On the Utilization of an Ensemble of Meta-Heuristics for Simulating Energy Consumption in Buildings

Eslam Mohammed Abdelkader, Structural Engineering Department, Faculty of Engineering, Cairo University, Egypt*
Nehal Elshaboury, Housing and Building National Research Centre, Egypt
Abobakr Al-Sakkaf, Department of Architecture and Environmental Planning, College of Engineering and Petroleum, Yemen

ABSTRACT

Predicting energy consumption has been a substantial topic because of its ability to lessen energy wastage and establish an acceptable overall operational efficiency. Thus, this research aims at creating a meta-heuristic-based method for autonomous simulation of heating and cooling loads of buildings. The developed method is envisioned on two tiers, whereas the first tier encompasses the use of a set of meta-heuristic algorithms to amplify the exploration and exploitation of Elman neural network through both parametric and structural learning. In this regard, 10 meta-heuristic were utilized, namely differential evolution, particle swarm optimization, invasive weed optimization, teaching-learning optimization, ant colony optimization, grey wolf optimization, grasshopper optimization, moth-flame optimization, antlion optimization, and arithmetic optimization. The second tier is designated for evaluating the meta-heuristic-based models through performance evaluation and statistical comparisons. An integrative ranking of the models is achieved using average ranking algorithm.

KEYWORDS

Elman Neural Network, Energy Consumption, Grey Wolf Optimization, Meta-Heuristic, Operational Efficiency, Particle Swarm Optimization

1 INTRODUCTION

The generation of greenhouse gases is directly proportional to energy consumption and climate change (Baldock et al., 2012; Al-Sakkaf et al., 2020). The worldwide increase in urbanization and industrialization, particularly in the building sector, has served as a major contributing factor. In this regard, the increases in energy consumption and related carbon dioxide emissions witnessed significant bumps in the last few decades, and it is expected to experience escalating increases due to rapid expansions in commercial and residential regions, and higher cooling demands in hot weather.
countries (Eom et al., 2020; Conevska et al., 2020). The amount of fuel consumed by this sector is equivalent to 2 billion tons of oil equivalent (TOE), representing 31% of fuel for worldwide energy use (International Energy Agency, 2015). More specifically, in terms of electricity and heating, about 0.84 billion TOE is consumed by the building sector. Moreover, the building sector consumes 12% of fresh water and contributes to 40% of global solid waste and 40% of CO2 emissions. Approximately 20 – 25% and 30 – 40% have been reported as consumed energy for developing and developed countries, respectively (Akande et al., 2015; Al-Sakkaf et al., 2019). In the United States and European Union, up to 40% energy consumption is attributed to the building sector (Cao et al., 2016).

Based on the aforementioned points, improved sustainability in buildings is essential to reduce environmental impacts and improve the wellbeing of individuals. This is also dependent on the efficiency of several elements in buildings. For example, consumed energy in the heating and cooling processes could be managed by heating, ventilation, and air conditioning (HVAC) systems to ensure a comfortable indoor environment in buildings. Therefore, proper design of HVAC systems, based on climate and building attributes, is critical to the energy efficiency of buildings (McQuiston & Parker, 1982; Castaldo et al., 2018; Bui et al., 2019). In this regard, energy prediction models are required for facility managers to better understand sustainability in buildings (Urge-Vorsatz et al., 2007; Egan et al., 2018; Fant et al., 2018; Al-Sakkaf et al., 2019). Forward and inverse models are typically applied to evaluate the energy performance of buildings (Zhao et al., 2012). For forward modeling, the building attributes are utilized despite being not easily determined, decreasing the computation accuracy, increasing the computation time, and hindering the application for occupied buildings (Yezioro et al., 2008; Park et al., 2016). Inverse modeling, on the other hand, is a machine learning model that can be applied as an alternative to identify the correlation between energy consumption and building attributes or factors (Catalina et al., 2008; Yu et al., 2010; O’Neill & O’Neill, 2016). This modeling technique is easy and has a fast computation speed (Bui et al., 2019).

The application of machine learning has emerged as a powerful tool in building energy management (Bui et al., 2019). For example, Alvarez et al. (2018) predicted the energy performance of a house using artificial neural network (ANN) models. The models were developed using a dataset of different buildings located in Spain. The inputs to these models were the construction year and surface area while the model output comprised the U-opaque value. It was found that the constructed area factor was associated with the highest prediction power. The modeled U-Opaque values were compared against the real values for the same buildings, yielding a correlation coefficient of 0.967. The results confirmed the ability to provide an accurate estimate of the building’s energy efficiency using some modeled building characteristics. Jihad & Tahiri (2018) applied different ANN models to predict the heating and cooling loads of residential buildings in Morocco. The forecasting model included six input variables; orientation, glazing rate, relative compactness, wall surface area, and building height and surface area. The training samples were generated using parametric analysis by changing the construction mode and building usage. The developed model yielded an accuracy of 98.7% and 97.6% for the training and testing data, respectively. This model could be useful to determine the energy consumption of a new building without the need for using simulation software nor calculating a thermodynamic balance.

Moayedi et al. (2019) estimated the heating load in energy-efficient buildings using six machine learning techniques. These techniques were the multi-layer perceptron regressor, random forest, lazy locally weighted learning, alternating model tree, ElasticNet, and radial basis function regression. The calculated outcomes of the above-mentioned models were analyzed using five statistical indexes; root relative squared error (RRSE), root mean squared error (RMSE), mean absolute error (MAE), Pearson correlation coefficient (PCC), and relative absolute error (RAE). The random forest model yielded PCC, MAE, RMSE, RAE, and RRSE for the training dataset of 0.99, 0.19, 0.24, 2.08, and 2.38, respectively. Besides, it determined the PCC, MAE, RMSE, RAE, and RRSE for the testing dataset to be 0.99, 0.34, 0.46, 3.68, and 4.59, respectively. Therefore, it was nominated as the most powerful predictive network. Amber et al. (2018) compared the performances of five intelligent
system techniques in the prediction of electricity consumption of buildings. The intelligent techniques comprised multiple regression, genetic programming, deep neural network, ANN, and support vector machines (SVM). The input parameters were daily electricity consumption, mean surrounding temperature, mean global irradiance, mean humidity, and mean wind velocity. It was concluded that the ANN model outperformed other prediction models achieving RMSE, MAE, mean relative error (MRE), mean absolute percentage error (MAPE), and normalized RMSE of 26, 17, 6%, 6%, and 10%, respectively.

Seyedzadeh et al. (2019) investigated the accuracy of several machine learning models for predicting heating and cooling loads of buildings. The simulated energy data generated in EnergyPlus and Ecotect were utilized to compare the results. The combinations of model parameters were examined using a grid-search method based on cross-validation. Besides, a sensitivity analysis was performed to examine the influence of input variables on the performance of the forecasting models. The outcomes of the analysis demonstrated the relative importance of input variables, leading to faster model fitting. Finally, the proposed models resulted in satisfactory accuracy compared to the existing models. Mohammed Abdelkader et al. (2020) conducted a comprehensive analysis among five machine learning models to predict heating and cooling loads in residential buildings. These models included the generalized regression neural network, back-propagation neural network, radial basis neural network, radial kernel SVM, and analysis of variance kernel SVM. The above-mentioned models were assessed using three statistical measures; RMSE, MAE, and MAPE. Besides, the statistical significance of these models was evaluated using the student’s t-test. The results affirmed that the radial basis function network outperformed the remaining models.

