor temperature in this storage, LSTM and MTL, alongside RFR and XGBR, produced the lowest prediction errors when the embedded-lasso FS method was applied. In understanding the most significant findings across the two different FDCSs, it's important to recognise that these storage systems differ in layout, size, and operations. Consequently, a direct comparison may not be entirely appropriate. Nevertheless, some significant patterns emerge. First, the ensemble methods XGBR and RFR consistently achieve top-tier performance with the lowest errors compared to other models in both environments. The success of these two algorithms likely stems from their advanced

TABLE 7. ML algorithms used in this work.

Model	Description	Ref.
K-Nearest Neighbours Regression (KNR)	A non-parametric algorithm that predicts the target based on the average of the K closest training examples in the input feature space.	[78]
Random Forest Regression (RFR)	An ensemble method that fits multiple decision trees on randomly sampled subsets of the data and combines their predictions.	[79]
Extreme Gradient Boosting Regression (XGBR)	An ensemble method that trains decision trees sequentially, each time fitting the residual errors of the previous tree.	[80]
Multi-Layer Perceptron (MLP) Regressor	A feedforward artificial neural network with multiple layers of nodes between input and output. Uses backpropagation to train the network weights and biases.	[81]
Long Short-Term Memory (LSTM)	A class of recurrent neural networks that can learn long-term dependencies by using a memory cell and three gating mechanisms.	[82]
Neural Network Multi-Task Learning (MTL)	A learning method where multiple tasks are handled concurrently, using a shared representation. By leveraging the similarities and variations across tasks, what is learnt for one task can aid in the learning of other tasks.	[64]

TABLE 8. Grid search results for hyperparameter tuning of models predicting electricity consumption. Training used a 4-core Intel CPU with 32GB RAM, Python 3.11, Scikit-learn, and TensorFlow frameworks.

	Model	Best hyperparameters
FDCS 1	KNNR	'leaf_size': 10, 'n_neighbours': 10, 'weights': 'distance'
	RFR	'max_depth': 10, 'min_samples_leaf': 2, 'min_samples_split': 10, 'n_estimators': 500
	XGBR	'learning_rate': 0.1, 'max_depth': 10, 'min_child_weight': 2, 'n_estimators': 500, 'subsample': 0.8
	MLP	'activation': 'relu', 'alpha': 0.0001, 'batch_size': 16, 'hidden_layer_sizes': (50, 100), 'learning_rate_init': 0.01, 'max_iter': 500, 'solver': 'adam'
	LSTM	LSTM_Layers: [20,10], Optimizer: Adam, Learning_Rate: 0.01, Epochs: 100, batch_size: 32
	MTL	Base_Layers: 10, 40; Activation: ReLU; Optimizer: Adam; Learning_Rate: 0.005; Loss_Function: MSE; Epochs: 100; Batch_Size: 32; Loss_Weights: 1
FDCS 2	KNNR	'leaf_size': 10, 'n_neighbours': 20, 'weights': 'distance'
	RFR	'max_depth': 10, 'min_samples_leaf': 10, 'min_samples_split': 2, 'n_estimators': 500
	XGBR	'learning_rate': 0.01, 'max_depth': 20, 'min_child_weight': 10, 'n_estimators': 500, 'subsample': 0.8
	MLP	'activation': 'relu', 'alpha': 0.001, 'batch_size': 32, 'hidden_layer_sizes': (50, 100), 'learning_rate_init': 0.001, 'max_iter': 500, 'solver': 'sgd'
	LSTM	LSTM_Layers: [20,10], Optimizer: Adam, Learning_Rate: 0.01, Epochs: 100, batch_size: 32
	MTL	Base_Layers: 30, 30; Activation: ReLU; Optimizer: Adam; Learning_Rate: 0.01; Loss_Function: MSE; Epochs: 100; Batch_Size: 32; Loss_Weights: 1

feature selection and ensemble techniques - bagging for RFR and boosting for XGBR. Secondly, the neural network-based models demonstrate considerable variability in their performance metrics across different FS methods. This fluctuation could be attributed to the models' sensitivity to specific features and/or the relatively small dataset sizes used for training.

Additionally, these results shed light on the predictability of energy consumption and indoor variables in FDCSs. The models consistently demonstrate the lowest errors in FDCS 1, suggesting that its energy consumption and indoor variables are more predictable than those in FDCS 2. This observation aligns with the consistent daily trends observed in electricity consumption, indoor temperature, and humidity depicted in Figure 5 for FDCS 1 compared to FDCS 2. While high predictability can facilitate planning and management, thereby boosting operational efficiency and cost savings, it should not be equated with efficiency. For example, an FDCS with high but predictable energy consumption may not be as efficient as one with less predictable but lower energy consumption. Therefore, these findings should be integrated into a broader strategy for energy efficiency understanding and improvement in FDCS.

B. EVALUATING FEATURE IMPORTANCE USING SHAP

Building on the insights gained from the FS analysis in subsection III-A, this section aims to delve deeper into understanding feature importance. This study employs Shapley Additive Explanations (SHAP) [83] for feature importance evaluation, as illustrated in Figure 10 and Figure 11. SHAP was selected due to its model-agnostic properties, local accuracy, and game-theoretic foundation, which ensures a fair and consistent distribution of predictive power across features. This method ranks features by their impact on the model's predictions, with the top feature being the most influential and data points spread along a horizontal axis showing the direction and magnitude (negative or positive) of their impact. For this experiment, XGBR was chosen over RFR due to its faster training speed.

