Sustainable Cities and Society 118 (2025) 105570

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models

Yang-Seon Kim [a,1], Moon Keun Kim [b,1,*], Nuodi Fu [c], Jiying Liu [d], Junqi Wang [c,e], Jelena Srebric [f]

[a] Mechanical Engineering Department, College of Engineering, Wichita State University, Wichita, KS 67260, USA
[b] Department of Built Environment, Oslo Metropolitan University, N-0130, Oslo, Norway
[c] School of Architecture, Southeast University, 2 Sipailou, Nanjing 210096, China
[d] School of Thermal Engineering, Shandong Jianzhu University, Jinan 250101, China
[e] School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou 215009, China
[f] Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Keywords:
Data normalization
Artificial neural networks
Electricity prediction
Occupancy rates

ABSTRACT

The study investigates the impact of data normalization on the prediction of electricity consumption in buildings using four multilayer Artificial Neural Networks (ANN) algorithms: Long Short-Term Memory Networks (LSTM), Levenberg–Marquardt Back-propagation (LMBP), Recurrent Neural Networks (RNN), and General Regression Neural Network (GRNN). Four data normalization approaches, Min-Max Scaling, Mean, Z-score, and Gaussian function were assessed on experimental datasets. The LSTM algorithm, when combined with Min-Max normalization, showed the most favorable predictive capabilities, with a low Coefficient of Variation of the Root Mean Square Error (CVRMSE) of 10.3 and Normalized Mean Bias Error (NMBE) of 0.6. The remaining three normalization approaches showed satisfactory concordance with empirical data, but with slight disparities in precision. The LMBP model, when using Z-score normalization, had favorable performance in forecasting electricity consumption, but the discrepancies across the models were not significant. The Recurrent Neural Network (RNN) model, when used with Gaussian normalization, exhibited the most favorable performance, with the lowest Coefficient of Variation of Root Mean Square Error (CVRMSE) at 11.8 and Normalized Mean Biased Error (NMBE) at 0.6. The Generalized Regression Neural Network (GRNN) model, trained on unprocessed data, exhibited superior performance, with the lowest Coefficient of Variation of Root Mean Square Error (CVRMSE) at 19.2 and NMBE at 1.0. In conclusion, the study highlights the significant influence of data normalization on the predictive capabilities of various ANN models, suggesting that careful use of data normalization techniques can significantly improve the accuracy of electricity consumption forecasting in buildings.

1. Introduction

Buildings represent about 36 percent of global final energy consumption and 39 percent of energy-related CO_2 emissions (IEA, 2023b). Buildings play a dominant role in global energy savings. In recent years, numerous studies have focused on building energy-saving technologies, and policies to overcome global climate change and national energy security challenges. Still, global final energy consumption in buildings increased by more than 5 percent between 2010 and 2017, as energy efficiency gains were outpaced by continued growth in the building sector size and demands (IEA, 2023a). If they are implemented in the design process, current technologies can result in significant energy and financial savings as well as other advantages. One of the most difficult challenges is buildings are not consuming the energy as designed. Therefore, accurate estimations of building energy consumption after construction are important for developing energy-saving strategies, building management systems, and energy distribution planning. And precise prediction of building energy consumption is essential for enhancing the utilization of renewable energy, offering dependable scientific backing for smart grid schedule, and safeguarding grid security (Yang et al., 2021). Moreover, forecasting building energy consumption

* Corresponding author.
E-mail address: Moon.Kim@oslomet.no (M.K. Kim).
[1] These authors contributed equally to the research and manuscript.

https://doi.org/10.1016/j.scs.2024.105570
Received 18 October 2023; Received in revised form 2 June 2024; Accepted 2 June 2024
Available online 3 June 2024

Nomenclature

Acronyms

b_c	cell candidate bias vector
b_f	forget bias vector
b_i	input bias vector
b_o	output bias vector
$\boldsymbol{e}$	error vector
$\boldsymbol{g}$	gradient vector
Gk	Gaussian Kernel's proximity
$\boldsymbol{I}$	unit matrix
$\boldsymbol{J}$	Jacobian matrix
σ	sigmoid function tanh hyperbolic tangent
$\boldsymbol{\mu}$	constant coefficient
h	hidden layer
t	time step
x	input layer
y	output layer
y_i	actual value
$\bar{y}_i$	predicted value
n	total number of data value.
W_c	cell candidate weight matrices
W_f	forget weight matrices
W_i	input weight matrices
W_o	output weight matrices

Abbreviation

ANN	artificial neural networks
ASHRAE	American society of heating, refrigerating, and air-conditioning engineers
CVRMSE	coefficient of variation of the root mean square error
DRNN	deep recurrent neural network
GRNN	general regression neural network
HVAC	heating ventilation and air conditioning
LMBP	Levenberg–Marquardt back-propagation
LSTM	long short-term memory networks
NMBE	normalized mean bias error
Min	minimum value
max	maximum value
mean	mean value
Std	standard deviation
RNN	recurrent neural networks

serves as the cornerstone for various building energy system, including building safety, energy demand calculation, and sensor control optimization (Antonopoulos et al., 2020; Ilbeigi et al., 2020; Qian et al., 2020).

Traditionally, building energy has been estimated with engineering software packages (e.g., EnergyPlus, DOE) that rely on an in-depth understanding of structural, geometrical, and material of building properties. These "engineering" models are computationally expensive and require very detailed input data (Robinson et al., 2017). However, this level of detailed information is not always available to use, and it isn't easy to use in wide-scale energy forecasting. Therefore, data-based building energy prediction approaches are a growing area of research. The data-based building energy prediction approaches is comparable, and sometimes superior, to traditional engineering-based energy forecasting while requiring significantly fewer input data sets (Feng et al., 2022; Jain et al., 2014; Kim et al., 2020b; Runge & Zmeureanu, 2019).

Various data-driven algorithms have been used for building electricity consumption prediction modeling, which includes the mathematical-based model, a statistical approach or a machine learning algorithm (Kim et al., 2020b; Runge & Zmeureanu, 2019; Zhu et al., 2022). The statistical approach typically uses a pre-set mathematical function and has shown a good results for medium to long term energy forecasting (Robinson et al., 2017). Machine learning-based models can be dealing with non-linear patterns, which conventional methods' lack flexibility (Daut et al., 2017). Therefore, multilayer Artificial Neural Network (ANN) models have been widely used in practical applications, including forecasting and predicting energy loads, ascertaining the current energy performance of buildings, and predicting potential for energy-saving strategies (Kim et al., 2020b; Lu et al., 2022; Wang et al., 2023; Wang & Srinivasan, 2017). In the building science community, many different algorithms in ANN models were used.

Time series data may contain temporal dependencies and the Recurrent Neural Networks (RNN) algorithm is using feedback connections to recall the value at prior time steps to estimate the building energy consumption. Many previous studies proved that RNN can be a good tool to estimate the time series building energy prediction (Abdulrahman et al., 2021; Kreider et al., 1995). Kreider et al. found that with a various inputs such as current hour's outside temperature, two hours' previous outside temperature, current hour's inside temperature and two hours' previous inside temperature, they could successfully estimate the current hours' heat loss Q and convert this into the building energy loads (Kreider et al., 1995).

Long Short-Term Memory Networks (LSTM) algorithm is an extended form of the deep recurrent neural network (DRNN). This method has been used in several different building types (commercial buildings and residential buildings) to predict the energy consumption of the buildings due to its ability to solve long-term and short-term dependencies (Faiq et al., 2023; Kong et al., 2019; Somu et al., 2020) LSTM can forecasts the lower frequency variations and it can predict one step to three step ahead with other predictors. Sendra-Arranz and Gutierrez developed an LSTM-based prediction model to forecasting the energy consumption of an HVAC systems and the model showed outstanding results with the normalized root mean square error (NRMSE) of 0.052 (Sendra-Arranz & Gutierrez, 2020) . Faiq et al. used LSTM to predict the daily day-ahead energy consumption in the school buildings and the proposed model achieved the best Root Mean Square Error (RMSE) of 561.692–592.319 (kWh) when compared to other types of algorithms (Faiq et al., 2023). Kong et al. proposed LSTM approach to estimate the short-term load forecasting for individual residential homes. The model successfully forecasting the load with mean absolute percentage error (MAPE) score range from 8.58 % to 9.14 % (Kong et al., 2019).

Levenberg–Marquardt Back-propagation (LMBP) algorithm is a well-liked alternative strategy to the Gauss-Newton method when seeking minimize the sum of squares for nonlinear functions. Authors previous paper found that LMBP algorithms has better accuracy and stability in relatively small scale networks to predict the electricity consumption in buildings (Kim et al., 2020a). Lei and Yin used LMBP algorithms to improve the model prediction for the energy consumption in high-rise building within from 2.2 to 8.7 on RMSE (Lei & Yin, 2022).

General Regression Neural Network (GRNN) has been used to investigate the HVAC thermal energy optimization in various buildings. Ben-Nakhi and Mahmoud used GRNN for the three different buildings hourly cooling loads estimation (Ben-Nakhi & Mahmoud, 2004). This algorithms can also be used to fault diagnosis of the building's HVAC system and it shows that the GRNN models are accurate and reliable estimators of highly non-linear and complex systems (Lee et al., 2004).

This study chose four primary ANN algorithms commonly employed in building energy prediction modeling for HVAC systems, emphasizing their strong accuracy and relatively quick prediction times. Many previous studies used different ANN algorithms in building energy modeling and systems. However, many of them have analyzed preselected algorithms without comparing it with different algorithms. Only few studies used different algorithms in building energy consumption forecasting (Kim et al., 2020b; Moore et al., 1991; Qiong et al., 2010). Even though they tried different algorithms, most previous studies have yet to try to adapt different data normalization methods to find the best model.

In this study, we investigated the novel analysis strategy how data normalizations using ANNs impact to predict building electricity consumption. With a following steps, we could identify the most suitable data normalization method for each ANN algorithm to enhance the accuracy and stability of electricity consumption predictions.

