

A Review of Data-Driven Building Energy Prediction

Huiheng Liu ¹, Jinrui Liang ¹, Yanchen Liu ^{1,2,*} and Huijun Wu ^{1,2}

¹ School of Civil Engineering, Guangzhou University, Guangzhou 510006, China

² Guangdong Provincial Key Laboratory of Building Energy Efficiency and Application Technologies, Guangzhou University, Guangzhou 510006, China

* Correspondence: liuych@gzhu.edu.cn

Abstract: Building energy consumption prediction has a significant effect on energy control, design optimization, retrofit evaluation, energy price guidance, and prevention and control of COVID-19 in buildings, providing a guarantee for energy efficiency and carbon neutrality. This study reviews 116 research papers on data-driven building energy prediction from the perspective of data and machine learning algorithms and discusses feasible techniques for prediction across time scales, building levels, and energy consumption types in the context of the factors affecting data-driven building energy prediction. The review results revealed that the outdoor dry-bulb temperature is a vital factor affecting building energy consumption. In data-driven building energy consumption prediction, data preprocessing enables prediction across time scales, energy consumption feature extraction enables prediction across energy consumption types, and hyperparameter optimization enables prediction across time scales and building layers.

Keywords: data driven; building energy prediction; machine learning; influencing factors

1. Introduction

Buildings accounted for 36% of global energy demand and 37% of energy-related CO₂ emissions in 2020 [1], with the operational phase accounting for greater than 80% of the total building energy consumption [2]. The contribution of building carbon emissions in developing countries can reach 52% [3]. For example, the world's largest developing country, China, produced 9467 million tons of energy-related CO₂ in 2018, accounting for approximately 29% of global CO₂ emissions [4], with the building sector accounting for 20% of total energy consumption and 30% of total CO₂ emissions [5]. Thus, reducing energy consumption and carbon emissions from buildings is urgently required in the context of carbon neutrality.

Improving the energy efficiency of buildings can save energy consumption by 30–80% and significantly reduce the corresponding building carbon emissions [1]. Building energy consumption prediction (BECP), an effective initiative to improve building energy efficiency, has played an important role in building energy control, building design optimization, building retrofit evaluation, energy price guidance, and COVID-19 prevention and control. At the building operation stage, BECP is combined with end-user group requirements to personalize the operation of energy-consuming equipment to optimize its energy efficiency without affecting the thermal comfort of its occupants [6]. Mariam et al. [7] proposed a neural network (NN)-based model predictive control for a heating, ventilation, and air conditioning (HVAC) system in the Qatar University Sports gymnasium, along with the management planning control system, which achieved energy savings of up to 46% while jointly optimizing the thermal comfort and air quality of the indoor environment. Tomasz [8] performed energy consumption prediction and energy efficiency control for a multi-family house and an office building in Poland, and the actual energy

Citation: Liu, H.; Liang, J.; Liu, Y.; Wu, H. A Review of Data-Driven Building Energy Prediction. *Buildings* **2023**, *13*, 532. <https://doi.org/10.3390/buildings13020532>

Academic Editor: Antonio Caggiano

Received: 18 January 2023

Revised: 7 February 2023

Accepted: 8 February 2023

Published: 15 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

savings exceeded 15% and 24%, respectively. At the building design stage, Kim [9] explored the effects of air permeability, solar heat gain coefficient (SHGC), and thermal conductivity on building energy consumption by varying the design variables based on the developed residential energy prediction model. Air permeability has a significant influence on heating load, while the SHGC has the greatest impact on cooling load. Increasing the thermal conductivity reduced the cooling energy consumption by 8–26%. Adding the heat transfer coefficient of the envelope leads to an increase in the heating energy requirement by approximately 27–29%. At the building retrofit stage, Seo et al. [10] assessed the feasibility of projects and helped decision makers predict the heating energy demand of low-income households in Korea. In addition, BECP can guide the price of energy supplied to end customers [11]. Jinseok and Ki-Il [12] proposed a long- and short-term memory (LSTM)-based BECP model and a time of use (TOU)-based operation algorithm that can theoretically reduce the peak demand cost by 22%. In response to future energy crises, building energy consumption forecasting can also provide facility managers, electricity suppliers, and decision makers with beneficial information for planning energy usage data, monitoring energy consumption anomalies, regulating energy costs, and responding to demand strategies to reshape building load profiles and reduce peak demand [13,14]. After the COVID-19 pandemic, building air-conditioning energy consumption predictions can be used to improve the health status of residents. Michael et al. [15] integrated the energy consumption prediction of an HVAC system with air-transmission risk to determine the required space and the lowest energy cost to reduce infection risk. Li et al. [16] combined the energy consumption forecast caused by COVID-19 with an econometric model to integrate the long- and short-term impacts under full statistical verification. Simultaneously, BECP also plays an important role in building operation, energy efficiency assessment, fault detection and diagnosis [17], demand-side management [18,19], and maintenance [20].

Currently, BECP methods are mainly divided into three types: white box, black box, and gray box [21,22]. The white box, also called a physical model, is built and analyzed based on thermodynamic principles and detailed building energy characteristic information. Commonly used building energy consumption simulation software programs include EnergyPlus, DesignBuilder, DeST, eQuest, TRNSYS, and DOE-2. These software programs can establish energy consumption prediction models according to the specific structural parameters of the buildings, outdoor meteorological data, and air-conditioning system performance. However, the disadvantage of a physical model is its requirement of multiple complex verification and adjustments to ensure satisfactory reliability of the predicted results. Therefore, scholars have developed a black-box model, also known as the data-driven method, which does not rely on thermodynamic principles and building energy characteristics. The data-driven method uses specific algorithms to analyze the collected large datasets and mine the logic between the data to achieve automatic decision-making. Compared with the physical model, data-driven methods are only based on large datasets and machine learning (ML) algorithms [23,24], which require a large data sample size. The gray-box model is a combination of a physical model and data-driven method to predict building energy consumption. Although the model offers improved prediction accuracy and reduced calculation difficulty compared with the physical model, the process of establishing the model may incorporate inaccurate assumptions [22]. In general, black boxes provide greater practical convenience than white or gray boxes. The use of data analysis and mining of large datasets circumvents the necessity of building a physical model to forecast energy consumption. As presented in Figure 1 (data source: Web of science), research on data-driven BECP has grown significantly since 2019. Owing to the development of building intelligence and wide application of high-precision data collectors, abundant historical data are available for studying energy consumption prediction. By searching “data driven” and “building energy consumption prediction”, we found greater than 400 related research papers in this review. Considering the correlation of journal impact factors, paper citations, and BECP, 116 research results strongly related to BECP, with

journal impact factors of greater than 6.0 or having higher than five citations, were selected for this literature review.

In this study, 116 research papers on data-driven BECP were reviewed from the perspective of data and ML algorithms, analyzed for the factors affecting data-driven building energy consumption prediction, and summarized for key techniques for prediction across time scales, building levels, and energy consumption types. Section 2 summarizes the research and application status of energy consumption prediction from the perspective of data and machine learning algorithms. Section 3 outlines the factors influencing data-driven BECP. Section 4 discusses the current situation and proposes directions for future research. Section 5 summarizes the main results of this study.

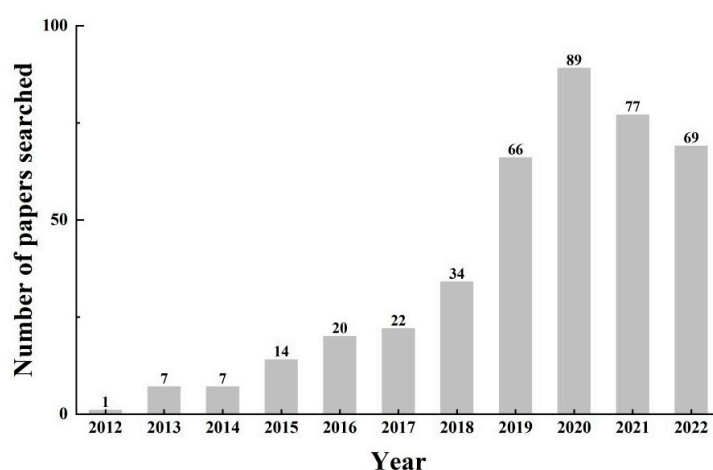


Figure 1. Research statistics on application of the data-driven method in BECP.

2. Application of Data-Driven Methods in Building Energy Consumption Prediction

As depicted in Figure 2, the data-driven BECP process comprises five main parts: data acquisition, data preprocessing, data selection, data training, and application. Data are generally collected from building (energy) management systems and sensor records. Data preprocessing involves integrating the data for missing values, outliers, noisy data, or numerical discrepancies in the original data. Data selection involves extracting the building energy consumption phase characteristics as the main input data for prediction. Data training seeks the predicted and actual values of energy consumption to achieve a high prediction performance of the model. The application part forecasts the corresponding building energy consumption as required. Therefore, the core of data-driven BECP comprises the data and algorithms. This section reviews the current research status of data-driven BECP from the perspectives of both data and ML algorithms.

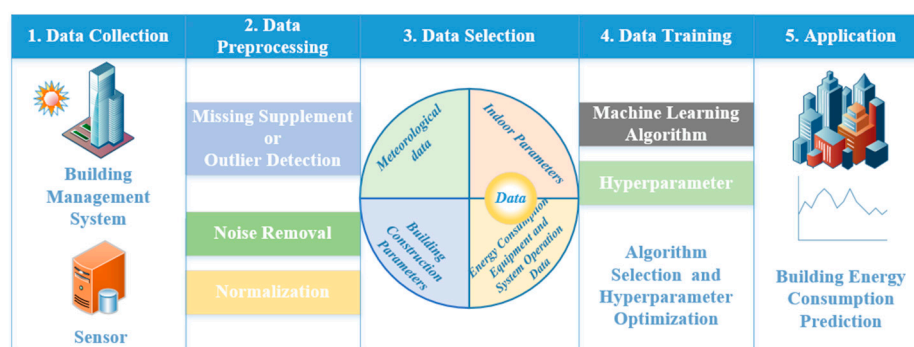


Figure 2. Data-driven method process in BECP.

2.1. Dataset Application

The datasets (energy consumption features) required for BECP are mainly divided into meteorological data, building energy consumption equipment and system operation data (BESD), indoor environmental parameters, and building construction parameters (Table 1).