For FDCS 1, the most critical feature for predicting electricity consumption, and indoor temperature was the hour of the day, indicating a daily cyclical pattern, potentially

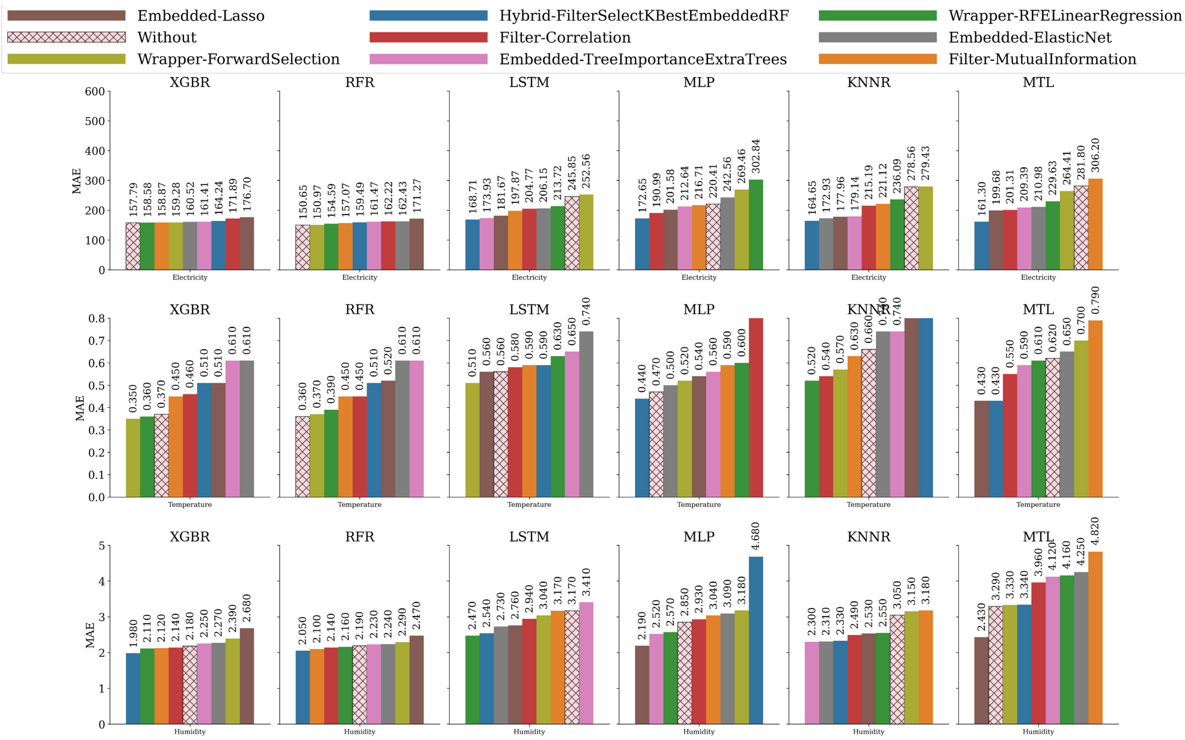


FIGURE 7. **Model performance evaluation using MAE of the test set of different feature selection methods in FDCS 1.**

influenced by operational routines in such an environment. Notably, in FDCS 1, among the top ten features for predicting electricity, four were extracted features, while the remaining six were weather-related features. Similarly, for predicting indoor temperature and humidity, some features were extracted features, highlighting the effectiveness of these methods in capturing complex patterns and trends in FDCS environments, thus improving the model's predictive accuracy.

Likewise, in FDCS 2, key features for predicting the same variables include time-lags, cyclic, and rolling-window statistical features, alongside weather-related features. These results across both systems underscore the importance of feature extraction in capturing FDCS complexities, thereby enhancing the model's prediction capability. The noticeable variance in SHAP values for electricity in FDCS 2, indicated by the more widely dispersed dots, implies that the features affecting its model predictions demonstrate greater variability. This could be due to its irregular usage patterns, aligning with earlier findings discussed in subsection III-A and Figure 5, unlike the predictable trends seen in FDCS 1. The weather impact on both FDCSs is noticeable, particularly the outdoor temperature, which directly influences the cooling demand, thereby affecting energy consumption, underscoring the considerable influence of weather-related features in energy forecasting strategies for such systems.

C. DATASET SIZE IMPLICATIONS

As demonstrated in Figure 12, the evaluation of how the volume of the training data affects forecasting performance was conducted using XGBR and RFR, as they had the most consistently superior performance as shown in previous analyses. In this experiment, the models were trained starting with a baseline of a single day's worth of hourly data. From this baseline, the dataset was expanded in one-day increments, each comprising 24 hourly data points, to systematically assess the impact of the train dataset size on prediction performance in the test set. While the best-performing models can provide valuable insights into the implications of dataset size, it is important to note that this approach is computationally expensive, as it involves iterative retraining of the models on an ever-expanding data corpus, a constraint that precluded a comprehensive investigation of dataset size implications across all examined algorithms.

In predicting electricity consumption, the XGBR and RFR models showed fluctuating yet overall declining MAE in the test set as the train dataset size increased in both FDCSs. The most notable improvements (i.e., reduction in prediction errors) for both models occurred at 1344 hours (56 days' worth of data) in FDCS 1 and at 1680 hours (70 days) in FDCS 2. After those levels, and as the dataset size continued to grow, both models showed a trend towards stability with minor MAE fluctuations, signalling performance plateaus. For temperature, both models showed signs of stabilisation around 1560 hours' worth of data with minor MAE changes in FDCS 1 compared to FDCS 2. In analysing the dataset size impact on humidity prediction, MAE decreased as data grew; however, there were more noticeable fluctuations in prediction errors.