Das et al. (2019) investigated the heating and cooling loads of residential buildings using Elman neural network (ENN), recurrent neural network, and backpropagation neural network. The models accounted for eight input factors, namely relative compactness, height, orientation, surface area, wall area, roof area, glazing area, glazing area distribution. The performance of the models was evaluated using MAE, MSE, and MRE. The back-propagation neural network yielded the highest accuracy compared to the existing methods. The computation results of the model reinforced the fact that ANN could be utilized to predict and analyze the energy performance of a building. Moon et al. (2019) proposed an ANN-based model for forecasting building energy consumption. The authors studied two different hyperparameters of ANN which were: the number of hidden layers and type of activation function. Some input factors were used for short-term load forecasting, namely calendar data, weather information, and historical electric loads. Ten different numbers of hidden layers in addition to five activation functions were examined. It was concluded that five hidden layers and scaled exponential linear unit function provided a better average performance when compared against other prediction models.

Shapi et al. (2021) developed a predictive model for energy consumption using ANN, SVM, and k-nearest neighbor. The model was conducted in a cloud-based machine learning platform. It was examined using two tenants from a commercial building in Malaysia. The performances of the models were compared using RMSE, MAPE, and normalized RMSE. It was proved that the tenant’s energy consumption had different distribution characteristics. Li et al. (2019) used ANN to emulate the energy consumption of complex architectural forms at the early design stage. The input variables encompassed building length, building width, window to wall ratio, story height, floor number, and room number. The output variables are cooling, heating, and lighting energy consumption. In this regard, the developed ANN model utilized 20 hidden neurons and hyperbolic tangent sigmoid transfer function. In the developed method, two architectural form decompositions were proposed which were: method of characterization decomposition (MCD) and method of spatial homogenization decomposition (MSHD). Results manifested that under the MCD method, the relative deviations were ±5% for cooling and heating energy consumption and ±10% for lighting energy consumption. Under the MSHD method, the relative deviations of total energy consumption were ±10%.
Yuan et al. (2018) adopted an ANN to simulate the seasonal hourly electricity consumption for three areas of the university compass in Japan. Six input parameters were used namely, day of the week, hour of the day, hourly dry-bulb temperature, hourly relative humidity, hourly global irradiance, and recorded hourly electricity consumption at the same hour and day of the week in the previous weeks. The Levenberg-Marquardt algorithm was used to train the developed feed-forward ANN model. The developed model accomplished a correlation coefficient that ranged from 96% to 99% at the training phase and from 95% to 99% at the testing phase. Zeng et al. (2020) presented a Gaussian regression process-based model for simulating electricity consumption of different types of buildings. It was derived that the developed model obtained higher prediction accuracies in office buildings than shopping malls and hotels attaining average deviations below 15%. It was also highlighted that the developed model could achieve a short testing time of 0.02 seconds per prediction.

In view of the above, it can be noticed that most of previous studies relied on classical machine learning models such as artificial neural network, multilayer perceptron and support vector machines to forecast building energy consumption. However, these conventional machine learning models are vulnerable to local minima entrapment and over fitting issues (Khan et al., 2020; Li et al., 2019). In addition, manual tuning-based prediction models suffer from inferior prediction capacity and long computational time originated from the need to calibrate the hyper parameters of the machine learning models (Jiang et al., 2020; Al-Allaf, 2011). Artificial neural network is the most widely utilized supervised learning algorithm that is usually trained using gradient descent back propagation algorithm. Its architecture is composed of input layer, output layer and hidden layers. The latter is composed of neurons which are responsible of analyzing the pattern of the fed dataset (Wang et al., 2021; Han et al., 2019). Multilayer perceptron is a feed forward artificial neural network that is composed of input layer, output layer and one or more hidden layers that are interposed between input and output layers, and it is normally trained using gradient descent algorithm (Hicham et al., 2017; Demirci et al., 2015). Support vector machines is one of the supervised learning algorithms that was proposed by Cortes and Vapnik (1995) and it can be used for either classification or regression problems. It is empowered based on mapping the original data into higher dimensional feature space through kernel functions, whereas the type of kernel function and its respective parameters are the most influential factors on the performance of support vector machines (Yetilmezsoy et al., 2021; Tezel & Buyukyildiz, 2016). In the light of foregoing, the main objective of the present study is to build a meta-heuristic-based Elman neural network model that an accurately simulate the heating and cooling loads of buildings. The developed model deploys utilizes a set of ten meta-heuristic algorithms for boosting the exploration-exploitation search capabilities of the ENN through optimizing both of its parameters and hyper parameters. The ten different types of meta-heuristics are: differential evolution (DE), particle swarm optimization (PSO), invasive weed optimization (IWO), teaching learning-based optimization (TLBO), ant colony optimization (ACO), grey wolf optimization (GWO), grasshopper optimization (GO), moth-flame optimization (MFO), antlion optimization (ALO), and arithmetic optimization (AO) algorithms. The performance of hybrid machine learning models is evaluated based on split validation using a set of performance metrics. The average ranking algorithm is adopted to create an integrative ranking of the hybrid models according to their performance with respect to several indicators. Besides, a correlation matrix is built in an attempt to examine the degree of correlation of the explanatory input variables with each other, and between the input and output variables.

2 PROPOSED MODEL

The main objective of the present study is to build optimized Elman neural network model that can accurately estimate building energy consumption. As illustrated in Figure 1, the framework of the developed model is composed of two main components, namely hybrid machine learning model and model assessment. The first step of the first component is to construct the energy efficiency dataset. In this regard, the developed model is established based on a dataset published by the
UCI machine learning repository (Asuncion & Newman, 2007) for a work carried out by Tsanas & Xifara (2012). The utilized dataset encompasses 768 data instances generated using the Ecotect energy analysis platform. The input explanatory features of the developed model incorporate a set of residential building characteristics including glazing area, glazing area distribution, orientation, overall height, roof area, wall area, surface area, and relative compactness. The output predicted variables of the developed model comprise the amounts of heating and cooling loads. The simulated heating and cooling loads are produced from different scenarios and combinations of the input variables. The numbers of possible values of glazing area, glazing area distribution, orientation, overall height, roof area, wall area, surface area, and relative compactness are 4, 6, 4, 2, 4, 7, 12, and 12, respectively.