Table 1. Building energy consumption features.

| Dataset Type | Meteorological Data | BESD | Indoor Parameters | Building Construction Parameters |
|-----------------|---|---|---|---|
| Dataset content | Air dry-bulb temperature, air relative humidity, solar radiation intensity, dew-point temperature, wind speed, air pressure, etc. | Air dry-bulb temperature, air relative humidity, solar radiation intensity, dew-point temperature, wind speed, air pressure, etc. | Air dry-bulb temperature, air relative humidity, light intensity, pollutant concentration, temperature of inner building envelopes, occupancy, etc. | Surface-to-volume ratio, area-of-enclosure structure, building height, orientation, overall heat transfer coefficient of building envelopes, solar radiation heat gain coefficient of exterior wall, window–wall ratio, sun-shading coefficient, etc. |

Meteorological data include outdoor air dry-bulb temperature, outdoor wet-bulb temperature, dew-point temperature, solar radiation intensity, wind speed, and air pressure [25]. Li et al. [26] selected the outdoor dry-bulb temperature, relative humidity, and solar radiation intensity to predict the cooling load of an office building in Guangzhou. Ding et al. [27] used the outdoor dry-bulb temperature, wind speed, and solar radiation as input parameters to predict the heating load. Zhang et al. [28] used historical load days, and outdoor dry-bulb temperatures to predict the air-conditioning cooling load. Fumo et al. [29] used parameters such as maximum, minimum, and average values of temperature and humidity, dew point, average atmospheric pressure, average wind speed, and maximum wind speed to forecast the energy consumption of residential buildings. Gao et al. [30] used similar meteorological parameters to predict the energy consumption of three office buildings in Shandong Province, China. Bünninga et al. [31] used the actual ambient air temperature (measured on the roof) to predict the residential heat load at a site in Switzerland. Yang et al. [32] used mutual information (MI) and principal component analysis (PCA) for feature selection and dimensionality reduction of multidimensional weather influences to avoid the interference of extraneous factors and to improve the computational speed. Meteorological data have also been applied to predict the electricity consumption of water source heat pump systems [33,34] and energy consumption of screw chillers [35].

BESD are mainly based on historical building energy consumption data, HVAC system operation data, lighting and socket power, and the power of other electrical appliances. Shi et al. [36] used the energy consumption of sockets, lights, and air conditioners measured in real time every hour as the input data of the Echo State Network to predict office energy consumption. Fekri et al. [37] used three years of online smart-meter data on energy consumption and meteorological parameters to predict the BECP of five family homes. Kim et al. [38] used chilled water flow, cooling water temperature, external dry- and wet-bulb temperatures, dew-point temperature, and external relative humidity to predict the cooling load of a building. Nisa and Kun [39] used equipment-operation data such as condensate supply temperature, condenser return temperature, condenser water flow rate, evaporator supply and return temperatures, and power consumption to predict the power consumption of water-cooled chillers. Fan et al. [40] predicted the building cooling load from the water flow and temperature of the evaporator and condenser of the unit, which are part of the outdoor weather data. Fan et al. [41] used chiller operation data, such as water flow rate, total freezing water flow rate, chilled water supply and return

temperature, and outdoor weather data, as input data for the cooling load prediction model. Jeong et al. [42] used only historical data of the electric load for next-day load forecasting where the daily load shows certain patterns and jumps. Ahmad et al. [43] used the hourly electricity consumption of a building's lighting system, air conditioning, elevators, and other equipment as input data for the BECP.

Indoor parameters include indoor air dry-bulb temperature, air relative humidity, occupancy, light intensity, pollutant concentration, wall-surface temperature inside the envelope, and occupancy rate. Ding et al. [44] used indoor environmental data, such as CO₂ concentration and PM2.5 concentration as well as meteorological data as input variables for the prediction model to predict the energy consumption of a green building in Shenzhen. Ding et al. [45] applied data on indoor parameters, such as regional air temperature, regional relative humidity, number of occupants, and outdoor weather parameters, to predict the heating load of ground source heat pump units. Because of difficulties in measuring and quantifying real-time human activity, Sha et al. [46] substituted temporal index characteristics for personnel activity, such as hours in a day, days in a month or week, and months in a year, to represent actual occupancy changes and activity characteristics. Li and Yao [47] incorporated occupant behavior into energy consumption characteristics when predicting the spatial heating and cooling loads of a dwelling.

Building construction parameters mainly include the surface-to-volume ratio, envelope area, building height, orientation, heat transfer coefficient of the envelope, solar radiation heat gain coefficient of exterior walls, window-to-wall ratio, and sun-shading coefficient. João et al. [48] predicted the energy consumption of diverse residences in terms of the relative room compactness, building surface area, wall area, roof area, total height, orientation, glazing area, and glazing area distribution. Joseph et al. [49] used the heat transfer coefficient of the envelope, window-shading coefficient, window-to-wall ratio, and BESD as input parameters to develop a multiple regression prediction model for energy consumption in office buildings under different climatic conditions in China. Qi et al. [50] predicted abnormally high energy consumption in urban buildings based on their construction and local meteorological data.

The results of energy consumption prediction are limited by several input data and uncertainty among diverse types of input data, and the multi-collinearity of the input variables affects the prediction accuracy when a dependent variable is used directly as the input variable without feature extraction [51,52]. Ding et al. [27] randomly formed eight combinations of input variables and compared them, and found that different combinations of input variables resulted in prediction models with completely different accuracies. Fan et al. [53] determined the dry-bulb temperature, direct solar radiation, occupancy rate, lighting power, equipment power, and ventilation rate as the input characteristics for the cooling load prediction of an office building in Guangzhou by calculating the influence coefficient (IC) of numerous input parameters and energy consumption; the results revealed that the outdoor dry-bulb temperature and direct radiation were closely related to the cooling load. Sha et al. [54] used the Pearson correlation coefficient (PCC) to determine the daily cooling and heating degree day as the most significant parameter for predicting the daily electricity consumption of an HVAC. Li et al. [55] found that the outdoor air temperature and solar radiation intensity had a significant effect on the electricity of a teaching building by performing a PCA of the external meteorological parameters. Huang et al. [56] used the relative building compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution to predict the heating and cooling loads of residential buildings.

2.2. Machine Learning Algorithm Application

This section summarizes the current state of application of major ML algorithms in the literature by describing its role not only in energy consumption prediction but also in data preprocessing, feature extraction, and hyperparameter optimization.

2.2.1. Prediction Stage

ML algorithms for prediction include LSTM, artificial neural network (ANN), back propagation (BP) neural network, multilayer perceptron (MLP), support vector regression (SVR), support vector machine (SVM), multiple linear regression (MLR), random forest (RF), and extreme gradient boosting (XGB).

LSTM (Long- and Short-Term Memory)

LSTM is a special recurrent neural network (RNN) that can effectively cope with gradient-disappearance and gradient-explosion problems during the training of long sequences and performs better in long sequences compared with RNN. Sendra-Arranz and Gutierrez [57] used LSTM to predict the daily energy consumption of an HVAC system. Pitti et al. [58] proposed an LSTM-based model to predict the daily energy consumption of a heat pump in Teatro Real, Spain. Rosemary et al. [59] developed an encoder–decoder LSTM model for hourly, day-ahead forecasting of residential high-voltage alternating current use and photovoltaic generation from load history data and outdoor meteorological parameters. Wang et al. [60] applied LSTM to predict miscellaneous electricity, lighting loads, number of occupants, and internal heat gains for double-office buildings in the United States. Jogunola et al. [61] extracted important features from a dataset using a CNN and used LSTM to predict the consumption for various buildings. Das et al. [62] proposed a bidirectional LSTM (Bi-LSTM) model to forecast electricity consumption for one day and one week. Ullah et al. [63] used LSTM, Bi-LSTM, and multilayer LSTM (M-LSTM) to predict dwelling energy consumption. Li et al. [64] applied K-means in combination with LSTM for load prediction at the span scale of building floors. Kim and Cho [65] combined CNN and LSTM to extract the spatiotemporal features of building energy consumption to effectively predict residential energy consumption. He and Tsang [66] proposed a hybrid network based on an improved fully integrated empirical modal decomposition with adaptive noise (iCEEMDAN) and LSTM for accurate short-term load predictions in colleges and universities. Ijaz et al. [67] used convolutional LSTM to extract and encode the spatial features of the data, and Bi-LSTM to decode and learn the sequence patterns, thus reducing the error of energy consumption prediction. He et al. [68] proposed a particle filter fusion method using LSTM and BP to predict the HVAC energy consumption in commercial buildings. Li et al. [69] predicted short-term peak demand forecasts for buildings based on LSTM, and Mughees et al. [70] used Bi-LSTM-based sequence-to-sequence (S2S) regression methods for one-day peak demand forecasting in emergency power outages. Chalapathy et al. [71] proposed an LSTM-based RNN-multiple-input multiple-output (MIMO) structure that performed well at both 1 h and one-day multi-step prediction levels for office buildings, hospitals, and shopping malls. In terms of average absolute error, RNN-MIMO exhibited 33% greater average accuracy than the present state-of-the-art shallow ML models (SVR and XGB). Hyperparameters and energy consumption data with anomalies determined through limited empirical and discounted experimental data can lead to a poor LSTM prediction performance. Salah et al. [72] chose two evolutionary metaheuristics (genetic algorithm [GA] and particle swarm optimization [PSO]) to optimize the performance of the LSTM model for power load prediction, and demonstrated that the model significantly outperformed SVR, RF, ANN, and manually tuned parametric LSTM.

ANN (Artificial Neural Network)

ANNs have performed well in various complex and difficult tasks with high temporal resolutions in the prediction of short-term heating loads in buildings [31]. Zhu et al. [73] used an improved differential evolutionary algorithm to optimize the hyperparameters (initial weights and thresholds) of an ANN to forecast the energy consumption of an HVAC. Yaser et al. [74] used an ANN to predict the daily energy consumption of a laboratory fan coil system. Byeongmo et al. [75] proposed an ANN-based control method for a DSF office

building in a humid-heat environment with 4.5% cost savings. Muralitharan et al. [76] used a GA and PSO to optimize an ANN that automatically adjusted the hyperparameters.

Other special ANNs include MLP, BP neural network, Elman neural network, and echo state network (ESN). Andrew et al. [77] used an MLP to optimize the energy consumption of an HVAC while maintaining the thermal comfort of a building under uncertain occupancy levels. Mitali et al. [78] predicted the HVAC energy consumption of residential buildings using BP. Ruiz et al. [79] established a method based on the Elman neural network to forecast the energy consumption of buildings at the University of Granada to improve their energy use efficiency without compromising on comfort and health. Shi et al. [36] used the ESN to predict office energy consumption.