Our study includes main objectives:

- Analyze the variations in four multilayer Artificial Neural Networks (ANNs) algorithms using four different data normalization methods based on experimental data collected in a campus building, including occupancy ratio, weather data, and electricity consumption data.
- Predict electricity consumption in a campus building using four ANN algorithms: Long Short-Term Memory Networks (LSTM), Levenberg–Marquardt Back-propagation (LMBP), Recurrent Neural Networks (RNN), and General Regression Neural Network (GRNN) for comparative analysis.
- Investigate the effect of four data normalization techniques: Min-Max scaler, Mean normalization, Z-score, and Gaussian normalization, in comparison with raw data set as a predictive analysis using ANN models with data normalization to understand their performance in electricity consumption prediction.

The main contribution of this paper is providing the valuable insights to improve the effectiveness of ANN-based predictive models with different data normalization methods for energy consumption forecasting in buildings, aiding in real-world applications.

2. ANNs models to predict electricity consumption

In the realm of built environment research, Artificial Neural Network (ANN) models have emerged as a specific choice for data-driven analysis and prediction modeling (Chae et al., 2016; Tian et al., 2021). What sets ANNs apart is their ability to effectively handle complex and non-linear relationships within experimental data, resulting in relatively high accuracy compared to conventional analytical methods (Andrean et al., 2018; Kim et al., 2019; Samy et al., 2022; Wang & Srinivasan, 2017; Wazirali et al., 2023; Ye & Kim, 2018). This advantage is attributed to the ANN's design, which draws inspiration from the human brain Lek and Guégan (1999). For this study, four specific algorithms were carefully chosen: the Long Short-Term Memory (LSTM), Levenberg–Marquardt Back-propagation (LMBP), recurrent neural network (RNN) and the Generalized Regression Neural Network (GRNN). The LSTM neural network is particularly favored for forecasting building energy usage due to its capability to retain crucial training information from the beginning of the data sequence, thus resolving the issue of gradient vanishing commonly encountered in recurrent neural networks (RNNs) (Hochreiter & Schmidhuber, 1997; Sendra-Arranz & Gutierrez, 2020). The Levenberg–Marquardt Back-propagation (LMBP) algorithm was specifically developed to optimize and minimize loss functions in neural networks (Ye & Kim, 2018; Zhang et al., 2022). A recurrent neural network (RNN) is a type of artificial neural network in which the connections between neurons allow feedback loops to exist between the layers during the training process (Babalhavaeji et al., 2023; Fang & He, 2023; Weerakody et al., 2021). And the GRNN exhibits remarkable potential in identifying intricate patterns while offering improved training speed, classification, and optimization for a wide array of problems when compared to conventional linear regression techniques (Park & Kim, 2017; Wang et al., 2021; Zhang et al., 2023). The selection of these algorithms in the research design enhances the overall predictive capacity and performance, making it a significant and suitable approach for addressing the research objectives. Fig. 1 depicts the prediction methodology employing four distinct neural network algorithms. These models consist of 6 input layer neurons, 14 hidden layer neurons, and 1 output layer neuron. The learning rate is chosen within the range of 0.01 to 0.7. The error target and the decision on the number of trainings have been determined. This model selects 0.01 as the error target and conducts training 50,000 times. These values are derived from ongoing testing.

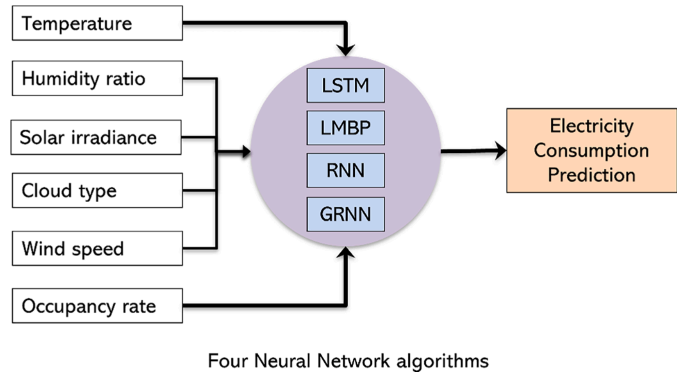

Fig. 1. The prediction method with four neural network algorithms.

2.1. Long short-term memory (LSTM) neural networks

Unlike feed-forward neural networks, recurrent neural networks (RNNs) establish connections that allow each neuron to transmit information between layers by looping back during the training process. As a result, RNNs demonstrate proficiency in recognizing and compressing data patterns. The feedback process in RNNs refines inputs, enhancing the accuracy of output results by leveraging estimated outputs to improve the input data, as supported by references (Ahmad & Chen, 2019; Calderano et al., 2023; Meng-Hock & Hagan, 1996; Moller, 1993; Ribeiro et al., 2006). LSTM networks, a unique type of RNN, stand out due to their special memory cell feature that facilitates the retention of information over extended durations (Ballabio & Vasighi, 2012; Lapuschkin, Binder, Montavon, Muller, & Samek, 2016; Van Houdt et al., 2020). To tackle the challenge of vanishing and exploding gradients, LSTM networks incorporate innovative gates like input and forget gates (Cowan, 2019; Van Houdt et al., 2020). These gates offer enhanced control over gradient flow, resulting in a superior preservation of long-term dependencies.

The LSTM algorithm is defined by the following equations (Agga et al., 2022; Li et al., 2022; Meng et al., 2022; Van Houdt et al., 2020):

$$\text{Inputgate}: \mathrm{i_t} = \sigma(\mathrm{W_i}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_i}) \tag{1}$$

$$\text{Forgetgate}: \mathrm{f_t} = \sigma(\mathrm{W_f}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_f}) \tag{2}$$

$$\text{Cellstate}\widehat{C}_\mathrm{t} = \tanh(\mathrm{W_c}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_c}) \tag{3}$$

$$\mathrm{C_t} = \mathrm{f_t} \times \mathrm{C_{t-1}} + \mathrm{i_t} \times \widehat{C}_\mathrm{t}) \tag{4}$$

$$\text{Outputgate}: \mathrm{o_t} = \sigma(\mathrm{W_o}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_o}) \tag{5}$$

$$\text{HiddenState}: \mathrm{h_t} = \mathrm{o_t} \times \tanh(\mathrm{C_t}) \tag{6}$$

The LSTM algorithm employs weight matrices, W_i, W_f, W_o and W_c to govern the input, forget, output gate, and cell candidate, Correspondingly the bias vectors; b_i, b_f, b_o and b_c are associated with these gates. The activation functions utilized in the LSTM are the sigmoid function, σ (*) and hyperbolic tangent, tanh(*).

Fig. 2 illustrates the LSTM network models. In the figure, the architecture and components of the LSTM networks are visually represented, highlighting the input, forget, output gates, and cell state, along with their interconnections and information flow.

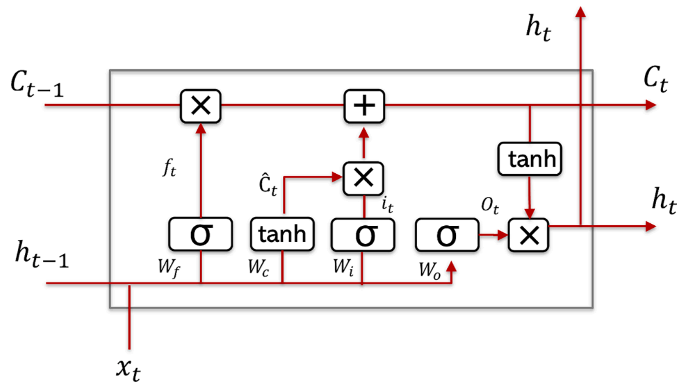

Fig. 2. Structure of the LSTM model.

2.2. Levenberg–Marquardt back-propagation (LMBP)

The backpropagation (BP) network is extensively applied in diverse fields like pattern recognition, data processing, error detection and intelligent control (Rumelhart et al., 1986; Xu et al., 2015). Enhancements in BP neural networks primarily focus on refining network structure and learning algorithms. Proposed methods for structural improvement encompass the additional momentum method, elastic BP algorithm, and adaptive learning rate. BP neural networks possess robust fault tolerance and associative memory capabilities, enabling them to restore complete information and maintain system operation even in the event of partial loss or damage. The Levenberg–Marquardt (LM) algorithm (Hao et al., 2013; Kim et al., 2020b; Ye & Kim, 2018; Zhang et al., 2022), a pivotal nonlinear training algorithm, combines gradient descent and Quasi-Newton methods to achieve swift local convergence and superior overall performance by exploring error in multiple directions during iteration.

The Levenberg–Marquardt Back-propagation (LMBP) algorithm was specifically developed for optimizing loss functions, providing a viable alternative to the Gauss-Newton method for minimizing the sum of squares of nonlinear functions (Bocheng et al., 2015; Yoo & Seul, 2017). In the Levenberg–Marquardt (LM) algorithm, a sequence of optimizations is performed, resulting in the formulation of weight approximations for the Hessian matrix and threshold in Eqs. (7, 8, and 9) (Azadeh et al., 2008; Hao et al., 2013; Ye & Kim, 2018). These equations play a pivotal role in efficiently navigating the optimization landscape and updating the network's weights during the training process, ultimately leading to improved convergence and accuracy.

$$H = \boldsymbol{J}^T\boldsymbol{J} \tag{7}$$

$$g = \boldsymbol{J}^T\boldsymbol{e} \tag{8}$$

$$x(k+1) = x(k) - \left[\boldsymbol{J}^T\boldsymbol{J} + \mu\boldsymbol{I}\right]^{-1}\boldsymbol{J}^T e \tag{9}$$

In these equations $\boldsymbol{J}$ represents the Jacobian matrix, $\boldsymbol{g}$ is a vector representing the gradient and the coefficient $\boldsymbol{\mu}$ represents a constant that is greater than zero. $\boldsymbol{I}$ represents a unit matrix and $\boldsymbol{e}$ denotes the error vector. When $\boldsymbol{\mu}$ is near zero, this method is equivalent to the Gauss-Newton method.

Some studies have reported that LMBP exhibits superior accuracy and stability, particularly in relatively small-scale networks. However, it is important to note that this algorithm requires more memory compared to other training methods (Kim et al., 2022, 2020a, 2020b). Despite this memory requirement, LMBP's advantages in training speed and performance make it a valuable option for various neural network applications.