The first component is designated for designing meta-heuristic-based Elman neural network model that can predict the amounts of heating and cooling loads. The developed model employs...
a set of meta-heuristic algorithms for amplifying the search capabilities of the ENN through both parametric and structural learning process. The learning process involves optimizing both the architecture and weighted connection of the ENN. This is expected to address the exploration-exploitation trade-off dilemma which usually hinders the training process of artificial neural networks (Chong et al., 2021; Chiroma et al., 2016). The ENN is trained based on formulating a variable length single-objective optimization function designated for minimizing the mean absolute percentage error of heating and cooling loads. Elman neural network is one of the recurrent neural networks that was first proposed by the scientist Elman (1990). Its basic configuration comprises input layer, hidden layer, output layer, and context layer. Elman neural network is augmented by additional context layers and feedback loop of unit delay element (Z^-1) which aid in memorizing the computations of the previous time step. In this regard, preserving states between succeeding time steps creates a non-linear dynamic behavior of Elman neural network which improves its learning capability over feed-forward neural network (Bianchi et al., 2017). The computations in the conventional feed-forward neural networks depend solely on the current time step and don’t account for preceding events which diminishes its learning abilities (kurach & Pawlowski, 2016; Ghasemi & Rasekhi, 2016). Meta-heuristics proved themselves as an efficient global search algorithms to improve the prediction performances of machine learning models in diverse civil engineering applications such as daily stream flow simulation (Malik et al., 2020), estimation of vertical settlement of raft-pile foundation (Liu et al., 2020), conceptual cost estimation of power plants (Hashemi et al., 2019) and prediction of structural failure of multistore buildings (Chatterjee et al., 2017). The developed model investigates the implementation of ten meta-heuristics such as DE, PSO, IWO, TLBO, ACO, GWO, GO, MFO, ALO and AO. In this regard, the applicability of ten different types of meta-heuristics are investigated as a result of the case dependency nature of meta-heuristics such that there is no absolute rule on how these algorithms can perform in a designated optimization problems. In this context, some meta-heuristics can provide better solutions than other implemented algorithms in a certain optimization problem. However, it may yield inferior solutions in other optimization problems (Teimouri et al., 2016; Martinez-Álvarez et al., 2014). There are two types of outputs from the first component of the model, whereas the first output includes the optimum transfer function, optimum numbers of hidden and context layers, optimum numbers of hidden and context neurons, and optimum weights. The second type involves the minimum mean absolute percentage error accomplished by the trained models.

Selection of the best performing heating and cooling load prediction model is a vital and complicated process (Shahsavar et al., 2020; Dutta et al., 2019). Hence, the second component of the developed model aims at evaluating the accuracy levels achieved by the hybrid meta-heuristic-based Elman neural network model as per two folds of assessment. In the first fold, the developed models are assessed based on split validation using a set of performance metrics. The utilized performance indicators are MAPE, RMSE, PCC, Nash-Sutcliffe coefficient of efficiency (NSCE), and Willmott’s index of agreement (WI). The average ranking algorithm is adopted to create an integrative ranking of the hybrid machine learning models according to their performance with respect to several indicators. The second fold involves building a correlation matrix for the heating and cooling loads in an attempt to examine the degree of correlation and association of the input variables with each other, and between the input and output variables. All the algorithms and mathematical operations in this research paper are implemented using Matlab R2017a.

3 MODEL DEVELOPMENT

This section describes the background of the utilized meta-heuristics and the designed automated training optimization function.
3.1 Overview of the Meta-Heuristics

3.1.1 Differential Evolution
This algorithm starts by initiating different individuals (i.e., candidate solutions) in the population. These old solutions are updated to produce new solutions only one time in each iteration. Each solution is updated using three different random solutions and the mutation factor. One of the random solutions searches around the space, while the two other solutions create a distance from the first one. There must be a mixture between the previous and current solutions to generate new solutions in the next iteration. The selection operation phase is performed to maintain better solutions in the current iteration. Finally, the solution associated with the lowest fitness function is chosen as the best solution (Storn & Price, 1997).

3.1.2 Particle Swarm Optimization
This algorithm, which was introduced by Eberhart & Kennedy (1995), imitates the movement of a school of fish or flocks of birds searching for their food. This algorithm commences by forming the particles, whereas each particle in the swarm is characterized by its current position, best position, and velocity. The position of a particle is updated by considering randomized values in some directions. Meanwhile, the velocity of any particle relies on particle best position, global best position, and the random function (Shi & Eberhart, 1998). This process is repeated until reaching a stopping criterion (e.g., satisfactory solution, maximum number of iterations, and constant fitness for a certain number of iterations). Finally, the current position of the best particle in the last iteration is defined as the global best solution (Elshaboury et al., 2020).

3.1.3 Invasive Weed Optimization
This algorithm, which was proposed by Mehrabian & Lucas (2006), simulates the invasive behavior of weeds in finding the most suitable place for growth and reproduction (Elshaboury et al., 2021). The randomness, resistance, and adaptability of weeds are imitated in this algorithm. By definition, weeds are powerful herbs that grow unintentionally and cause a serious threat to crops. The basic features of this algorithm are considered as follows: (1) population initialization where the seeds are distributed in the search space, (2) reproduction each flowering plant produces seeds based on its fitness value, (3) spectral spread that eliminates inappropriate plants, (4) competitive deprivation that ensures removing the grass with worst fitness from the colony to keep a constant number of herbs in the colony, (5) repeating the previous steps until reaching the maximum number of iterations, and (6) storing the minimum cost function of the grasses (Misaghi & Yaghoobi, 2019).

3.1.4 Teaching Learning-Based Optimization
It is a recent meta-heuristics algorithm which is inspired by the behaviors of teachers and learners in a classroom. The teacher is determined to be the learner with the best grade in the population. The teacher is responsible for training the learners and improving the mean grade of the class (known as the teacher phase). Meanwhile, each learner randomly interacts with another learner (known as the learner phase). In this algorithm, a group of students (i.e., learner) is considered as the population. Besides, the different subjects resemble the optimization design variables. The fitness function of the optimization problem is determined based on the results of the learner. This process is repeated iteratively until the termination condition is met and the best solution in the entire population is considered as the teacher (Rao et al., 2011).

3.1.5 Ant Colony Optimization
This algorithm, which was proposed by Dorigo et al. (1999), was inspired by the behavior of ant colonies searching for their food sources. The ants communicate with each other using the pheromone, such that the ant that finds a food source deposits pheromone along the path to be followed by other ants. In this way, ants could find the shortest distance between their nests and the food source (Tran et al.,
The procedures of this algorithm are described as follows: (1) deriving solutions to the problem by using greedy information and pheromone values and (2) updating the pheromone values using the best solutions. Therefore, this optimization technique relies on learning from positive examples (i.e., positive learning) (Nurcahyadi & Blum, 2021).

3.1.6 Grey Wolf Optimization

This algorithm, which was introduced by Mirjalili et al. (2014), is inspired by the leadership and hunting process of grey wolves (Panda & Das, 2019). The leadership hierarchy is simulated using four types of grey wolves: alpha, beta, delta, and omega. The alphas are responsible for deciding the hunting time and resting place for the whole group. The beta wolves maintain discipline in the group and advise the alphas in their decisions. The delta wolves are authorized to dominate omegas while following the orders of alphas and betas. The last group of wolves follows the orders of all other dominant wolves. The hunting mechanism of grey wolves comprises tracking and chasing the prey, harassing the target, and attacking the prey target (Jitkongchuen et al., 2016).