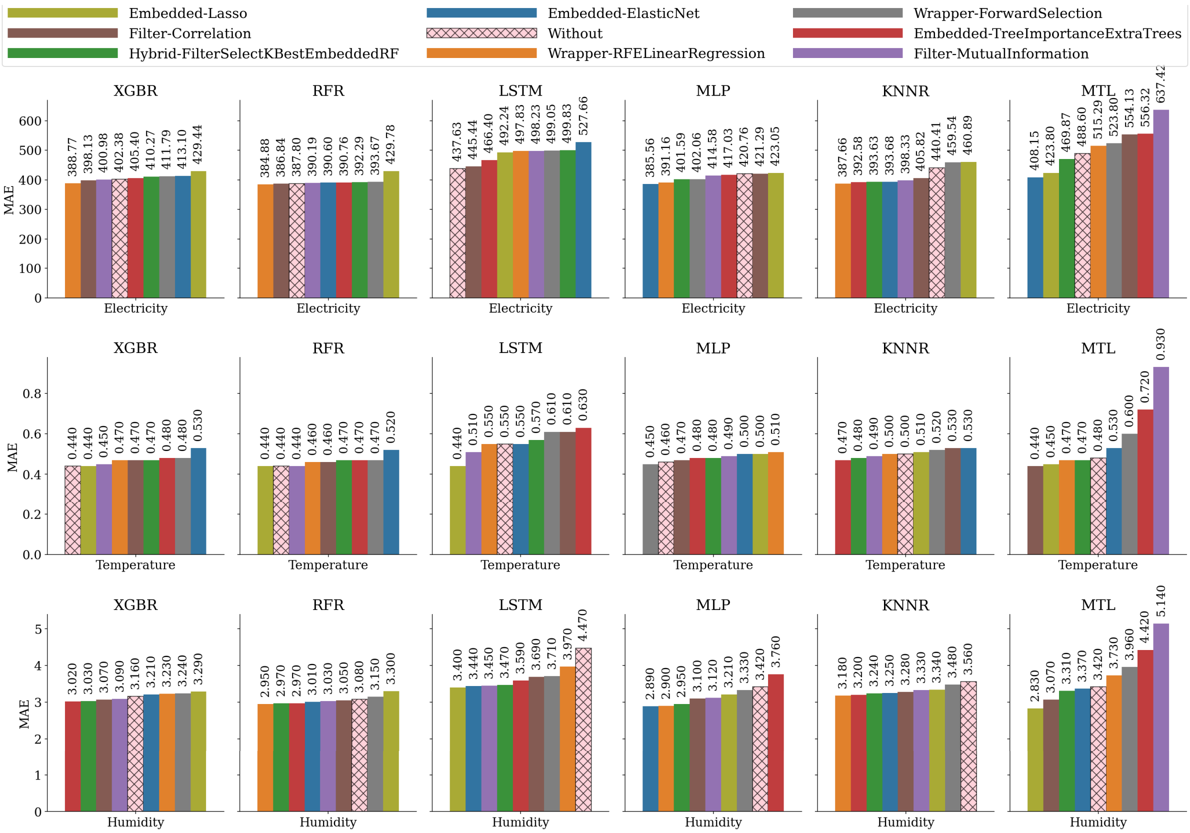


FIGURE 8. Model performance evaluation using MAE of the test set of different feature selection methods in FDCS 2.

TABLE 9. Summary of key recommendations for ML applications in FDCS systems, offering guidelines and insights to enhance the performance and efficiency of forecasting models in real-world applications.

Focus Area	Recommendations
Enhancement of input features	Use feature extraction (e.g., the hour of the day, and cyclic features).
Feature selection techniques	Hybrid FS methods generally enhance model performance. Avoid wrapper methods if computational resources are limited.
Algorithms selection	Ensemble-based learning methods (XGBR, RFR) were superior to traditional ML, NN-based, and deep learning models in FDCS predictions. XGBR is particularly recommended for its computational efficiency.
Dataset collection	Contrary to popular belief, larger datasets were not necessary for accurate prediction in this work. Smaller datasets from real-world FDCS facilities may yield reliable predictions when enhanced with robust feature engineering.

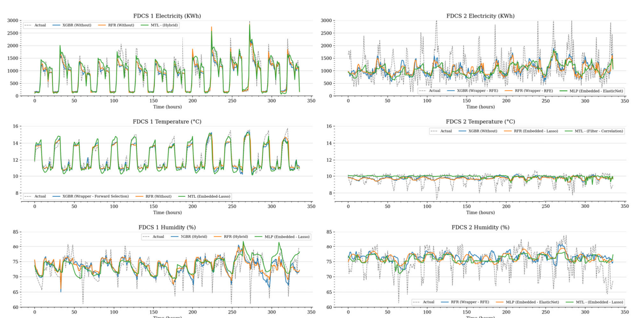


FIGURE 9. Comparison of the three lowest model errors in predicting electricity, temperature, and humidity in FDCS 1 and 2. Plots show actual versus predicted values on the test set.

