A multi-energy load prediction of a building using the multi-layer perceptron neural network method with different optimization algorithms

Zhongzhen Yan¹, Xinyuan Zhu¹, Xianglong Wang¹, Zhiwei Ye¹, Feng Guo², Lei Xie³,⁴ and Guiju Zhang⁵

Abstract
Since cooling and heating loads are recognized as key characteristics for evaluating the energy efficiency of buildings, it appears indisputable that they must be predicted and analyzed for residential structures. Accordingly, the multi-layer perceptron neural network is applied for predicting the heating and cooling loads using the experimental dataset. The used dataset is obtained by monitoring the impact of the building’s dimensions on energy consumption. To optimize the training process of the multi-layer perceptron neural network, several optimizers are employed. Besides, different statistical performance indicators are considered to see which selected optimizer outperforms the rest in terms of accuracy and authenticity. The obtained results emphasize the remarkable performance of adaptive chaotic grey wolf optimization, which can be used to train the multi-layer perceptron neural network and forecast the buildings’ energy consumption with the highest accuracy. According to the obtained results, the hybrid multi-layer perceptron neural network- adaptive chaotic grey wolf optimization method demonstrates the best performance. The optimum number of neurons in the hidden layer is obtained to be 15. Also, based on the statistical performance indicators, the selected method reveals an R² of 0.9123 and 0.9419 for cooling and heating loads, respectively.

¹School of Computer Science and Technology, Hubei University of Technology, Wuhan, China
²China Railway Seventh Bureau Group Electrical Engineering Co. Ltd, Zhengzhou, China
³Intelligent Transport Systems Research Center (ITSC), Wuhan University of Technology, Wuhan, China
⁴National Engineering Research Centre for Water Transport Safety (WTSC), Wuhan, China
⁵Key Laboratory of Hunan Province for Efficient Power System and Intelligent Manufacturing, Hunan, China

Corresponding author:
Lei Xie, Intelligent Transport Systems Research Center (ITSC), Wuhan University of Technology, Wuhan, China.
National Engineering Research Centre for Water Transport Safety (WTSC), Wuhan, China.
Email: xielei@whut.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Keywords
Building energy prediction, multi-layer perceptron neural network, chaotic grey wolf optimization, statistical performance indicators

Introduction
Nowadays, transportation, industrial sectors, and buildings have the main contribution to global energy consumption (Al-Homoud, 2001). About one-third of the total global energy is consumed by buildings yearly (Levesque et al., 2018). Overall, the buildings consume a considerable share of global energy, which is likely to rise to 32.4% by 2024 (Ngo et al., 2022). The energy consumption in buildings for the countries such as Australia, New Zealand, the United Kingdom, and the United States that are members of the Organization of Economic Cooperation and Development (OECD) has increased by 1.5% yearly from 2012 to 2040 (Bui et al., 2020). This rate is about 2.1% per year for the same period for the developing countries (Moldovan and Slowik, 2021).

Accordingly, the development of energy-efficient building systems is of great importance, which has been examined by many studies as well as the current paper (Ahmadi-Karvigh et al., 2018; Zanchini and Naldi, 2019). For instance, for a two-story office building in Cyprus, an organic Rankine cycle (ORC) powered by solar thermal collectors is used to power the air conditioning system using a Variable-Refrigerant-Flow (VRF) unit (Gilani et al., 2022). The study revealed that the system could achieve an energy saving of 7700 kWh, which is equal to 97% of the yearly energy demand of the VRF system. The energy, exergy, economics, and environmental impact of a photovoltaic thermal system incorporated into a building are examined by Shakouri et al. (2022). To ensure that it can be implemented as a feasible solution, the scenario that yielded the maximum income (55,157 USD) and the shortest payback period (1.58 years) was chosen as the most practical one for retrofitting. In another study, Sohani et al. (2022) studied the real-time dynamic multi-objective optimization of a building integrated photovoltaic thermal system improved by phase change materials (PCM). According to obtained results, it was revealed that the utilization of PCM instead of the air has some advantages, such as 17.69% lower CO₂ emission, 9.93% higher power generation, and 22.24% higher yearly energy storage. Gilani et al. (2021) examined the energy efficiency and power production potential of a VRF heat pump system coupled with photovoltaic (PV) technology in a five-story residential building in Cyprus. Also, different energy systems were suggested for providing heating and cooling loads (Mahmoudan et al., 2022; Mahmoudan et al., 2021).

According to most of the studies presented in this regard, the need for heating and cooling energy is mainly related to building energy consumption. Hence, it is essential to predict and minimize the heating and cooling loads when designing an energy-efficient building (Pérez-Lombard et al., 2008). Intending to specify an optimal solution for reducing energy consumption, the designers need to have access to some building designs or Heating Ventilation and Air Conditioning (HAVC) system control options (Bui et al., 2020). For example, Frayssinet et al. (2018) aimed to estimate the heating and cooling energy demand for the buildings at the city scale. Also, in this study, the related modeling methods that consist of detailed approaches were examined. The concept of the fourth generation of district heating (4DH), including the combination of heating, cooling, and power sector, was investigated by Dorotić et al. (2019), and acceptable results were presented. Sedighizadeh et al. (2018) proposed an energy management model on a residential microgrid in which the combined cooling, heating, and power, plug-in hybrid electric...
vehicles, photovoltaic unit, and battery energy storage systems were considered. Besides, the efficient scheduling state of such units was specified considering the uncertainty of distributed energy resources. A model-based predictive control method was also proposed by Ferreira et al. (2012) to control the HAVC system in a building, and the energy consumption was reduced by 50%. In another study, the energy consumption of HVAC was precalculated to optimize the setpoint and deadband parameter of the HVAC system considering a systematic technique (Ghahramani et al., 2016). Considering such studies, the total energy consumption of a building can be significantly minimized when the heating and cooling loads are predicted beforehand.

It should be noted that parameters such as temperature, sunlight equipment, occupant behavior, wall materials, glazing area, surface, height, and volume of the building are important to specify the interactions between the general energy requirements of the building (Barkokebas et al., 2019; Ghahramani et al., 2016). For example, according to the study of Ihara et al. (2015), it was revealed that the whole façade properties consisting of solar reflectance, U-value, and solar heat gain coefficient (SHGC) have a variety of impacts on the energy efficiency of buildings. The authors stated that the energy consumption is minimized when the SHGC is significantly reduced. Liu et al. (2019) have an opposing idea and concluded that the SHGC reduction could not be the main reason for enhancing the energy efficiency in the building without a suitable U-value. Since the weather conditions, building dimensions, and the behavior of residents are also influential in the whole energy consumption, calculating the energy consumption is significantly difficult. Hence, using an artificial neural network (ANN) can be a good idea to specify the heating and cooling loads for forecasting the energy consumption in the buildings.