2.3. Recurrent neural network (RNN)

A recurrent neural network (RNN) possesses the unique characteristic of allowing connections between neurons to create feedback loops, facilitating information transmission between layers during the training process. This inherent capability endows RNNs with the potential for data pattern recognition and compression. The feedback mechanism plays a crucial role in refining inputs and improving the accuracy of output results by utilizing estimated outputs to enhance the input data (Brezak et al., 2012; Cossu et al., 2021; Karim & Rivera, 1992; Kurnaz & Demir, 2022; Moore et al., 1991)

The RNN algorithm can be described by the following equations (Karim & Rivera, 1992; Moore et al., 1991):

$$\mathrm{y}^t = \mathrm{g}\left(\mathrm{h}^t\right) \tag{10}$$

$$\mathrm{h}^t = f\left(\mathrm{x}^t, \mathrm{h}^{t-1}\right) \tag{11}$$

where, x,y and h represent the input, output and hidden layer respectively at time step t. Eq. (10) computes the output y at time step t by applying the function g to the hidden state h^t. Eq. (11) describes how the hidden layer h is updated at time step t using the input x^t and the previous hidden state h^{t-1}.

2.4. Generalized regression neural networks (GRNN)

The Generalized Regression Neural Network (GRNN) model is widely employed as a time series forecasting tool for estimating energy performance and efficiency associated with specific components due to its rapid analysis capabilities (Park & Kim, 2017; Zhang et al., 2023). GRNN facilitates the comparison of two data elements to determine their similarity and presents the results as a regression model, incorporating a coefficient to elucidate the relationship (Li et al., 2011; Mirza et al., 2022). One of the remarkable features of GRNN is its single pass learning pattern from input data, requiring only the setting or fitting of a single parameter, which eliminates the need for multiple training iterations to achieve accurate and reliable results (Li et al., 2011; Mirza et al., 2022; Zhang et al., 2023).

In the work of Specht (Specht, 1991), a variant of GRNN was designed using a radial basis neural network. The radial basis neural network adopts the multivariate Gaussian Kernel formula as follows:

Eq. (12) represents the multivariate Gaussian Kernel formula used in the Generalized Regression Neural Network (GRNN) (Specht, 1991):

$$\mathrm{Gk}(\mathrm{x},\ \mathrm{x_i}) = \exp\left(-\frac{\|\mathrm{x} - \mathrm{x_i}\mathrm{AptCommand2016;}^2}{2\sigma^2}\right) \tag{12}$$

where, x and xi represent the center and smoothing parameter, respectively, which determine the output of a hidden layer neuron based on the input vector x. The smoothing parameter, y adjusts the degree of proximity between the input vector and the center.

GRNN's training patterns can be divided into two parts: the input vectors (x1, x2, x3, …, xn) and the target vector (y1, y2, y3, …, yn). The prediction value Yx for input x is computed using Eq. (13):

$$y_x = \frac{\sum_{i=1}^{N} y_i Gk\ (x,\ x_i)}{\sum_{i=1}^{n} Gk\ (x,\ x_i)} \tag{13}$$

where, $\mathrm{Y_x}$ represents the prediction value for input x, and $\mathrm{y_i}$ denotes the activation weight for the pattern layer at i. This equation computes the weighted average of the target vector values based on their respective activation weights and the Gaussian Kernel's proximity to the input x.

3. Data normalization methods

Data normalization is one of the pre-processing approaches used to scale or transform data, ensuring that each feature makes an equal contribution to the analysis (Liu et al., 2019; Swift et al., 2023). By normalizing the data, the values of different features are adjusted to a common scale, preventing any particular feature from dominating the analysis simply due to its larger magnitude (Elkhouly et al., 2023; Singh & Singh, 2020). This normalization process aids in improving the performance and convergence of various machine learning algorithms, as it ensures that each feature's influence is balanced and contributes effectively to the overall analysis (Blokhintsev & Savin, 2021; Gutierrez-Osuna & Nagle, 1999). Indeed, the importance of data normalization in constructing accurate predictive models has been extensively examined across various machine learning algorithms. Data normalization plays a critical role in preparing the input data for modeling, as it helps to mitigate the impact of differing scales and distributions of features (Elkhouly et al., 2023; Liu et al., 2019).

In this study, four different normalization schemes widely used are proposed for utilizing ANN models and for improving prediction performance of energy consumption. These normalization approaches are intended to preprocess the input data and enhance the performance and effectiveness of the ANN models. Each normalization scheme is designed to address specific characteristics of the data and tailor the input features to better suit the neural network architecture.

The four normalization schemes may include, but are not limited to, techniques such as: Min-Max scaling, mean normalization, Z-score normalization, and Gaussian normalization.

3.1. Min-Max scaling normalization

Min-Max Scaling: Scaling the data to a specific range (e.g., [0, 1]) to preserve the relative relationships between feature values.

$$x' = \frac{x - \min(\max_n - \min_n)}{\max - \min} + \min_n \tag{14}$$

where x is the original value, x' is the normalized value, min and max are the minimum and maximum values of the feature, respectively. min-max data normalization is a technique used to equalize the weight of different features and make them have the same effect on the decision-making process in machine learning and data analysis (Ali, 2022).

3.2. Mean normalization

Mean normalization represents a commonly applied data normalization technique in the fields of machine learning and data analysis (Almehrizi, 2016; Mortensen et al., 2018). It involves scaling the data so that the mean (average) of the feature becomes zero. This process centers the data around the mean, effectively removing any bias in the feature values.

The mean normalization formula for a feature x' is as follows:

$$x' = \frac{x - \text{mean}(x)}{\max(\mathrm{x}) - \min(\mathrm{x})} \tag{15}$$

where mean(x) represents the mean of the feature x, max (x) is the maximum value of x and min (x) is the minimum value of x.

3.3. Z-score normalization

Z-score normalization, also known as standardization, is a widely used data normalization technique in machine learning and statistical analysis. It involves transforming the data in such a way that the mean of the feature becomes zero and the standard deviation becomes one (Henn et al., 2014; Zhang et al., 2020). This process centers the data around zero and scales it to have a unit standard deviation.

The formula for z-score normalization for a feature x' is as follows:

$$x' = \frac{x - \text{mean}(x)}{\mathrm{Std(x)}} \tag{16}$$

where x is the original value, x' is the normalized value, mean(x) is the mean value of the feature x, and std(x) is the standard deviation of x. one of the significant advantages of z-score normalization (standardization) is its robustness to outliers compared to other normalization methods. Outliers are extreme values that differ significantly from the majority of the data points and can distort the statistical properties of the dataset (Henn et al., 2014; Zhang et al., 2020).

3.4. Gaussian function normalization

The Gaussian function representing the probability density function of a normally distributed random variable values follows (Bergman et al., 2023; Elkhouly et al., 2023):

$$x' = \frac{1}{\mathrm{std(x)}\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x - mean(x))^2}{std(x)^2} \right) \tag{17}$$

where std(x) is the standard deviation of x.

The Gaussian function is widely used in various different fields, including statistics, probability theory, and data analysis, to model continuous random variables that exhibit normal distribution characteristics (Bergman et al., 2023; Delattre & Roquain, 2016; Elkhouly et al., 2023).

Fig. 3 visually demonstrates the four neural network training and prediction processes, each utilizing a distinct normalization method, employed in this study. The figure presents a clear and systematic representation of the step-by-step procedures involved in both the training and testing stages for each neural network model with their respective normalization techniques. The processes encompass data preprocessing, input data feeding, weight optimization, and output prediction.

The four ANN models are separately trained and tested using their corresponding normalization methodologies. This approach enables a comprehensive comparison of the performance and accuracy of each neural network model in addressing the specific tasks outlined in the study. By using different normalization methods, the study aims to evaluate how each technique impacts the efficiency and effectiveness of the neural network models.

Overall, Fig. 3 provides a concise and informative visual summary, offering a better understanding of the experimental setup and the significance of using various normalization approaches in this study.

4. Collected data and analysis

The study conducted data collection on building occupancy rates and electricity consumption for a campus building in University Park, PA, USA. The building's area was around 7000 m^2 and comprised three stories, with specific proportions dedicated to lecture rooms, offices, common areas, and laboratories. Occupancy rates were measured using an infrared thermal sensor placed at the main entrance, which detected movement and counted occupants entering and leaving its field of view with a 5 % error margin (Kim et al., 2017).

Hourly data on building occupancy rates and electricity consumption were measured and gathered for a total of 157 days between August 14, 2013, and April 28, 2014. Dates with missing data or not measured were excluded. Additionally, local weather data from the National Solar Radiation Database (NSRDB) (Laboratory, 2023) was incorporated into the analysis. The dataset was divided into two groups: working days (109 days, 2616 hourly data) and non-working days (48 days, 1152 hourly data). This division allowed for a comprehensive investigation of how occupancy rates and weather conditions influenced electricity consumption in different operational scenarios. To simulate electricity energy consumption using ANN models, the working day data was further divided into training data, consisting of 95 days, and test data, comprising 14 days. These test days were fairly distributed across the seasons (spring, summer, autumn, and winter), with approximately 3–4 days allocated to each season, but selected randomly from the dataset.