SVM and SVR (Support Vector Machine and Support Vector Regression)

In the context of BECP, SVM (or SVR) is based on nonlinear mapping that maps input data to a high-dimensional space for linear regression, and finally obtains the effect of nonlinear regression on the original input space [80]. Zhong et al. [81] used SVR to predict the cooling load of a large office building in Tianjin. Li et al. [26] developed a time-by-time building cooling load prediction model based on SVM and BP, and applied it to the cooling load prediction of an office building in Guangzhou. The prediction results revealed that SVM was more accurate than BP. Ding et al. [80] developed GA-SVR and GA-WD-SVR models to predict the cooling load of an office building at different time scales and found that GA-SVR predicted the next-day cooling load better, whereas GA-WD-SVR was better at predicting the 1-h cooling load. Paudel et al. [82] used an SVM to predict the thermal loads of a low-energy building and found that the method using relevant data as input has a better accuracy (root mean square error [RMSE] = 3.4) than the full data modeling method (RMSE = 7.1). Seyedzadeh et al. [83] compared SVM, RF, RNN, XGB, and other algorithms for predicting the cooling and heating loads of commercial and residential buildings, and reported that SVM is the best choice for relatively simple data. To refine the effects of outdoor meteorological parameters on cooling and heating loads, Zhao and Liu [84] first performed wavelet transform (WT) noise reduction on historical energy consumption data, and then used features with low correlation with loads for partial least squares (PLS) prediction and features with high correlation for SVM prediction to significantly improve the prediction accuracy. Ngo et al. [85] developed a novel time series wolf-inspired optimization SVR model (WIO-SVR) to predict cross-building energy consumption.

RF (Regression Trees)

RF is an integration predictive model comprising several regression trees, and it uses them to train and predict the samples. It can automatically perform feature selection to determine the interaction between different variables without feature selection, and can still maintain high prediction accuracy in the case of missing features. Wang et al. [86] used the RF model to forecast the hourly electricity consumption of two teaching buildings in Florida. Seyedzadeha et al. [87] used RF to predict the cooling and heating loads of multiple residential and commercial buildings. Rana et al. [88] predicted the one-month cooling load of a large retail shopping center and an office building in Australia based on a divided number regression forest. Ahmad et al. [89] used a binomial decision tree, tight regression Gaussian process model, stepwise Gaussian process regression, and generalized linear regression models to predict monthly, quarterly, and annual electricity consumption.

XGB (Extreme Gradient Boosting)

XGB can handle nonlinear relationships well without considerable adjustment, and is a gradient-boosting decision tree designed for speed and performance [29]. João et al. [48] proposed a hyperparametric adaptive XGB model based on the Jaya algorithm to forecast the energy consumption of residential buildings. Lu et al. [90] used XGB to fore-

cast the energy consumption of a water tower because of its ability to optimize the prediction by smoothing raw data with large fluctuations. Feng et al. [91] used XGB to predict the cooling loads of three houses in the United States in hot, humid, cold, and dry climates.

MLR (Multiple Linear Regression)

MLR models establish a linear relationship between building energy and input data to predict energy consumption [52]. Nelson and Biswas [92] used linear regression to predict residential 1 h versus one-day HVAC energy consumption and found that the quadratic regression model provided better results at shorter time scales (1 h) and not at longer ones (1 d). Fan and Ding [93] predicted the hourly cold load of a large library using a simplified multiple nonlinear regression (MNR) model. Chen et al. [94] developed a PB-MLR model to predict the time-wise cold load of office buildings in response to the problem of weak generalization of prediction models trained on small samples.

Other Machine Learning Algorithms

A few emerging or still developing algorithms have also been proposed in addition to the aforementioned mainstream algorithms. For example, Munkhammar et al. [95] used a Markov chain mixed distribution model for short-term forecasting of residential electricity load. Gonzaga et al. [96] used nonlinear autoregressive (NAR) and nonlinear autoregressive neural networks with exogenous inputs (NARX) to forecast the future two-year energy consumption of public buildings. Lyes et al. [97] proposed an ML model to predict household electricity consumption using a smooth wavelet transform and transformer-based model.

2.2.2. Data Preprocessing

ML algorithms commonly used for data preprocessing include the following: normalization, Monte Carlo method (MCM), sliding window, wavelet decomposition (WD), wavelet transform (WT), clustering algorithms, and generative adversarial networks (GAN). Kim et al. [38] normalized the data, removed missing values, and used a NARX model with diverse hyperparameters to forecast building cooling load. The results revealed higher prediction accuracy compared with no data preprocessing. Zhao et al. [98] used MCM to preprocess the data, and the predicted results of the cold load were closer to the actual values compared with those obtained without preprocessing. Fan et al. [51] performed an MCM simulation for offline calibration and stochastic processing of input variables, selected significant variables for predicting the cooling load, and used SVM to predict the cooling load of a library in Guangzhou. They found that the uncertainty of all data were reduced to different degrees after correction. Zhao and Liu [84] segmented office building cooling load data into high and low frequencies based on WT to distinguish weekday and non-weekday loads. Tian et al. [99] used density-based spatial clustering noise for the clustering of electricity demand to classify buildings into different types to facilitate energy consumption prediction. To address the problem of insufficient raw data, Tian et al. [100] used GAN to generate artificial data supplements and found after comparing the overall residential energy consumption predictions that the accuracy of data-driven predictions based on hybrid data was better than that based on pure historical data.

2.2.3. Feature Extraction

Commonly used ML algorithms for feature extraction include the residential PCA, K-means clustering algorithm, PCC, Spearman correlation coefficient (SCC), and Taguchi method (TM). Li et al. [55] analyzed the impact of meteorological parameters on the electrical energy of an academic building using PCA. Guo et al. [52] determined the input parameters for the daily average cold load prediction of an office building using PCA. Khan et al. [101] used the K-means algorithm to extract all typical load profiles to infer information related to the year-round operational behavior of the building. Sha et al. [54]

calculated the PCC between meteorological parameters and HVAC system power consumption, and determined that the dry-bulb temperature was the most relevant factor for energy variation in HVAC systems. Huang and Li [56] used SCC to analyze building construction parameters and explore the main factors affecting the prediction of thermal and cooling loads in residential buildings. Sholahudin and Han [102] used TM to study the most significant meteorological parameters affecting the prediction of building heat loads. Tesfaye and Matti [103] used a binary GA and Gaussian process regression to extract the input features of the prediction model. Wang et al. [104] used ResNet to extract the complex and significant features of building load sequences.

2.2.4. Hyperparameter Optimization

ML algorithms used for hyperparameter search include the following: GA, PSO, partial least squares (PLS), Bayesian optimization (BO), K-fold cross-validation, sine cosine optimization algorithm (SCOA), gray wolf heuristic algorithm (GWO), taught optimization approach (TLBO), imperialist competitive algorithm (ICA), and ant colony search algorithm (ACO). Luo [105] used GA to determine the optimal structure of each deep neural network sub-model as a suitable feature dataset to accurately predict the energy consumption for outdoor weather conditions in different seasons. Chen [94] used PSO to optimize PB-MLR for office building cooling load prediction. Li et al. [106] used PLS to optimize the weights of functionally weighted single-input rule modules to connect the fuzzy inference systems prediction model. He and Tsang [66] used BO for the automatic optimization of hyperparameters in LSTM during prediction. Nivethitha et al. [107] used SCOA to optimize the learning rate, weight decay, momentum, and number of hidden layers of KCNN-LSTM for a better prediction of building energy consumption. Dan et al. [108] used an elite GA (EGA) to optimize LSTM to predict the electricity consumption in office and commercial buildings. Ding et al. [80] used GA and K-fold cross-validation algorithms to optimize the SVR for one-day and 1 h predictions of office building cooling loads. Seyedzadeh et al. [83] used K-fold cross-validation to test the combination of model hyperparameters. Nikhil and Ahn [109] proposed the shuffled frog-leaping algorithm (SFLA)-optimized regression tree integration to predict HVAC system energy consumption. Chitsaz et al. [110] trained an autoregressive wavelet neural network using the Levenberg-Marquardt (LM) algorithm to forecast electricity consumption in schools. Huang and Li [56] used ACO to optimize the predictive capability of a wavelet neural network for cooling and heating loads in residential buildings. Li et al. [55] used a new hybrid method, TLBO-ANN, to optimize neural network parameters to enhance the 1 h prediction accuracy of ANN for building energy. Hyperparameter optimization not only improves the accuracy of prediction, but also provides the possibility of prediction across time scales, building levels, and energy consumption types.

In the 116 papers, classified based on energy consumption feature data, Figure 3 illustrates that meteorological data, BESD, indoor environmental parameters, and building construction parameters accounted for 42%, 39%, 12%, and 7% studies, respectively. In studies based on the ML algorithms for predicting energy consumption, LSTM and ANN accounted for 16% and 10%, respectively, with XGB, SVR, and MLR each used in 6% of the studies.

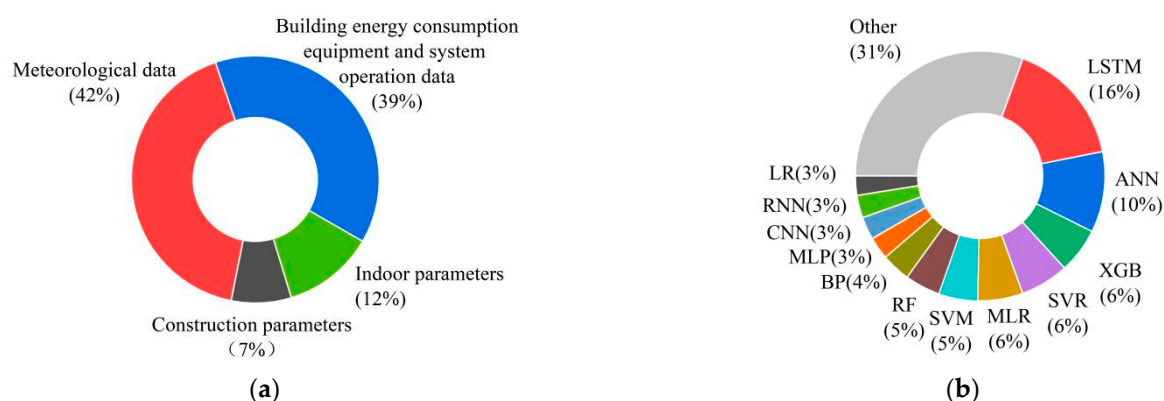


Figure 3. Distribution of the reviewed papers according to (a) type of the energy consumption feature data; and (b) type of ML algorithm used for BECP.

As illustrated in Figure 4, according to the building type (Figure 4a), office buildings, residential buildings, academic buildings, commercial buildings, and retail stores accounted for 29%, 21%, 16%, 6%, and 4%, respectively, with libraries, laboratories, and hospitals accounting for 3% each. Among the different components influencing building energy consumption (Figure 4b), overall electrical energy contributes 41%, HVAC 22%, refrigeration 21%, heating 12%, and lighting and outlets 5%. According to the time scale of prediction (Figure 4c), the one-day scale was used in the maximum studies, accounting for 34%, followed by 28% for 1 h, and other time scales such as one month and one week accounting for less than 5%.