ANN is one of the most significant artificial intelligence algorithms utilized in various applications. This approach was developed to simulate the architecture and activities of human brain cells during learning and thinking. These cells are known as neurons, and there are more than 50 billion of them in the average human brain (AlShabi and Assad, 2021). ANN approaches are among the methodologies used to estimate the performance of renewable energy systems. In applications involving renewable and sustainable energy, ANN approaches are widely used. For instance, Rashidi et al. (2022) studied the thermal conductivity of EG-Water-based nano-fluids containing alumina particles using Multi-Layer Perceptron (MLP) neural network and Group Method of Data Handling (GMDH) as two effective intelligent techniques. According to the obtained results, the coefficient of determination ($R^2$) for MLP neural network with tansig and radial basis functions and GMDH were obtained to be 0.9998, 0.9998, and 0.9996, correspondingly. In another application of ANN models, ANN with two transfer functions, including normalized radial basis and Tansig, was utilized by Komeili Birjandi et al. (2022) to estimate CO$_2$ emission from Southeast Asian countries. According to the results, using a network with a normalized radial basis and 11 neurons in the hidden layer yielded the most accurate model with an $R^2$ of 0.9997.

Moreover, ANN models have acquired popularity in the area of energy forecasting of buildings owing to the extraordinary outcomes they deliver. Neto and Fiorelli (2008) indicated that the energy consumption could be accurately predicted through the simulation method and ANN; however, using ANN minimizes the computational cost and is efficient in terms of the expedition. There exists a considerable body of literature on using the neural network method in this regard. To mention a few, Biswas et al. (2016) showed how ANN models might be developed and validated in the TxAIRE Research homes. The number of days, temperature, and solar radiation are all input factors utilized from the home data, whereas house energy use and heat pump energy usage are output variables. Promising findings with $R^2$ in the range of 0.87–0.91 were obtained using the Levenberg-Marquardt and OWO-Newton algorithms. Li et al. (2018) proposed a hybrid...
teaching-learning artificial neural network to predict the main parameters for electrical energy. In the cutting-edge paper, the issue of changes in the energy consumption using ANN for forecasting the user-based energy consumption was examined by Lee et al. (2019), and acceptable results were presented. The obtained results imply a relationship between the characteristics and energy usage. Also, Ilbeigi et al. (2020) attempted to introduce a trustworthy method for optimizing the energy consumption of buildings in 2020. In order to simulate the energy consumption, a robust ANN based on the MLP model was created, trained, and tested. In a major advance of 2021, Wang et al. (2021) used a deep convolutional neural network to predict the building load and presented remarkable results. Notably, a novel feature fusion was presented to improve the learning ability of the model. Despite such interest, there are still many concerns regarding the accuracy and ability of the ANNs since the training process is not always efficient. To tackle this problem, optimizers can be used to optimize the training procedure.

A large number of existing studies in the broader literature have examined the capabilities of the various optimizers for training the ANNs in the field of energy consumption prediction. For instance, the issue of optimization in the heating, HAVC system operations, and other building parameters using ANN and multi-objective genetic algorithm (MOGA) was examined by Satrio et al. (2019). The aim was to minimize the annual energy consumption and maximize thermal comfort. Notably, the two-chiller system operation of a building was optimized through the applied techniques, and acceptable results were presented. Furthermore, in a major advance of 2021, Chou and Truong (2021) used the nonlinear machine learning models and specified the historical pattern of regional energy consumption to propose a new prediction system that optimizes the linear time series. Also, the capabilities of the machine learning models in predicting the HL and CL of the building were exploited by Sauer et al. (2021). Zhou et al. (2021) evaluated the energy performance of building using the multi-layer perceptron neural network modified by teaching-learning-based metaheuristic method. As a result, the proposed method’s prediction error decreased by around 20%, and the correlation between measured and predicted cooling loads increased from 0.8875 to 0.9207. To boost energy efficiency without deteriorating comfort, Ruiz et al. (2018) proposed Elman neural networks with the genetic algorithm for forecasting energy demand in public buildings. According to the obtained results, an average improvement of 61% was obtained.

**Main contributions and novelties**

Using energy prediction tools is important to improve decision-making when it comes to reducing the amount of energy used in buildings since they can analyze a wide range of building designs and methodologies. However, several important factors contribute to a building’s energy consumption, including weather conditions, the behavior of its inhabitants, and the installed technology and equipment. This makes energy prediction a difficult scientific topic to solve. However, progress has been made in the area of energy consumption evaluation in sustainable buildings. Artificial Intelligence-based methods have recently captivated the interest of scientists and have been applied in a broad variety of applications.

Despite the significance of the above-mentioned reviewed papers, there are some research gaps:

- The review studies do not concentrate on the impact of “building characteristics” on energy efficiency. Various features/characteristics, including temperature, weather, and date, have been used to estimate energy consumption. However, there aren’t many studies on building energy consumption prediction research that concentrates on the specifics of the building (surface...
area, wall area, roof area, relative compactness, overall height, orientation, glazing area, and glazing area distribution), which are the most important aspects to take into account during the design phase.

- The reviewed studies do not pay much attention to the categorization/combination of prediction techniques and hybrid optimization algorithms. The current work aims at giving a new combination of prediction techniques and optimizers with thorough descriptions, statistics, and comparisons of the used approaches.

- Existing works do not offer a mapping between energy consumption forecast methodologies and appropriate performance indicators in buildings.

To bridge these gaps, this research aims at using different optimization approaches to improve the multi-layer perceptron neural network in order to forecast the heating and cooling of residential buildings. Surface area, wall area, roof area, relative compactness, overall height, orientation, glazing area, and glazing area distribution are the eight independent parameters employed in this research. Test and training data are also evaluated to demonstrate the dependability and accuracy of the findings. Therefore, the MLP neural network is considered for the process of prediction in which genetic algorithm (GA), multi-verse optimizer (MVO), chaotic grey wolf optimization (CGWO), adaptive chaotic grey wolf optimization (ACGWO), augmented grey wolf optimizer (AGWO), grey wolf optimizer (Lee et al.), particle swarm optimization (PSO) are employed to optimize the hyperparameters of MLP. Additionally, the statistical analysis considering standard error (SE), mean squared error (MSE), root mean squared errors (RMSE), mean absolute percent error (MAPE), mean absolute error (MAE), relative absolute error (RAE), correlation coefficient (R), Coefficient of determination (R²), and normalized mean squared errors (NMSE) as the statistical performance indicators were used to see which model is the best in terms of accuracy, expedition, and authenticity.