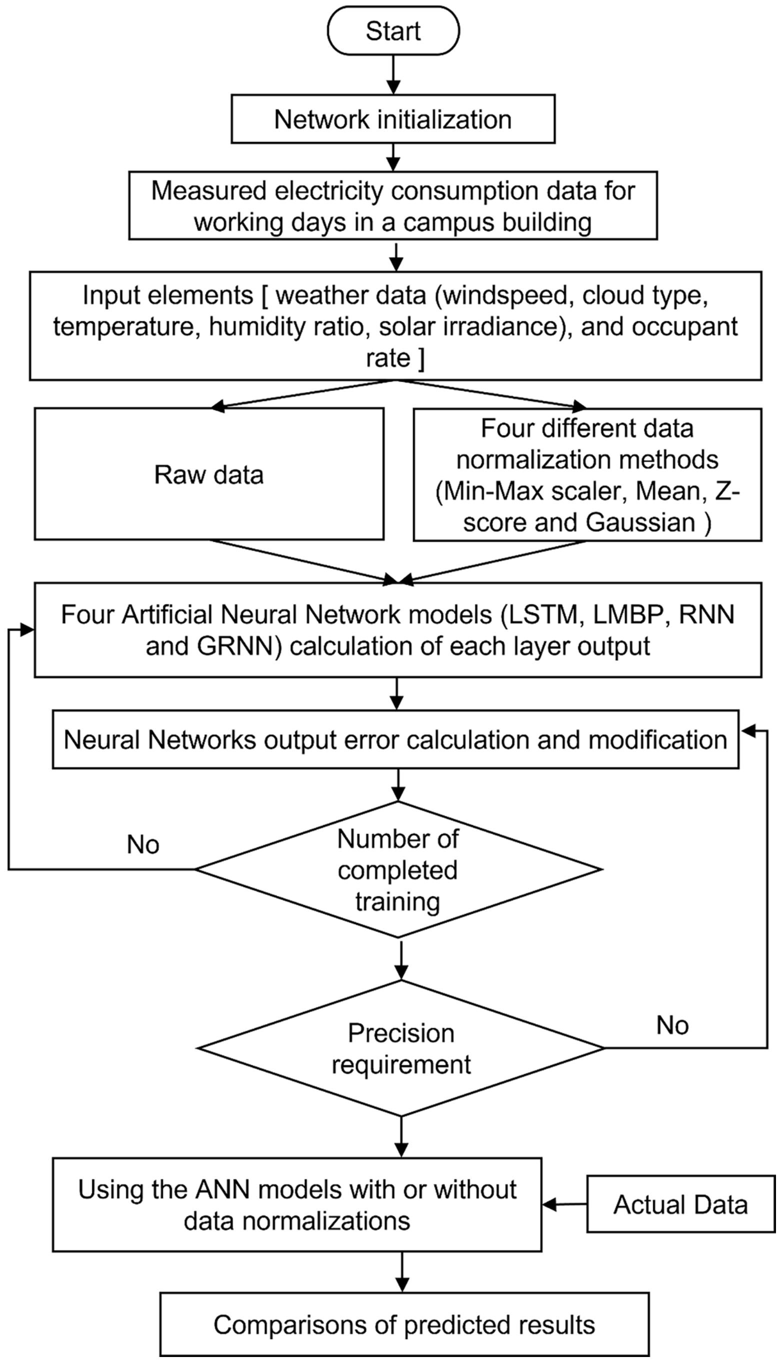

Fig. 3. Training process for the ANN models with data normalizations.

The decision to exclude non-working day data from the study was made because the energy consumption patterns on non-working days significantly differ from those observed on working days. Following the utilization of each ANN algorithm on the testing data, the validations are assessed based on measurements obtained from real data.

Given that the study aims to understand and predict energy consumption patterns during normal working days when the building is fully operational, including non-working day data may introduce unstable patterns into the analysis, as there might not be sufficient data collected in comparison to that of working days. Non-working days, such as weekends or holidays, often exhibit distinct energy usage patterns due to reduced occupancy and altered building operations. By excluding non-working day data, the study can concentrate on analyzing the electricity consumption behavior during typical working days, providing more accurate and focused insights for energy management and optimization strategies specific to regular operational scenarios. The training and test data sets, consisting of working day data, will allow the ANN models to learn and generalize patterns specific to these typical operating conditions, enhancing the accuracy of the predictive models and their practical applicability.

In the model, input nodes that affected building electricity energy consumption (kW) were defined, including occupant number, temperature (°C), humidity ratio (g/kg), direct normal irradiance (DNI; W/m2), wind speed (m/s), and cloud type (ranging from 0 for clear to 12 for smoke).

4.1. Comparison methodology

To evaluate the accuracy and error rate of the four ANN algorithms with and without data normalization, this study employed two performance metrics: the coefficient of variation of the root mean square error (CVRMSE) and the normalized mean bias error (NMBE). These metrics are commonly used in the field of energy prediction and heating, refrigerating, and air-conditioning engineering, as indicated by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Guideline 14–2002 (ASHRAE, 2002). The equations for these two metrics are as follows:

$$CVRMSE\ (\%) = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2\right]^{1/2}}{\bar{y}} \times 100 \tag{18}$$

$$NMBE\ (\%) = \frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)}{N \times \bar{y}} \times 100 \tag{19}$$

where y_i is the actual value for the *i*th data point. $\bar{y}_i$ is the predicted value and n is the total number of data value.

5. Results

To predict electricity energy consumption in a campus building, we used four ANN methods—LSTM, LMBP, RNN and GRNN algorithm. The data were collected with a total of 3778 hourly data representing 157 days of occupancy rates and weather parameters, temperature (°C), humidity ratio (g/kg), normal direct irradiation (W/m^2), cloud type (0–12), and wind speed (m/s). This study denotes the accuracy of each model and also estimates how significantly the data normalization method influence to predict electricity consumed in a building. Moreover, using data collected and normalized, we can predict long-term electricity consumption rates correctly based on occupants' diversity and weather conditions. We used four prediction methods—LSTM, LMBP, RNN and GRNN neural network with four different normalization methods-Min-Max, Mean, Z-score and Gaussian function and predicted electricity consumption as the output layer.

The training data for the simulation set contained measured and historical data for 90 working days and excluding non-working days. The electricity consumption results were predicted for 14 working days using four different ANNs to validate each algorithm compared with the actual measured values. The four ANN models were evaluated based on accuracy and error rate with four different data normalization methods; the results are shown in Figs. 4–8.

Figs. 4 and 8 showcase the prediction performance of the LSTM model, both with and without data normalization. The results demonstrate that in both cases, the predicted values align closely with the actual values. The study's findings reveal that the Min-Max and Mean normalization methods outperform both raw data and other normalization techniques in terms of prediction performance using the LSTM model. The Min-Max scaler offers distinct advantages in improving prediction accuracy by minimizing various error metrics when used with the LSTM method. It helps in reducing the predicted interquartile range (IQR) value, leading to a more concentrated and precise prediction interval. By scaling the data to a specific range, Min-Max normalization ensures that the model's predictions are confined within a narrower interval, resulting in a more accurate representation of prediction uncertainty (Ali, 2022; Swift et al., 2023). And the Min-Max scaler contributes to minimizing both the maximum and minimum error rates between predicted and actual values. By equalizing the feature weights and enhancing model generalization, it reduces the magnitude of errors, making the predictions more consistent and balanced across different data points.

Figs. 5 and 8 illustrates the performance using LMBP algorithm with comparisons between data normalization, raw data and measured data. In the results of the study, the Z-score normalization method exhibited higher accuracy performance and minimized maximum and minimum error ranges compared to the raw data method. Although the difference between the two methods was relatively small, the Z-score normalization method demonstrated a slight advantage in improving prediction accuracy and reducing prediction errors. On the other hand, the Gaussian normalization method showed less favorable prediction performance when used with the LMBP algorithm. The predictions made using the Gaussian normalization method were less accurate and had higher error rates compared to the other normalization methods and the raw data. Overall, the findings suggest that Z-score normalization is a more effective preprocessing technique for enhancing the prediction performance of the LMBP algorithm compared to using raw data directly. However, the Gaussian normalization method did not yield similarly favorable results in this specific context.

Figs. 6 and 8 present the performance of the RNN algorithm with comparisons between data normalization, raw data, and measured data. The results of the study indicate that data normalization methods exhibited higher accuracy performance compared to using raw data directly. Specifically, the four normalization techniques, Min-Max,

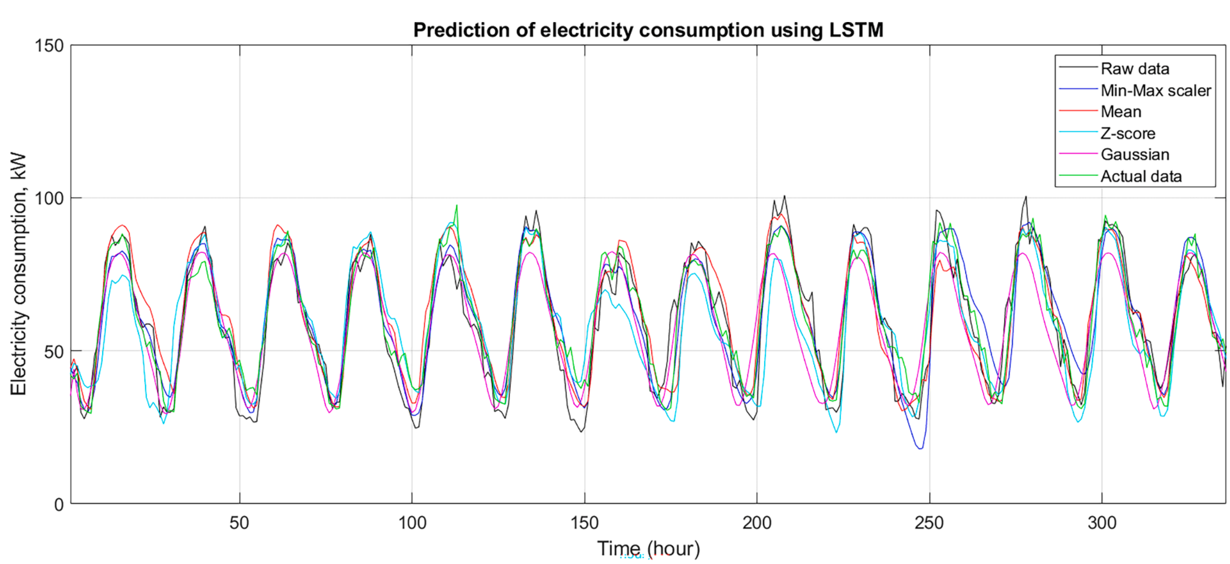

Fig. 4. Prediction of electricity consumption using LSTM algorithm with comparisons between data normalization, raw data and measured data.