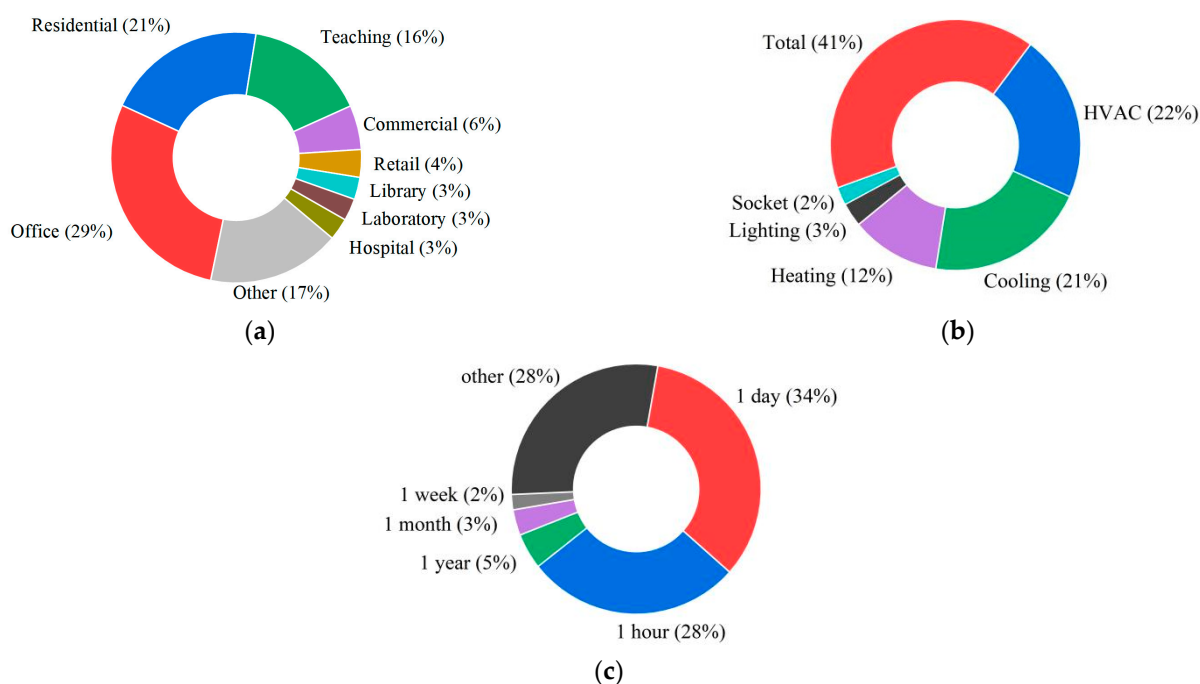


Figure 4. Distribution of the reviewed papers according to (a) type of building; (b) type of building energy consumption; and (c) type of prediction time scale.

3. Factors Affecting Data-Driven Building Energy Consumption Prediction

This review summarized the factors affecting data-driven BECP based on the variability of accuracy metrics. The accuracy metrics were divided into: RMSE, coefficient of variation (CV), mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (MSE), and coefficient of determination (R^2).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_{1,i} - y_{2,i})^2}{n}} \quad (1)$$

$$\text{CV}(\%) = \frac{\text{RMSE}}{\bar{y}_2} \times 100 \quad (2)$$

$$\text{MAPE}(\%) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{1,i} - y_{2,i}}{y_{2,i}} \right| \times 100 \quad (3)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_{1,i} - y_{2,i}| \quad (4)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_{1,i} - y_{2,i})^2 \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{1,i} - y_{2,i})^2}{\sum_{i=1}^n (y_{1,i} - \bar{y}_2)^2} \quad (6)$$

Here $y_{1,i}$ is the i th predicted data, $y_{2,i}$ is the i th measured data, n is the total number of data, and \bar{y}_2 is the average value of the total measured data.

3.1. Data

3.1.1. Data Preprocessing

In addition to the ML algorithms introduced in Section 2.2.2 for data processing, such as noise reduction, decomposition, and smoothing, data processing methods for anomalies. Quartile spacing rules is used to reject outliers [111]. The interpolation method can be used to supplement missing data [112–115] to ensure that the time series is complete. However, the interpolation method is subject to errors, and certain scholars have chosen to directly eliminate the missing values [65]. In addition, the data have to be normalized for differences in values and magnitudes among different energy consumption features [116,117].

Data preprocessing can reduce the computation time of ML algorithms, improving their prediction efficiency. Kim et al. [38] used NARX to forecast the cold load of offices after normalizing the data and removing missing values; experimental results revealed that although the CV reached a maximum of 27.6% before removing the missing values, it decreased to 11.1% after the removal. Zhao et al. [98] used SVM to predict the cold load of an office building for 24 h after preprocessing the meteorological parameters using MCM, and obtained improved accuracy compared with the unprocessed case, with the MAPE reducing from 11.54% to 10.92% (Table 2). The GA-SVR model proposed by Ding et al. [80] successfully predicted the next one-day cold load. An improved model (GA-WD-SVR) based on GA-SVR was combined with wavelet operators to decompose the historical load data to obtain a multiband load signal, and predicted the cold load for the next 1 h with greater efficiency, demonstrating that data preprocessing affects the prediction on different time scales. Chou et al. [118] developed LSTM models based on ensemble empirical mode decomposition (EEMD) and WT with data noise reduction, respectively, and found that EEMD-LSTM had the lowest MAPE mean (7.6%), compared with LSTM and WT-LSTM, in predicting the energy consumption of 20 buildings (industrial, education, commercial, government, and residential); further, SPSS statistical tests revealed that EEMD-LSTM is significantly different compared with the other two models, whereas LSTM and WT-LSTM are similar to each other.

Table 2. Accuracy comparison of different data preprocessing methods.

| Reference | Building Type | Prediction Algorithm | Time Scale | Data Preprocessing Method | Metric |
|-----------|---------------|----------------------|------------|---------------------------|---------------|
| [98] | Office | SVR | 24 h | MCM | 10.92% (MAPE) |
| | | | | - | 11.54% |
| [80] | Office | GA-SVR | 1 day | - | 11.77 (RMSE) |
| | | | | WD | 12.06 |
| | | | 1 h | - | 10.80 |
| | | | | WD | 8.18 |

MAPE: mean absolute percentage error; RMSE: root mean square error.

3.1.2. Feature Extraction

Excessive data input can lead to long training times and reduced model efficiency [119]. Appropriate data selection (energy consumption feature extraction) is key to the accurate prediction of building energy consumption, not only reducing the prediction time but also ensuring the consistency of the prediction accuracy [120]. Table 3 presents the effective improvement in the accuracy of the prediction models after feature extraction and the prediction results achieved using different feature extraction methods. While predicting the daily average cooling load of two office buildings in Tianjin, the use of PCA-extracted features resulted in a MAPE of less than 8.0%. Huang and Li [56] used SCC to analyze the eight most influencing factors in building construction and found that roof area and overall height had the greatest influence on the heating and cooling loads of residential buildings. Liang [121] found that although outdoor air temperature is a vital data source for predicting HVAC energy consumption, occupancy has a greater effect on the total building energy consumption than HVAC energy consumption because of its direct effect on plug and lighting loads.

Outdoor dry-bulb temperature and direct radiation are closely related to office cooling loads [53]. The outdoor daily maximum, average, minimum, and dew-point temperatures are the most important features influencing the overall energy consumption of office buildings [122]. Relative humidity ratio, wind speed, solar radiation, and dry-bulb temperature are the most relevant factors affecting the variation in HVAC energy consumption [54]. The outdoor air dry-bulb temperature and intensity of solar radiation have a major influence on the electrical energy of an academic building [55]. Outdoor dry-bulb temperature and wind speed have great effects on dwellings than dew-point temperature, direct normal radiation, and diffuse horizontal radiation [102]. It can be observed in Table 4 that outdoor temperature is a vital meteorological parameter affecting building energy consumption, and the prediction accuracy can be effectively improved after feature extraction.

Table 3. Comparative prediction accuracy of different energy consumption feature selection methods.

| Reference | Energy Type | Feature Selection Method | Prediction Algorithm | Metric |
|-----------|-------------|--------------------------|----------------------|-------------|
| [53] | HVAC | IC | ARX | 28.3 (MAE) |
| | | - | ARX | 51.1 |
| | | - | MLR | 91.6 |
| | | - | AR | 112.9 |
| [45] | Total | PCA + CA + WD | SVR | 6.00 (MAPE) |
| | | CA + WD | | 9.50 |
| | | CA + WD | | 9.50 |
| | | CA | | 11.40 |
| [40] | Total | PCC | MLR | 32.9 (CV) |

| | | | |
|------|--|-----|------|
| | | SVR | 21.6 |
| | | ANN | 25.1 |
| | | XGB | 19.3 |
| STAT | | MLR | 34.3 |
| | | SVR | 25.3 |
| | | ANN | 24.9 |
| | | XGB | 19.3 |
| AE | | MLR | 30.2 |
| | | SVR | 20.2 |
| | | ANN | 23.5 |
| | | XGB | 19.0 |
| CAE | | MLR | 30.1 |
| | | SVR | 20.7 |
| | | ANN | 22.5 |
| | | XGB | 18.3 |
| GAN | | MLR | 26.4 |
| | | SVR | 21.3 |
| | | ANN | 20.1 |
| | | XGB | 17.7 |

MAE: mean absolute error; MAPE: mean absolute percentage error; CV: coefficient of variation.

Table 4. Correlation between energy and meteorological parameters.

| Reference | Feature Selection Method | Energy Type | Building Type and Location | Outdoor Temperature | Outdoor Humidity | Wind Speed | Solar Radiation |
|-----------|--------------------------|-----------------|--------------------------------------|---------------------|------------------|------------|-----------------|
| [104] | PCC | Total | Laboratory A | 0.75 | −0.21 | 0.03 | − |
| | | | Laboratory B | 0.69 | −0.09 | 0.05 | − |
| | | | Office | 0.43 | −0.14 | 0.16 | − |
| [54] | Correlation coefficient | Cooling Heating | Retail building | 0.91 | −0.21 | −0.05 | 0.47 |
| | | | | −0.83 | −0.15 | 0.21 | 0.07 |
| [123] | SCC | HVAC | Office | 0.89 | 0.16 | − | − |
| [124] | Influence coefficient | Cooling | Library, Guangzhou, China | 1.84 | 0.351 | −0.017 | 0.291 |
| [125] | PCC | Total | Academic building, Bangkok, Thailand | 0.50 | −0.36 | − | 0.56 |
| | | | Laboratory, Hyderabad, India | 0.57 | −0.35 | − | 0.46 |
| | | | Office, Virginia, USA | 0.74 | −0.26 | −0.02 | − |
| | | | School, New York, USA | 0.51 | −0.07 | −0.04 | − |
| | | | Store, Massachusetts, USA | 0.76 | −0.10 | 0.02 | − |

3.2. Prediction Model

3.2.1. Hyperparameter

ML algorithms for hyperparameter optimization can be divided into empirical methods and grid random search.

