
Le Thi Le 1,*, Hoang Nguyen 2,*, Jie Dou 3 and Jian Zhou 4

1 Thanh Hoa University of Culture, Sports and Tourism, Thanh Hoa 440000, Vietnam
2 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
3 Civil and Environmental Engineering, Nagaoka University of Technology, 1603-1, Kami-Tomioka, Nagaoka, Niigata 940-2188, Japan
4 School of Resources and Safety Engineering, Central South University, Changsha 410083, China

* Correspondence: lethiledhvhtdl@gmail.com (L.T.L.); nguyenhoang23@duytan.edu.vn (H.N.)

Received: 3 June 2019; Accepted: 27 June 2019; Published: 28 June 2019

Abstract: Energy-efficiency is one of the critical issues in smart cities. It is an essential basis for optimizing smart cities planning. This study proposed four new artificial intelligence (AI) techniques for forecasting the heating load of buildings’ energy efficiency based on the potential of artificial neural network (ANN) and meta-heuristics algorithms, including artificial bee colony (ABC) optimization, particle swarm optimization (PSO), imperialist competitive algorithm (ICA), and genetic algorithm (GA). They were abbreviated as ABC-ANN, PSO-ANN, ICA-ANN, and GA-ANN models; 837 buildings were considered and analyzed based on the influential parameters, such as glazing area distribution (GLAD), glazing area (GLA), orientation (O), overall height (OH), roof area (RA), wall area (WA), surface area (SA), relative compactness (RC), for estimating heating load (HL). Three statistical criteria, such as root-mean-squared error (RMSE), coefficient determination ($R^2$), and mean absolute error (MAE), were used to assess the potential of the aforementioned models. The results indicated that the GA-ANN model provided the highest performance in estimating the heating load of buildings’ energy efficiency, with an RMSE of 1.625, $R^2$ of 0.980, and MAE of 0.798. The remaining models (i.e., PSO-ANN, ICA-ANN, ABC-ANN) yielded lower performance with RMSE of 1.932, 1.982, 1.878; $R^2$ of 0.972, 0.970, 0.973; MAE of 1.027, 0.980, 0.957, respectively.

Keywords: smart building; meta-heuristic algorithm; heating load; smart city; hybrid model

1. Introduction

One of the indispensable components for smart cities is energy and the applications of artificial intelligence (AI) [1]. Nowadays, smart cities are becoming more popular and the first choice for those who want a comfortable and productive life [2–5]. This includes intelligent, modern, energy efficient utilities, as well as sustainable environmental protection [6–8]. Of those components, heating load (HL) and cooling load (CL) systems are a part of energy efficiency. Many studies were conducted to predict and optimize the use of buildings’ energy efficiency (EEB) as well as building energy consumption [9]. For instance, Catalina et al. [10] used multiple regression method to estimate the demand for heating energy of the building. The south equivalent surface, global heat loss coefficient of building, and the difference between the sol-air and the indoor temperatures, were used as the input variables to estimate the demand of heating energy in their study. Their positive results were confirmed with a determination coefficient ($R^2$) of 0.987. Chou, Bui [11] also developed an ensemble model based on...
support vector regression (SVR) and an artificial neural network (ANN) to predict HL and CL for building design, called ANN-SVR, using the datasets of 17 buildings. A variety of the other models were also considered and developed to investigate and compare with their proposed ANN-SVR model, including SVR, ANN, chi-squared automatic interaction detector, classification and regression tree, and general linear regression. Their results confirmed the feasibility of AI techniques designing and optimizing EEB systems, especially the ANN-SVR model with a mean absolute percentage error (MAPE) below 4% and root-mean-squared error (RMSE) lower than 39%–65.9% in a comparison of the previous works [12,13]. In another study, Castelli et al. [14] applied a genetic programming (GP) model for evaluating the energy efficiency of EEB systems. Three forms of GP were investigated and compared, such as geometric semantic GP (GSGP), GSGP with local search (HYBRID), and HYBRID with linear scaling (HYBRID-LIN). Their results indicated that the HYBRID-LIN technique provided better results than the other techniques (i.e., GP, HYBRID). Deep learning techniques in AI have also been developed to estimate the energy efficiency of EEB systems (i.e., CL) by Fan et al. [15]. The potential of deep learning was exploited and interpreted for a variety of AI models in predicting CL of EEB systems during 24 h, including multiple linear regression (MLR), elastic net, random forest (RF), gradient boosting trees (GBT), SVR, extreme gradient boosting (XGB), and deep learning (DNN). Their results showed that their XGB model with deep learning technique yielded the highest accuracy with an RMSE of 106.5 and MAE of 71.6. The efforts for optimizing an ANN model using uncertainty means and sensitivity analyses were conducted to predict the energy demand of buildings by Ascione et al. [16]. Its performance was proven in a short-term prediction of the energy demand of buildings. As a result, their findings showed the powerful potential of the optimized ANN model in predicting the energy demand of buildings with an $R^2$ of 0.995, and the average relative error is between 2.0% and 11%. Ngo [17] also developed an ensemble machine learning model to predict the CL of EEB systems with high accuracy (e.g., RMSE = 158.77, MAE = 112.07, MAPE = 6.17%, and $R^2 = 0.990$). By the use of a hybrid model (M5Rules-particle swarm optimization (PSO)), Nguyen et al. [18] predicted the CL of EEB systems with a promising result. A similar study for predicting the HL of EEB systems was also performed by Bui et al. [19], using a novel hybrid approach, i.e., M5Rules-genetic algorithm (GA). By the use of the meta-heuristic algorithms (i.e., PSO, GA) to optimize the M5Rules model, Nguyen et al. [18] and Bui et al. [19] provided two new hybrid intelligent techniques (i.e., M5Rules-PSO and M5Rules-GA) to predict the CL and HL of EEB systems with high accuracy, i.e., RMSE of 0.0066, 0.0548, and $R^2$ of 0.999, 0.998, for the M5Rules-PSO and M5Rules-GA, respectively. Additionally, many other studies used/developed AI techniques for evaluating and predicting energy consumption as well as its efficiency [20–24].