**Methodology**

The methods employed to analyze the problem are described in this section. In fact, the heating and cooling loads are predicted based on the experimental data presented in this paper. For this purpose, six meta-heuristic algorithms, namely GA, MVO, CGWO, ACGWO, AGWO, GWO, and PSO, are considered to optimize the MLP neural network that are explained here. Notably, a statistical analysis is also conducted here to introduce the best prediction technique among the selected ones.

**Multi-layer perceptron (MLP) neural network**

Artificial neural networks (ANNs) have received much attention due to their remarkable applications for modeling engineering parameters and are widely used to solve complicated problems. In fact, ANNs are popular due to their striking features such as adaptive learning, self-organization, real-time operators, generalization, stability and flexibility, and parallel processing. ANN is mainly used to estimate functions, make predictions, pattern recognition, control, and so on (Bui et al., 2018).

A multi-layer perceptron (MLP) which is considered a subbranch of a feedforward artificial neural network (ANN), is employed in the present work to specify the unknown parameters of HL and CL. The aim here is to predict the energy consumption in buildings and minimize it significantly according to the values of HL and CL. Accordingly, the MLP is utilized to estimate any nonlinear function with the highest accuracy and identify the system’s unknown parameters.
However, the outputs are sometimes not efficient as expected, and the training process needs to be optimized using the optimizers in this case. Hence, several optimizers are used in this study to reach the best parameters and examine the HL and CL that are explained in the following.

**Optimization methods**

The optimization methods considered to specify the best parameters of the CL and HL and train the MLP neural network are as follows: GA, MVO, CGWO, ACGWO, AGWO, GWO, and PSO.

As shown in Figure 1, the optimization process is conducted for eight inputs, namely relative compactness, surface area, wall area, roof area, overall height, glazing area, orientation, and glazing area distribution. Also, cooling and heating loads are considered the outputs in this analysis.

According to Figure 1, the input and output data are pre-processed and normalized in the beginning, and then the network structure is determined. After that, the activation function, parameters, hidden layers, maximum number of neurons in the hidden layer, and the performance function are selected. Then, the main processes of training for the network based on the appropriate algorithms are conducted. Notably, the best algorithm in terms of accuracy and expedition is specified through the statistical analyses.

*Genetic algorithm and particle swarm optimization.* The genetic algorithm has a vast range of applications for optimizing the training process in the MLP neural network. Due to the remarkable ability to create high-quality solutions for optimization and search problems, this algorithm is employed...
here, relying on biologically inspired operators such as mutation, crossover, and selection (Gerges et al., 2018).

Moreover, PSO has significantly drawn the researchers’ attention for efficiently identifying the unknown parameters and optimizing the network training process (Askarzadeh and Rezazadeh, 2011; Sedighizadeh et al., 2011). Basically, the emulation of birds’ movements constituted the PSO, which mainly suggests some initial solutions (Takagi and Sugeno, 1985). As the name of this optimizer implies, the swarm means a that is also indicated by the set of particles. The particles of the swarm are considered for searching the space to specify the efficient solution and best location. Accordingly, the swarm’s particle constantly corrects its location. Eqs. 1 and 2 indicate the best swarm location and particle location that Gbest and Pbest can show, respectively (Diab and Rezk, 2017; Mohamed et al., 2019).

\[
P^d_i(k + 1) = P^d_i(k) + v^d_i(k + 1)
\]

\[
v^d_i(k + 1) = w v^d_i(k) + c_1 r_1(P_{\text{best}_i}^d - P^d_i(k)) + c_2 r_2(G_{\text{best}_i}^d - P^d_i(k))
\]

According to the equations above, \( k \) is the iteration, and a particle \( i \) changes its position in a dimension \( d \) from location \( P(k) \) to location \( P(k + 1) \), which depends on these equations.

**Figure 2.** Hierarchy and schematic of grey wolf optimizer.
Figure 3. Pseudocode of the GWO algorithm (Mirjalili et al., 2014).

Figure 4. The nonlinear convergence factor $a$ proposed in the AGWO algorithm.
Figure 5. The flowchart of ACGWO.
weight of inertia is indicated by $w$. Besides, $c_1$ and $c_2$ are the two scaling constants based on which the local experience and the global experience weights are determined. Furthermore, $r_1$ and $r_2$ are considered to represent the two random parameters according to their variations in the closed range of $[0, 1]$, which indicates the stochastic behavior of the optimizer.

**Multi-verse optimizer.** The current paper has also exploited the MVO as a new stochastic population-based algorithm to optimize the process of training in the MLP neural network (Mirjalili et al., 2016). This algorithm is basically inspired by the theory of multi-verse in physics. Notably, the MVO is based on three concepts of the multi-verse theory, including the white hole, black hole, and wormhole, that are mathematically modeled to construct this optimizer.

**Grey wolf optimizer and chaotic grey wolf optimization.** Many studies emphasize the benefits of GWO, which is inspired by grey wolves (Canis lupus) (Mirjalili et al., 2014). As shown in Figure 2, the leadership hierarchy and hunting mechanism of grey wolves in nature is the base of this optimizer according to which the optimization process is conducted.

Grey wolf or Canis lupus which belongs to the Canidae family, is based on the mechanism saying that grey wolves are regarded as apex predators and are at the top of the food chain. These animals mainly tend to live in packs with a population of 5–12 wolves on average. According to Figure 2, a considerably strict dominant hierarchy is considered for them. The male and female can be the leaders in this pack and are categorized in the alphas group responsible for deciding hunting, sleeping place, wake-up time, etc. Hence the members of the pack are supposed to follow the alphas’ orders. Accordingly, the beta category belongs to the subordinate wolves, including males and females, that aid the alpha in decision-making and other pack activities. Also, the omega group, which includes the lowest-ranked wolves, is related to the ones that must always (Mech, 1999) follow the rules and mainly eat the leftovers. Lastly, the delta group is related to the ones that are not alpha, beta, or omega. Based on the study of Muro et al. (2011), the major steps of grey wolf hunting can be described as follows:

![Figure 6. The buildings’ shapes and values of rc considered in this study.](image)
Table 1. The values of the main parameters considered in CSGWO.

| Parameter                                      | Value     |
|------------------------------------------------|-----------|
| Size of the initial population in GWO          | 50        |
| Number of GWO iteration                        | 50        |
| Range of scale value (k)                       | 0 < k < 1 |
| \(\vec{r}_1\) and \(\vec{r}_2\)              | dynamic   |
| Range of \(\vec{r}_1\) and \(\vec{r}_2\)     | 0 < \(\vec{r}_1\) < 1, 0 < \(\vec{r}_2\) < 1 |
| Number of nests (CS)                           | 15        |
| Number of iterations in CS                     | 15        |
| Pa (probability of nest rebuilding)            | dynamic   |

Figure 7. The comparison results between the obtained data and experimental data are based on the ACGWO algorithm for the various neurons. a) Cooling b) Heating. (continued)
Tracking, pursuing, and coming close to the prey.
Chasing, surrounding, and harassing unless the prey stops running.
Attack towards the prey.