Although the empirical method is easily operable by manually setting the hyperparameters, a certain professional knowledge of the algorithm and data is required, limiting the generalization of the model. Grid random search addresses this point well by searching the parameter domain completely or randomly for tuning. However, because checking all combinations is generally impractical [126], an optimization-seeking algorithm is required to for tuning the hyperparameters. Luo [105] used GA to determine the optimal structure of each deep neural network sub-model as a suitable feature dataset for accurate energy consumption prediction based on outdoor weather conditions in different seasons. Dan et al. [108] used EGA to optimize the LSTM, resulting in better accuracy, robustness, and generalization ability compared with those of base models such as LSTM and SVR. As presented in Table 5, Salah et al. [72] found that LSTM, improved by GA and PSO optimization, provided higher-accuracy predictions than the original LSTM. João et al. [48] used the modified Jaya (mJaya) hyperparametric adaptive XGB model to predict residential building energy consumption and demonstrated better prediction accuracy over other algorithms. Nivethitha et al. [127] optimized the learning rate, weight decay, momentum, and number of hidden layers of LSTM using an improved SCOA, and achieved a reduction in the average absolute percentage error of the prediction results from 4.3221 to 3.3159. Muralitharan et al. [76] used GA and PSO to optimize ANN and found that GA-ANN and PSO-ANN have greater suitability for short-term and long-term energy forecasting, respectively. In summary, hyperparameter tuning not only improves the accuracy of forecasting, but also facilitates forecasting across time, buildings, and energy scales.

Table 5. Comparative accuracy of different hyperparameter optimization algorithms.

| Reference | Energy Type | Period | Hyperparameter Optimization Algorithm | Prediction Algorithm | Metric |
|-----------|-------------|--------|---------------------------------------|----------------------|----------------------|
| [72] | Total | 1 day | GA | LSTM | 0.621 (CV) |
| | | | PSO | LSTM | 0.622 |
| | | | - | Multi-sequence LSTM | 0.643 |
| | | | - | RF | 0.805 |
| | | | - | SVR | 0.835 |
| | | | - | Extra TR | 0.889 |
| | | | - | ANN | 1.311 |
| [48] | Heating | - | mJaya | XGB | 0.437 (average RMSE) |
| | | | PSO | | 0.454 |
| | | | Jaya | | 0.473 |
| | | | BHO | | 0.475 |
| | | | GWO | | 0.482 |
| | | | ALO | | 0.485 |
| | | | DA | | 0.517 |
| | | | CS | | 0.523 |
| | | | DE | | 0.525 |
| | | | GA | | 0.607 |
| | Cooling | - | mJaya | XGB | 1.182 (average RMSE) |
| | | | PSO | | 1.191 |
| | | | Jaya | | 1.260 |
| | | | BHO | | 1.226 |

| | | | | | |
|-------|-------|---------|-------|------------|---------------|
| | | | GWO | 1.256 | |
| | | | ALO | 1.253 | |
| | | | DA | 1.229 | |
| | | | CS | 1.352 | |
| | | | DE | 1.371 | |
| | | | GA | 1.446 | |
| [127] | Total | 1 year | ISCOA | LSTM | 3.3159 (MAPE) |
| | | | PSO | LSTM | 4.115 |
| | | | SCA | LSTM | 4.3221 |
| | | | GA | LSTM | 5.3272 |
| | | | DBN | Regression | 10.3215 |
| | | | SV | Regression | 13.107 |
| | | | - | ARINA | 27.8963 |
| [66] | Total | 1 h | BO | LSTM | 4.6458 (MAPE) |
| | | | - | LSTM | 6.6107 |
| | | | - | GRU | 5.5063 |
| | | | - | RNN | 9.9572 |
| | | | - | SARIMA | 12.7770 |
| [76] | Total | 1 day | - | CNN | 0.541 (MSE) |
| | | | PSO | ANN | 0.495 |
| | | | GA | ANN | 0.391 |
| | | 1 month | - | CNN | 0.492 (MSE) |
| | | | PSO | ANN | 0.479 |
| | | | GA | ANN | 0.387 |
| | | 1 year | - | CNN | 0.513 (MSE) |
| | | | PSO | ANN | 0.408 |
| | | | GA | ANN | 0.429 |

CV: coefficient of variation; average RMSE: average root mean square error; MAPE: mean absolute percentage error; MSE: root mean squared error.

3.2.2. Machine Learning Algorithm

The accuracy of ML algorithms that drive the BECP is influenced by the energy consumption type, building type, data capacity, and time scale. In other words, ML algorithms are selective for different energy consumption requirements, building types, data volumes, and time scales. Nikhil [109] proposed an SFLA optimized with regression tree integration, and compared it with stepwise regression (STR) and Gaussian process regression (GPR) based on different kernel functions. The proposal method performed better than STR and GPR in predicting residential cooling, heat load, and HVAC system energy consumption.

Regarding the use of ML algorithms for the selection of building types: Gao et al. [128] used ANN, LSTM, and CNN for hourly prediction of multiple-source energy consumption in multiple building types and found that LSTM outperformed the ANN and CNN models for hospitals and large hotels whereas it underperformed for offices, retail stores, shopping streets, and supermarkets. Across all tasks, better results were observed in the task of predicting total gas and electricity consumption than in predicting segment-wise energy consumption, such as fans, cooling, heating, and other indoor equipment. In addition, electricity-related predictions produced better results than natural gas-related predictions, and LSTM models outperformed CNN models for most building types, with better suitability for such multivariate predictions.

Regarding ML algorithms for the choice of time scale: Fan [124] proposed a multivariate nonlinear regression model (MNR) for fast prediction of cold loads in large public buildings and compared the prediction performance of MLR, AR, ARX, and BP at diverse

time scales, and found that with less training data, simple hardware, and lower computational complexity, MNR can predict cold loads quickly in the short term, as shown in Table 6. The RNN-MIMO-LSTM [71] model performed well in predicting the 1 h and one-day cold loads in office buildings, hospitals, and shopping malls compared with traditional LSTM models, and performed better than shallow learning methods over a longer time horizon. Wang and Hong [129] discussed the effects of forecast period and input features on load forecasting accuracy and algorithms and found that LSTM performs better in 1 h forecasting and exhibits greater stability to input uncertainty, and that the temporal information retained by LSTM helps isolate forecast and inaccurate weather parameters. XGB performs better in 24 h forecasting because the sequence information captured by LSTM may be less relevant and therefore less helpful for 24 h prediction.

Regarding the use of ML algorithms for data selection: Fan et al. [41] compared the prediction performance of recursive methods, direct methods, and MIMO models, and demonstrated that direct methods based on recursive models yield the most accurate prediction results without a significant increase in computational effort, and that both LSTM and GRU can better preserve temporal correlation in long time series compared with traditional recursive units, and that GRU has a shorter computational time than LSTM. As presented in Table 6, Zhang et al. [28] compared the prediction accuracy of the ARX and MLP models using the cooling load data of a single-story factory in Shanghai for two cooling seasons (2017 and 2018) as a dataset, and found that ARX was comparable with and simpler than MLP for large sample data and small variable dimensions. Kim [123] found that changing the number of neurons in an ANN has no significant effect on the prediction accuracy, whereas adding to the number of input variables and adjusting the proportion of training data can effectively improve the prediction accuracy.

In addition, combining several prediction models can offset random errors and achieve more accurate predictions in certain cases [130]. Bedi et al. [131] compared the performance of two developed models (the Elman RNN model and exponential model) for real-time and near-term power consumption estimation and prediction in the laboratory, and reported that the Elman RNN model outperformed the exponential model. Ding et al. [44] proposed an EDA-LSTM model to forecast the energy consumption of a green building for one year, and reported a reduction in the RMSE of the model by 89% compared with that using LSTM. Wang et al. [104] proposed a ResNet-based DCNN model that outperformed ResNet, GRU, and LSTM in BECP for two experimental buildings and an office in Switzerland. Therefore, the prediction performance of integrated algorithms is generally better than that of a single classical algorithm.

Table 6. Comparative accuracy of different ML algorithms.

| Reference | Building | Time Scale | Model | Metric |
|-----------|----------------------------|------------|---------------|--------------|
| [28] | One-story factory workshop | 1 h | MLP | 11.5% (CV) |
| | | | ARX | 11.60% |
| | | 2 h | MLP | 15.4% (CV) |
| | | | ARX | 15.80% |
| | | 3 h | MLP | 18.0% (CV) |
| | | | ARX | 18.90% |
| | | 4 h | MLP | 19.2% (CV) |
| | | | ARX | 20.70% |
| [71] | Market | 1 day | XGB | 24.13 (RMSE) |
| | | | SVR | 29.19 |
| | | | RNN-MIMO | 34.8 |
| | | | S2S | 36.66 |
| | | | RNN-Recursive | 37.45 |
| | | | Naive-LE | 320.02 |

| | | | |
|----------|-------|---------------|--------------|
| Office | 1 day | XGB | 15.38 (RMSE) |
| | | SVR | 16.47 |
| | | RNN-Recursive | 27.65 |
| | | RNN-MIMO | 28.52 |
| | | S2S | 28.88 |
| | | Naive-LE | 69.23 |
| Hospital | 1 day | SVR | 31.95 (RMSE) |
| | | S2S | 32.97 |
| | | RNN-Recursive | 34.25 |
| | | RNN-MIMO | 34.4 |
| | | XGB | 37.21 |
| | | Naive-LE | 203.11 |

CV: coefficient of variation; RMSE: root mean square error.

4. Discussion and Outlook

Energy consumption features are closely related to climatic conditions, building energy consumption types, building functions, structures, etc. When selecting energy consumption features, multiple covariance between features (i.e., weak correlation) should be avoided, and a strong correlation between features and energy consumption should be ensured to improve the prediction performance. In Figure 3, indoor environmental parameters account for only 12%; on the one hand, because data-driven BECP is a dynamic process and indoor environmental factors (indoor temperature, humidity, light intensity, etc.) are basically in a relatively stable state, they do not significantly influence the energy consumption compared with meteorological parameters and BESD. However, on the other hand, individual indoor parameters (e.g., indoor occupancy) are not precise and directly collectible, and they need to be mined, for example, by using temporal index features [46] or personnel heat disturbance [132]. However, the two methods mentioned above are insufficient to represent the real situation, and the building occupancy rate may vary across buildings; therefore, the actual occupancy rate should be considered according to the building function. Second, the building construction parameters account for 7% of the total factors. Because building construction parameters are essentially constant and less relevant to short-term energy consumption forecasting than meteorological data and BESD are, this results in the low use of current hour-wise and day-wise forecasting. However, building construction parameters are potentially key energy consumption features in cross-building level forecasting. In addition, if sufficient data are available from different buildings, cross-building layer prediction can be achieved using feature extraction [64] and by finding the optimal hyperparameters in the training and testing datasets [85].

Among the total classified building types in this study (Figure 3), office buildings, residential buildings, academic buildings, commercial buildings, and retail stores accounted for 29%, 21%, 16%, 6%, and 4%, respectively, with libraries, laboratories, and hospitals accounting for 3% each. The building is more functional, and its energy composition and data type are more complex (e.g., hospitals [133]). Direct training of prediction models would lead to large prediction errors and more time costs, so building energy consumption data needs to be clustered, noise reduced, or classified. As discussed in Section 3.1.1, suitable preprocessing of energy consumption data can improve prediction performance and even facilitate cross-time-scale prediction. When processing collected raw data, considering the most suitable preprocessing method for the type of energy consumption predicted, prediction time scale, and type of algorithm for prediction is necessary.