3.1.7 Grasshopper Optimization

This algorithm, which was developed by Saremi et al. (2017), mimics the social behavior of grasshoppers. Grasshoppers are considered harmful insects and cause damage to crops. The lifecycle of grasshopper consists of three stages: egg, nymph, and adult. The nymph grasshopper is born when the eggs hatch. These young insects start rolling and eat everything that comes in their way. Then, they become adults and start swarming in the air (Gad et al., 2020). Each insect in the swarm is characterized by two important motions; cooperation between grasshoppers in larvae and insect phases and foraging for the food movement. A swarm is formed of randomly generated grasshopper agents. The fitness value evaluation determines the best search agent, and the search agents start moving toward the best grasshopper search agent (Ullah et al., 2020).

3.1.8 Moth-Flame Optimization

This algorithm, which was introduced by Mirjalili (2015a), is inspired by the navigational mechanisms of moths during the night. The moth maintains a fixed angle of the moonlight to be able to travel long distances in a straight path. Since the moon is far away from the moth, it flies in a straight line by using the near-parallel light near the surface. Despite the effectiveness of lateral orientation, moths often fly spirally around lights. The same case applies when the moths preserve a stable angle with the light source in the case of human-made artificial light. This algorithm commences by generating random moths within the search space, then evaluating the position (i.e., fitness value) of each moth, and determining the best position by flame. The moths’ positions are updated based on a spiral movement function, and the previous processes are updated until the termination criteria are met.

3.1.9 Antlion Optimization

This algorithm mimics the hunting mechanism of antlions (Mirjalili, 2015b). Antlions act as search agents that hunt preys by building traps, entrapping ants, catching preys, and rebuilding traps (Tian et al., 2018). This algorithm is characterized by a good exploration capability with the help of random walk and roulette wheel to build traps. Besides, the exploitation efficiency of this algorithm is highlighted by the time-varying boundary shrinking mechanism and elitism. The major advantages of this algorithm are as follows: ease of implementation, high precision, avoidance of local optima, and reduced need for parameter adjustment (Horng & Lee, 2021).

3.1.10 Arithmetic Optimization

This algorithm is motivated by the application of arithmetic operators (i.e., multiplication, division, subtraction, and addition) in solving arithmetic problems. This optimization algorithm starts by generating
a set of candidate solutions. The arithmetic operators determine the feasible positions of the near-optimal solution. The position of each solution is updated based on the exploration and exploitation mechanisms. Eventually, the algorithm stops by satisfying the end criterion (Abualigah et al., 2021).

3.2 Automated Training of Elman Neural Network

The developed meta-heuristic-based Elman neural network model is characterized by its both parametric and structural learning nature. It aims at the autonomous optimization of the numbers of hidden layers, hidden neurons, context neurons, context layers, type of transfer function, and values of weighted connections in the Elman neural network. The developed model explores the implementation of eight activation functions, namely radial basis function, normalized radial basis, triangular basis function, linear function, positive linear function, log-sigmoid function, hyperbolic tangent sigmoid function and Elliot symmetric sigmoid function. As a result of the structural learning features of the developed model, the number of weighted connections varies adaptively during each training epoch according to the numbers of the hidden layer, hidden neurons, and context layers. Hence, the number of weighted connections should be computed during each training epoch to design an efficient variable-length optimization model. The developed estimator is created for the sake of handling the variability in the length of the optimization model, and it can be mathematically expressed as follows.

\[
\text{Num} = \left( (I + 1) \times N \right) + \left( (N \times C \times P + \left( (N + 1) \times N \times (P - 1) \right) \right) + \left( (N + 1) \times O \right)
\]

Where; \(I\) and \(N\) denote the numbers of input neurons and hidden neurons, respectively. \(C\) represents the number of context neurons. \(P\) denotes the number of hidden and context layers. \(O\) stands for the number of output neurons. It is with mentioning that architecture of Elman neural network is triggered by the type of simulation problem (heating or cooling), and the search exploration ability of the employed meta-heuristic optimizer.

4 MODEL ASSESSMENT

This section describes the scoring metrics used to evaluate the developed prediction models and the average ranking algorithm.

4.1 Performance Comparison

The present research study utilizes five performance indicators to evaluate the developed models. In this context, the evaluation metrics of MAPE, RMSE, NSCE, WI, and PCC can be computed using Equations 2-6, respectively (Altunkaynak, 2019; Guo et al., 2019; Elshaboury & Marzouk, 2020; Natarajan & Nachimuthu, 2020).

\[
\text{MAPE} = \frac{100}{k} \times \sum_{i=1}^{k} \frac{|P_i - O_i|}{O_i}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{K} \sum_{i=1}^{K} \left( O_i - P_i \right)^{2}}
\]
\[ NSCE = 1 - \frac{\sum_{i=1}^{K}(P_i - O_i)^2}{\sum_{i=1}^{K}(O_i - \bar{O})^2} \] (4)

\[ WI = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{K}(|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \] (5)

\[ PCC = \frac{\sum_{i=1}^{K}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{K}(O_i - \bar{O})^2 \sum_{i=1}^{K}(P_i - \bar{P})^2}} \] (6)

where; \( P_i \) and \( O_i \) indicate the predicted and observed heating or cooling load, respectively. \( \bar{O} \) and \( \bar{P} \) are the average observed and predicted heating or cooling load, respectively. \( k \) stands for the size of data observations. A smaller value of MAPE and RMSE indicates a smaller deviation of the predicted values from the actual values. The values of NSCE and WI range from 0 to 1 while the values of PCC vary from -1 to 1. In this regard, the closer values of NSCE, WI, and PCC to 1 imply an increasing agreement between the predicted and actual heating or cooling loads.

4.2 Unified Ranking

Each one of the developed meta-heuristic-based Elman neural network models behaves differently according to the tackled type of the performance indicator. As such, the average ranking algorithm is utilized to generate a final sorting of the meta-heuristic-based Elman neural network models according to their scores with respect to the utilized performance evaluation metrics. The mean and standard deviation of the rankings can be obtained using Equations 7 and 8, respectively (Yu et al., 2018; Mohammed Abdelkader et al., 2021).

\[ \frac{1}{A} = \frac{\sum_{b=1}^{n} R_b}{B} \] (7)

\[ \bar{A} = \frac{\sum_{b=1}^{n} (R_b - \frac{1}{A})^2}{B} \] (8)

where; \( \frac{1}{A} \) and \( \bar{A} \) indicate the mean and standard deviation of the rankings of the meta-heuristic-based Elman neural network models, respectively. \( B \) denotes the number of performance metrics. A lower value of \( \frac{1}{A} \) and \( \bar{A} \) imply a high-performing and more robust prediction model.
5 RESULTS AND DISCUSSION

The utilized dataset is composed of 768 observations, whereas 614 (80%) and 154 (20%) data instances are randomly selected and deployed for training and testing purposes. Figures 2, 3, 4 and 5 describe the relationship between each of the influential independent variables and the output dependent variables. Table 1 summarizes a set of statistical indicators for the input and output variables. The statistical analysis includes minimum, maximum, average, median, standard deviation, mode, skewness, sample variance, range, kurtosis, and coefficient of variation. For instance, the highest value of standard deviation is for glazing area distribution (88.03) followed by overall height (45.14) while the glazing area has the least standard deviation (0.11). With regards to the coefficient of variation, the surface area (56.8%) is associated with the highest coefficient of variation and then relative compactness (55.11%). The least coefficients of variation are for glazing area (13.83%), glazing area distribution (13.11%) and orientation (13.69%).