According to the aforementioned explanations, the hunting technique and the social hierarchy of grey wolves are mathematically modeled for designing GWO and conducting optimization. GWO algorithm can theoretically solve the optimization problems whose pseudocode is provided in Figure 3 (Mirjalili et al., 2014). This algorithm is employed in the present paper since it can solve complicated function optimization and engineering problems (Yu et al., 2016).

Despite the fact that GWO has an acceptable convergence rate, GWO cannot always be efficient in specifying the global optima with a considerable effect on the convergence rate of the algorithm (Kohli and Arora, 2018). In order to tackle this problem, the CGWO can be considered. When chaos is added to the GWO, the CGWO is developed. This algorithm is actually a deterministic, random-like method found in a nonlinear, dynamical system, and it is non-period, non-converging,
and bounded. Accordingly, a variety of chaotic maps with different mathematical equations can be employed. Hence, it is able to explore the search space more dynamically and globally during the optimization process.

**Augmented grey wolf optimizer.** In spite of the striking benefits of GWO, including simplicity, flexibility, and globalism, the Augmented AGWO has a better performance in hunting (Qais et al., 2018). This algorithm which is more useful for the low number of search agents like electric power system applications, is based on raising the possibility of the exploration process over the exploitation process by correcting the behavior of the control parameter and position updating. Figure 4 illustrates the number of iterations based on the convergence factor \(a\) for the linear regressive and nonlinear regressive curves. Accordingly, the highest value of \(a\) and the iterations number are 2 and 100, respectively.

**Figure 8.** Residual and histogram of the proposed approach for the various neurons. a) Cooling b) Heating. (continued)
Adaptive chaotic grey wolf optimization. The fundamental GWO algorithm also has several shortcomings, including that it easily falls into local optimization and slow convergence in the later phase of the search. In order to deal with such drawbacks, an ACGWO with the capabilities of tackling the uncertainties can be a good idea which is considered in this paper. In addition, the flowchart of this algorithm method proposed here is outlined in Figure 5.

**Statistical analysis**

The main statistical parameters, including SE, MSE, RMSE, MAPE, MAE, RAE, R, $R^2$, and NMSE, are considered to examine the best algorithm in terms of accuracy and authenticity. More details regarding these performance criteria are presented in the previous studies and are computed as follows (Bui et al., 2018; Chou and Bui, 2014; Chou et al., 2016):
Figure 9. The comparison results between the predicted data based on the proposed method and experimental data for the various neurons. a) Cooling  b) Heating. (continued)
Figure 9. Continued.
\[ R = \frac{\sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} y_{n,o} \bar{y}_{n,o} - (\sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} y_{n,o}) (\sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} \bar{y}_{n,o})}{\sqrt{\sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} y_{n,o}^2} - (\sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} y_{n,o})^2} \]

\[ R^2 = RR \]

\[ \text{RMSE} = \sqrt{\frac{1}{NN_{\text{out}}} \sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} (\bar{y}_{n,o} - y_{n,o})^2} \]

\[ \text{MAE} = \frac{1}{NN_{\text{out}}} \sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} |y_{n,o} - \bar{y}_{n,o}| \]

\[ \text{MAPE} = \frac{1}{NN_{\text{out}}} \sum_{n=1}^{N} \sum_{o=1}^{N_{\text{out}}} \left| \frac{y_{n,o} - \bar{y}_{n,o}}{y_{n,o}} \right| \]

Figure 10. The obtained RMSE values for the proposed and other well-known optimization. a) Cooling b) Heating. (continued)
Figure 10. Continued.

Figure 11. Comparison between the parameters of se, RAE, and MAE values for proposed and other well-known optimization. a) Cooling b) Heating.
According to Eqs. (3) to (7), the number of instances and the number of outputs are indicated by $N$ and $N_{out}$, respectively. It should be noted that $y$ and $\bar{y}$ is the actual output and the predicted output through ANN (Bui et al., 2020).

**Data collection**

The total dataset considered in this paper includes 770 inputs, of which 75% is used for the training process, and the rest is allocated for testing (Tsanas and Xifara, 2012). Employing Ecotect simulation software, this dataset has been obtained from twelve building types which were indicated by 18 simple cubes ($3.5 \times 3.5 \times 3.5 \text{ m}$), and their shape is demonstrated in Figure 6. Accordingly, the building with the same volume and various surface areas and dimensions is considered. The simulated buildings were residential buildings in Athens, Greece. Considering this dataset, the impact of dimension on the CL of a building was examined. Thus, the material properties of the façade employed for the whole twelve buildings consist of the U-value of the wall (1.78 $W/\text{m}^2\text{K}$), window (2.26 $W/\text{m}^2\text{K}$) floors (0.86 $W/\text{m}^2\text{K}$), and roofs (0.50 $W/\text{m}^2\text{K}$) considered for this analysis are the same. Notably, the values of 300 lux and 2 $W/\text{m}^2$ were set for the lighting level and latent heat. Employing eight features consisting of the relative compactness (RC), surface area, wall area, roof area, overall height, orientation, glazing area, and glazing distribution were considered for simulating the CL. Accordingly, using Eq. 8, the RC is computed as follows (Bui et al., 2020):

$$RC = 6V^{2/3}A^{-1}$$  \hspace{1cm} (8)
Figure 12. Continued.

Figure 13. The convergence rates of the proposed method compared to the other ones.
According to Eq. (8), the volume and surface area of the building are denoted by $V$ and $A$.

The building was simulated according to the experiments, considering a glazing system and a system without glazing. Three glazing-to-floor area ratios, which are 10\%, 25\%, and 40\%, are employed in the glazing system. In this situation, five glazing distributions are examined which

Figure 14. Comparison between obtained data by the various optimizers and experimental data. a) Cooling b) Heating.