According to the best review of the authors, meta-heuristics algorithms with a combination of ANN model was considered and developed in many areas with high reliability [25–35]; however, they are not still considered and prepared for estimating the HL of EEB systems. Therefore, this study developed and proposed four novel hybrid models based on four meta-heuristics algorithms and ANN model, for estimating the HL of EEB systems, namely PSO-ANN, GA-ANN, imperialist competitive algorithm (ICA)-ANN, and artificial bee colony (ABC)-ANN models. Four meta-heuristics algorithms were considered in this study, including artificial bee colony (ABC) optimization, particle swarm optimization (PSO), imperialist competitive algorithm (ICA), and genetic algorithm (GA). They were abbreviated as ABC-ANN, PSO-ANN, ICA-ANN, and GA-ANN models.

2. Data Collection and Its Characteristics

For data collection, twelve types of buildings were investigated and simulated by Ecotect computer software [13]. Accordingly, 768 experimental datasets were simulated and collected by Tsanas, Xifara [13]. To ensure the diverse of the dataset, 69 other buildings (during the winter of 2018) were also considered and investigated in Vietnam with similar conditions and materials. Finally, a total of 837 experimental datasets were considered and analyzed for estimating the HL of EEB systems in this work. Floor/surface area (SA), roof area (RA), wall area (WA), and overall height (OH),
were considered as the main components of the buildings, as illustrated in Figure 1. Additionally, glazing area distribution (GLAD), relative compactness (RC), glazing area (GLA), and orientation (O) were also extended investigated for estimating the HL of EEB systems. Table 1 summaries the heating load of the energy efficiency database used herein. Also, Figure 2 illustrates the properties of the dataset used for estimating the HL of EEB systems in this study.

Figure 1. Illustrating the components of building [11].

Table 1. Summary of the heating load of the energy efficiency database used.

<table>
<thead>
<tr>
<th>Elements</th>
<th>GLAD</th>
<th>GLA</th>
<th>O</th>
<th>OH</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>1.000</td>
<td>0.00</td>
<td>1.000</td>
<td>1.040</td>
<td>138.2</td>
</tr>
<tr>
<td>Mean</td>
<td>3.016</td>
<td>22.54</td>
<td>2.581</td>
<td>5.509</td>
<td>180.5</td>
</tr>
<tr>
<td>Max.</td>
<td>5.000</td>
<td>50.00</td>
<td>4.000</td>
<td>8.479</td>
<td>223.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elements</th>
<th>WA</th>
<th>SA</th>
<th>RC</th>
<th>HL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>234.2</td>
<td>488.6</td>
<td>0.4194</td>
<td>5.353</td>
</tr>
<tr>
<td>Mean</td>
<td>350.7</td>
<td>659.4</td>
<td>0.7954</td>
<td>29.575</td>
</tr>
<tr>
<td>Max.</td>
<td>459.7</td>
<td>825.0</td>
<td>1.1960</td>
<td>65.034</td>
</tr>
</tbody>
</table>