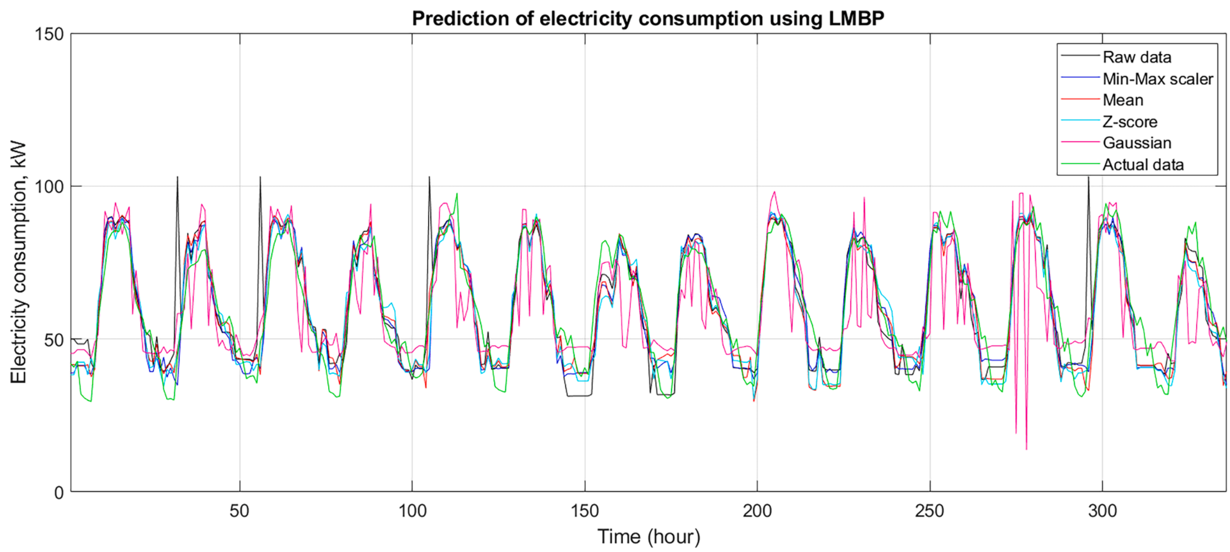

Fig. 5. Prediction of electricity consumption using LMBP algorithm with comparisons between data normalization, raw data and measured data.

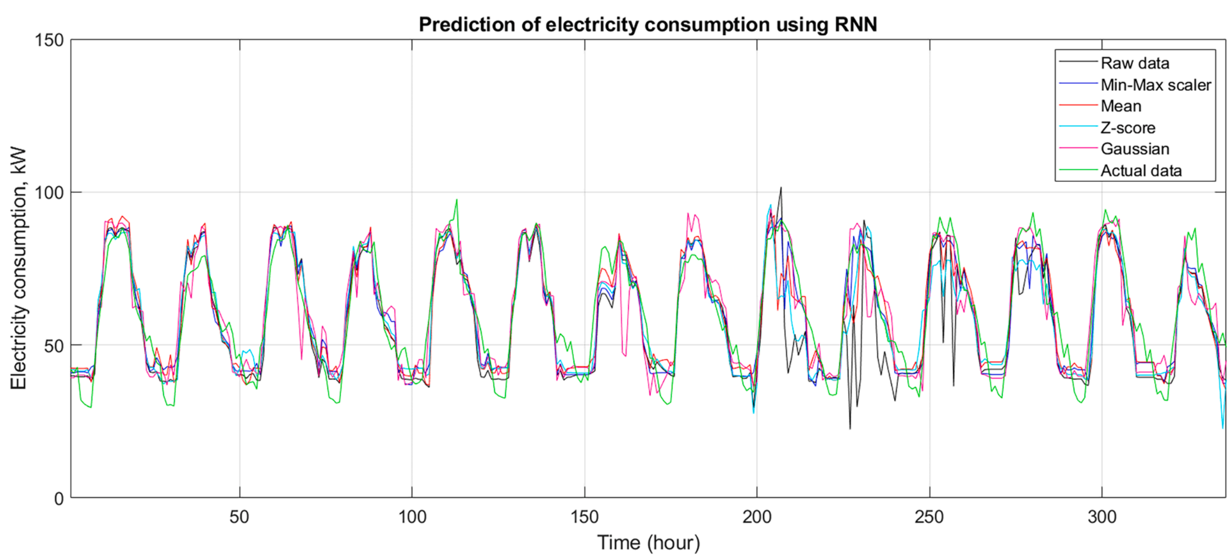

Fig. 6. Prediction of electricity consumption using RNN algorithm with comparisons between data normalization, raw data, and measured data.

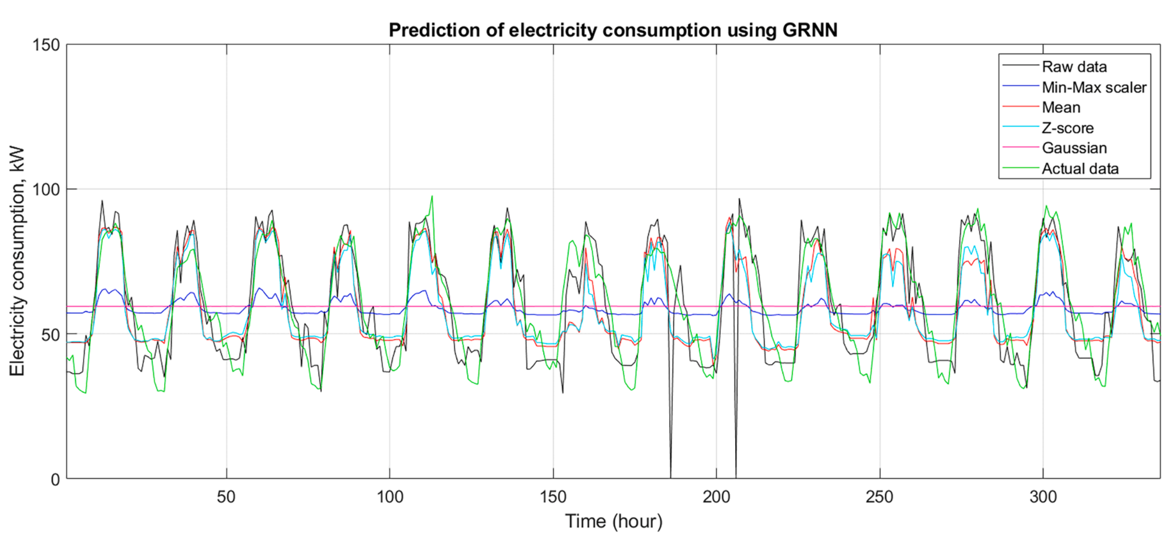

Fig. 7. Prediction of electricity consumption using GRNN algorithm with comparisons between data normalization, raw data and measured data.

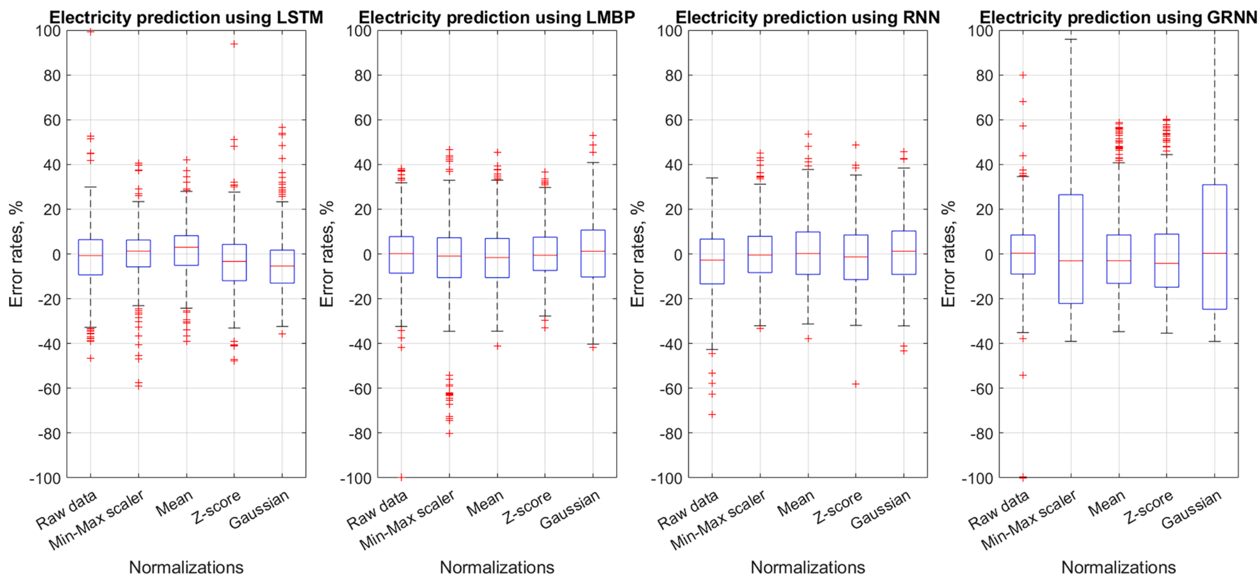

Fig. 8. Error rates of the four prediction ANN models using four data normalization methods and raw data.

Mean, Z-score and Gaussian normalization, showed superior performance in minimizing maximum and minimum error ranges, indicating their effectiveness in reducing prediction errors. The advantages of data normalization methods in improving prediction accuracy and reducing errors underscore the importance of proper data preprocessing in the RNN algorithm. By normalizing the data, the RNN model can handle disparate feature scales more effectively, leading to better prediction performance. Overall, the findings suggest that data normalization techniques are a crucial preprocessing step for enhancing the prediction performance of the RNN algorithm. When compared to using raw data directly, the normalization methods provide a clear advantage in achieving more accurate and reliable predictions.

Figs. 7 and 8 illustrate the performance of the GRNN algorithm with comparisons between data normalization, raw data, and measured data. The results of the study indicate that data normalization methods exhibited less accuracy performance compared to using raw data directly. Specifically, the two normalization techniques, Min-Max and Gaussian normalization, showed poor performance in minimizing maximum and minimum error ranges, indicating that they were not as effective in reducing prediction errors. Contrary to the findings in previous models (RNN and LSTM), there were no special advantages observed for data normalization methods in improving prediction accuracy and reducing errors in the context of the GRNN algorithm. In fact, using raw data directly in the GRNN model appeared to be more effective in handling disparate feature scales and achieving better prediction performance.