Based on the statistics of the algorithm driving the prediction, LSTM, ANN, XGB, SVR, MLR, SVM, RF, and other algorithms have been widely applied to BECP. To date, no one algorithm can achieve superlative prediction performance because of being limited by the type of data, time scale of prediction, and hyperparameters. Choosing the right forecasting model is not simple, and highly complex models are not always the best [134].

For example, the most widely used LSTM relies on a large amount of training data to determine the hyperparameters, resulting in a prediction performance that depends on the effects of the training set. The data-driven model is suitable only when the prediction conditions are approximately within the range of the training data. Declining equipment or building performance, changes in residents or outdoor environment, and other reasons, the running performance of the building energy system will change accordingly. The data-driven model is not applicable, and the model accuracy decreases with an increase in the difference between the test and training data [125]. Therefore, the model needs to be retrained and updated to enhance its accuracy [119]. Although the prediction accuracy of classical algorithms can be improved by integrating multiple algorithms, the time cost of prediction will change accordingly; therefore, accuracy should not be overemphasized by ignoring speed. To ensure a balanced improvement in the prediction performance, the models proposed in future studies need to consider both accuracy and speed.

According to the forecast time-scale statistics in the reviewed studies, one-day and 1 h predictions accounted for 34% and 28%, respectively, whereas long-term predictions (e.g., one week, one month, one year) were used in less than 5% of studies. The energy consumption prediction for a single building is mainly used for fast response control of the energy system and improving the energy efficiency. In addition, the short-term forecast is in line with actual needs. For example, short-term heating load prediction is conducive to the optimization of HVAC operation, whereas ultra-short-term heating load prediction is conducive to the detection of large load fluctuations [45]. In addition, long-term forecasts may encounter cross-season conditions, and both the structure and system operation of building energy consumption vary, leading to large fluctuations in data and insignificant characteristics of energy consumption, resulting in inaccurate predictions. Luo [105] found noticeably different outdoor climate conditions in different seasons and proposed seasonal modeling, which produced better prediction results than for continuous prediction throughout the year. In addition, different prediction models can be used for different time periods to distinguish between the prediction of hot and cold loads.

5. Conclusions

In this study, we reviewed 116 research papers on data-driven prediction of building energy consumption from the perspectives of data and ML, analyzed the factors affecting data-driven building energy consumption prediction, and summarized and discussed key techniques for prediction across time scales, building levels, and energy consumption types. The overview results revealed that meteorological data are the most used feature set (up to 42%) in building energy consumption prediction compared with building energy consumption equipment and system operation data, indoor parameters, and building construction parameters, while outdoor dry-bulb temperature is the most important parameter affecting building energy consumption. Among the three segments of data preprocessing, selection, and training, it was found that data preprocessing and energy consumption feature extraction can improve the prediction performance of the model, whereas ML algorithms can improve the prediction accuracy of the model through multiple algorithm integration and hyperparameter optimization. In addition, data preprocessing can achieve cross-time-scale prediction, energy consumption feature extraction can achieve cross-energy consumption type prediction, and hyperparameter search can achieve cross-time-scale and cross-building layer prediction.

Author Contributions: Conceptualization, H.L. and J.L.; methodology, H.L.; validation, Y.L.; formal analysis, Y.L. and H.W.; writing—original draft preparation, H.L. and J.L.; writing—review and editing, Y.L. and H.W.; supervision, H.W.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [the Science and Technology Program of Guangzhou] grant number [202002030267], [the National Natural Science Foundation of China] grant number

[52108074], and [the Guangdong Basic and Applied Basic Research Foundation] grant number [2020A1515111023].

Data Availability Statement: Data openly available in a public repository.

Acknowledgments: This study is supported by [the Science and Technology Program of Guangzhou] grant number [202002030267], [the National Natural Science Foundation of China] grant number [52108074], and [the Guangdong Basic and Applied Basic Research Foundation] grant number [2020A1515111023].

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

References

1. Global Alliance for Buildings and Construction. *2020 Global Status Report for Buildings and Construction*; Global Alliance for Buildings and Construction: Paris, France: 2020.
2. United Nations Environment Programme (UNEP). *Buildings Can Play Key Role in Combating Climate Change*; UNEP News Centre: Nairobi, Kenya, 2007.
3. Subramanyam, V.; Ahiduzzaman, M.; Kumar, A. Greenhouse gas emissions mitigation potential in the commercial and institutional sector. *Energy Build.* **2017**, *140*, 295–304.
4. Outlook, E. *Global Energy Statistical Yearbook*; United Nations Publications: New York, NY, USA, 2019. Available online: <https://desapublications.un.org/publications/energy-statistics-yearbook-2019> (accessed on 18 January 2023).
5. Xu, G.; Wang, W. China's energy consumption in construction and building sectors: An outlook to 2100. *Energy* **2020**, *195*, 117045.
6. Song, K.; Kwon, N.; Anderson, K.; Park, M.; Lee, H.; Lee, S. Predicting hourly energy consumption in buildings using occupancy-related characteristics of end-user groups. *Energy Build.* **2017**, *156*, 121–133.
7. Elnour, M.; Himeur, Y.; Fadli, F.; Mohammedsherif, F.; Meskin, N.; Ahmad, M.A.; Petri, L.; Rezgui, Y.; Hodorog, A. Neural network-based model predictive control system for optimizing building automation and management systems of sports facilities. *Appl. Energy* **2022**, *318*, 119153.
8. Cholewa, T.; Siuta-Olcha, A.; Smolarz, A.; Muryjas, P.; Wolszczak, P.; Guz, U.; Bocian, M.; Balaras, C.A. An easy and widely applicable forecast control for heating systems in existing and new buildings: First field experiences. *J. Clean. Prod.* **2022**, *352*, 131605.
9. Kim, D.D.; Suh, H.S. Heating and cooling energy consumption prediction model for high-rise apartment buildings considering design parameters. *Energy Sustain. Dev.* **2021**, *61*, 1–14.
10. Seo, J.; Kim, S.; Lee, S.; Jeong, H.; Kim, T.; Kim, J. Data-driven approach to predicting the energy performance of residential buildings using minimal input data. *Build. Environ.* **2022**, *214*, 108911.
11. Dawood, N. Short-term prediction of energy consumption in demand response for blocks of buildings: DR-BoB approach. *Buildings* **2019**, *9*, 221.
12. Kim, J.; Kim, K.-I. Data-driven hybrid model and operating algorithm to shave peak demand costs of building electricity. *Energy Build.* **2020**, *229*, 110493.
13. Jota, P.R.; Silva, V.R.; Jota, F.G. Building load management using cluster and statistical analyses. *Int. J. Electr. Power Energy Syst.* **2011**, *33*, 1498–1505.
14. Atalay, S.D.; Calis, G.; Kus, G.; Kuru, M. Performance analyses of statistical approaches for modeling electricity consumption of a commercial building in France. *Energy Build.* **2019**, *195*, 82–92.
15. Risbeck, M.J.; Bazant, M.Z.; Jiang, Z.; Lee, Y.M.; Douglas, J.D. Modeling and multiobjective optimization of indoor airborne disease transmission risk and associated energy consumption for building HVAC systems. *Energy Build.* **2021**, *253*, 111497.
16. Li, Z.; Ye, H.; Liao, N.; Wang, R.; Qiu, Y.; Wang, Y. Impact of COVID-19 on electricity energy consumption: A quantitative analysis on electricity. *Int. J. Electr. Power Energy Syst.* **2022**, *140*, 108084.
17. Fan, C.; Xiao, F.; Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. *Appl. Energy* **2017**, *195*, 222–233.
18. Mohammed, A.; Alshibani, A.; Alshamrani, O.; Hassanain, M. A regression-based model for estimating the energy consumption of school facilities in Saudi Arabia. *Energy Build.* **2021**, *237*, 110809.
19. Sala-Cardoso, E.; Delgado-Prieto, M.; Kampouropoulos, K.; Romeral, L. Activity-aware HVAC power demand forecasting. *Energy Build.* **2018**, *170*, 15–24.
20. Yan, D.; Xia, J.; Tang, W.; Song, F.; Zhang, X.; Jiang, Y. DeST—An integrated building simulation toolkit Part I: Fundamentals. *Building Simul.* **2008**, *1*, 95–110.
21. Wei, Y.; Zhang, X.; Shi, Y.; Xia, L.; Pan, S.; Wu, J.; Han, M.; Zhao, X. A review of data-driven approaches for prediction and classification of building energy consumption. *Renew. Sustain. Energy Rev.* **2018**, *82*, 1027–1047.
22. Himeur, Y.; Alsalemi, A.; Bensali, F.; Amira, A. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Appl. Energy* **2021**, *287*, 116601.
23. Afram, A.; Janabi-Sharifi, F. Review of modeling methods for HVAC systems. *Appl. Therm. Eng.* **2014**, *67*, 507–519.