Note: glazing area distribution (GLAD), glazing area (GLA), orientation (O), overall height (OH), roof area (RA), wall area (WA), surface area (SA), relative compactness (RC), heating load (HL).
3. Methods

3.1. Particle Swarm Optimization (PSO) Algorithm

PSO is a swarm algorithm inspired by the behavior of the particles/social animals, such as fish, or birds. It was introduced and developed by Eberhart, Kennedy [36] and classified as one of the metaheuristic techniques. It was considered as an evolutionary computation technique in the statistical community with many advantages [29,37–39]. This method attempts to take a strong point of the information-sharing procedure from the cluster that affects the overall swarm behavior. Thus, PSO works with the potential solution of a population rather than a single separate item. The best solution is found out based on the experiences of all individuals in the swarm during searching. The PSO algorithm implements six steps for optimal searching as the following pseudo-code [40]:
Algorithm: The particle swarm optimization (PSO) pseudo-code for the optimization process

1. for each particle $i$
2. for each dimension $d$
3. Initialize position $x_{id}$ randomly within permissible range
4. Initialize velocity $v_{id}$ randomly within permissible range
5. end for
6. end for
7. Iteration $k = 1$
8. do
9. for each particle $i$
10. Calculate fitness value
11. if the fitness value is better than $p_{best_id}$ in history
12. Set current fitness value as the $p_{best_id}$
13. end if
14. end for
15. Choose the particle having the best fitness value as the $g_{best_id}$
16. for each particle $i$
17. for each dimension $d$
18. Calculate velocity according to the following equation
19. $v_{i+1}^{j} = w v_{i}^{j} + (c_1 r_1 (local\ best_{j} - x_{i}^{j}) + (c_2 r_2 \times (global\ best_{j} - x_{i}^{j})), v_{min} \leq v_{i}^{j} \leq v_{max}$
20. Update particle position according to the following equation
21. $x_{i+1}^{j} = x_{i}^{j} + v_{i+1}^{j}, j = 1, 2, \ldots, n$
22. end for
23. end for
24. $k = k+1$
25. while maximum iterations or minimum error criteria are not attained

3.2. Genetic Algorithm (GA)

Genetic algorithm (GA) is an optimization algorithm based on Darwin’s theory of natural selection to find the optimal values of a function [41,42]. GA represents one branch of evolutionary computation [43]. It applies the principles: genetics, mutation, natural selection, and crossover. A set of initial candidates is created, and their corresponding fitness values are calculated [44–46]. In GA, many processes are random, like in evolution. However, this optimization technique allows setting random levels and levels of control. In this way, GA is considered as a robust and comprehensive search algorithm. The executable GA may be specified as following (Figure 3):

- Population origination: randomly generates a population of $n$ individuals.
- Calculate the adaptive values: Estimating the adaptation of each individual.
- Stop condition: check the state to finish the algorithm.
- Selection: select two parents from the old population according to their adaptation (the higher the individual is, the more likely they are to be selected).
- Crossover: with each probability selected, a crossover between two parents is made to create a new individual.
- Mutation: for each potential variation selected, new individuals are formed.
- Select the result: if the stopping condition is satisfied, the algorithm ends, and the best solution is found in the current population. When the stopping conditions are not met, the new society will be continually created by repeating three steps: selection, crossover, and mutation.
GA has two necessary stopping conditions:

1. Based on the chromosome structure, controlling the number of genes that are converging, if the number of genes is united at a point or beyond that point, the algorithm ends.
2. Based on the special meaning of the chromosome, examine the change of the algorithm after each generation. If the difference is less than a constant, then the algorithm ends.

Figure 3. Flow chart of a genetic algorithm (GA).