(continued)
consisting of 25% glazing for the faces, 55% on the north side and 15% for the outer faces; 55% for the east face, 15% for the other face; 55% in the west face and 15% for the other face (Tsanas and Xifara, 2012). Lastly, the whole building shapes were rotated to four orientations: north, south, east, and west. The necessary information regarding the selected parameters is highlighted in Table 1.
Figure 15. Comparison between obtained power error data by the proposed optimizer and experimental data. a) Cooling b) Heating. (continued)
Figure 15. Continued.
Results and discussion

The main results obtained in this study are presented and discussed in this section. The selected optimization methods illustrated in the second section are applied to train the MLP neural network, and the obtained results are analyzed.

Figure 16. Residual and histogram diagram for the selected optimizers. a) Cooling b) Heating.

(continued)
Neuron number selection

According to the MLP neural network, the number of neurons is selected for the CL and HL parameters that plays a prominent role inaccuracy of the final results. Accordingly, the possibility of having the highest accuracy for eight, fifteen, twenty, twenty-four, and thirty neurons are examined here.

Figure 16. Continued.

Neuron number selection

According to the MLP neural network, the number of neurons is selected for the CL and HL parameters that plays a prominent role inaccuracy of the final results. Accordingly, the possibility of having the highest accuracy for eight, fifteen, twenty, twenty-four, and thirty neurons are examined here.
According to Figures 7 to 9, fifteen neurons are the best choice for the analysis of this study. As shown in Figures 7 and 8, the accuracy generally increases when the neuron raises. But this trend is not good after fifteen neurons, and the data are not well organized when the number of neurons is higher than fifteen. In this case, the accuracy is not desirable, and the computational cost is high as well. Besides, the values of $R^2$ are generally high, as shown in Figure 9.

| Number of neurons | SE   | MSE  | RMSE | MAPE | MAE  | RAE  | R    | $R^2$  | NMES |
|-------------------|------|------|------|------|------|------|------|--------|------|
| Train             |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1108 | 8.8410 | 2.9734 | 0.0869 | 2.1233 | 0.2531 | 0.9485 | 0.8997 | 0.0122 |
| $N = 15$          | **0.1080** | **7.8867** | **2.8083** | **0.0792** | **1.9242** | **0.2297** | **0.9542** | **0.9106** | **0.0117** |
| $N = 20$          | 0.1085 | 7.4810 | 2.7351 | 0.0821 | 1.9656 | 0.2330 | 0.9566 | 0.9152 | 0.0116 |
| $N = 24$          | 0.1099 | 8.2205 | 2.8671 | 0.0863 | 2.1011 | 0.2498 | 0.9523 | 0.9068 | 0.0120 |
| $N = 30$          | 0.1081 | 7.7515 | 2.7841 | 0.0820 | 1.9742 | 0.2327 | 0.9551 | 0.9122 | 0.0116 |
| Test              |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1150 | 8.9126 | 2.9854 | 0.0945 | 2.3284 | 0.2781 | 0.9567 | 0.9153 | 0.0121 |
| $N = 15$          | **0.0847** | **5.5333** | **2.3523** | **0.0676** | **1.7053** | **0.1868** | **0.9761** | **0.9528** | **0.0070** |
| $N = 20$          | 0.1172 | 8.9540 | 2.9923 | 0.0942 | 2.3301 | 0.2756 | 0.9549 | 0.9118 | 0.0125 |
| $N = 24$          | 0.1163 | 9.4428 | 3.0729 | 0.0928 | 2.2992 | 0.2698 | 0.9510 | 0.9045 | 0.0133 |
| $N = 30$          | 0.1535 | 12.4376 | 3.5267 | 0.1217 | 2.8671 | 0.3670 | 0.9421 | 0.8875 | 0.0199 |
| Total             |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1115 | 8.8548 | 2.9757 | 0.0883 | 2.1628 | 0.2582 | 0.9499 | 0.9023 | 0.0122 |
| $N = 15$          | **0.1062** | **7.9468** | **2.8190** | **0.0788** | **1.9481** | **0.2311** | **0.9551** | **0.9123** | **0.0114** |
| $N = 20$          | 0.1101 | 7.7648 | 2.7865 | 0.0845 | 2.0358 | 0.2413 | 0.9562 | 0.9142 | 0.0118 |
| $N = 24$          | 0.1111 | 8.4561 | 2.9079 | 0.0876 | 2.1393 | 0.2537 | 0.9522 | 0.9067 | 0.0122 |
| $N = 30$          | 0.1181 | 8.6543 | 2.9419 | 0.0897 | 2.1463 | 0.2570 | 0.9510 | 0.9044 | 0.0132 |
| MPL different neurons (ACGWO) - Heating |      |      |      |      |      |      |      |        |      |
| Number of neurons | SE   | MSE  | RMSE | MAPE | MAE  | RAE  | R    | $R^2$  | NMES |
| Train             |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1361 | 7.1245 | 2.6692 | 0.0962 | 1.8950 | 0.2138 | 0.9630 | 0.9274 | 0.0161 |
| $N = 15$          | **0.1211** | **6.0497** | **2.4596** | **0.0864** | **1.7607** | **0.1987** | **0.9687** | **0.9384** | **0.0130** |
| $N = 20$          | 0.1201 | 5.9613 | 2.4416 | 0.0885 | 1.7811 | 0.2015 | 0.9692 | 0.9394 | 0.0132 |
| $N = 24$          | 0.1262 | 7.2058 | 2.6844 | 0.0961 | 1.9766 | 0.2211 | 0.9626 | 0.9266 | 0.0145 |
| $N = 30$          | 0.1254 | 6.5459 | 2.5585 | 0.0937 | 1.8860 | 0.2115 | 0.9661 | 0.9334 | 0.0145 |
| Test              |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1780 | 10.9957 | 3.3160 | 0.1342 | 2.6887 | 0.3314 | 0.9673 | 0.9356 | 0.0248 |
| $N = 15$          | **0.0846** | **5.533291** | **2.35229** | **0.0676** | **1.7052** | **0.1867** | **0.9761** | **0.9528** | **0.0070** |
| $N = 20$          | 0.0975 | 6.2589 | 2.5018 | 0.0792 | 1.8967 | 0.2119 | 0.9729 | 0.9465 | 0.0088 |
| $N = 24$          | 0.1143 | 8.6276 | 2.9373 | 0.0930 | 2.2578 | 0.2699 | 0.9668 | 0.9347 | 0.0123 |
| $N = 30$          | 0.1155 | 8.2387 | 2.8703 | 0.0909 | 2.1235 | 0.2184 | 0.9658 | 0.9328 | 0.0147 |
| Total             |      |      |      |      |      |      |      |        |      |
| $N = 8$           | 0.1450 | 7.8705 | 2.8054 | 0.1035 | 2.0480 | 0.2348 | 0.9609 | 0.9234 | 0.0178 |
| $N = 15$          | **0.1150** | **5.9502** | **2.4393** | **0.0827** | **1.7500** | **0.1963** | **0.9705** | **0.9419** | **0.0119** |
| $N = 20$          | 0.1160 | 6.0186 | 2.4533 | 0.0867 | 1.8033 | 0.2033 | 0.9701 | 0.9412 | 0.0123 |
| $N = 24$          | 0.1239 | 7.4798 | 2.7349 | 0.0955 | 2.0308 | 0.2298 | 0.9628 | 0.9269 | 0.0141 |
| $N = 30$          | 0.1235 | 6.8722 | 2.6215 | 0.0932 | 1.9317 | 0.2128 | 0.9659 | 0.9330 | 0.0145 |
Table 3. The obtained statistical results for the various optimizers with neurons = 15.