Figure 2. Spatial distributions of glazing area and glazing area distribution against heating and cooling loads
Different initializations of parameters' configuration were experimented in order to search for their optimum setting. In order to create a fair comparison between the ten meta-heuristics, the number of iterations and population size are assumed in all meta-heuristics to be 200 and 50, respectively. In the differential evolution algorithm, the mutation rate is assumed to be uniform distribution ranging between 0.2 and 0.8. In addition, the crossover probability rate is set as 0.2. In the particle swarm optimization algorithm, the social parameter and cognitive learning parameter are assumed as two, and inertia weight is assumed as 0.5. In the invasive weed optimization algorithm, the non-linear modulation index is assumed as two, and the initial and final standard deviations are 0.5 and 0.001, respectively. The maximum and minimum numbers of seeds are equal to five and zero, respectively. With regards to the ant colony optimization algorithm, the sample size, intensification factor and deviation-distance ratio are assumed as 40, 0.5 and 1, respectively. For the grey wolf optimization algorithm, the trade-off parameter which manages the balance between exploitation and exploration.

Figure 3. Spatial distributions of orientation and overall height against heating and cooling
is assumed to be linearly varying from two to zero. In the grasshopper optimization, the maximum and minimum deceleration values of grasshoppers when approaching food source and eating it are assumed 1 and 0.00004, respectively. In the moth-flame optimization algorithm, the convergence constant is assumed to be decreasing from -1 and -2, and the logarithmic spiral motion constant is set as 1. The search control parameter and exploitation control parameter in the arithmetic optimization algorithm are assumed 0.5 and 5, respectively. Each meta-heuristic was then run for five times in an independent manner to circumvent unstable solutions and the best solution associated with the run of the least prediction error was appended to be used for the forthcoming performance comparison analysis.

As mentioned earlier, ten hybrid meta-heuristic-based Elma neural network models are developed for the sake of predicting the heating and cooling loads. The convergence of the models for predicting heating loads is demonstrated in Figure 6. ENN-PSO accomplished the least MAPE and ENN-GWO
achieved the second-lowest MAPE followed by ENN-MFO while ENN-AO provided the highest MAPE. In this context, ENN-PSO, ENN-GWO, ENN-MFO and ENN-AO yielded MAPE of 8.62%, 8.98%, 9.26% and 21.81%, respectively. It is worth mentioning that ENN-TLBO, ENN-IWO, and ENN-ALO attained MAPE of 9.09%, 9.67%, and 12.76%, respectively. With respect to the cooling loads, the convergence of the models is depicted in Figure 7. ENN-IWO achieved the lowest prediction error with a MAPE of 8.08%. On the other hand, ENN-AO provided the highest error rate with MAPE of 16.32%. ENN-PSO, ENN-GWO, and ENN-TLBO attained MAPE of 8.19%, 8.2%, and 8.61%, respectively.

The optimum structures of the developed meta-heuristic-based Elman neural network models for the prediction of heating and cooling loads are reported in Tables 2 and 3, respectively. It can be noticed that different architectures are obtained for either the prediction of the amount of heating or cooling loads, which imply different optimal solutions are generated by the ten meta-heuristic Elman neural
Table 1. Summary of the descriptive statistics of the input and output parameters

<table>
<thead>
<tr>
<th>Descriptive statistic</th>
<th>Glazing area</th>
<th>Glazing area distribution</th>
<th>Orientation</th>
<th>Overall height</th>
<th>Roof area</th>
<th>Wall area</th>
<th>Surface area</th>
<th>Relative Compactness</th>
<th>Heating load</th>
<th>Cooling load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0.62</td>
<td>514.5</td>
<td>245</td>
<td>110.25</td>
<td>3.5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>6.01</td>
<td>10.90</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.98</td>
<td>808.5</td>
<td>416.5</td>
<td>220.5</td>
<td>7</td>
<td>5</td>
<td>0.4</td>
<td>5</td>
<td>43.1</td>
<td>48.03</td>
</tr>
<tr>
<td>Average</td>
<td>0.76</td>
<td>671.71</td>
<td>318.5</td>
<td>176.6</td>
<td>5.25</td>
<td>3.5</td>
<td>0.23</td>
<td>2.81</td>
<td>22.31</td>
<td>24.59</td>
</tr>
<tr>
<td>Median</td>
<td>0.75</td>
<td>673.75</td>
<td>318.5</td>
<td>183.75</td>
<td>5.25</td>
<td>3.5</td>
<td>0.25</td>
<td>3</td>
<td>18.95</td>
<td>22.08</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.11</td>
<td>88.03</td>
<td>43.6</td>
<td>45.14</td>
<td>1.75</td>
<td>1.12</td>
<td>0.13</td>
<td>1.55</td>
<td>10.08</td>
<td>9.51</td>
</tr>
<tr>
<td>Mode</td>
<td>0.98</td>
<td>514.5</td>
<td>294</td>
<td>220.5</td>
<td>7</td>
<td>2</td>
<td>0.1</td>
<td>1</td>
<td>15.16</td>
<td>21.33</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.5</td>
<td>-0.13</td>
<td>0.53</td>
<td>-0.16</td>
<td>0</td>
<td>0</td>
<td>-0.06</td>
<td>-0.09</td>
<td>0.36</td>
<td>0.4</td>
</tr>
<tr>
<td>Sample variance</td>
<td>0.01</td>
<td>7749.06</td>
<td>1900.79</td>
<td>2037.31</td>
<td>3.06</td>
<td>1.25</td>
<td>0.02</td>
<td>2.40</td>
<td>101.68</td>
<td>90.39</td>
</tr>
<tr>
<td>Range</td>
<td>0.36</td>
<td>294</td>
<td>171.50</td>
<td>110.25</td>
<td>3.50</td>
<td>3.00</td>
<td>0.4</td>
<td>5</td>
<td>37.09</td>
<td>37.13</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.71</td>
<td>-1.06</td>
<td>0.12</td>
<td>-1.78</td>
<td>-2.01</td>
<td>-1.36</td>
<td>-1.33</td>
<td>-1.15</td>
<td>-1.25</td>
<td>-1.15</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>13.83%</td>
<td>13.11%</td>
<td>13.69%</td>
<td>25.56%</td>
<td>33.33%</td>
<td>31.94%</td>
<td>56.80%</td>
<td>55.11%</td>
<td>45.20%</td>
<td>38.67%</td>
</tr>
</tbody>
</table>

Figure 6. Convergence of the meta-heuristic-based Elman neural network models for forecasting heating loads
network models. This exemplifies that the complexity of optimization problem and the exploration-exploitation search capabilities play a principal role in the quality of generated optimal solutions, which evinces the need for studying the need for studying the implementation of five different types of meta-heuristics. In the prediction of heating loads, the optimum structure of ENN-PSO consists of 3 hidden and context layers, 4 hidden and context neurons, and hyperbolic tangent sigmoid as the optimum transfer function. The optimum structure of ENN-GO is composed of 4 hidden and context layers, 2 hidden and context neurons, and linear function as the optimum transfer function. At the level of predicting cooling loads, the optimum structure of ENN-GWO comprises 3 hidden and context layers, 6 hidden and context neurons, and hyperbolic tangent sigmoid as the optimum transfer function. The optimum topology of ENN-GO encompasses 2 hidden and context layers, 8 hidden and context neurons, and log-sigmoid as the optimum transfer function. In addition to that, it is found that ENN-DE obtained varying topologies in the interpretation of heating and loading patterns. For instance, ENN-DE obtained one hidden layer, one context layer, two hidden and context neurons and hyperbolic tangent sigmoid function in heating loads assessment. In cooling loads evaluation, ENN-DE achieved an optimum topology of two hidden layers, two context layers, seven hidden and context neurons, and Elliot symmetric sigmoid activation function.