24. Chen, Y.; Chen, Y.; Guo, M.; Chen, Z.; Chen, Z.; Ji, Y. Physical energy and data-driven models in building energy prediction: A review. *Energy Rep.* **2022**, *8*, 2656–2671.
25. Wang, Z.; Ravi, S. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renew. Sustain. Energy Rev.* **2017**, *75*, 796–808.
26. Li, Q.; Meng, Q.; Cai, J.; Yoshino, H.; Mochida, A. Applying support vector machine to predict hourly cooling load in the building. *Appl. Energy* **2009**, *86*, 2249–2256.
27. Ding, Y.; Zhang, Q.; Yuan, T.; Yang, F. Effect of input variables on cooling load prediction accuracy of an office building. *Appl. Therm. Eng.* **2018**, *128*, 225–234.
28. Zhang, Q.; Tian, Z.; Ding, Y.; Lu, Y.; Niu, J. Development and evaluation of cooling load prediction models for a factory workshop. *J. Clean. Prod.* **2019**, *230*, 622–633.
29. Kamel, E.; Sheikh, S.; Huang, X. Data-driven predictive models for residential building energy use based on the segregation of heating and cooling days. *Energy* **2020**, *206*, 118045.
30. Gao, Y.; Ruan, Y.; Fang, C.; Yin, S. Deep learning and transfer learning models of energy consumption forecasting for a building with poor information data. *Energy Build.* **2020**, *223*, 110156.
31. Bünnig, F.; Heer, P.; Smith, R.S.; Lygeros, J. Improved day ahead heating demand forecasting by online correction methods. *Energy Build.* **2020**, *211*, 109821.
32. Yang, W.; Shi, J.; Li, S.; Song, Z.; Zhang, Z.; Chen, Z. A combined deep learning load forecasting model of single household resident user considering multi-time scale electricity consumption behavior. *Appl. Energy* **2022**, *307*, 118197.
33. Ahmad, T.; Chen, H. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy Build.* **2018**, *166*, 460–476.
34. Sun, S.; Chen, H. Data-driven sensitivity analysis and electricity consumption prediction for water source heat pump system using limited information. In *Building Simulation*; Tsinghua University Press: Beijing, China, 2021; Volume 14.
35. Kim, C.H.; Kim, M.; Song, Y.J. Sequence-to-sequence deep learning model for building energy consumption prediction with dynamic simulation modeling. *J. Build. Eng.* **2021**, *43*, 102577.
36. Shi, G.; Liu, D.; Wei, Q. Energy consumption prediction of office buildings based on echo state networks. *Neurocomputing* **2016**, *216*, 478–488.
37. Fekri, M.N.; Patel, H.; Grolinger, K.; Sharma, V. Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network. *Appl. Energy* **2021**, *282*, 116177.
38. Kim, J.-H.; Seong, N.-C.; Choi, W. Cooling load forecasting via predictive optimization of a nonlinear autoregressive exogenous (NARX) neural network model. *Sustainability* **2019**, *11*, 6535.
39. Chaerun, N.E.; Yean-Der, K. Comparative Assessment to Predict and Forecast Water-Cooled Chiller Power Consumption Using Machine Learning and Deep Learning Algorithms. *Sustainability* **2021**, *13*, 744.
40. Fan, C.; Sun, Y.; Zhao, Y.; Song, M.; Wang, J. Deep learning-based feature engineering methods for improved building energy prediction. *Appl. Energy* **2019**, *240*, 35–45.
41. Fan, C.; Wang, J.; Gang, W.; Li, S. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Appl. Energy* **2019**, *236*, 700–710.
42. Jeong, D.; Park, C.; Ko, Y.M. Short-term electric load forecasting for buildings using logistic mixture vector autoregressive model with curve registration. *Appl. Energy* **2021**, *282*, 116249.
43. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy Build.* **2017**, *147*, 77–89.
44. Ding, Z.; Chen, W.; Hu, T.; Xu, X. Evolutionary double attention-based long short-term memory model for building energy prediction: Case study of a green building. *Appl. Energy* **2021**, *288*, 116660.
45. Ding, Y.; Zhang, Q.; Yuan, T.; Yang, K. Model input selection for building heating load prediction: A case study for an office building in Tianjin. *Energy Build.* **2018**, *159*, 254–270.
46. Sha, H.; Xu, P.; Yan, C.; Ji, Y.; Zhou, K.; Chen, F. Development of a key-variable-based parallel HVAC energy predictive model. *Build. Simul.* **2022**, *15*, 1193–1208.
47. Li, X.; Yao, R. A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour. *Energy* **2020**, *212*, 118676.
48. Sauer, J.; Mariani, V.C.; Coelho, L.S. Extreme gradient boosting model based on improved Jaya optimizer applied to forecasting energy consumption in residential buildings. *Evol. Syst.* **2022**, *13*, 577–588.
49. Lam, J.C.; Wan, K.; Liu, D.; Tsang, C.L. Multiple regression models for energy use in air-conditioned office buildings in different climates. *Energy Convers. Manag.* **2010**, *51*, 2692–2697.
50. Lin, Q.; Liu, K.; Hong, B.; Xu, X.; Chen, J.; Wang, W. A data-driven framework for abnormally high building energy demand detection with weather and block morphology at community scale. *J. Clean. Prod.* **2022**, *354*, 131602.
51. Fan, C.; Liao, Y.; Zhou, G.; Zhou, X.; Ding, Y. Improving cooling load prediction reliability for HVAC system using Monte-Carlo simulation to deal with uncertainties in input variables. *Energy Build.* **2020**, *226*, 110372.
52. Qiang, G.; Zhe, T.; Yan, D.; Neng, Z. An improved office building cooling load prediction model based on multivariable linear regression. *Energy Build.* **2015**, *107*, 445–455.
53. Fan, C.; Liao, Y.; Ding, Y. Development of a cooling load prediction model for air-conditioning system control of office buildings. *Int. J. Low-Carbon Technol.* **2019**, *14*, 70–75.

54. Sha, H.; Xu, P.; Hu, C.; Li, Z.; Chen, Y.; Chen, Z. A simplified HVAC energy prediction method based on degree-day. *Sustain. Cities Soc.* **2019**, *51*, 101698.
55. Li, K.; Xie, X.; Xue, W.; Dai, X.; Chen, X.; Yang, X. A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction. *Energy Build.* **2018**, *174*, 323–334.
56. Huang, Y.; Li, C. Accurate heating, ventilation and air conditioning system load prediction for residential buildings using improved ant colony optimization and wavelet neural network. *J. Build. Eng.* **2021**, *35*, 101972.
57. Sendra-Arranz, R.; Gutiérrez, A. A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy Build.* **2020**, *216*, 109952.
58. Mendoza-Pitti, L.; Calderón-Gómez, H.; Gómez-Pulido, J.M.; Vargas-Lombardo, M.; Castillo-Sequera, J.L.; Blas, C.S. Developing a Long Short-Term Memory-Based Model for Forecasting the Daily Energy Consumption of Heating, Ventilation, and Air Conditioning Systems in Buildings. *Appl. Sci.* **2021**, *11*, 6722.
59. Alden, R.E.; Gong, H.; Jones, E.; Ababei, C.; Ionel, D.M. Artificial intelligence method for the forecast and separation of total and hvac loads with application to energy management of smart and nze homes. *IEEE Access* **2021**, *9*, 160497–160509.
60. Wang, Z.; Hong, T.; Piette, M.A. Data fusion in predicting internal heat gains for office buildings through a deep learning approach. *Appl. Energy* **2019**, *240*, 386–398.
61. Jogunola, O.; Adebisi, B.; Hoang, K.V.; Tsado, T.; Popoola, S.I.; Hammoudeh, M.; Nawaz, R. CBLSTM-AE: A Hybrid Deep Learning Framework for Predicting Energy Consumption. *Energies* **2022**, *15*, 810.
62. Das, A.; Annaqeeb, M.K.; Azar, E.; Novakovic, V.; Kjrgaard, M.B. Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Appl. Energy* **2020**, *269*, 115135.
63. Ullah, F.; Khan, N.; Hussain, T.; Lee, M.; Baik, S. Diving Deep into Short-Term Electricity Load Forecasting: Comparative Analysis and a Novel Framework. *Mathematics* **2021**, *9*, 611.
64. Li, W.; Gong, G.; Peng, P.; Liang, C.; Fan, H. A clustering-based approach for cross-scale load prediction on building level in HVAC systems. *Appl. Energy* **2021**, *282*, 116223.
65. Kim, T.-Y.; Cho, S.-B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72–81.
66. He, Y.; Tsang, K.F. Universities power energy management: A novel hybrid model based on iCEEMDAN and Bayesian optimized LSTM. *Energy Rep.* **2021**, *7*, 6473–6488.
67. Haq, I.; Ullah, A.; Khan, S.U.; Khan, N.; Baik, S.W. Sequential learning-based energy consumption prediction model for residential and commercial sectors. *Mathematics* **2021**, *9*, 605.
68. He, N.; Liu, L.; Qian, C.; Zhang, L.; Yang, Z.; Li, S. A closed-loop data-fusion framework for air conditioning load prediction based on LBF. *Energy Rep.* **2022**, *8*, 7724–7734.
69. Li, G.; Li, F.; Xu, C.; Fang, X. A spatial-temporal layer-wise relevance propagation method for improving interpretability and prediction accuracy of LSTM building energy prediction. *Energy Build.* **2022**, *271*, 112317.
70. Chakraborty, D.; Alam, A.; Chaudhuri, S.; Baaolu, H.; Langar, S. Scenario-based prediction of climate change impacts on building cooling energy consumption with explainable artificial intelligence. *Appl. Energy* **2021**, *291*, 116807.
71. Chalapathy, R.; Khoa, N.L.D.; Sethuvenkatraman, S. Comparing multi-step ahead building cooling load prediction using shallow machine learning and deep learning models. *Sustain. Energy Grids Netw.* **2021**, *28*, 100543.
72. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting. *Energies* **2020**, *13*, 391.
73. Zhu, Q.; Liu, M.; Liu, H.; Zhu, Y. Application of machine learning and its improvement technology in modeling of total energy consumption of air conditioning water system. *Math. Biosci. Eng.* **2022**, *19*, 4841–4855.
74. Alamin, Y.I.; Lvarez, J.D.; Castilla, M.D.M.; Ruano, A. An Artificial Neural Network (ANN) model to predict the electric load profile for an HVAC system. *IFAC-PapersOnLine* **2018**, *51*, 26–31.
75. Seo, B.; Yoon, Y.B.; Mun, J.H.; Cho, S. Application of artificial neural network for the optimum control of HVAC systems in double-skinned office buildings. *Energies* **2019**, *12*, 4754.
76. Muralitharan, K.; Sakthivel, R.; Vishnuvarthan, R. Neural network based optimization approach for energy demand prediction in smart grid. *Neurocomputing* **2018**, *273*, 199–208.
77. Kusiak, A.; Xu, G.; Zhang, Z. Minimization of energy consumption in HVAC systems with data-driven models and an interior-point method. *Energy Convers. Manag.* **2014**, *85*, 146–153.
78. Ray, M.; Samal, P.; Panigrahi, C.K. Implementation of a Hybrid Technique for the Predictive Control of the Residential Heating Ventilation and Air Conditioning Systems. *Eng. Technol. Appl. Sci. Res.* **2022**, *12*, 8772–8776.
79. Ruiz, L.; Rueda, R.; Cuéllar, M.; Pegalajar, M. Energy consumption forecasting based on Elman neural networks with evolutive optimization. *Expert Syst. Appl.* **2018**, *92*, 380–389.
80. Ding, Y.; Zhang, Q.; Yuan, T. Research on short-term and ultra-short-term cooling load prediction models for office buildings. *Energy Build.* **2017**, *154*, 254–267.
81. Zhong, H.; Wang, J.; Jia, H.; Mu, Y.; Lv, S. Vector field-based support vector regression for building energy consumption prediction. *Appl. Energy* **2019**, *242*, 403–414.
82. Paudel, S.; Elmitri, M.; Couturier, S.; Nguyen, P.N.; Kamphuis, R.; Lacarrière, B.; Corre, O.L. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. *Energy Build.* **2017**, *138*, 240–256.
83. Seyedzadeh, S.; Rahimian, F.P.; Rastogi, P.; Glesk, I. Tuning machine learning models for prediction of building energy loads. *Sustain. Cities Soc.* **2019**, *47*, 101484.