Neuron = 15 - Cooling

| Optimizer | SE     | MSE     | RMSE   | MAPE   | MAE    | RAE    | R     | R²     | NMES  |
|-----------|--------|---------|--------|--------|--------|--------|-------|--------|-------|
| ACGWO     | 0.1080 | 7.8867  | 2.8083 | 0.0792 | 1.9242 | 0.2297 | 0.9542| 0.9106 | 0.0117|
| AGWO      | 0.1556 | 12.1861 | 3.4909 | 0.1226 | 2.6787 | 0.3588 | 0.9385| 0.8808 | 0.0207|
| CGWO      | 0.1247 | 11.1557 | 3.3400 | 0.0992 | 2.4078 | 0.3014 | 0.9350| 0.8743 | 0.0163|
| GWO       | 0.1432 | 13.9291 | 3.7322 | 0.1112 | 2.7105 | 0.3052 | 0.9192| 0.8450 | 0.0195|
| MVO       | 0.1440 | 13.0233 | 3.6088 | 0.1157 | 2.7383 | 0.3145 | 0.9263| 0.8580 | 0.0211|
| PSO       | 0.1544 | 14.2930 | 3.7806 | 0.1174 | 2.7833 | 0.3078 | 0.9207| 0.8477 | 0.0210|
| GA        | 0.2794 | 42.6634 | 6.5317 | 0.2274 | 5.2699 | 0.4350 | 0.8854| 0.7840 | 0.0676|

Neuron = 15 - Heating

| Optimizer | SE     | MSE     | RMSE   | MAPE   | MAE    | RAE    | R     | R²     | NMES  |
|-----------|--------|---------|--------|--------|--------|--------|-------|--------|-------|
| ACGWO     | 0.0847 | 5.5333  | 2.3523 | 0.0676 | 1.7053 | 0.1868 | 0.9761| 0.9528 | 0.0070|
| AGWO      | 0.1448 | 13.0068 | 3.6065 | 0.1121 | 2.7652 | 0.3814 | 0.9482| 0.8990 | 0.0183|
| CGWO      | 0.1329 | 10.3272 | 3.2136 | 0.1092 | 2.5583 | 0.2973 | 0.9533| 0.9088 | 0.0155|
| GWO       | 0.0990 | 10.0225 | 3.1658 | 0.0759 | 2.1452 | 0.2332 | 0.9499| 0.9023 | 0.0111|
| MVO       | 0.1547 | 19.4755 | 4.4131 | 0.1102 | 2.9872 | 0.3757 | 0.9007| 0.8113 | 0.0246|
| PSO       | 0.1095 | 11.0398 | 3.3226 | 0.0868 | 2.4175 | 0.2655 | 0.9536| 0.9094 | 0.0139|
| GA        | 0.2774 | 70.3481 | 8.3874 | 0.2472 | 6.7771 | 0.8926 | 0.8000| 0.6401 | 0.1159|

Neuron = 15 - Heating

| Optimizer | SE     | MSE     | RMSE   | MAPE   | MAE    | RAE    | R     | R²     | NMES  |
|-----------|--------|---------|--------|--------|--------|--------|-------|--------|-------|
| ACGWO     | 0.1211 | 6.0497  | 2.4596 | 0.0864 | 1.7607 | 0.1987 | 0.9687| 0.9384 | 0.0130|
| AGWO      | 0.1627 | 10.0341 | 3.1677 | 0.1088 | 2.3186 | 0.2813 | 0.9518| 0.9059 | 0.0216|
| CGWO      | 0.1769 | 12.1700 | 3.4886 | 0.1468 | 2.8225 | 0.2757 | 0.9535| 0.9092 | 0.0387|
| GWO       | 0.1874 | 12.2733 | 3.5033 | 0.1519 | 2.7700 | 0.2650 | 0.9491| 0.9009 | 0.0373|
| MVO       | 0.1361 | 9.7451  | 3.1217 | 0.1112 | 2.3551 | 0.2625 | 0.9505| 0.9035 | 0.0185|
| PSO       | 0.1729 | 10.0563 | 3.1712 | 0.1282 | 2.4989 | 0.3016 | 0.9478| 0.8983 | 0.0239|
| GA        | 0.3123 | 30.0479 | 5.4815 | 0.2557 | 4.5794 | 0.4821 | 0.8358| 0.6986 | 0.0769|

(continued)
As regards Figure 10, the statistical results of RMSE based on the selected optimization methods for HL and CL parameters are presented. Accordingly, the ACGWO method has the best performance since its RMSE value is lower than the rest and includes 10% of the whole. On the other hand, the genetic algorithm (GA) is considered the weakest optimization technique in terms of the RMSE performance criterion. Besides, CGWO for the parameter of CL has an acceptable performance after ACGWO. Notably, the performance of PSO, MVO, and AGWO for the parameter of HL is also considerable.

Furthermore, the results shown in Figure 11 imply that ACGWO has the best performance compared to the others. The results of RAE, SE, and MAE performance criteria are drawn for the test data, training data, and total data based on HL and CL. Accordingly, there is a considerable gap between the values of these parameters obtained for GA and the values obtained for the other optimizers. As a result, GA has the weakest performance and is not suitable for optimizing the MLP neural network. The values of RAE, SE, and MAE are not a lot which means that the performance of MVO, CGWO, AGWO, GWO, and PSO is acceptable, especially for examining CL.

The results of R, R2, MSE, and MAPE are also indicated in Figure 12, according to which ACGWO has the best performance and GA has a significant difference from the rest, which is not desirable.

As shown in Figure 13, AGWO has the highest speed in the convergence rate and shows a value of 0.238 s based on 157 iterations. Notably, the least value of convergence rate belongs to ACGWO with 3000 iterations at 0.201 s.