3.3. Imperialist Competitive Algorithm (ICA)

Inspired by the simulation of a computer of human social evolution, the ICA was proposed by Atashpaz-Gargari, Lucas [47] to solve optimization problems. It is one of the swarm intelligence techniques that can effectively solve continuous functions [48–50]. Briefly, ICA is a global search algorithm inspired by imperialistic competition and based on a social policy of imperialism. Accordingly, the most potent empire will dominate many colonies and their sources of use. If an empire collapses, other realms will compete for the territory. The core of the ICA can be described by the following steps:

1. Create random search spaces and initial empires;
2. Assimilation of colonies: the colonies moved in different directions to the realms;
3. Revolution: random changes occur in the characteristics of each country;
4. Exchange the position of the territory for the empire. A colony with a better place than the realm will have the opportunity to rise and control the empire, replacing the existing empire;
5. Imperial competition: competition and conquest occurs among the empires to possess each other’s colonies;
6. Eliminate weaker empires. Natural selection rules are applied. Weak empires will collapse and lose the entire colonies;
7. If the stop condition is satisfied, stop, otherwise return to step 2;
8. End.

3.4. Artificial Bee Colony (ABC)

Optimization algorithms are one of the branches of AI which have been researched and developed based on nature’s inspiration, and swarm intelligence is one of them. Inspired by the bees’ search
for food, Karaboga [51] introduced the ABC optimization algorithm as a robust tool for optimization problems. Although it is pure swarm intelligence, valid for both discrete optimization problems and continuous are significant [52–54]. In the ABC algorithm, the bees are divided into three groups in the population, including employed bees, onlookers, and scouts. Employed bees get food from the found food sources and send information to the onlooker bees. The onlooker bees get information from the employed bees and make choices for better food sources. When the source of the food is exhausted by the employed bees, the onlooker bees will become scouting bees looking for random food sources. The framework of ABC optimization is shown in Figure 4.

For initialization of the swarm, each food source \( x_i \) is a D-dimensional vector with D is the number of variables; \( i = 1, 2, \ldots N \). It can be created using the uniform distribution in Equation (1):

\[
x_{i,j} = x_{\text{min}j} + \text{rand}[0,1](x_{\text{max}j} - x_{\text{min}j})
\]  

(1)

where \( \text{rand}[0,1] \) is a uniformly distributed random number in the range [0,1]; \( x_{\text{min}j} \) and \( x_{\text{max}j} \) are the bounds of \( x_i \) in \( j^{th} \) dimension. After initialization of the swarm, ABC performed cycles of three phases, including employed, onlooker bees, and scouts.

For the employed bees phase, the position of the \( i^{th} \) food source is updated as follows:

\[
v_{ij} = x_{i,j} + \rho_{i,j}(x_{i,j} - x_{t,j})
\]  

(2)

where \( t \in \{1, 2, \ldots N\} \) and \( t \neq i; j \in \{1, 2 \ldots D\} \); \( \rho_{i,j} \) lies in the range \([-1,1]\).

For the onlooker bees phase, the food source can be chosen depending on the probability value associated, i.e., \( p_i \), can be computed by the following equation:

\[
p_{i,j} = \frac{\text{fit}_i}{\sum_{n=1}^{N} \text{fit}_n}
\]  

(3)

where \( \text{fit}_i \) is the solution fitness value \( i^{th} \) evaluated by employed bees. Based on the probability, the onlooker bees select a better position for the food source.

In the scouting phase, the feed will be dropped if no location is updated according to Equation (2) in a predetermined cycle. Now, the onlooker will become a scout. A scout will perform a search for new food sources randomly in the search space, as described in Equation (1). In ABC, the number of cycles a food source is then dropped is called limit. It is an important parameter used to assess the quality of the model.
3.5. Artificial Neural Network (ANN)

Based on the human brain operation principle, ANN has been researched and developed as an alternative tool for different social purposes. It is even smarter than the human in some cases, with substantial computing power. In real-life, ANN was studied and applied to solve many problems, such as prediction of self-compacting concrete strength [55], anisotropic masonry failure criterion [56], prediction of the mechanical properties of sandcrete materials [57], blasting issues [58–64], landslide assessment [65–67], to name a few [68–75]. They operate based on data analysis from input neurons, where the input data of the dataset is contained. Here, the information is analyzed and transmitted through hidden layers containing hidden neurons, via the transfer function. In the hidden layers, data is encrypted, analyzed, and calculated through weights. The biases are also estimated to ensure a balanced level of data. Finally, the outcome is computed on the output layer. Figure 5 illustrates the framework of ANN model for predicting the HL of EEB systems in this study based on the eight input variables and one output variable.
4. Evaluation Performance Indices

To evaluate the quality of the PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN models, $R^2$, RMSE, and MAE, were used as the indicators of the model’s performances. They were computed as Equations (4–6):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$  \hspace{1cm} (4)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2}$$  \hspace{1cm} (5)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$  \hspace{1cm} (6)

$n$ stands for the number of instances; $\overline{y}$, $y_i$, and $\hat{y}_i$ are considered as average, calculated, and modeled amounts of the response variable.