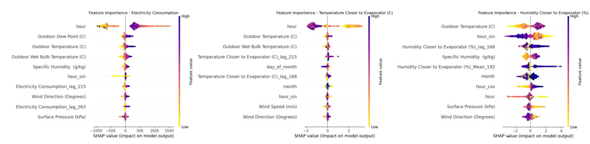


FIGURE 10. Analysis of the top ten features in terms of their impact on model output, as explained by SHAP for electricity, temperature, and humidity variables in FDCS 1.

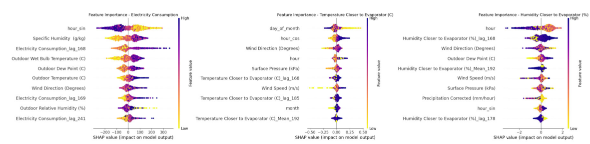


FIGURE 11. Analysis of the top ten features in terms of their impact on model output, as explained by SHAP for electricity, temperature, and humidity variables in FDCS 2.

It is important to acknowledge that these conclusions are drawn from the available data, and additional research is needed to confirm these findings in broader applications. Yet, such observations could be valuable in scenarios where

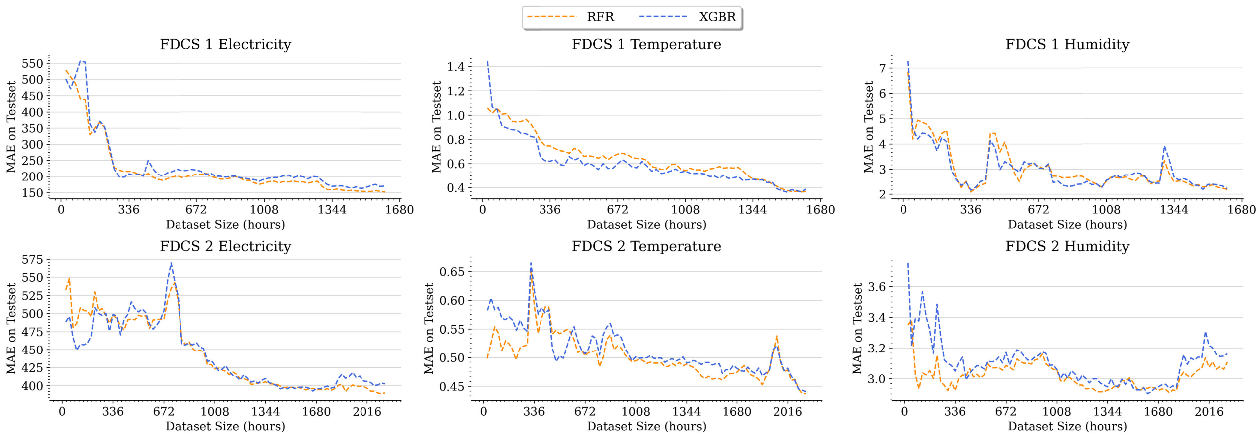


FIGURE 12. Variations of MAE in the test set occur when predicting electricity, temperature, and humidity in both FDCSs, as a function of increasing the training dataset size for the XGBR and RFR models.

acquiring extensive historical data is impractical due to time, cost, or data availability limitations.

IV. CONCLUSION, RECOMMENDATIONS, AND FUTURE WORK

This study proposes an ML pipeline tailored for predicting electricity consumption, indoor temperature, and humidity one week (hourly) into the future in FDCS settings, addressing the unique challenges of such environments compared to previous building-focused ML studies and also suitable for small dataset sizes. Two real-world datasets of FDCSs have been employed for training, validation, and testing of the developed models. The results show that ensemble-based methods (RFF and XGBR) outperformed other models in both examined FDCS datasets, evidenced by the lowest MAE values while neural network-based and deep learning models showed varied performance. Eight FS methods have been investigated, and the results from the datasets used in this study indicate that hybrid methods generally enhance model performance, while wrapper methods are computationally expensive. The conducted feature importance analysis underscores the importance of feature extraction, given that the extracted features have a noticeable impact on model outputs, as evidenced by SHAP analysis. The implication of dataset size on model accuracy was analysed, providing some insights into estimating the minimum dataset size needed for forecasting electricity in FDCS (1344 hours' worth of data for FDCS 1 and 1680 hours for FDCS 2). Nevertheless, these conclusions are drawn from the available datasets, and further research is needed.

This study, while providing insightful findings, acknowledges some limitations that pave the way for future research. The data collected predominantly during colder months (end of October to end of January) may not fully capture FDCS dynamics during warmer periods. Future studies should expand data collection across different seasons, particularly warmer months, to better understand the impact of varying outdoor conditions on FDCS environments and the performance of ML models. Another notable limitation is the untested generalisability of these ML models across a broader range of FDCS settings. Although benefitted from comparative analysis using datasets from two different FDCS environments, this does not fully address potential variations in operational scales and geographic locations. Consequently, future research should focus on validating and enhancing the universality of these models in various environments, further exploring their applicability across a wider spectrum of FDCS operational contexts.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Nasser Alkhulaifi: Conceptualization, Data curation, Formal analysis, Validation, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Alexander Bowler**: Supervision, Conceptualization, Methodology, Formal analysis, Software, Validation, Investigation, Writing – review & editing. **Direnc Pekaslan**: Supervision, Writing – review & editing. **Gulcan Serdaroglu**: Investigation. **Steve Closs**: Investigation. **Nicholas Watson**: Supervision, Conceptualization, Methodology, Investigation, Writing – review & editing, Funding acquisition. **Isaac Triguero**: Supervision, Conceptualization, Methodology, Writing – review & editing.