84. Zhao, J.; Liu, X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. *Energy Build.* **2018**, *174*, 293–308.
85. Ngo, N.T.; Truong, T.T.H.; Truong, N.S.; Pham, A.D.; Huynh, N.T.; Pham, T.M.; Pham, V.H.S. Proposing a hybrid metaheuristic optimization algorithm and machine learning model for energy use forecast in non-residential buildings. *Sci. Rep.* **2022**, *12*, 1065.
86. Wang, Z.; Wang, Y.; Zeng, R.; Ravi, S.S.; Sherry, A. Random Forest based hourly building energy prediction. *Energy Build.* **2018**, *171*, 11–25.
87. Seyedzadeh, S.; Rahimian, F.P.; Oliver, S.; Glesk, I.; Kumar, B. Data driven model improved by multi-objective optimisation for prediction of building energy loads. *Autom. Constr.* **2020**, *116*, 103188.
88. Rana, M.; Sethuvenkatraman, S.; Goldsworthy, M. A data-driven approach based on quantile regression forest to forecast cooling load for commercial buildings. *Sustain. Cities Soc.* **2022**, *76*, 103511.
89. Ahmad, T.; Chen, H.; Huang, R.; Guo, Y.; Wang, J.; Shair, J.; Akram, H.M.A.; Mohsan, S.A.H.; Kazim, M. Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment. *Energy* **2018**, *158*, 17–32.
90. Lu, H.; Cheng, F.; Ma, X.; Hu, G. Short-term prediction of building energy consumption employing an improved extreme gradient boosting model: A case study of an intake tower. *Energy* **2020**, *203*, 117756.
91. Feng, Y.; Duan, Q.; Chen, X.; Yakkali, S.S.; Wang, J. Space cooling energy usage prediction based on utility data for residential buildings using machine learning methods. *Appl. Energy* **2021**, *291*, 116814.
92. Fumo, N.; Rafe Biswas, M.A. Regression analysis for prediction of residential energy consumption. *Renew. Sustain. Energy Rev.* **2015**, *47*, 332–343.
93. Fan, C.; Ding, Y. Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model. *Energy Build.* **2019**, *197*, 7–17.
94. Chen, S.; Zhou, X.; Zhou, G.; Fan, C.; Ding, P.; Chen, Q. An online physical-based multiple linear regression model for building's hourly cooling load prediction. *Energy Build.* **2022**, *254*, 111574.
95. Munkhammar, J.; van der Meer, D.; Widén, J. Very short term load forecasting of residential electricity consumption using the Markov-chain mixture distribution (MCM) model. *Appl. Energy* **2021**, *282*, 116180.
96. Ruiz, L.G.B.; Cuéllar, M.P.; Calvo-Flores, M.D.; Jiménez, M.D.C.P. An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. *Energies* **2016**, *9*, 684.
97. Saoud, L.S.; Al-Marzouqi, H.; Hussein, R. Household Energy Consumption Prediction Using the Stationary Wavelet Transform and Transformers. *IEEE Access* **2022**, *10*, 5171–5183. <https://doi.org/10.1109/ACCESS.2022.3140818>.
98. Zhao, J.; Duan, Y.; Liu, X. Uncertainty analysis of weather forecast data for cooling load forecasting based on the Monte Carlo method. *Energies* **2018**, *11*, 1900.
99. Tian, C.; Ye, Y.; Lou, Y.; Zuo, W.; Zhang, G.; Li, C. Daily power demand prediction for buildings at a large scale using a hybrid of physics-based model and generative adversarial network. In *Building Simulation*; Tsinghua University Press: Beijing, China, 2022; Volume 15.
100. Tian, C.; Li, C.; Zhang, G.; Lv, Y. Data driven parallel prediction of building energy consumption using generative adversarial nets. *Energy Build.* **2019**, *186*, 230–243.
101. Khan, W.; Liao, J.Y.; Walker, S.; Zeiler, W. Impact assessment of varied data granularities from commercial buildings on exploration and learning mechanism. *Appl. Energy* **2022**, *319*, 119281.
102. Sholahudin, S.; Han, H. Simplified dynamic neural network model to predict heating load of a building using Taguchi method. *Energy* **2016**, *115*, 1672–1678.
103. Eseye, A.T.; Lehtonen, M. Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models. *IEEE Trans. Ind. Inform.* **2020**, *16*, 7743–7755.
104. Wang, J.; Chen, X.; Zhang, F.; Chen, F.; Xin, Y. Building load forecasting using deep neural network with efficient feature fusion. *J. Mod. Power Syst. Clean Energy* **2021**, *9*, 160–169.
105. Luo, X.; Oyedele, L.O.; Ajayi, A.O.; Akinade, O.O.; Owolabi, H.A.; Ahmed, A. Feature extraction and genetic algorithm enhanced adaptive deep neural network for energy consumption prediction in buildings. *Renew. Sustain. Energy Rev.* **2020**, *131*, 109980.
106. Li, C.; Tang, M.; Zhang, G.; Wang, R.; Tian, C. A hybrid short-term building electrical load forecasting model combining the periodic pattern, fuzzy system, and wavelet transform. *Int. J. Fuzzy Syst.* **2020**, *22*, 156–171.
107. Somu, N.; R, G.R.M.; Ramamritham, K. A deep learning framework for building energy consumption forecast. *Renew. Sustain. Energy Rev.* **2021**, *137*, 110591.
108. Dan, Z.; Wang, B.; Zhang, Q.; Wu, Z.; Fan, H.; Liu, L. Fitting multiple temporal usage patterns in day-ahead hourly building load forecasting under patch learning framework. *Neural Comput. Appl.* **2022**, *34*, 16291–16309.
109. Pachauri, N.; Ahn, C.W. Regression tree ensemble learning-based prediction of the heating and cooling loads of residential buildings. *Build. Simul.* **2022**, *15*, 2003–2017.
110. Chitsaz, H.; Shaker, H.; Zareipour, H.; Wood, D.; Amjadi, N. Short-term electricity load forecasting of buildings in microgrids. *Energy Build.* **2015**, *99*, 50–60.
111. Liu, T.; Tan, Z.; Xu, C.; Chen, H.; Li, Z. Study on deep reinforcement learning techniques for building energy consumption forecasting. *Energy Build.* **2020**, *208*, 109675.

112. Li, Z.; Friedrich, D.; Harrison, G.P. Demand Forecasting for a Mixed-Use Building Using Agent-Schedule Information with a Data-Driven Model. *Energies* **2020**, *13*, 780.
113. Chitalia, G.; Pipattanasomporn, M.; Garg, V.; Rahman, S. Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. *Appl. Energy* **2020**, *278*, 115410.
114. Khan, P.W.; Byun, Y.-C. Adaptive error curve learning ensemble model for improving energy consumption forecasting. *CMC-Comput. Mater. Contin* **2021**, *69*, 1893–1913.
115. Fang, X.; Gong, G.; Li, G.; Chun, L.; Li, W.; Peng, P. A hybrid deep transfer learning strategy for short term cross-building energy prediction. *Energy* **2021**, *215*, 119208.
116. Li, A.; Xiao, F.; Fan, C.; Hu, M. Development of an ANN-based building energy model for information-poor buildings using transfer learning. *Build. Simul.* **2020**, *14*, 89–101.
117. Luo, X. A novel clustering-enhanced adaptive artificial neural network model for predicting day-ahead building cooling demand. *J. Build. Eng.* **2020**, *32*, 101504.
118. Chou, S.-Y.; Dewabharata, A.; Zulvia, F.E.; Fadil, M. Forecasting Building Energy Consumption Using Ensemble Empirical Mode Decomposition, Wavelet Transformation, and Long Short-Term Memory Algorithms. *Energies* **2022**, *15*, 1035.
119. Xiao, Z.; Gang, W.; Yuan, J.; Chen, Z.; Li, J.; Wang, X.; Feng, X. Impacts of data preprocessing and selection on energy consumption prediction model of HVAC systems based on deep learning. *Energy Build.* **2022**, *258*, 111832.
120. Wang, Z.; Wang, Y.; Srinivasan, R.S. A novel ensemble learning approach to support building energy use prediction. *Energy Build.* **2018**, *159*, 109–122.
121. Liang, X.; Hong, T.; Shen, G.Q. Improving the accuracy of energy baseline models for commercial buildings with occupancy data. *Appl. Energy* **2016**, *179*, 247–260.
122. Gao, Y.; Ruan, Y. Interpretable deep learning model for building energy consumption prediction based on attention mechanism. *Energy Build.* **2021**, *252*, 111379.
123. Kim, J.-H.; Seong, N.-C.; Choi, W. Modeling and optimizing a chiller system using a machine learning algorithm. *Energies* **2019**, *12*, 2860.
124. Fan, C.; Ding, Y.; Liao, Y. Analysis of hourly cooling load prediction accuracy with data-mining approaches on different training time scales. *Sustain. Cities Soc.* **2019**, *51*, 101717.
125. Zhang, C.; Li, J.; Zhao, Y.; Li, T.; Zhang, X. A hybrid deep learning-based method for short-term building energy load prediction combined with an interpretation process. *Energy Build.* **2020**, *225*, 110301.
126. Pałasz, P.; Przysowa, R. Using Different ML Algorithms and Hyperparameter Optimization to Predict Heat Meters' Failures. *Appl. Sci.* **2019**, *9*, 3719.
127. Somu, N.; R, G.R.M.; Ramamritham, K. A hybrid model for building energy consumption forecasting using long short term memory networks. *Appl. Energy* **2020**, *261*, 114131.
128. Gao, L.; Liu, T.; Cao, T.; Hwang, Y.; Radermacher, R. Comparing deep learning models for multi energy vectors prediction on multiple types of building. *Appl. Energy* **2021**, *301*, 117486.
129. Wang, Z.; Hong, T.; Piette, M.A. Building thermal load prediction through shallow machine learning and deep learning. *Appl. Energy* **2020**, *263*, 114683.
130. Mariano-Hernández, D.; Hernández-Callejo, L.; Solís, M.; Zorita-Lamadrid, A.; Duque-Perez, O.; Gonzalez-Morales, L.; Santos-García, F. A Data-Driven Forecasting Strategy to Predict Continuous Hourly Energy Demand in Smart Buildings. *Appl. Sci.* **2021**, *11*, 7886.
131. Bedi, G.; Ganesh, K.V.; Rajendra, S. Development of an IoT-driven building environment for prediction of electric energy consumption. *IEEE Internet Things J.* **2020**, *7*, 4912–4921.
132. Zhou, C.; Yao, Z.; Hu, Y.; Cui, W. Study On The Application Of BP Neural Network In The Prediction Of Office Building Energy Consumption. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *546*, 052021.
133. Cao, L.; Li, Y.; Zhang, J.; Jiang, Y.; Han, Y.; Wei, J. Electrical load prediction of healthcare buildings through single and ensemble learning. *Energy Rep.* **2020**, *6*, 2751–2767.
134. Maltais, L.-G.; Gosselin, L. Forecasting of short-term lighting and plug load electricity consumption in single residential units: Development and assessment of data-driven models for different horizons. *Appl. Energy* **2022**, *307*, 118229.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.