Overall, the findings suggest that data normalization techniques might not have a significant impact as a preprocessing step for enhancing the prediction performance of the GRNN algorithm. In contrast to the other models, where data normalization proved beneficial, it appears that the GRNN algorithm performed better with raw data, without the need for normalization. These results highlight the importance of considering the characteristics of the specific algorithm and dataset when choosing data preprocessing techniques. While data normalization has shown advantages in certain cases, it may not always lead to improved performance, as observed with the GRNN algorithm in this study. In a discussion, it's noteworthy that the main distinguishing feature of GRNN is its single learning pattern, which sets it apart from other ANN patterns. Consequently, normalization methods may have a lesser impact on the performance of predictions.

Table 1 presents the results of CVRMSE and NMBE for the four ANNs using raw data and four normalization models. For the LSTM model, Min-Max normalization achieved the lowest CVRMSE (10.3) and NMBE (0.6) values, indicating its superior performance compared to other ANN methods. The LMBP model with Z-score normalization (CVRMSE: 20.0, NMBE: 1.1) also showed good performance in predicting electricity consumption. However, the differences in CVRMSE and NMBE among the LMBP models were relatively small, ranging from 0.6 to 11.7 and 0 to 0.6, respectively. In contrast, the RNN model with Gaussian normalization demonstrated the lowest CVRMSE (11.8) and NMBE (0.6) values, making it a favorable choice for electricity consumption prediction. Additionally, using raw data with the GRNN model resulted in the lowest CVRMSE (19.2) and NMBE (1.0) values compared to the four normalization methods. And The differences in prediction performance between the LMBP and GRNN models were not significant. The Z-score normalization showed some advantage for the LMBP models compared to raw data, although the improvement was small. In contrast, for the GRNN model, the effect of normalization was negligible, as raw data performed better than all four data normalization methods. Based on these results, data normalization is suggested for the RNN model, as all four normalization methods indicated better prediction performance than using raw data. However, for the LMBP and GRNN models, the choice of normalization method may depend on the specific application and the trade-off between improved performance and computational

Table 1
Comparison of Performance of four ANNs using raw data and four data normalization methods (R: raw data, MM: Min-Max, M: Mean, Z: Z-score and G: Gaussian normalization).

ANNs	Average error rates, %					CVRMSE, %					NMBE, %				
	R	MM	M	Z	G	R	MM	M	Z	G	R	MM	M	Z	G
LSTM	−1.8	−0.4	1.7	−3.3	−4.2	24.5	10.3	21.1	72.7	98.3	1.3	0.6	−1.2	4.0	5.4
LMBP	0.21	−2.4	−1.1	0.19	1.2	20.6	67.2	40.0	20.0	8.9	1.1	3.7	2.2	1.1	0.5
RNN	−3.7	0.4	0.9	−0.7	1	94.6	22.5	16.2	48.1	11.8	5.2	1.2	0.9	2.6	0.6
GRNN	0.3	6.0	0.9	0.8	8.4	19.2	68.8	60.6	74.0	44.7	1.0	3.8	3.3	4.0	2.4

complexity. In conclusion, data normalization has a noticeable impact on the prediction performance of different ANN models, and its effectiveness varies depending on the specific algorithm and dataset. The results from Table 1 provide valuable insights into the advantages and limitations of data normalization in enhancing prediction accuracy for electricity consumption forecasting using different ANN models.

This study aimed to evaluate the correlation between data normalization and its impact on actual electricity consumption in a building. To achieve this, four ANN models and four different data normalization methods were employed to predict electricity consumption profiles. Based on the results of the predicted data, it can be observed that the LSTM model with Min-Max normalization exhibits higher prediction accuracy compared to other data normalization methods and using raw data directly. The Min-Max normalization technique, which scales the data to a specific range, has proven to be particularly effective in improving the LSTM model's ability to capture patterns and relationships within the dataset. By addressing the varying feature scales and bringing the data to a standardized form, Min-Max normalization enhances the model's generalization capabilities and reduces the impact of outliers, resulting in more accurate predictions. And this study also recommends the use of Min-Max normalization for the LSTM and RNN algorithms. This normalization method demonstrated superior performance in these models, leading to improved prediction accuracy and reduced error rates. For the LMBP algorithm, Z-score normalization is suggested, as it showed good performance in predicting electricity consumption compared to using raw data. However, the improvements achieved with Z-score normalization were relatively small, indicating that the choice of normalization method in the LMBP model may depend on the specific application and requirements. In the case of the GRNN model, the study found that none of the data normalization methods were recommended. Using raw data directly yielded better performance in the GRNN model compared to all four normalization methods. This suggests that data normalization may not have a significant impact on the prediction accuracy of the GRNN algorithm. Overall, the study highlights the varying effects of data normalization on different ANN models. Gaussian normalization method, in particular, showed less favorable performance in the LSTM and RNN models but exhibited better results in the GRNN model. The findings of this study provide valuable insights into the importance of data normalization in improving prediction accuracy for electricity consumption forecasting. By recommending specific normalization methods for each ANN model, the study contributes to a better understanding of the role of data preprocessing in enhancing the performance of machine learning algorithms for energy consumption prediction. The pre-processing methods of data normalization, which aim to adjust or scale the data, guaranteeing that each feature has a uniform impact on the analysis, can increase accuracy when utilizing machine learning prediction modeling. It is essential to consider these results in practical applications and select the most suitable data normalization method based on the characteristics of the dataset and the specific requirements of the prediction task.

This study has successfully shed light on the impact of data normalization on electricity consumption prediction using different ANN models. However, there are several limitations that warrant further exploration in future research. Firstly, the study focused on a specific campus building for data collection, which may not fully represent the diverse range of building types and occupancy patterns. The accuracy and error rates of electricity consumption prediction could vary for other building types due to differences in weather conditions, occupant behavior, plug load and thermal energy demand. To address this limitation, future studies could include data from a wider variety of building types to obtain a more comprehensive understanding of the generalizability of the prediction models with variable normalization methods. Secondly, while the ANN models used in this study have shown promising results, there are other advanced ANN algorithms that could potentially further improve prediction accuracy. Exploring the performance of different ANN algorithms in conjunction with various normalization methods could provide valuable insights into finding the best combination for electricity consumption prediction tasks. Additionally, recent developments in data normalization methods could be investigated to determine their impact on prediction accuracy. Furthermore, validating the predictive models with data normalizations on an external dataset or conducting long-term monitoring in real-world scenarios can provide a more robust evaluation of the model's performance under different conditions. In conclusion, while this study has made significant contributions to understanding the role of data normalization in electricity consumption prediction using ANN models, there are ample opportunities for future research to address the limitations and further enhance prediction accuracy and model generalizability.

By expanding the scope of data collection, exploring advanced ANN algorithms, and considering recent normalization methods, researchers can continue to make strides towards more accurate and reliable electricity consumption prediction models for various building types and applications.

In this study, several limitations warrant further exploration through subsequent research. The prediction of electricity consumption was based on less than one year of data. However, long-term electricity consumption rates may vary depending on factors such as building type and climatic data fluctuations. Therefore, future studies should incorporate additional long-term data and parameters to reduce prediction error rates.

6. Conclusion

This study highlights the novel analysis strategy developed to understand how data normalization methods using ANNs impact the prediction of electricity consumption in a building. The study focused on evaluating the correlation between each data normalization method and its combination with four ANN algorithms on the electricity energy consumption in a campus building. The four ANN models - LSTM, LMBP, RNN, and GRNN - were designed with input nodes from weather data (direct normal irradiation, wind speed, cloud type, temperature, and humidity ratio) and occupancy rates. These models were then evaluated using four different data normalization methods - Min-Max, Mean, Z-score, and Gaussian normalization - on a training dataset and a test set. The results showed that the LSTM algorithm with Min-Max normalization achieved the lowest CVRMSE (10.3) and NMBE (0.6) values, demonstrating superior performance and stability compared to the other ANN methods with data normalization. The other three methods also showed good agreement with the actual experimental data, although the accuracy differences were relatively small. The LMBP model with Z-score normalization (CVRMSE: 20.0, NMBE: 1.1) showed good performance in predicting electricity consumption, but the differences among the LMBP models were not substantial. The RNN model with Gaussian normalization demonstrated the lowest CVRMSE (11.8) and NMBE (0.6) values, making it a favorable choice for electricity consumption prediction. The GRNN model using raw data yielded the lowest CVRMSE (19.2) and NMBE (1.0) values compared to the four normalization methods, and there were no significant differences between the LMBP and GRNN models in predicting electricity consumption. Overall, data normalization had a notable impact on the prediction performance of the different ANN models, and the effectiveness varied depending on the specific algorithm and dataset. The findings underscore the importance of selecting appropriate data normalization methods based on the specific application and the trade-off between improved performance and computational complexity. The proposed ANN models and data normalization strategies are valuable in predicting long-term energy consumption in buildings. They offer insights into building performance evaluation and can be used to assess the efficiency of different normalization methods and ANN models for specific building systems and data types. Future research could further enhance the accuracy of predictions by developing additional input elements and exploring more effective combinations of ANN models with specific data normalization strategies. In conclusion, the study presents a comprehensive analysis of the impact of data normalization on electricity consumption prediction using ANN models and offers valuable insights for future research in building energy consumption forecasting.