5. Prediction of Heating Load (HL) by the Genetic Algorithm-Artificial Neural Network (GA-ANN) Model

Before predicting the HL of EEB systems by the stated models, the dataset was split into two clusters, i.e., training and testing. According to the previous studies, the original dataset should be divided into two parts by randomly according to the 80/20 ratio [76,77]. Thus, for the training process, 80% of the whole dataset (672 experimental datasets) was selected randomly to develop the models. The remaining 20% (165 experimental datasets) was used for the testing process, which is the method for evaluating the quality/performance of the GA-ANN, PSO-ANN, ICA-ANN, and ABC-ANN models.

For the prediction of HL of EEB systems by the GA-ANN model, an initialization ANN model was developed first; then, the GA was used to optimize the developed ANN model, where the weights and biases were optimized. According to Nguyen et al. [68], one or two hidden layers of the ANN model
can implement very well all regression problems. Therefore, a “trial and error” (TAE) procedure was conducted with one and two hidden layers of ANN models. To avoid overfitting of the initial ANN model, the min-max scale method was applied with the scale lies in the range of $[-1,1]$. Ultimately, the ANN model 8-24-18-1 was defined as the best ANN technique for predicting HL of EEB systems in this study. This was the moment for the optimization of the weights and biases of the ANN 8-24-18-1 model by the GA. The number of populations ($p$), crossover probability ($P_c$), mutation probability ($P_m$), and the number variable ($n$), are the parameters of the GA, that needed to be set up before optimizing herein. In this study, the TAE procedure of $p$ with different values was conducted, i.e., $p = 100, 200, 300, 400, 500$; $P_m$ was set equal to 0.1; $P_c$ was set equal to 0.9; $n = 4$. To evaluate the performance of the optimization process, RMSE was used as the fitness function according to Equation (4). The searching operation were performed in 1000 iterations to ensure the optimal searching for the weights and biases of the selected ANN model. The optimal values of weight and bias for the ANN 8-24-18-1 model after optimizing by the GA (i.e., GA-ANN model), were corresponding to the lowest RMSE. The performance of the optimization process by the GA for the ANN 8-24-18-1 model is shown in Figure 6. The final ANN model, after optimized by the GA (i.e., GA-ANN model), is shown in Figure 7.

![Figure 6. Genetic algorithm-artificial neural network (GA-ANN) performance for estimating HL of EEB systems.](image-url)
As stated above, 672 experimental datasets were investigated and analyzed to develop the models. The back-propagation algorithm was applied to training the GA-ANN model. Note that the min-max scale with the range \([-1,1]\) was used for all the models to avoid underfitting/overfitting. The performance of the training process for predicting HL of EEB systems is interpreted in Figure 8. Subsequently, 165 experimental datasets were used to evaluate GA-ANN performance as the new dataset. The results of HL prediction on the new data (i.e., 165 experimental datasets) were estimated by the developed GA-ANN model and are shown in Figure 9.
6. Prediction of HL by the Particle Swarm Optimization (PSO)-ANN Model

Like the GA-ANN model, the selected initialization ANN model was optimized by the PSO algorithm for predicting HL of EEB systems, called PSO-ANN model. In this regard, the parameters of the PSO algorithm were set up before optimization of the ANN model (i.e., ANN 8-24-18-1 model), including the number of particle swarms ($Sw$), maximum particle’s velocity ($V_{max}$), individual cognitive ($\phi_1$), group cognitive($\phi_2$), inertia weight ($w$), and maximum number of iteration ($m_i$). Then, the weights and biases of the initialization ANN model were optimized by the PSO algorithms, as those applied for the GA-ANN model above. Similar to the GA-ANN model, a TAE procedure of $Sw$ was implemented, with $Sw$ of 100, 200, 300, 400, 500, respectively; $V_{max} = 1.8$; $\phi_1 = \phi_2 = 1.7$; $w = 1.8$, and $m_i = 1000$. The similar techniques as those used for the GA-ANN model were also applied for the PSO-ANN model in developing the model (i.e., back-propagation algorithm, min-max scale $[-1,1]$). Finally, the best PSO-ANN model was determined with the lowest RMSE. Figure 10 shows the performance of the PSO-ANN model in the training process. Figure 11 illustrates the structure of the PSO-ANN model. Note that, although the number of input neurons, hidden layers, and neurons, as well as the output layer, is the same as Figure 8; however, the weights and biases of them are different. Eventually, the HL predictions on the training dataset and testing dataset were conducted based on the developed PSO-ANN model, as shown in Figures 12 and 13, respectively.