CODE AND DATA AVAILABILITY

The code developed for this work can be found online at https://github.com/Nasser-Alkhulaifi/FDCS_Paper. The authors do not have permission to share the dataset.

ACKNOWLEDGMENT

The authors would like to thank the Energy for Business Team, University of Nottingham, for helping with data collection.

REFERENCES

[1] P. D. Gaspar, P. D. Silva, L. P. Andrade, and J. Nunes, "Computational tool for the analysis of the energy performance of cold storage systems in the wine industry," in *Trends in Beverage Packaging*. Cambridge, MA, USA: Academic, 2019, pp. 243–278, doi: 10.1016/B978-0-12-816683-3.00009-8.

[2] A. G. Reynolds, *Managing Wine Quality: Viticulture and Wine Quality*. Oxford, U.K.: Elsevier, 2010.

[3] J. Walther and M. Weigold, "A systematic review on predicting and forecasting the electrical energy consumption in the manufacturing industry," *Energies*, vol. 14, no. 4, p. 968, Feb. 2021.

[4] H.-M. Hoang, S. Duret, D. Flick, and O. Laguerre, "Preliminary study of airflow and heat transfer in a cold room filled with apple pallets: Comparison between two modelling approaches and experimental results," *Appl. Thermal Eng.*, vol. 76, pp. 367–381, Feb. 2015.

[5] J. A. Evans, E. C. Hammond, A. J. Gigiel, A. M. Fostera, L. Reinholdt, K. Fikiin, and C. Zilio, "Assessment of methods to reduce the energy consumption of food cold stores," *Appl. Thermal Eng.*, vol. 62, no. 2, pp. 697–705, Jan. 2014.

[6] O. J. Fisher, N. J. Watson, J. E. Escrig, R. Witt, L. Porcu, D. Bacon, M. Rigley, and R. L. Gomes, "Considerations, challenges and opportunities when developing data-driven models for process manufacturing systems," *Comput. Chem. Eng.*, vol. 140, Sep. 2020, Art. no. 106881.

[7] F. J. Montáns, F. Chinesta, R. Gómez-Bombarelli, and J. N. Kutz, "Data-driven modeling and learning in science and engineering," *Comp. Rendus. Mécanique*, vol. 347, no. 11, pp. 845–855, Nov. 2019.

[8] H.-X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, Aug. 2012.

[9] A. Mosavi and A. Bahmani. (Mar. 2019). *Energy Consumption Prediction Using Machine Learning: A Review*. [Online]. Available: http://dx.doi.org/10.20944/preprints201903.0131.v1

[10] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192–1205, Jan. 2018.

[11] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas, and P. Liatsis, "Electricity demand loads modeling using AutoRegressive moving average (ARMA) models," *Energy*, vol. 33, no. 9, pp. 1353–1360, Sep. 2008.

[12] W. T. Ho and F. W. Yu, "Predicting chiller system performance using ARIMA-regression models," *J. Building Eng.*, vol. 33, Jan. 2021, Art. no. 101871.

[13] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2010.

[14] P. Sen, M. Roy, and P. Pal, "Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization," *Energy*, vol. 116, pp. 1031–1038, Dec. 2016.

[15] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *Jpn. Architectural Rev.*, vol. 3, no. 1, pp. 62–76, Jan. 2020.

[16] G. R. Newsham and B. J. Birt, "Building-level occupancy data to improve ARIMA-based electricity use forecasts," in *Proc. 2nd ACM Workshop Embedded Sens. Syst. Energy-Efficiency Building*, pp. 13–18.

[17] C. M. Pereira, N. N. D. Almeida, and M. L. F. Velloso, "Fuzzy modeling to forecast an electric load time series," *Proc. Comput. Sci.*, vol. 55, pp. 395–404, 2015.

[18] H. Cui and X. Peng, "Short-term city electric load forecasting with considering temperature effects: An improved ARIMAX model," *Math. Problems Eng.*, vol. 2015, pp. 1–10, Dec. 2015.

[19] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 74, pp. 902–924, Jul. 2017.

[20] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, "Machine learning approaches for estimating commercial building energy consumption," *Appl. Energy*, vol. 208, pp. 889–904, Dec. 2017.

[21] S. Seyedzadeh, F. P. Rahimian, I. Glesk, and M. Roper, "Machine learning for estimation of building energy consumption and performance: A review," *Visualizat. Eng.*, vol. 6, no. 1, pp. 1–20, Dec. 2018.

[22] H. M. Hoang, M. Akerma, N. Mellouli, A. L. Montagner, D. Leducq, and A. Delahaye, "Development of deep learning artificial neural networks models to predict temperature and power demand variation for demand response application in cold storage," *Int. J. Refrigeration*, vol. 131, pp. 857–873, Nov. 2021.

[23] C. Deb, L. S. Eang, J. Yang, and M. Santamouris, "Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks," *Energy Buildings*, vol. 121, pp. 284–297, Jun. 2016.

[24] C. Fan, F. Xiao, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Appl. Energy*, vol. 195, pp. 222–233, Jun. 2017.

[25] L. Xu, S. Wang, and R. Tang, "Probabilistic load forecasting for buildings considering weather forecasting uncertainty and uncertain peak load," *Appl. Energy*, vol. 237, pp. 180–195, Mar. 2019.