Illustrations of the performances of ENN-PSO, ENN-GWO, ENN-ACO, and ENN-ALO in the prediction of heating loads are depicted in Figures 8 and 9. X-axis represents a certain data instance
which corresponds to a specific scenario for a combination of the input variables. Y-axis denotes the corresponding heating load, whereas five hundred data instances that were randomly sampled from training and testing partitions are plotted in these figures. The black and blue curves denote the actual and predicted heating values that are obtained by the meta-heuristic-based Elman neural network models. It can be inferred that ENN-PSO, ENN-GWO have successfully predicted the heating loads pattern, whereas they produce very close predicted values to the actual heating loads. However, ENN-ACO and ENN-ALO failed to predict the actual heating loads efficiently providing considerable

<table>
<thead>
<tr>
<th>Model</th>
<th>Hidden Layers</th>
<th>Hidden Neurons</th>
<th>Context Layers</th>
<th>Context Neurons</th>
<th>Transfer Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN-DE</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-PSO</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-IWO</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>Elliot symmetric sigmoid</td>
</tr>
<tr>
<td>ENN-TLBO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-ACO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-GWO</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Elliot symmetric sigmoid</td>
</tr>
<tr>
<td>ENN-GO</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>Linear</td>
</tr>
<tr>
<td>ENN-MFO</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-AO</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-ALO</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Hidden Layers</th>
<th>Hidden Neurons</th>
<th>Context Layers</th>
<th>Context Neurons</th>
<th>Transfer Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN-DE</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>Elliot symmetric sigmoid</td>
</tr>
<tr>
<td>ENN-PSO</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>Linear</td>
</tr>
<tr>
<td>ENN-IWO</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>Elliot symmetric sigmoid</td>
</tr>
<tr>
<td>ENN-TLBO</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-ACO</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-GWO</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
<tr>
<td>ENN-GO</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>Log-sigmoid</td>
</tr>
<tr>
<td>ENN-MFO</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>Log-sigmoid</td>
</tr>
<tr>
<td>ENN-AO</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>Elliot symmetric sigmoid</td>
</tr>
<tr>
<td>ENN-ALO</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>Hyperbolic tangent sigmoid</td>
</tr>
</tbody>
</table>
mismatch with the observed heating patterns. Figures 10 and 11 show the performances of ENN-PSO, ENN-GWO, ENN-ACO, and ENN-ALO in the prediction of cooling loads. Y-axis represents the cooling load in these figures, as five hundred data instances were randomly picked from the training and testing partitions, and visualized. The black and blue curves depict the actual and predicted amounts of cooling loads generated from the meta-heuristic-based Elman neural network models. Graphical comparisons reveal that the cooling pattern is better recognized by ENN-PSO, ENN-GWO than by ENN-ACO and ENN-ALO. In this regard, the prediction trend of the cooling loads provided by the models of ENN-PSO, ENN-GWO models resembled the measured actual values. However, explicit variations are encountered between the actual and predicted values by the models of ENN-ACO and ENN-ALO. Figure 12 describes the correlation between the expected and measured heating loads for all the training and testing instances based on ENN-PSO. The regression chart shows that the determination coefficient \( \left( R^2 \right) \) is 92.6%, which demonstrates a very good agreement between the predicted and measured heating loads. A graphical comparison between the predicted and actual cooling loads for all data instances based on ENN-GWO is shown in Figure 13. It can be found that an acceptable consistency \( \left( R^2 = 88.3\% \right) \) is obtained between the predicted and actual cooling loads.

A performance comparative analysis between the meta-heuristic-based Elman neural network models for predicting heating loads is reported in Table 4. It can be concluded that the ENN-PSO model

Figure 8. Actual and predicted heating loads using ENN-PSO and ENN-GWO models
achieved the lowest prediction error followed by ENN-GWO and then ENN-IWO. On the other hand, ENN-AO provided the highest prediction error. In this regard, the ENN-PSO model accomplished MAPE, RMSE, NSCE, WI, and PCC of 8.335%, 2.732, 0.927, 0.98, and 0.927, respectively. ENN-GWO provided MAPE, RMSE, NSCE, WI, and PCC of 8.788%, 2.947, 0.915, 0.976, and 0.92, respectively. It could be also observed that the ENN-AO model yielded MAPE, RMSE, NSCE, WI, and PCC of 21.972%, 5.881, 0.66, 0.916, and 0.746, respectively. With respect to cooling loads, their prediction performances are recorded in Table 5. It can be noticed that ENN-GWO achieved the highest prediction performance, whereas it provided MAPE, RMSE, NSCE, WI, and PCC of 8.154%, 3.343, 0.876, 0.956, and 0.884, respectively. The developed models of ENN-PSO and ENN-IWO yielded the second highest prediction performance. In addition, it is observed that ENN-AO attained the lowest prediction performance, such that MAPE, RMSE, NSCE, WI, and PCC are equal to 17.342%, 6.156, 0.581, 0.844, and 0.781, respectively. As such, it can be interpreted that the developed ENN-PSO and ENN-GWO provided lower error rate than ENN-AO by 37.46% and 35.18% in the prediction of amounts of heating and cooling loads, respectively. At the level of both heating and cooling loads, it can be argued that the developed modes of ENN-PSO, ENN-GWO, ENN-IWO and ENN-TLBO demonstrated satisfactory good prediction performance. This can be explained by the high exploration and local search behaviors of these models which allowed them to efficiently investigate the full range of entire search space meanwhile avoiding local minima stagnation. On the
other hand, the developed models of ENN-AO, ENN-GO and ENN-ALO failed to adequately inspect the search space of parameters and hyper parameters of Elman neural network as a result of their low exploration abilities and slow convergence rate.