![Figure 9](image9.png)

**Figure 9.** HL predictions on the testing dataset of the GA-ANN model.

![Figure 10](image10.png)

**Figure 10.** Particle swarm optimization (PSO-ANN) performance for estimating HL of EEB systems in the training process.
Figure 11. Structure of the PSO-ANN model for estimating HL of EEB systems.

Positive weights are represented by black lines, and negative weights by grey lines. The thickness of the lines is proportional to the magnitude of the weights relative to all others. The input variables (GLAD, GLA, O, OH, RA, WA, SA, RC) are presented by I1 to I8, respectively. The number of hidden neurons in each hidden layer are presented by H1 to H24, respectively. The HL is the output parameter, which illustrated by O1. The biases applied as constant values to the nodes are presented by B1, B2, and B3.

Figure 12. HL predictions on the training dataset of the PSO-ANN model.

$R^2 = 0.972$
The best ICA-ANN model is associated with the lowest RMSE. Figure 14 shows the performance of the models were also applied for the ICA-ANN model in developing the model (i.e., back-propagation algorithm, min-max scale \([-10,10]\)). Additionally, the similar techniques as those used for the GA-ANN and PSO-ANN models were also applied for the ICA-ANN model in developing the model (i.e., back-propagation algorithm, min-max scale \([-1,1]\)).

Figure 13. HL predictions on the testing dataset of the PSO-ANN model.

7. Prediction of HL by the Imperialist Competitive Algorithm (ICA)-ANN Model

In this section, the HL of EEB systems was predicted by the ICA-ANN model. As those applied for the GA-ANN and PSO-ANN model, the ICA was used to optimize the weights and biases of the selected initialization ANN model (i.e., ANN 8-24-18-1 model). The parameters of ICA are also needed to be set up before optimization of the ANN model, including the number of initial countries \((N_{\text{country}})\), initial imperialists \((N_{\text{imper}})\), maximum number of iterations \((N_i)\), lower-upper limit of the optimization region \((L)\), assimilation coefficient \((A_s)\), and revolution of each country \((r)\). For implementing this task, a TAE procedure was also applied for \(N_{\text{country}}\), with \(N_{\text{country}}\) set equal to 100, 200, 300, 400, 500, respectively; \(N_{\text{imper}}\) was set equal to 10, 20, 30, respectively; \(L\) was set in the rage of \([-10,10]\); \(A_s\) equal to 3; \(r\) as to 0.5, and \(N_i\) was set equal to 1000. Afterward, the emperies perform a global search for the colonies (e.g., weights and biases). The fitness of the emperies was assessed through RMSE. The best ICA-ANN model is associated with the lowest RMSE. Figure 14 shows the performance of the optimization process by the ICA for the ANN model. Ultimately, the final ICA-ANN model was found, as shown in Figure 15. Note that the structure of the developed ICA-ANN model is the same with the GA-ANN and PSO-ANN models; however, the weights and biases (e.g., black and grey lines) of them are different. Additionally, the similar techniques as those used for the GA-ANN and PSO-ANN models were also applied for the ICA-ANN model in developing the model (i.e., back-propagation algorithm, min-max scale \([-1,1]\)).

Figure 14. Imperialist competitive algorithm (ICA)-ANN performance for estimating HL of EEB systems in the training process.
Figure 15. Structure of the ICA-ANN model for estimating HL of EEB systems.

Based on the ICA-ANN model developed, the outcome of HL predictions was performed. Figure 16 shows the HL predictions of the training dataset when the development of the ICA-ANN model. Applying the ICA-ANN model developed, the new dataset includes 165 experimental datasets on the testing dataset was used to check the quality of the model, like those tested for the GA-ANN and PSO-ANN models. The results of the HL predictions on the new dataset (testing dataset) are shown in Figure 17.