[26] W. Cao, J. Yu, M. Chao, J. Wang, S. Yang, M. Zhou, and M. Wang, "Short-term energy consumption prediction method for educational buildings based on model integration," *Energy*, vol. 283, Nov. 2023, Art. no. 128580.

[27] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Appl. Energy*, vol. 212, pp. 372–385, Feb. 2018.

[28] Y. Ding, Q. Zhang, T. Yuan, and F. Yang, "Effect of input variables on cooling load prediction accuracy of an office building," *Appl. Thermal Eng.*, vol. 128, pp. 225–234, Jan. 2018.

[29] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72–81, Sep. 2019.

[30] C.-L. Liu, C.-J. Tseng, T.-H. Huang, J.-S. Yang, and K.-B. Huang, "A multi-task learning model for building electrical load prediction," *Energy Buildings*, vol. 278, Jan. 2023, Art. no. 112601.

[31] Y. Ding, Q. Zhang, and T. Yuan, "Research on short-term and ultra-short-term cooling load prediction models for office buildings," *Energy Buildings*, vol. 154, pp. 254–267, Nov. 2017.

[32] Y. Gao and Y. Ruan, "Interpretable deep learning model for building energy consumption prediction based on attention mechanism," *Energy Buildings*, vol. 252, Dec. 2021, Art. no. 111379.

[33] G. Hong, G.-S. Choi, J.-Y. Eum, H. S. Lee, and D. D. Kim, "The hourly energy consumption prediction by KNN for buildings in community buildings," *Buildings*, vol. 12, no. 10, p. 1636, Oct. 2022.

[34] G. Li, F. Li, T. Ahmad, J. Liu, T. Li, X. Fang, and Y. Wu, "Performance evaluation of sequence-to-sequence-attention model for short-term multi-step ahead building energy predictions," *Energy*, vol. 259, Nov. 2022, Art. no. 124915.

[35] J. Moon, S. Park, S. Rho, and E. Hwang, "A comparative analysis of artificial neural network architectures for building energy consumption forecasting," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 9, Sep. 2019, Art. no. 155014771987761.

[36] J. Kim, J. Moon, E. Hwang, and P. Kang, "Recurrent inception convolution neural network for multi short-term load forecasting," *Energy Buildings*, vol. 194, pp. 328–341, Jul. 2019.

[37] L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, and W. Livingood, "A review of machine learning in building load prediction," *Appl. Energy*, vol. 285, Mar. 2021, Art. no. 116452.

[38] M. Bourdeau, X. Q. Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," *Sustain. Cities Soc.*, vol. 48, Jul. 2019, Art. no. 101533.

[39] Y. Wei, L. Xia, S. Pan, J. Wu, X. Zhang, M. Han, W. Zhang, J. Xie, and Q. Li, "Prediction of occupancy level and energy consumption in office building using blind system identification and neural networks," *Appl. Energy*, vol. 240, pp. 276–294, Apr. 2019.

[40] J. Moon, S. Jung, J. Rew, S. Rho, and E. Hwang, "Combination of short-term load forecasting models based on a stacking ensemble approach," *Energy Buildings*, vol. 216, Jun. 2020, Art. no. 109921.

[41] R. Wang, S. Lu, and W. Feng, "A novel improved model for building energy consumption prediction based on model integration," *Appl. Energy*, vol. 262, Mar. 2020, Art. no. 114561.

[42] Y. Ding, Q. Zhang, T. Yuan, and K. Yang, "Model input selection for building heating load prediction: A case study for an office building in Tianjin," *Energy Buildings*, vol. 159, pp. 254–270, Jan. 2018.

[43] F. Wahid, R. Ghazali, M. Fayaz, and A. S. Shah, "Statistical features based approach (SFBA) for hourly energy consumption prediction using neural network," *Int. J. Inf. Technol. Comput. Sci.*, vol. 9, no. 5, pp. 23–30, May 2017.

[44] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, and S. Ahrentzen, "Random forest based hourly building energy prediction," *Energy Buildings*, vol. 171, pp. 11–25, Jul. 2018.

[45] T. Chen and Y.-C. Wang, "Long-term load forecasting by a collaborative fuzzy-neural approach," *Int. J. Electr. Power Energy Syst.*, vol. 43, no. 1, pp. 454–464, Dec. 2012.

[46] K. Li, C. Hu, G. Liu, and W. Xue, "Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis," *Energy Buildings*, vol. 108, pp. 106–113, Dec. 2015.

[47] L. Peng, L. Wang, D. Xia, and Q. Gao, "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory," *Energy*, vol. 238, Jan. 2022, Art. no. 121756.

[48] C. Fan, Y. Sun, Y. Zhao, M. Song, and J. Wang, "Deep learning-based feature engineering methods for improved building energy prediction," *Appl. Energy*, vol. 240, pp. 35–45, Apr. 2019.

[49] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Mar. 2003.

[50] S. Jurado, À. Nebot, F. Mugica, and N. Avellana, "Hybrid methodologies for electricity load forecasting: entropy-based feature selection with machine learning and soft computing techniques," *Energy*, vol. 86, pp. 276–291, Jun. 2015.

[51] Y. Sun, F. Haghighat, and B. C. M. Fung, "A review of the-state-of-the-art in data-driven approaches for building energy prediction," *Energy Buildings*, vol. 221, Aug. 2020, Art. no. 110022.

[52] A. Bagnasco, F. Fresi, M. Saviozzi, F. Silvestro, and A. Vinci, "Electrical consumption forecasting in hospital facilities: An application case," *Energy Buildings*, vol. 103, pp. 261–270, Sep. 2015.