The average ranking algorithm is utilized to blend the performances of the ten meta-heuristic-based Elman neural network models with respect to the five evaluation metrics. The results generated from the average ranking algorithm for the prediction of heating and cooling loads are presented in Tables 6 and 7, respectively. In the prediction of heating loads, it is found that ENN-PSO accomplished the first ranking and the most stable performance followed by ENN-TLBO while ENN-AO provided the tenth ranking. In this regard, $\mu_a$ and $\tilde{A}_a$ of ENN-PSO are equal to 1 and 0, respectively. For ENN-TLBO, $\mu_a$ and $\tilde{A}_a$ are equal to 2.4 and 0.49, respectively. For cooling loads, ENN-GWO accomplished the first ranking followed by ENN-PSO and then ENN-IWO. ENN-GWO achieved $\mu_a$ and $\tilde{A}_a$ of 1.6 and 1.2, respectively. It is also shown that $\mu_a$ and $\tilde{A}_a$ of ENN-IWO are equal to 2.4 and 0.49, respectively. Furthermore, it is worth noting that the performances of meta-heuristics vary from one optimization problem due to their case dependent nature, whereas some meta-heuristics can outperform their counterparts in some sort of problem. However, they may underperform in other optimization problems. For instance, ENN-ACO accomplished higher rank than ENN-ALO in heating loads prediction. However, it is outranked by it in cooling loads prediction. Moreover, ENN-TLBO
Figure 11. Actual and predicted cooling loads using ENN-ACO and ENN-ALO models

Figure 12. Correlation between the observed and predicted heating loads using ENN-PSO model
obtained higher ranking than ENN-GWO in predicting amounts of heating loads. Nevertheless, ENN-GWO outranks ENN-TLBO in cooling loads assessment.

Further comparative performance evaluations of meta-heuristic-based Elman neural network models against classical machine learning models are reported in Table 8. In the back propagation artificial neural network (BPANN) model, the numbers of hidden layers and hidden neurons are assumed 5 and 2, respectively. The momentum coefficient and learning rate are assumed 0.8 and 0.01, respectively. The activation function is log sigmoid and gradient descent algorithm is used for
training the network. In the support vector machines (SVM) model, Gaussian radial basis function is utilized with sigma and scaling factor equal to 1. In both heating and cooling loads, the developed ENN-PSO and ENN-GWO performed better than BPANN and SVM. In this context, ENN-PSO and ENN-GWO provided higher prediction accuracies than BPANN and SVM in heating and cooling,
respectively. It is also found that ENN-PSO and ENN-GWO can improve the prediction performance of SVM by 27.68% and 14.49% in heating and cooling, respectively. With reference to BPANN, it is derived that ENN-PSO and ENN-GWO managed to improve the prediction accuracies by 29.75% and 40.84% in heating and cooling loads prediction, respectively. BPANN obtained higher error rate than the SVM in both heating and cooling prediction. In heating behavior prediction, SVM yielded MAPE, RMSE, NSCE, WI, and PCC of 16.15%, 4.869, 0.767, 0.933, and 0.771, respectively. With regards to cooling loads, it is found that SVM exhibited MAPE, RMSE, NSCE, WI, and PCC of 13.809%, 4.184, 0.806, 0.943, and 0.872, respectively.

Table 7. Average and standard deviation of the rankings of the meta-heuristic-based Elman neural network models for predicting cooling loads

<table>
<thead>
<tr>
<th>Model</th>
<th>Average ranking</th>
<th>Standard deviation of ranking</th>
<th>Final ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN – DE</td>
<td>6.8</td>
<td>0.98</td>
<td>7</td>
</tr>
<tr>
<td>ENN – PSO</td>
<td>2.2</td>
<td>0.75</td>
<td>2</td>
</tr>
<tr>
<td>ENN – IWO</td>
<td>2.4</td>
<td>0.49</td>
<td>3</td>
</tr>
<tr>
<td>ENN – TLBO</td>
<td>3.80</td>
<td>0.4</td>
<td>4</td>
</tr>
<tr>
<td>ENN – ACO</td>
<td>8.6</td>
<td>0.49</td>
<td>9</td>
</tr>
<tr>
<td>ENN – GWO</td>
<td>1.6</td>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>ENN – GO</td>
<td>8.4</td>
<td>1.02</td>
<td>8</td>
</tr>
<tr>
<td>ENN – MFO</td>
<td>5.2</td>
<td>0.4</td>
<td>5</td>
</tr>
<tr>
<td>ENN – AO</td>
<td>9.8</td>
<td>0.4</td>
<td>10</td>
</tr>
<tr>
<td>ENN – ALO</td>
<td>6.2</td>
<td>0.4</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 8. Comparative performance evaluation against classical machine learning models

### Heating Loads

<table>
<thead>
<tr>
<th>Model</th>
<th>MAPE</th>
<th>RMSE</th>
<th>NSCE</th>
<th>WI</th>
<th>PCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN – PSO</td>
<td>8.335%</td>
<td>2.732</td>
<td>0.927</td>
<td>0.98</td>
<td>0.927</td>
</tr>
<tr>
<td>BPANN</td>
<td>17.9%</td>
<td>4.967</td>
<td>0.757</td>
<td>0.924</td>
<td>0.761</td>
</tr>
<tr>
<td>SVM</td>
<td>16.15%</td>
<td>4.869</td>
<td>0.767</td>
<td>0.933</td>
<td>0.771</td>
</tr>
</tbody>
</table>

### Cooling Loads

<table>
<thead>
<tr>
<th>Model</th>
<th>MAPE</th>
<th>RMSE</th>
<th>NSCE</th>
<th>WI</th>
<th>PCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN – GWO</td>
<td>8.154%</td>
<td>3.343</td>
<td>0.876</td>
<td>0.956</td>
<td>0.884</td>
</tr>
<tr>
<td>BPANN</td>
<td>14.831%</td>
<td>5.938</td>
<td>0.61</td>
<td>0.665</td>
<td>0.69</td>
</tr>
<tr>
<td>SVM</td>
<td>13.809%</td>
<td>4.184</td>
<td>0.806</td>
<td>0.943</td>
<td>0.872</td>
</tr>
</tbody>
</table>
Computational time is an important aspect when comparing the prediction models. Table 9 reports the average training time and average testing time of the meta-heuristic-based Elman neural network models. The used laptop is of specifications: Intel Core i7 CPU, 2.21 GHz and 16 GB of memory. It can be observed that ENN-GWO and ENN-MFO are the most time efficient prediction models, whereas they consumed lesser training times than the remainder of other prediction models. In this regard, the training times of ENN-GWO and ENN-MFO are 10.88 and 14.26 minutes, respectively. It can be also interpreted that ENN-IWO and ENN-AO have relatively short training times of 16.34 and 19.42 minutes, respectively. On the other hand, ENN-GO has the longest training time with 94.49 minutes. In addition, it can be noticed that most of the developed meta-heuristic-based Elman neural network models nearly consume the same testing times, whereas all of their testing times vary from 1.25 to 2.45 seconds.

Table 10 describes the correlation between the input and output variables. As can be seen, the glazing area has a negative correlation with orientation and overall height. The roof area has a positive strong correlation with heating load and cooling load. The overall height has a negative strong correlation with heating load and cooling load. In this regard, the correlation coefficients of the roof area with heating load and cooling load are 0.889 and 0.896, respectively. Additionally, the correlation coefficients of overall height with heating and cooling loads are -0.862 and -0.863, respectively. It is also revealed that wall area and relative compactness manifested low levels of dependencies with heating and cooling loads. It is also evinced that the amount of heating load is highly correlated with cooling load (correlation coefficient is 0.976).