According to Figures 14 to 16, since the values of R and R² are more than the rest in ACGWO, this optimizer is more beneficial and is likely to predict the energy consumption with the highest accuracy. It is also noteworthy that the error rate in this optimizer is considerably low, and this result is not observed for the other parameters.

In general, the statistical results that are highlighted in Tables 2 and 3 emphasize the remarkable capabilities of ACGWO for optimizing the training process in the MLP neural network. The values of R and R² are significantly high for ACGWO, while this trend is vice versa for GA. Table 3 organizes the selected optimizers in order of accuracy according to which GA has the weakest performance and ACGWO is strikingly accurate.

### Conclusion

In summary, the issue of energy consumption has been comprehensively examined by calculating the HL and CL of a building. Using the MLP neural network for a dataset, the heating load (HL) and
cooling load (Zanchini and Naldi) were calculated for the building. Since the training process of the MLP network is not desirable to bring the optimal results, the various optimizers have been considered, including GA, MVO, CGWO, ACGWO, AGWO, GWO, and PSO. In addition, the statistical analysis was conducted to find the best optimizer, which was able to obtain the most accurate results. For this purpose, the performance criteria, namely SE, MSE, RMSE, MAPE, MAE, RAE, R, R², and NMSE, were analyzed in this study. However, MVO, CGWO, AGWO, GWO, and PSO turned out to be practical for the optimization process; the performance of ACGWO outperforms the rest according to the statistical analyses. On the other hand, the performance of GA was not acceptable since it has the highest values of R and R² and the least values of SE, MSE, RMSE, MAPE, MAE, and RAE compared to the other selected optimizers. Notably, AGWO has the highest speed in the convergence rate and shows a value of 0.238 s based on 157 iterations. Overall, ACGWO is introduced as the best optimizer in terms of accuracy and authenticity for predicting energy consumption. More research is needed to apply and test other methods and parameters in future research. According to the obtained results, MLP-ACGWO demonstrates the best performance when the number of neurons in the hidden layer is considered to be 15. In this condition, for the total dataset, the coefficient of performance is obtained to be 0.94198 and 0.9123 for the heating load and cooling load prediction, respectively.

Acknowledgements
This paper was supported by the key area research and development program of Guangdong province 2020B1111420002, and national key technologies research & development program (2019YFB1600600, 2019YFB1600604).

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key Technologies Research & Development Program, by the key area research and development program of Guangdong province, (grant number 2019YFB1600600, 2019YFB1600604, 2020B1111420002).

ORCID iD
Guiju Zhang https://orcid.org/0000-0002-1951-1500

References
Ahmadi-Karvigh S, Ghahramani A, Becerik-Gerber B, et al. (2018) Real-time activity recognition for energy efficiency in buildings. Applied Energy 211: 146–160.
Al-Homoud MS (2001) Computer-aided building energy analysis techniques. Building and Environment 36: 421–433.
AlShabi M and Assad MEH (2021) Artificial intelligence applications in renewable energy systems. Design and Performance Optimization of Renewable Energy Systems 16: 251–295.
Askarzadeh A and Rezazadeh A (2011) Optimization of PEMFC model parameters with a modified particle swarm optimization. International Journal of Energy Research 35: 1258–1265.
Barkokebas RD, Chen Y, Yu H, et al. (2019) Achieving housing energy-efficiency requirements: methodologies and impacts on housing construction cost and energy performance. *Journal of Building Engineering* 26: 100874.

Biswas MAR, Robinson MD and Fumo N (2016) Prediction of residential building energy consumption: A neural network approach. *Energy* 117: 84–92.

Bui D-K, Nguyen T, Chou J-S, et al. (2018) A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. *Construction and Building Materials* 180: 320–333.

Bui D-K, Nguyen TN, Ngo TD, et al. (2020) An artificial neural network (ANN) expert system enhanced with the electromagnetism-based firefly algorithm (EFA) for predicting the energy consumption in buildings. *Energy* 190: 116370.

Chou J-S and Bui D-K (2014) Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings* 82: 437–446.

Chou J-S, Chong WK and Bui D-K (2016) Nature-inspired metaheuristic regression system: Programming and implementation for civil engineering applications. *Journal of Computing in Civil Engineering* 30: 4016007.

Chou JS and Truong DN (2021) Multistep energy consumption forecasting by metaheuristic optimization of time-series analysis and machine learning. *International Journal of Energy Research* Wiley Online Library 45: 4581–4612.

Diab AAZ and Rezk H (2017) Global MPPT based on flower pollination and differential evolution algorithms to mitigate partial shading in building integrated PV system. *Solar Energy* 157: 171–186.

Dorotic H, Pukšec T and Duć N (2019) Multi-objective optimization of district heating and cooling systems for a one-year time horizon. *Energy* 169: 319–328.

Ferreira PM, Ruano AE, Silva S, et al. (2012) Neural networks based predictive control for thermal comfort and energy savings in public buildings. *Energy and Buildings* 55: 238–251.

Frayssinet L, Merlier L, Kuznik F, et al. (2018) Modeling the heating and cooling energy demand of urban buildings at city scale. *Renewable and Sustainable Energy Reviews* 81: 2318–2327.

Gerges F, Zouein G and Azar D (2018) Genetic algorithms with local optima handling to solve sudoku puzzles. Proceedings of the 2018 international conference on computing and artificial intelligence. 19–22.

Ghahramani A, Zhang K, Dutta K, et al. (2016) Energy savings from temperature setpoints and deadband: quantifying the influence of building and system properties on savings. *Applied Energy* 165: 930–942.

Gilani HA, Hoseinzadeh S, Esmaeilion F, et al. (2022) A solar thermal driven ORC-VFR system employed in subtropical Mediterranean climatic building. *Energy* 250: 123819.

Gilani HA, Hoseinzadeh S, Karimi H, et al. (2021) Performance analysis of integrated solar heat pump VRF system for the low energy building in Mediterranean island. *Renewable Energy* 174: 1006–1019.

Ihara T, Gustavsen A and Jelle BP (2015) Effect of facade components on energy efficiency in office buildings. *Applied Energy* 158: 422–432.

Ilbeigi M, Ghomeishi M and 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: 102325.

Kohli M and Arora S (2018) Chaotic grey wolf optimization algorithm for constrained optimization problems. *Journal of Computational Design and Engineering* 5: 458–472.

Komeili Birjandi A, Fahim Alavi M, Salem M, et al. (2022) Modeling carbon dioxide emission of countries in southeast of Asia by applying artificial neural network. *International Journal of Low-Carbon Technologies* 17: 321–326.