CRediT authorship contribution statement

Yang-Seon Kim: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Moon Keun Kim:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nuodi Fu:** Writing – review & editing, Methodology, Formal analysis. **Jiying Liu:** Writing – review & editing, Validation, Supervision. **Junqi Wang:** Writing – review & editing, Validation, Methodology. **Jelena Srebric:** Writing – review & editing, Supervision, Project administration, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Abdulrahman, M. L., Ibrahim, K. M., Gital, A. Y., Zambuk, F. U., Ja'afaru, B., Yakubu, Z. I., & Ibrahim, A. (2021). A review on deep learning with focus on deep recurrent neural network for electricity forecasting in residential building. *Procedia Computer Science, 193*, 141–154. https://doi.org/10.1016/j.procs.2021.10.014

Agga, A., Abbou, A., Labbadi, M., Houm, Y. E., & Ou Ali, I. H. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electric Power Systems Research, 208*, Article 107908. https://doi.org/10.1016/j.epsr.2022.107908

Ahmad, T., & Chen, H. X. (2019). Deep learning for multi-scale smart energy forecasting. *Energy, 175*, 98–112. https://doi.org/10.1016/j.energy.2019.03.080

Ali, P. J. M. (2022). Investigating the impact of min-max data normalization on the regression performance of k-nearest neighbor with different similarity measurements. *Aro-the Scientific Journal of Koya University, 10*(1), 85–91. https://doi.org/10.14500/aro.10955

Almehrizi, R. (2016). Normalization of mean squared differences to measure agreement for continuous data. *Statistical Methods in Medical Research, 25*(5), 1955–1974. https://doi.org/10.1177/0962280213507506

Andrean, V., Chang, P., & Lian, K. (2018). A review and new problems discovery of four simple decentralized maximum power point tracking algorithms—Perturb and observe, incremental conductance, golden section search, and Newton's quadratic interpolation. *Energies, 11*(11), 2966. http://www.mdpi.com/1996-1073/11/11/2966.

Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S., & Wattam, S. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews, 130*, Article 109899. https://doi.org/10.1016/j.rser.2020.109899

ASHRAE, A. (2002). Ashrae guideline 14: Measurement of energy and demand savings. *American Society of Heating, Refrigerating and Air-Conditioning Engineers, 35*, 41–63.

Azadeh, A., Ghaderi, S., & Sohrabkhani, S. (2008). Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion and Management, 49*(8), 2272–2278.

Babalhavaeji, A., Radmanesh, M., Jalili, M., & Gonzalez, S. A. (2023). Photovoltaic generation forecasting using convolutional and recurrent neural networks. *Energy Reports, 9*, 119–123. https://doi.org/10.1016/j.egyr.2023.09.149

Ballabio, D., & Vasighi, M. (2012). A MATLAB toolbox for self organizing maps and supervised neural network learning strategies. *Chemometrics and Intelligent Laboratory Systems, 118*, 24–32. https://doi.org/10.1016/j.chemolab.2012.07.005

Ben-Nakhi, A. E., & Mahmoud, M. A. (2004). Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management, 45*(13–14), 2127–2141. https://doi.org/10.1016/j.enconman.2003.10.009

Bergman, D., Cardonha, C., Imbrogno, J., & Lozano, L. (2023). Optimizing the expected maximum of two linear functions defined on a multivariate Gaussian distribution. *Informs Journal on Computing, 35*(2), 304–317. https://doi.org/10.1287/ijoc.2022.1259

Blokhintsev, L. D., & Savin, D. A. (2021). Study of the influence of different methods of taking into account the coulomb interaction on determining asymptotic normalization coefficients within the framework of exactly solvable model. *Physics of Atomic Nuclei, 84*(4), 401–407. https://doi.org/10.1134/S1063778821040098

Bocheng, Z., Kuo, L., Dinghao, L., Jing, L., & Xuan, F. (2015). Short-term prediction of building energy consumption based on GALM neural network. In *International conference on advances in mechanical engineering and industrial informatics (AMEII)*.

Brezak, D., Bacek, T., Majetic, D., Kasac, J., & Novakovic, B. (2012). A comparison of feed-forward and recurrent neural networks in time series forecasting. In *2012 IEEE conference on computational intelligence for financial engineering & economics (CIFEr), 29–30 March 2012*.

Calderano, P. H. S., Gheorghe, D. R. M., Teixeira, R. S., Amaral, R. P. F., & Menezes, I. F. M. (2023). Type-1 and singleton fuzzy logic system binary classifier trained by BFGS optimization method. *Fuzzy Optimization and Decision Making, 22*(1), 149–168. https://doi.org/10.1007/s10700-022-09387-y

Chae, Y. T., Horesh, R., Hwang, Y., & Lee, Y. M. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings, 111*, 184–194. https://doi.org/10.1016/j.enbuild.2015.11.045

Cossu, A., Carta, A., Lomonaco, V., & Bacciu, D. (2021). Continual learning for recurrent neural networks: An empirical evaluation. *Neural Networks, 143*, 607–627. https://doi.org/10.1016/j.neunet.2021.07.021

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin, 145*(8). https://doi.org/10.1037/bul0000199

Daut, M. A. M., Hassan, M. Y., Abdullah, H., Rahman, H. A., Abdullah, M. P., & Hussin, F. (2017). Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. *Renewable and Sustainable Energy Reviews, 70*, 1108–1118. https://doi.org/10.1016/j.rser.2016.12.015

Delattre, S., & Roquain, E. (2016). On empirical distribution function of high-dimensional Gaussian vector components with an application to multiple testing. *Bernoulli, 22*(1), 302–324. https://doi.org/10.3150/14-Bej659

Elkhouly, A., Andrew, A. M., Rahim, H. A., Abdulaziz, N., Malek, M. F. A., & Siddique, S. (2023). Data-driven audiogram classifier using data normalization and multi-stage feature selection. *Scientific Reports, 13*(1). https://doi.org/10.1038/s41598-022-25411-y

Faiq, M., Tan, K. G., Liew, C. P., Hossain, F., Tso, C. P., Lim, L., Wong, A. Y. K., & Shah, Z. M. (2023). Prediction of energy consumption in campus buildings using long short-term memory. *Alexandria Engineering Journal, 67*, 65–76. https://doi.org/10.1016/j.aej.2022.12.015

Fang, L., & He, B. (2023). A deep learning framework using multi-feature fusion recurrent neural networks for energy consumption forecasting. *Applied Energy, 348*, Article 121563. https://doi.org/10.1016/j.apenergy.2023.121563

Feng, Z., Luo, X., Wang, J., & Cao, S.-J. (2022). Energy-efficient preservation environment control for enclosed exhibition hall of earthen relics. *Energy and Buildings, 256*, Article 111713. https://doi.org/10.1016/j.enbuild.2021.111713

Gutierrez-Osuna, R., & Nagle, H. T. (1999). A method for evaluating data-preprocessing techniques for odor classification with an array of gas sensors. *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics, 29*(5), 626–632. https://doi.org/10.1109/3477.790446

Hao, Y., Li, H., & Wang, H. (2013). Analysis of energy consumption and energy-saving of residential buildings based on LM-BP algorithm. *International Journal of Applied Mathematics and Statistics, 51*(24), 348–356.

Henn, M. C., Brady, B. D., Kar, J., Cupps, B. P., Wallace, K., Braverman, A. C., & Pasque, M. K. (2014). Quantification and Z-score "normalization" of regional myocardial contractile data in ischemic coronary artery disease patients. *Circulation, 130*(suppl_2). https://doi.org/10.1161/circ.130.suppl_2.20288

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

IEA, I.E.A. (2023a). *2018 Global Status Report—Analysis*. https://www.iea.org/reports/2018-global-status-report.

IEA, I.E.A. (2023b). *Buildings—Energy System*. https://www.iea.org/energy-system/buildings.

Ilbeigi, M., Ghomeishi, M., & Dehghanbanadaki, A. (2020). Prediction and optimization of energy consumption in an office building using artificial neural network and a genetic algorithm. *Sustainable Cities and Society, 61*, Article 102325. https://doi.org/10.1016/j.scs.2020.102325

Jain, R. K., Smith, K. M., Culligan, P. J., & Taylor, J. E. (2014). Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Applied Energy, 123*, 168–178. https://doi.org/10.1016/j.apenergy.2014.02.057

Karim, M. N., & Rivera, S. L. (1992). Comparison of feed-forward and recurrent neural networks for bioprocess state estimation. *Computers and Chemical Engineering, 16*, S369–S377. https://doi.org/10.1016/S0098-1354(09)80044-6

Kim, M. K., Cha, J., Lee, E., Pham, V. H., Lee, S., & Theera-Umpon, N. (2019). Simplified neural network model design with sensitivity analysis and electricity consumption prediction in a commercial building. *Energies, 12*(7). https://doi.org/10.3390/en12071201

Kim, M. K., Cremers, B., Liu, J., Zhang, J., & Wang, J. (2022). Prediction and correlation analysis of ventilation performance in a residential building using artificial neural network models based on data-driven analysis. *Sustainable Cities and Society,, 83*, Article 103981. https://doi.org/10.1016/j.scs.2022.103981

Kim, M. K., Kim, Y.-S., & Srebric, J. (2020a). Impact of correlation of plug load data, occupancy rates and local weather conditions on electricity consumption in a building using four back-propagation neural network models. *Sustainable Cities and Society, 62*, Article 102321. https://doi.org/10.1016/j.scs.2020.102321

Kim, M. K., Kim, Y.-S., & Srebric, J. (2020b). Predictions of electricity consumption in a campus building using occupant rates and weather elements with sensitivity analysis: Artificial neural network vs. linear regression. *Sustainable Cities and Society, 62*, Article 102385. https://doi.org/10.1016/j.scs.2020.102385

Kim, Y.-S., Heidarinejad, M., Dahlhausen, M., & Srebric, J. (2017). Building energy model calibration with schedules derived from electricity use data. *Applied Energy, 190*, 997–1007. https://doi.org/10.1016/j.apenergy.2016.12.167

Kong, W. C., Dong, Z. Y., Jia, Y. W., Hill, D. J., Xu, Y., & Zhang, Y. (2019). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid, 10*(1), 841–851. https://doi.org/10.1109/Tsg.2017.2753802

Kreider, J. F., Claridge, D. E., Curtiss, P., Dodier, R., Habert, J. S., & Krarti, M. (1995). Building energy use prediction and system-identification using recurrent neural networks. *Journal of Solar Energy Engineering-Transactions of the ASME, 117*(3), 161–166. https://doi.org/10.1115/1.2847757

Kurnaz, G., & Demir, A. S. (2022). Prediction of SO2 and PM10 air pollutants using a deep learning-based recurrent neural network: Case of industrial city Sakarya. *Urban Climate, 41*, Article 101051. https://doi.org/10.1016/j.uclim.2021.101051

Laboratory, T.N.R.E. (2023). *NSRDB: National Solar Radiation Database*. https://nsrdb.nrel.gov/.