6 CONCLUSION

Analyzing energy performance is crucial to maintain sustainable energy consumption in residential buildings. As such, this research paper proposed a novel hybrid meta-heuristic-based Elman neural network model for autonomous forecasting of amounts of heating and cooling loads in residential buildings. Performance comparison demonstrated that the complexity of the designated optimization problem and search capability of meta-heuristic significantly implicate the performance of meta-

<table>
<thead>
<tr>
<th>Model</th>
<th>Training time (Sec)</th>
<th>Testing time (Sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENN – DE</td>
<td>53.31</td>
<td>2.45</td>
</tr>
<tr>
<td>ENN – PSO</td>
<td>37.80</td>
<td>1.34</td>
</tr>
<tr>
<td>ENN – IWO</td>
<td>16.34</td>
<td>1.59</td>
</tr>
<tr>
<td>ENN – TLBO</td>
<td>37.39</td>
<td>1.46</td>
</tr>
<tr>
<td>ENN – ACO</td>
<td>39.51</td>
<td>1.36</td>
</tr>
<tr>
<td>ENN – GWO</td>
<td>10.88</td>
<td>2.35</td>
</tr>
<tr>
<td>ENN – GO</td>
<td>94.49</td>
<td>2.25</td>
</tr>
<tr>
<td>ENN – MFO</td>
<td>14.26</td>
<td>2.36</td>
</tr>
<tr>
<td>ENN – AO</td>
<td>19.42</td>
<td>1.35</td>
</tr>
<tr>
<td>ENN – ALO</td>
<td>49.52</td>
<td>1.25</td>
</tr>
</tbody>
</table>
heuristics. In this context, different meta-heuristics obtained different optimum architectures in the same optimization problem. In addition, same meta-heuristic yielded different optimum topologies in different optimization problems (heating and cooling assessment). Comparative analysis also illustrated that the developed ENN-PSO and ENN-GWO accomplished the highest performance accuracies among the meta-heuristic-based Elman neural network models in heating and cooling loads prediction, respectively. In this regard, they managed to reduce the prediction error of ENN-AO by 37.46% and 35.18% in heating and cooling loads, respectively. Against the ENN-DE model, the developed ENN-PSO and ENN-GWO managed to improve the prediction accuracies by 13.44% and 19.55% in heating and cooling loads, respectively. The outcome of the average ranking algorithm manifested that ENN-PSO, ENN-AO, and ENN-ALO models exhibited the most stable ranking in the heating loads prediction $(\hat{A}_n = 0)$ meanwhile the ENN-TLBO, ENN-MFO, ENN-AO, and ENN-ALO models accomplished the lowest standard deviation of rankings $(\hat{A}_n = 0.4)$ in cooling loads. In addition to that, the average ranking algorithm exemplified that the relative importance (i.e., ranking) varied from heating to cooling loads prediction. For example, EN-PSO obtained a higher ranking (first) than ENN-GWO (third) in heating loads simulation. Nevertheless, ENN-GWO (first) outranked ENN-PSO (second) in cooling loads simulation.

When compared against the conventional machine learning models of BPANN and SVM, the developed ENN-PSO succeeded in diminishing the forecasting error of heating loads by 29.75% and 27.68%, respectively. Additionally, the developed ENN-GWO was able to reduce the prediction error of cooling loads estimation by 40.84% and 14.98%, respectively. Analysis of computation times of the meta-heuristic-based Elman neural network models revealed that ENN-GWO and ENN-MFO had the shortest training times. Besides, ENN-IWO and ENN-AO consumed comparatively short training times. It was also observed that ENN-GO requires substantially long training time although it produces high prediction error (eighth ranking in both heating and cooling). The constructed correlation matrices illustrated that the highest level of correlation were exhibited between the input factors of roof area and overall height, and the output variables of heating and cooling loads. It was also noticed that wall area and relative compactness are lowly correlated with heating and cooling loads. It is expected that the developed model can aid architectural designers through providing them with a platform to study the implications of several design parameters on buildings’ energy efficiency.

Table 10. Correlation matrix of the input and output variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Glazing area</th>
<th>Glazing area distribution</th>
<th>Orientation</th>
<th>Overall height</th>
<th>Roof area</th>
<th>Wall area</th>
<th>Surface area</th>
<th>Relative Compactness</th>
<th>Heating load</th>
<th>Cooling load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glazing area</td>
<td>1</td>
<td>-0.992</td>
<td>-0.204</td>
<td>-0.869</td>
<td>0.828</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.622</td>
<td>0.634</td>
</tr>
<tr>
<td>Glazing area distribution</td>
<td>-0.992</td>
<td>1</td>
<td>0.196</td>
<td>0.881</td>
<td>-0.858</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.658</td>
<td>-0.673</td>
</tr>
<tr>
<td>Orientation</td>
<td>-0.204</td>
<td>0.196</td>
<td>1</td>
<td>-0.292</td>
<td>0.281</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.456</td>
<td>0.427</td>
</tr>
<tr>
<td>Overall height</td>
<td>-0.869</td>
<td>0.881</td>
<td>-0.292</td>
<td>1</td>
<td>-0.973</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-0.862</td>
<td>-0.863</td>
</tr>
<tr>
<td>Roof area</td>
<td>0.828</td>
<td>-0.858</td>
<td>0.281</td>
<td>-0.973</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.889</td>
<td>0.896</td>
</tr>
<tr>
<td>Wall area</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-0.003</td>
<td>0.014</td>
</tr>
<tr>
<td>Surface area</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.213</td>
<td>0.27</td>
<td>0.087</td>
<td>0.051</td>
</tr>
<tr>
<td>Relative Compactness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.213</td>
<td>1</td>
<td>0.087</td>
<td>0.051</td>
</tr>
<tr>
<td>Heating load</td>
<td>0.622</td>
<td>-0.658</td>
<td>0.456</td>
<td>-0.862</td>
<td>0.889</td>
<td>-0.003</td>
<td>0.27</td>
<td>0.087</td>
<td>1</td>
<td>0.976</td>
</tr>
<tr>
<td>Cooling load</td>
<td>0.634</td>
<td>-0.673</td>
<td>0.427</td>
<td>-0.863</td>
<td>0.896</td>
<td>0.014</td>
<td>0.208</td>
<td>0.051</td>
<td>0.976</td>
<td>1</td>
</tr>
</tbody>
</table>
REFERENCES


---

Eslam Mohammed Abdelkader is a research assistant in Building Engineering- Concordia University, Canada. He worked as an assistant lecturer in Structural Engineering Department in Cairo University. He received his B.SC and M.SC degrees in Civil Engineering from Cairo University, Egypt in 2012, and 2017, respectively. The topic of his M.SC thesis is “Optimizing Construction Emissions for Sustainable Construction Projects”. Eslam attended a TOT workshop meeting for EU Tempus Project: Integrated Design Environment for Engineering Education (BimIDE3) at Chalmers University of Technology. His computer skills include programming, Autocad, Revit, Sap, Primavera. His research interests include sustainability, optimization, simulation, building information modeling, multi-criteria decision making, fuzzy set theory, programming, infrastructure management, ground penetrating radar, and digital imaging.