[53] C. Li, Y. Tao, W. Ao, S. Yang, and Y. Bai, "Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition," *Energy*, vol. 165, pp. 1220–1227, Dec. 2018.

[54] Y. Ding and X. Liu, "A comparative analysis of data-driven methods in building energy benchmarking," *Energy Buildings*, vol. 209, Feb. 2020, Art. no. 109711.

[55] H.-X. Zhao and F. Magoulès, "Feature selection for predicting building energy consumption based on statistical learning method," *J. Algorithms Comput. Technol.*, vol. 6, no. 1, pp. 59–77, Mar. 2012.

[56] C. Fan, F. Xiao, and S. Wang, "Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques," *Appl. Energy*, vol. 127, pp. 1–10, Aug. 2014.

[57] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, "Short-term load forecasting in a non-residential building contrasting models and attributes," *Energy Buildings*, vol. 92, pp. 322–330, Apr. 2015.

[58] Z. Wang, Y. Wang, and R. S. Srinivasan, "A novel ensemble learning approach to support building energy use prediction," *Energy Buildings*, vol. 159, pp. 109–122, Jan. 2018.

[59] L. Zhang and J. Wen, "A systematic feature selection procedure for short-term data-driven building energy forecasting model development," *Energy Buildings*, vol. 183, pp. 428–442, Jan. 2019.

[60] R. Olu-Ajayi, H. Alaka, I. Sulaimon, H. Balogun, G. Wusu, W. Yusuf, and M. Adegoke, "Building energy performance prediction: A reliability analysis and evaluation of feature selection methods," *Expert Syst. Appl.*, vol. 225, Sep. 2023, Art. no. 120109.

[61] Y. Li, Y. Peng, D. Zhang, Y. Mai, and Z. Ruan, "XGBoost energy consumption prediction based on multi-system data HVAC," 2021, *arXiv:2105.09945*.

[62] A. A. Bassi, A. Shenoy, H. Sharma, C. G. Sigurdson, and J. H. Chan, "Building energy consumption forecasting: A comparison of gradient boosting models," in *Proc. 12th Int. Conf. Adv. Inf. Technol.*, 2021, pp. 1–9.

[63] S. Seyedzadeh, F. Pour Rahimian, P. Rastogi, and I. Glesk, "Tuning machine learning models for prediction of building energy loads," *Sustain. Cities Soc.*, vol. 47, May 2019, Art. no. 101484.

[64] R. Caruana, "Multitask learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, 1997.

[65] L. Wang, M. Tan, J. Chen, and C. Liao, "Multi-task learning based multi-energy load prediction in integrated energy system," *Appl. Intell.*, vol. 53, pp. 1–17, Mar. 2022.

[66] Z. Tan, G. De, M. Li, H. Lin, S. Yang, L. Huang, and Q. Tan, "Combined electricity-heat-cooling-gas load forecasting model for integrated energy system based on multi-task learning and least square support vector machine," *J. Cleaner Prod.*, vol. 248, Mar. 2020, Art. no. 119252.

[67] A. R. Ajiboye, R. Abdullah-Arshah, H. Qin, and H. Isah-Kebbe, "Evaluating the effect of dataset size on predictive model using supervised learning technique," *Int. J. Comput. Syst. Softw. Eng.*, vol. 1, no. 1, pp. 75–84, Feb. 2015.

[68] W. Choi and S. Lee, "Performance evaluation of deep learning architectures for load and temperature forecasting under dataset size constraints and seasonality," *Energy Buildings*, vol. 288, Jun. 2023, Art. no. 113027.

[69] M. Cai, M. Pipattanasomporn, and S. Rahman, "Day-ahead building-level load forecasts using deep learning vs. Traditional time-series techniques," *Appl. Energy*, vol. 236, pp. 1078–1088, Feb. 2019.

[70] X. Wang, T. Zhao, H. Liu, and R. He, "Power consumption predicting and anomaly detection based on long short-term memory neural network," in *Proc. IEEE 4th Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2019, pp. 487–491.

[71] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine learning, deep learning and statistical analysis for forecasting building energy consumption—A systematic review," *Eng. Appl. Artif. Intell.*, vol. 115, Oct. 2022, Art. no. 105287.

[72] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao, "A review of data-driven approaches for prediction and classification of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 1027–1047, Feb. 2018.

[73] NASA. (2023). *NASA Langley Research Centre's POWER Project (NASA Power—Data Access Viewer)*. [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/

[74] R. Sendra-Arranz and A. Gutiérrez, "A long short-term memory artificial neural network to predict daily HVAC consumption in buildings," *Energy Buildings*, vol. 216, Jun. 2020, Art. no. 109952.

[75] N. Somu and K. Ramamritham, "A hybrid model for building energy consumption forecasting using long short term memory networks," *Appl. Energy*, vol. 261, Mar. 2020, Art. no. 114131.

[76] C. Zhou, Z. Fang, X. Xu, X. Zhang, Y. Ding, X. Jiang, and Y. Ji, "Using long short-term memory networks to predict energy consumption of air-conditioning systems," *Sustain. Cities Soc.*, vol. 55, Apr. 2020, Art. no. 102000.

[77] O. Valgaev, F. Kupzog, and H. Schmeck, "Building power demand forecasting using K-nearest neighbours model–practical application in smart city demo aspern project," *CIRED Open Access Proc. J.*, vol. 2017, no. 1, pp. 1601–1604, Oct. 2017.