Figure 16. HL predictions on the training dataset of the ICA-ANN model.
For the HL predictions by the ABC-ANN model, a process of the development of the hybrid model was conducted, similar to those models above (e.g., ICA-ANN, PSO-ANN, GA-ANN). Accordingly, the ABC algorithm was applied to optimize the parameters of the selected ANN model (i.e., ANN 8-24-18-1 model), for predicting HL of EEB systems. The initial setting for the ABC algorithm is necessary, as with those set for the previous models (e.g., ICA-ANN, PSO-ANN, GA-ANN), including the number of bees ($N_{\text{bees}}$), the number of food sources ($N_{\text{foodsource}}$), the limit of a food source ($M_{\text{foodsource}}$), the boundary of the parameters ($b$), and the maximum number of repetitions for optimization ($n_{\text{round}}$). Similar to the GA, PSO, and ICA, a TAE procedure for in the ABC algorithm was conducted, with $N_{\text{bees}} = 100, 200, 300, 400, 500$, respectively. The other parameters of the ABC algorithm were set as follow: $N_{\text{foodsource}} = 50; M_{\text{foodsource}} = 100; b = [-10;10]$, and $n_{\text{round}} = 1000$. Once the parameters of the ABC algorithms were established, the initialization ANN model 8–24–18–1 model was optimized by the global search of the bee colony. RMSE was also used to evaluate the efficiency of the optimization of the ABC-ANN model, with the optimal ABC-ANN model corresponding to the lowest RMSE. Figure 18 presents the performance of the optimization process of the ABC-ANN model in estimating the HL of EEB systems. Finally, the optimal ABC-ANN model was defined with the optimal weights and biases, as shown in Figure 19.
Figure 19. Structure of the ABC-ANN model for estimating HL of EEB systems.

It should be noted that although Figures 7, 11, 15 and 19 are the same, their structure is different since the weights and biases of the models are different. In addition, similar techniques as those used for the ICA-ANN, PSO-ANN, and GA-ANN models were also applied for the development of the ABC-ANN model (i.e., back-propagation algorithm, min-max scale $[-1,1]$). Figure 20 shows the HL predictions of the ABC-ANN model on the training dataset. Then, 165 experimental datasets were predicted based on the developed ABC-ANN models, as shown in Figure 21.

Figure 20. HL predictions on the training dataset of the ABC-ANN model.
9. Comparison and Evaluation of the Developed Models

After the models were developed and HL of EEB systems was predicted, their results were compared and evaluated together through the performance metrics (e.g., RMSE, $R^2$, and MAE), and the intensity of color and ranking methods. A comprehensive assessment of the developed models based on both training and the testing dataset was conducted in this section. Table 2 presents the prediction results of HL by the hybrid intelligent techniques (i.e., GA-ANN, ABC-ANN, PSO-ANN, and ICA-ANN), and their performance in the training process.

Table 2. Prediction results of the hybrid models and their performance (for the training process).

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>$R^2$</th>
<th>MAE</th>
<th>Rank for RMSE</th>
<th>Rank for $R^2$</th>
<th>Rank for MAE</th>
<th>Total Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA-ANN</td>
<td>1.701</td>
<td>0.972</td>
<td>0.784</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>PSO-ANN</td>
<td>1.822</td>
<td>0.972</td>
<td>0.872</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>ICA-ANN</td>
<td>1.847</td>
<td>0.971</td>
<td>0.860</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>ABC-ANN</td>
<td>1.833</td>
<td>0.972</td>
<td>0.813</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

From Table 2, the color intensity revealed that the GA-ANN model provided the most dominant performance in the training process. It obtained the lowest error with an RMSE of 1.701, $R^2$ of 0.972, and MAE of 0.784, and the total ranking of 10, on the training dataset. The ABC and PSO meta-heuristics algorithms yielded lower performance in the optimization of the ANN model in the training process, with RMSE of 1.833, 1.822, $R^2$ of 0.927, 0.972; MAE of 0.813, 0.872, and the total ranking of 7, and 6, respectively. The weakest model in this optimization process is the ICA-ANN model with an RMSE of 1.847, $R^2$ of 0.971, MAE of 0.860, and the total ranking of 4. To have a complete conclusion, the models’ performances were assessed on the testing dataset, where the dataset was considered as the new data and never used in the training process. Table 3 shows the results and the performance of the models in the testing process.

Table 3. Prediction results of the hybrid models and their performance (for the testing process).