Lapuschkin, S., Binder, A., Montavon, G., Muller, K. R., & Samek, W. (2016). The LRP toolbox for artificial neural networks. *Journal of Machine Learning Research, 17*(114), 1–5.

Lee, W. Y., House, J. M., & Kyong, N. H. (2004). Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks. *Applied Energy, 77*(2), 153–170. https://doi.org/10.1016/S0306-2619(03)00107-7

Lei, R. C., & Yin, J. (2022). Prediction method of energy consumption for high building based on LMBP neural network. *Energy Reports, 8*, 1236–1248. https://doi.org/10.1016/j.egyr.2022.02.071

Lek, S., & Guégan, J. F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling, 120*(2), 65–73. https://doi.org/10.1016/S0304-3800(99)00092-7

Li, H. Z., Tao, W., Gao, T., Li, H., Lu, Y. H., & Su, Z. M. (2011). Improving the accuracy of density functional theory (DFT) calculation for homolysis bond dissociation energies of Y-NO bond: Generalized regression neural network based on grey relational analysis and principal component analysis. *International Journal of Molecular Sciences, 12*(4), 2242–2261. https://doi.org/10.3390/ijms12042242

Li, Y., Tong, Z., Tong, S., & Westerdahl, D. (2022). A data-driven interval forecasting model for building energy prediction using attention-based LSTM and fuzzy information granulation. *Sustainable Cities and Society, 76*, Article 103481. https://doi.org/10.1016/j.scs.2021.103481

Liu, X. Y., Li, N., Liu, S., Wang, J., Zhang, N., Zheng, X. B., Leung, K. S., & Cheng, L. X. (2019). Normalization methods for the analysis of unbalanced transcriptome data: A review. *Frontiers in Bioengineering and Biotechnology, 7*, Article 358. https://doi.org/10.3389/fbioe.2019.00358

Lu, C. J., Li, S. H., & Lu, Z. J. (2022). Building energy prediction using artificial neural networks: A literature survey. *Energy and Buildings, 262*, Article 111718. https://doi.org/10.1016/j.enbuild.2021.111718

Meng-Hock, F., & Hagan, M. T. (1996). Levenberg–Marquardt training for modular networks. In *Proceedings of international conference on neural networks (ICNN'96), 3–6 June 1996*.

Meng, A., Wang, P., Zhai, G., Zeng, C., Chen, S., Yang, X., & Yin, H. (2022). Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization. *Energy, 254*, Article 124212. https://doi.org/10.1016/j.energy.2022.124212

Mirza, A. F., Haider, S. K., Ahmed, A., Rehman, A. U., Shafiq, M., Bajaj, M., Zawbaa, H. M., Szczepankowski, P., & Kamel, S. (2022). Generalized regression neural network and fitness dependent optimization: Application to energy harvesting of centralized TEG systems. *Energy Reports, 8*, 6332–6346. https://doi.org/10.1016/j.egyr.2022.05.003

Moller, M. F. (1993). A scaled conjugate-gradient algorithm for fast supervised learning. *Neural Networks, 6*(4), 525–533. https://doi.org/10.1016/S0893-6080(05)80056-5

Moore, J. E., Kim, M., Seo, J. G., Wu, Y., & Kalaba, R. (1991). Linear-programming, recurrent associative memories, and feedforward neural networks. *Computers and Mathematics with Applications, 22*(11), 71–90. https://doi.org/10.1016/0898-1221(91)90036-4

Mortensen, K. N., Gjedde, A., Thompson, G. J., Herman, P., Paret, M. J., Rothman, D. L., Kupers, R., Ptito, M., Stender, J., Laureys, S., Riedl, V., Alkire, M. T., & Hyder, F. (2018). Impact of global mean normalization on regional glucose metabolism in the human brain. In *Neural Plasticity, 2018*, Article 6120925. https://doi.org/10.1155/2018/6120925

Park, J., & Kim, K.-Y. (2017). Meta-modeling using generalized regression neural network and particle swarm optimization. *Applied Soft Computing, 51*, 354–369. https://doi.org/10.1016/j.asoc.2016.11.029

Qian, F., Gao, W., Yang, Y., & Yu, D. (2020). Potential analysis of the transfer learning model in short and medium-term forecasting of building HVAC energy consumption. *Energy, 193*, Article 116724. https://doi.org/10.1016/j.energy.2019.116724

Qiong, L., Peng, R., & Qinglin, M. (2010). Prediction model of annual energy consumption of residential buildings. In *2010 international conference on advances in energy engineering, 19–20 June 2010*.

Ribeiro, M. V., Duque, C. A., & Romano, J. M. T. (2006). An interconnected type-1 fuzzy algorithm for impulsive noise cancellation in multicarrier-based power line communication systems. *IEEE Journal on Selected Areas in Communications, 24*(7), 1364–1376. https://doi.org/10.1109/Jsac.2006.874417

Robinson, C., Dilkina, B., Hubbs, J., Zhang, W. W., Guhathakurta, S., Brown, M. A., & Pendyala, R. M. (2017). Machine learning approaches for estimating commercial building energy consumption. *Applied Energy, 208*, 889–904. https://doi.org/10.1016/j.apenergy.2017.09.060

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533–536. https://doi.org/10.1038/323533a0

Runge, J., & Zmeureanu, R. (2019). Forecasting energy use in buildings using artificial neural networks: A review. *Energies, 12*(17), Article 3254. https://doi.org/10.3390/en12173254

Samy, M. M., Almamlook, R. E., Elkhouly, H. I., & Barakat, S. (2022). Decision-making and optimal design of green energy system based on statistical methods and artificial neural network approaches. *Sustainable Cities and Society, 84*, Article 104015. https://doi.org/10.1016/j.scs.2022.104015

Sendra-Arranz, R., & Gutierrez, A. (2020). A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy and Buildings, 216*, Article 109952. https://doi.org/10.1016/j.enbuild.2020.109952

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing, 97*, Article 105524. https://doi.org/10.1016/j.asoc.2019.105524

Somu, N., Raman, M. R. G., & Ramamritham, K. (2020). A hybrid model for building energy consumption forecasting using long short term memory networks. *Applied Energy, 261*, Article 114131. https://doi.org/10.1016/j.apenergy.2019.114131

Specht, D. F. (1991). A general regression neural network. *IEEE Transactions on Neural Networks, 2*(6), 568–576.

Swift, D., Cresswell, K., Johnson, R., Stilianoudakis, S., & Wei, X. T. (2023). A review of normalization and differential abundance methods for microbiome counts data. *Wiley Interdisciplinary Reviews-Computational Statistics, 15*(1), Article e1586. https://doi.org/10.1002/wics.1586

Tian, X., Cheng, Y., & Lin, Z. (2021). Modelling indoor environment indicators using artificial neural network in the stratified environments. *Building and Environment*. , Article 108581. https://doi.org/10.1016/j.buildenv.2021.108581

Van Houdt, G., Mosquera, C., & Napoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review, 53*(8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Wang, H., Wang, J., Feng, Z., Haghighat, F., & Cao, S.-J. (2023). An intelligent anti-infection ventilation strategy: From occupant-centric control and computer vision perspectives. *Energy and Buildings, 296*, Article 113403. https://doi.org/10.1016/j.enbuild.2023.113403

Wang, L., Lee, T. J., Bavendiek, J., & Eckstein, L. (2021). A data-driven approach towards the full anthropometric measurements prediction via generalized regression neural networks. *Applied Soft Computing, 109*, Article 107551. https://doi.org/10.1016/j.asoc.2021.107551

Wang, Z. Y., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews, 75*, 796–808. https://doi.org/10.1016/j.rser.2016.10.079

Wazirali, R., Yaghoubi, E., Abujazar, M. S. S., Ahmad, R., & Vakili, A. H. (2023). State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques. *Electric Power Systems Research, 225*, Article 109792. https://doi.org/10.1016/j.epsr.2023.109792

Weerakody, P. B., Wong, K. W., Wang, G., & Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. *Neurocomputing, 441*, 161–178. https://doi.org/10.1016/j.neucom.2021.02.046

Xu, B.Y., Zhang, H.J., Wang, Z.T., Wang, H.X., & Zhang, Y.L. (2015). Model and algorithm of BP neural network based on expanded multichain quantum optimization. *Mathematical Problems in Engineering.* 362150 https://doi.org/10.1155/2015/362150.

Yang, M., Shi, C., & Liu, H. (2021). Day-ahead wind power forecasting based on the clustering of equivalent power curves. *Energy, 218*, Article 119515. https://doi.org/10.1016/j.energy.2020.119515

Ye, Z. Y., & Kim, M. K. (2018). Predicting electricity consumption in a building using an optimized back-propagation and Levenberg–Marquardt back-propagation neural network: Case study of a shopping mall in China. *Sustainable Cities and Society, 42*, 176–183. https://doi.org/10.1016/j.scs.2018.05.050

Yoo, S., & Seul, N. (2017). Design of a control cable condition prediction system for nuclear power plants using LMBP neural networks. *Journal of Institute of Control, Robotics and Systems, 23*(3), 194–198.

Zhang, C., Wang, J. B., Li, L., Wang, F. F., & Gang, W. J. (2020). Utilization of earth-to-air heat exchanger to pre-cool/heat ventilation air and its annual energy performance evaluation: A case study. *Sustainability, 12*(20), Article 8330. https://doi.org/10.3390/su12208330

Zhang, H., Srinivasan, R., Yang, X., Ahrentzen, S., Coker, E. S., & Alwisy, A. (2022). Factors influencing indoor air pollution in buildings using PCA-LMBP neural network: A case study of a university campus. *Building and Environment, 225*, Article 109643. https://doi.org/10.1016/j.buildenv.2022.109643

Zhang, X., Guo, Z., Pan, F., Yang, Y., & Li, C. (2023). Dynamic carbon emission factor based interactive control of distribution network by a generalized regression neural network assisted optimization. *Energy, 283*, Article 129132. https://doi.org/10.1016/j.energy.2023.129132

Zhu, H.-C., Ren, C., & Cao, S.-J. (2022). Dynamic sensing and control system using artificial intelligent techniques for non-uniform indoor environment. *Building and Environment, 226*, Article 109702. https://doi.org/10.1016/j.buildenv.2022.109702