[78] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[79] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, vol. 11, Aug. 2016, pp. 785–794.

[80] S. Afzal, B. M. Ziapour, A. Shokri, H. Shakibi, and B. Sobhani, "Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms," *Energy*, vol. 282, Nov. 2023, Art. no. 128446.

[81] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[82] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geoscientific Model Develop.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.

[83] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–12.

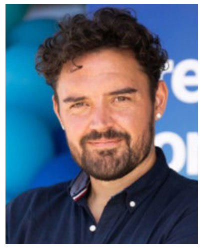
NASSER ALKHULAIFI received the B.Sc. (Hons.) and B.A. degrees from Al-Imam Muhammad Ibn Saud University and the M.Sc. degree (Hons.) in logistics and supply chain from the Operations Management and Information Systems Department, University of Nottingham, where he is currently pursuing the Ph.D. degree with the Computational Optimisation and Learning (COL) Laboratory, School of Computer Science. He is also a member of the Horizon CDT Institute funded by the Engineering and Physical Sciences Research Council (EPSRC). His Ph.D. research focuses on improving energy forecasting by integrating predictive machine learning, automated feature engineering, and decision-focused learning to enhance efficiency and sustainability in the energy domain. He is also a Certified Project Management Professional (PMP), accredited by the Project Management Institute, USA.

ALEXANDER L. BOWLER received the M.Eng. and Ph.D. degrees in chemical engineering from the University of Nottingham. He is currently a Research Fellow in artificial intelligence for sustainable food, specializing in machine learning, optimization, sensors, hybrid modeling, and carbon footprinting. His Ph.D. focusing on machine learning and ultrasonic sensors for process monitoring. After joining the University of Leeds, in 2023, he has worked on projects related to carbon footprint calculations, Bayesian optimization, and sensor-based process optimization. He is currently a PDRA on a BBSRC-funded project using AI to optimize microbial protein production from food side streams.

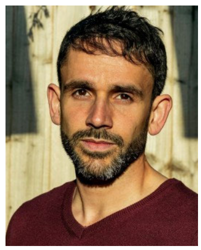
DIRENC PEKASLAN (Member, IEEE) is currently a Transitional Assistant Professor with the School of Computer Science, University of Nottingham, and a member of the Laboratory for Uncertainty in Data and Decision Making (LUCID). His research interests include uncertainty, adaptive systems, and interpretability in artificial intelligence and fuzzy logic systems. He contributes to developing novel approaches for capturing, modeling, and reasoning with uncertain data in decision support contexts. His work combines theoretical advances with practical applications in areas, such as smart cities, renewable energy forecasting, and transportation systems to addressing real-world challenges through AI techniques that effectively handle uncertainty.

GULCAN SERDAROGLU received the M.Sc. degree (Hons.) in chemical engineering and in energy and environmental engineering. She has three years of industry experience as a Research and Development Engineer, Ph.D. research in hydrogen generation, and five years of experience as a Low-Carbon Energy Consultant for Businesses, and currently the Programme Manager of the Sustainable Hydrogen Centre for Doctoral Training, Faculty of Engineering, University of Nottingham.

STEVE CLOSS received the M.Phil. degree from Sheffield Hallam University. He developed energy loss test techniques for large buildings with Sheffield Hallam University. He has a wealth of experience gained in various roles linked to sustainability and the pursuit of net zero. He joins the University of Nottingham Energy Institute as a Business Development Manager to undertake a pioneering airport decarburization study, as part of regional activity surrounding East Midlands Airport. He is also a Chartered Engineer and a BREEAM Accredited Professional.



NICHOLAS J. WATSON received the M.Eng. degree in mechanical engineering from the University of Hull and the Ph.D. degree in chemical engineering from the University of Leeds. After working as a Postdoctoral Research Assistant with the University of Leeds, where he developed acoustic sensing technologies, he joined the University of Nottingham, in 2014, as an Assistant Professor and was promoted to an Associate Professor, in 2020. His research interests include integrating in-process sensing (acoustic and optical) with machine learning for digital manufacturing, particularly in the food and drink sector. In 2023, he became a Professor in artificial intelligence in food with the University of Leeds. He has published over 70 articles, led major projects funded by Innovate U.K., EPSRC, and the Royal Academy of Engineering; and regularly speaks at industry events on digital manufacturing and AI. He also collaborates extensively with the food, pharmaceutical, and FMCG industries.



ISAAC TRIGUERO (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Granada, Spain, in 2009 and 2014, respectively. He is currently enjoying a Distinguished Senior Research Fellowship with the University of Granada. His work is mostly concerned with the research of novel methodologies for big data analytics. He has published more than 90 international publications in the fields of big data, machine learning, and optimization (H-index=35 and more than 5,700 citations on Google Scholar). He is the Section Editor-in-Chief of the *Machine Learning and Knowledge Extraction* journal, and an Associate Editor of the *Big Data and Cognitive Computing* journal and IEEE ACCESS journal. He has acted as the Program Co-Chair of the IEEE Conference on Smart Data, in 2016, the IEEE Conference on Big Data Science and Engineering, in 2017, and the IEEE International Congress on Big Data, in 2018. He is currently co-leading two projects on General Purpose Artificial Intelligence: A €1.6M University-Industry Research Grant funded by European Union-Next Generation EU and a €120K Knowledge Generation Project, funded by the Ministry of Science, Innovation and Universities of Spain.