Lee S, Jung S and Lee J (2019) Prediction model based on an artificial neural network for user-based building energy consumption in South Korea. *Energies*. 12: 1–18.

Levesque A, Pietzcker RC, Baumstark L, et al. (2018) How much energy will buildings consume in 2100? A global perspective within a scenario framework. *Energy* 148: 514–527.

Li K, Xie X, Xue W, et al. (2018) A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction. *Energy and Buildings* 174: 323–334.
Liu Z, Liu Y, He B-J, et al. (2019) Application and suitability analysis of the key technologies in nearly zero energy buildings in China. Renewable and Sustainable Energy Reviews 101: 329–345.

Mahmoudan A, Esmaeilion F, Hoseinzadeh S, et al. (2022) A geothermal and solar-based multigeneration system integrated with a TEG unit: development, 3E analyses, and multi-objective optimization. Applied Energy 308: 118399.

Mahmoudan A, Samadof P, Hosseinzadeh S, et al. (2021) A multigeneration cascade system using ground-source energy with cold recovery: 3E analyses and multi-objective optimization. Energy 233: 121185.

Mech LD (1999) Alpha status, dominance, and division of labor in wolf packs. Canadian Journal of Zoology. Canada: NRC Research Press Ottawa, 1196–1203.

Mirjalili S, Mirjalili SM and Hatamlou A (2016) Multi-verse optimizer: A nature-inspired algorithm for global optimization. Neural Computing and Applications 27: 495–513.

Mirjalili S, Mirjalili SM and Lewis A (2014) Grey wolf optimizer. Advances in Engineering Software. 69: 46–61.

Mohamed MA, Diab AAZ and Rezk H (2019) Partial shading mitigation of PV systems via different metaheuristic techniques. Renewable Energy 130: 1159–1175.

Moldovan D and Slowik A (2021) Energy consumption prediction of appliances using machine learning and multi-objective binary grey wolf optimization for feature selection. Applied Soft Computing 111: 107745.

Muro C, Escobedo R, Spector L, et al. (2011) Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations. Behavioural Processes 88: 192–197.

Neto AH and Fiorelli FAS (2008) Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy and Buildings 40: 2169–2176.

Ngo N-T, Truong TTH, Truong N-S, et al. (2022) Proposing a hybrid metaheuristic optimization algorithm and machine learning model for energy use forecast in non-residential buildings. Scientific Reports 12: 1–18.

Pérez-Lombard L, Ortiz J and Pout C (2008) A review on buildings energy consumption information. Energy and Buildings 40: 394–398.

Qais MH, Hasanien HM and Alghuwainem S (2018) Augmented grey wolf optimizer for grid-connected PMSG-based wind energy conversion systems. Applied Soft Computing 69: 504–515.

Rashidi M, Alhuyi Nazari M, Mahariq I, et al. (2022) Modeling and sensitivity analysis of thermal conductivity of ethylene glycol-water based nanofluids with alumina nanoparticles. Experimental Techniques 16: 1–8.

Ruiz LGB, Rueda R, Cuéllar MP, et al. (2018) Energy consumption forecasting based on elman neural networks with evolutive optimization. Expert Systems with Applications 92: 380–389.

Satrio P, Mahlia TMI, Giannetti N, et al. (2019) Optimization of HVAC system energy consumption in a building using artificial neural network and multi-objective genetic algorithm. Sustainable Energy Technologies and Assessments 35: 48–57.

Sauer J, Mariani VC, dos Santos Coelho L, et al. (2021) Extreme gradient boosting model based on improved jaya optimizer applied to forecasting energy consumption in residential buildings. Evolving Systems 25: 1–12.

Sedighizadeh M, Esmaïli M and Mohammadkhani N (2018) Stochastic multi-objective energy management in residential microgrids with combined cooling, heating, and power units considering battery energy storage systems and plug-in hybrid electric vehicles. Journal of Cleaner Production 195: 301–317.

Sedighizadeh M, Rezagadeh A, Khoddam M, et al. (2011) Parameter optimization for a PEMFC model with particle swarm optimization. International Journal of Engineering and Applied Sciences 3: 102–108.

Shakouri M, Ghadamian H, Hoseinzadeh S, et al. (2022) Multi-objective 4E analysis for a building integrated photovoltaic thermal double skin façade system. Solar Energy 233: 408–420.

Sohani A, Dehnavi A, Sayyaadi H, et al. (2022) The real-time dynamic multi-objective optimization of a building integrated photovoltaic thermal (BIPV/T) system enhanced by phase change materials. Journal of Energy Storage 46: 103777.

Takagi T and Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE transactions on Systems, man, and Cybernetics. IEEE 1: 116–132.

Tsanas A and Xifara A (2012) Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy and Buildings 49: 560–567.

Takagi T and Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE transactions on Systems, man, and Cybernetics. IEEE 1: 116–132.
Wang J, Chen X, Zhang F, et al. (2021) Building load forecasting using deep neural network with efficient feature fusion. *Journal of Modern Power Systems and Clean Energy* 9: 160–169.

Yu H, Yu Y, Liu Y, et al. (2016) Chaotic grey wolf optimization. 2016 International Conference on Progress in Informatics and Computing (PIC). IEEE, 103–113.

Zanchini E and Naldi C (2019) Energy saving obtainable by applying a commercially available M-cycle evaporative cooling system to the air conditioning of an office building in north Italy. *Energy* 179: 975–988.

Zhou G, Moayedi H and Foong LK (2021) Teaching–learning-based metaheuristic scheme for modifying neural computing in appraising energy performance of building. *Engineering with Computers* 37: 3037–3048.

**Acronym glossary**

ACGWO    
Adaptive Chaotic Grey Wolf Optimization

AGWO     
Augmented Grey Wolf Optimizer

ANN      
Artificial Neural Network

CL       
Cooling Load

CGWO     
Chaotic Grey Wolf Optimization

DH       
District Heating

GA       
Genetic Algorithm

GWO      
Grey Wolf Optimizer

HAVC     
Heating Ventilation and Air Conditioning

HL       
Heating Load

MAE      
Mean Absolute Error

MAPE     
Mean Absolute Percent Error

MLP      
Multi-Layer Perceptron

MOGA     
Multi-objective Genetic Algorithm

MSE      
Mean Square Error

MVO      
Multi-Verse Optimizer

NMSE     
Normalized Mean Square Error

OECD     
Economic Cooperation and Development

PSO      
Particle Swarm Optimization

R        
Correlation Coefficient

R²       
Coefficient of determination

RAE      
Relative Absolute Error

RMSE     
Root-Mean-Squared Error

SE       
Standard Error

SHGC     
Solar Heat Gain Coefficient