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>$R^2$</th>
<th>MAE</th>
<th>Rank for RMSE</th>
<th>Rank for $R^2$</th>
<th>Rank for MAE</th>
<th>Total Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA-ANN</td>
<td>1.625</td>
<td>0.980</td>
<td>0.798</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>PSO-ANN</td>
<td>1.932</td>
<td>0.972</td>
<td>1.027</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>ICA-ANN</td>
<td>1.982</td>
<td>0.970</td>
<td>0.980</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>ABC-ANN</td>
<td>1.878</td>
<td>0.973</td>
<td>0.957</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>

Based on the reports of Table 3, similar results to the training process were reflected. The color intensity of the red color indicated that the GA-ANN model was the best model in a comparison of
the other models. The corresponding performance values of the GA-ANN model also found with an RMSE of 1.625, $R^2$ of 0.980, MAE of 0.798, and the total ranking of 12. Whereas, the ABC-ANN, PSO-ANN, and ICA-ANN model proved lower performances, as like the training process, with RMSE of 1.878, 1.932, 1.982; $R^2$ of 0.973, 0.972, 0.970; MAE of 0.957, 1.027, 0.980; and the total ranking of 9, 5, 4, respectively.

10. Sensitivity Analysis

To get an overall conclusion and optimization solutions in building design aim to use energy-efficiency, the importance level of the input variables for predicting HL in the present work was conducted. The initial ANN model (i.e., ANN 8-24-18-1) was investigated using the Olden method [78] to analyze the importance of the input variables. This method enables the analysis of the importance of input variables for hidden multiple-layer ANN models [79]. Ultimately, the importance level of the input variables for predicting HL of EEB systems was determined, as shown in Figure 22. Based on the sensitivity analysis results of this study, it can be seen that GAD, SA, GA, RA, OH, and WA, were the most important variables in predicting the HL of EEB systems, especially SA and GA.

![Importance level of the input variables for predicting the HL of EEB systems.](image)

**Figure 22.** Importance level of the input variables for predicting the HL of EEB systems.

11. Conclusions

Energy efficiency is one of the essential requirements for smart cities. Artificial intelligence has also been considered as powerful support tools for these objectives in smart cities. This study developed and proposed four new hybrid models based on AI techniques for estimating the HL of EEB systems with high reliability, i.e., GA-ANN, PSO-ANN, ICA-ANN, and ABC-ANN models. A comprehensive comparison and assessment of the developed models were performed in this work. As a conclusion, the meta-heuristics algorithms performed very well in the optimization of the ANN model. Of the meta-heuristics algorithms used in this study, the GA provided the highest performance in optimizing the ANN model, to predict the HL of EEB systems, i.e., GA-ANN model. The remaining meta-heuristics algorithms (i.e., PSO, ICA, ABC) provided more unsatisfactory performance, corresponding to the performance of the PSO-ANN, ICA-ANN, and ABC-ANN models.

Based on the results of this study, the HL of EEB can be accurately predicted and controlled to ensure the energy efficiency of buildings in smart cities. Software or applications on computers and smartphones can be developed in the future based on the results of this study for the use of energy.
saving and efficiency of buildings in smart cities. Besides, it can also be integrated into smart houses to adjust and control the HL of the houses automatically. Furthermore, optimization techniques of building design, as well as smart city planning, can also be conducted based on the models developed in this study. Notably, the GAD, SA, GA, RA, OH, and WA are the main parameters which should be carefully concerned and calculated in designing buildings and smart cities. Based on the results of this study as well as software or applications on smartphones and computers, engineers can optimize the building parameters to use HL in smart cities effectively.


Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Thanh Hoa University of Culture, Sports and Tourism, Thanh Hoa City, Vietnam, for supporting this study.

Conflicts of Interest: The authors declare no conflict of interest.

References
3. Caragliu, A.; Del Bo, C.; Nijkamp, P. Smart cities in Europe. J. Urban Technol. 2011, 18, 65–82. [CrossRef]
10. Catalina, T.; Iordache, V.; Caracaleanu, B. Multiple regression model for fast prediction of the heating energy demand. Energy Build. 2013, 57, 302–312. [CrossRef]


57. Asteris, P.; Roussis, P.; Douvika, M. Feed-forward neural network prediction of the mechanical properties of sandcrete materials. Sensors 2017, 17, 1344. [CrossRef] [PubMed]


